

Tecnología y ciencia aplicada

TYCA

ENERO - JUNIO 2026

VOLUMEN 9 / NUMERO 1

ISSN: 2992-8338



Monitoreo de anemia

mediante POCT
microrreológico:
un gran avance.

 EDUCACIÓN
SECRETARÍA DE EDUCACIÓN PÚBLICA

 TECNOLÓGICO
NACIONAL DE MÉXICO

cenidet[®]
Centro Nacional de Investigación
y Desarrollo Tecnológico

TECNOLOGÍA Y CIENCIA APLICADA. Volumen 9, Núm. 1, Enero - Junio 2026

Tecnológico Nacional de México/Centro Nacional de Investigación y Desarrollo Tecnológico
Interior Internado Palmira s/n, C.P. 62490,
Cuernavaca, Morelos, México. <https://tycap.cenidet.tecnm.mx>, <http://cenidet.tecnm.mx/>

TECNOLOGÍA Y CIENCIA APLICADA, es una publicación semestral editada por el Tecnológico Nacional de México, Avenida Universidad 1200, Alcaldía Benito Juárez, C.P. 03330, Teléfono 5536002511 Ext. 65092, correo d_vinculacion05@tecnm.mx, a través del Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), Interior Internado Palmira s/n, Col. Palmira, C.P. 62490, Cuernavaca, Morelos, México, Contacto: +(52)7773627770, vin_cenidet@tecnm.mx, www.cenidet.tecnm.mx. Editores Responsables: Dr. Ricardo Fabricio Escobar Jiménez y Dr. Carlos Daniel García Beltrán. Reserva de Derechos al Uso Exclusivo: No. 04-2023-0 dor del Centro Nacional de Investigación y Desarrollo Tecnológico, Dr. Carlos Daniel García Beltrán, Interior Internado Palmira s/n, Colonia Palmira, C.P. 62490, Cuernavaca, Morelos, México. Fecha de última modificación, 26 de junio de 2026. Queda prohibida la reproducción parcial o total del contenido e imágenes de esta publicación en cualquier medio, sin previa autorización del Centro Nacional de Investigación y Desarrollo Tecnológico

EDITORES RESPONSABLES.

Dr. Ricardo Fabricio Escobar Jiménez

Editor en Jefe

Dr. Carlos Daniel García Beltrán

Editor ejecutivo

Mtro. Ramón Jiménez López

Director General del Tecnológico Nacional de México

Lic. Lorenzo O. Hernández

Director del TecNM/Centro Nacional de Investigación y Desarrollo Tecnológico

Dr. Carlos Manuel Astorga Zaragoza

Subdirector Académico

C.P. Erika Yaneli Bahena Flores

Subdirección de Planeación y Vinculación

Dra. Gloria Lilia Osorio Gordillo

Subdirección de Servicios Administrativos

Dr. Noé Alejandro Castro Sánchez

Departamento de Ciencias Computacionales

Dr. Jarniel García Morales

Departamento de Ingeniería Electrónica

Dr. Andrés Blanco Ortega

Departamento de Ingeniería Mecánica

Dr. Jesús Arce Landa

Coordinación de Ingenierías

Dr. Erik Rosado Tamariz

Coordinación de Mecatrónica

Dra. Rodolfo Amalio Vargaz Méndez

Departamento de Desarrollo Académico e Idiomas

Mtro. Emanuel Olmos Barba

Departamento de Organización y Seguimiento de Estudios

Mtra. Cecilia López Romero

Oficina de Centro de Cómputo y Telecomunicaciones

M.A. Daniela Hernández Castañeda

Centro de Información

Mtra. Maritza De La Sancha San Román

Departamento de Planeación

PROGRAMACIÓN Y PRESUPUESTACIÓN

Dr. Carlos Daniel García Beltrán

Departamento de Gestión Tecnológica y Vinculación

Dr. Vitervo López Caballero

Departamento de Comunicación y Eventos

Mtra. Verónica Sotelo Boyás

Departamento de Servicios Escolares

Ing. Christopher Garduño Corona

Departamento de Recursos Materiales y Servicios M.A.

Silvia Muñoz Moreno

Departamento de Recursos Humanos

Mtra. Honoria Martínez Pérez

Departamento Recursos Financieros.

CONSEJO EDITORIAL

Dr. Ricardo Fabricio Escobar Jiménez - **Editor en Jefe**

Dr. Carlos Daniel García Beltrán - **Editor Ejecutivo**

Dr. Arturo Abúndez Pliego - **Editor Asociado**

Dr. Andrés Blanco Ortega - **Editor Asociado**

Dr. Erik Rosado Tamariz - **Editor Asociado**

Dr. Miguel Ángel Chagolla Aranda - **Editor Asociado**

Dr. Gabriel González Serna - **Editor Asociado**

Dr. Dante Mújica Vargas - **Editor Asociado**

Dra. Gloria Lilia Osorio Gordillo - **Editora Asociada**

Mtra. Nayeli Hernández Barba - **Editora Asociada**

Dr. Rodolfo Vargas Méndez - **Editor Asociado**

Dra. Marisol Cervantes Bobadilla - **Editora Asociada.**

DISEÑO Y EDICIÓN

Juliana Cantún Cárdenas

ENCARGADO DE LA PÁGINA WEB

Dr. Víctor Josué Ruiz Martínez

COMITÉ DE ARBITRAJE

Oficina de Centro de Cómputo y Telecomunicaciones M.A. Daniela Hernández Castañeda Centro de Información Mtra. Maritza De La Sancha San Román Departamento de Planeación Programación y Presupuestación Dr. Carlos Daniel García Beltrán Departamento de Gestión Tecnológica y Vinculación Dr. Vitervo López Caballero Departamento de Comunicación y Eventos Mtra. Verónica Sotelo Boyás Departamento de Servicios Escolares Ing. Christopher Garduño Corona Departamentode Recursos Materiales y Servicios M.A. Silvia Muñoz Moreno Departamento deRecursos Humanos Mtra. Honoria Martínez Pérez Departamento Recursos Financieros. Consejo Editorial Dr. Ricardo Fabricio Escobar Jiménez -Editor en Jefe Dr. Carlos DanielGarcía Beltrán -Editor Ejecutivo Dr. Arturo Abúndez Pliego -Editor Asociado Dr. Andrés Blanco Ortega -Editor Asociado Dr. Erik Rosado Tamariz -Editor Asociado Dr. Miguel Ángel Chagolla Aranda -Editor Asociado Dr. Gabriel González Serna -Editor Asociado Dr. Dante Mújica Vargas -Editor Asociado Dra. Gloria Lilia Osorio Gordillo -Editora Asociada Mtra. Nayeli Hernández Barba -Editora Asociada Dr. Rodolfo Vargas Méndez -Editor Asociado Dra. Marisol Cervantes Bobadilla -Editora Asociada. Diseño y edición: D. G. Luis Fernando García B. Miguel Alvarado Flores, BUAP; Hugo Arnoldo Mitre Hernández, CIMAT; Mariana Gisela Peña Juárez, TecNM/CRODE_O; Héctor Miguel Buenabad Arias, CIICAp-UAEM; Susana Silva Martínez, CIICAp-UAEM; José Alfredo Hernández Pérez, CIICAp-UAEM; Gabriela Hernández Luna, CIICAp-UAEM; Víctor Manuel Zezatti Flore, CIICAp-UAEM; Mario Andrés Aguilar Orduña, CINVESTAV; José Hugo Arellano, ENES Juriquilla; Jesús Emmanuel Solís Pérez, ENES Juriquilla; Katya Rodríguez Vázquez, IIMAS, UNAM; Ana Odilia Martínez Hernández, IMSS; Sandra Eloisarangel Estrada, INIFAP; Lucero De Montserrat Ortiz Aguilar, TecNM/ITSPR; Jorge Misael Ruiz Martínez, TecNM/ITAO; Julio César González Juárez, TecNM/ITMor; Elieel Rafael Romero García, TecNM/ITMor; Héctor Ramón Azcaray Rivera, TecNM/ITOX; Giner Alor Hernández, TecNM/ITZ; Fredy Juárez Trujillo, TecNM/ITSA; Eligio Flores Rodríguez, TecNM/ITSI; Wendy Guadalupe Báez García, TecNM/ITSM; Javier Uriarte Flores, TecNM/ITST; Alfredo De La Barrera González, IPN; Edgar Alfredo Portilla Flores, IPN; Paola Andrea Niño Suárez, IPN; Jorge Bedolla Hernández, TecNM/ITA; Isay Obet Huicochea Ríos, TecNM/CENIDET; Máximo López Sánchez, TecNM/CENIDET; Abisai Acevedo Quiroz, TecNM/CENIDET; Alfredo Terrazas Porcayo, TecNM/CENIDET; Alicia Martínez Rebollar, TecNM/CENIDET; Andrea Magadán Salazar, TecNM/CENIDET; Andrés Blanco Ortega, TecNM/CENIDET; Ángel Rodrigo Guadarrama Estrada, TecNM/CENIDET; Antonio Juan Capistran Abundez, TecNM/CENIDET; Aristides Caballero Alfaro, TecNM/CENIDET; Arturo Abúndez Pliego, TecNM/CENIDET; Arturo Daniel Espinosa Pedroza, TecNM/CENIDET; Blanca Dina Valenzuela Robles, TecNM/CENIDET; Blanca Lucero Xopo Rodríguez, TecNM/CENIDET; Carlos Daniel García Beltrán, TecNM/CENIDET; Carlos Manuel Astorga Zaragoza, TecNM/CENIDET; Daniela Hernández Castañeda, TecNM/CENIDET; Dante Mujica Vargas, TecNM/CENIDET; Eddie Helbert Clemente Torres, TecNM/CENIDET; Efraín Simá Moo, TecNM/CENIDET; Eladio Martínez, TecNM/CENIDET; Elías Ramos, TecNM/CENIDET; Emmanuel Olmos Barba, TecNM/CENIDET; Enrique Alcudia Zacarías, TecNM/CENIDET; Erik Rosado Tamariz, TecNM/CENIDET; Ezequiel Domínguez Lagunas, TecNM/CENIDET; Fredi Alegría, TecNM/CENIDET; Gerardo Vicente Guerrero Ramírez, TecNM/CENIDET; Gloria Lilia Osorio Gordillo, TecNM/CENIDET; Iván Antonio Juárez Trujillo, TecNM/CENIDET; Jarniel García Morales, TecNM/CENIDET; Javier Ortiz Hernández, TecNM/CENIDET; Jesús Darío Mina Antonio, TecNM/CENIDET; Jorge Colín Ocampo, TecNM/CENIDET; Jorge Alberto Fuentes Pacheco, TecNM/CENIDET; José Luis Rodríguez Pliego, TecNM/CENIDET; Juan Carlos Rojas Pérez, TecNM/CENIDET; Juan Gabriel González-Serna, TecNM/CENIDET; Karime Baltazar Alcaraz, TecNM/CENIDET; Ma. Del Carmen Toledo Pérez, TecNM/CENIDET; Ma. Guadalupe López López, TecNM/CENIDET; Manuel Adam Medina, TecNM/CENIDET; María Fernanda De Jesús Ramírez, TecNM/CENIDET; Mario Ponce Silva, TecNM/CENIDET; Marisol Cervantes Bobadilla, TecNM/CENIDET; Marisol Coba Martínez, TecNM/CENIDET; Miguel Chagolla Aranda, TecNM/CENIDET; Mónica Borunda Pacheco, TecNM/CENIDET; Nayeli Del Milagro Hernández Barba, TecNM/CENIDET; Noé Alejandro Castro Sánchez, TecNM/CENIDET; Rafael Campos Amezcua, TecNM/CENIDET; Raúl Pinto Elías, TecNM/CENIDET; René Santaolaya Salgado, TecNM/CENIDET; Reynaldo Ángel Serrano Cruz, TecNM/CENIDET; Ricardo Fabricio Escobar Jiménez, TecNM/CENIDET; Rodolfo Amalio Vargas Méndez, TecNM/CENIDET; Roger Alejandro Cundapí Arroyo, TecNM/CENIDET; Samanta López Salazar, TecNM/CENIDET; Sara Lilia Moya Acosta, TecNM/CENIDET; Semadar Bojórquez Sánchez, TecNM/CENIDET; Vicente Borja Jaimes, TecNM/CENIDET; Víctor Hugo Olivares Peregrino, TecNM/CENIDET; Víctor Josué Ruiz Martínez, TecNM/CENIDET; Víctor Manuel Alvarado Martínez, TecNM/CENIDET; Víctor Miguel Salinas Cortés, TecNM/CENIDET; Vitervo López Caballero, TecNM/CENIDET; Yasmín Hernández, TecNM/CENIDET; Yvonne Chávez Chena, TecNM/CENIDET; Arturo Alvarado Lassmann, TecNM/CRODE_O; Melina Velasco Plascencia, TecNM/ITM; Cecilia Irene Loeza-Mejía, TecNM/ITM; Edgar Omar Bañuelos Lozoya, TecNM/ITP; Cruz Ernesto Aguilar Rodríguez, TecNM/ITSLR; Demetrio Pérez Vigueras, TecNM/ITSOEH; Luis Enrique Calatayud Velázquez, TecNM/ITT; Macdiel Emilio Acevedo Quiroz, TecNM/ITZ; Areli Marlen Salgado Delgado, TecNM/ITZ; Arturo Javier Martínez Mata, TecNM/ITZ; Enrique De Jesús Moreno Carpintero, TecNM/ITZ; Erika Darnely Rojas Ayala, TecNM/ITZ; Iris Lizeth Hernández Solís, TecNM/ITZ; Wendy Netzy Hernández Díaz, TecNM/ITZ; Zulay Vargas Galarza, TecNM/ITZ; Israel Isaac Zetina Ríos, TecNM/SMA; Jorge Edmundo Mastache Mastache, TecNM/TESJ; Luis Felipe Rivas Mendoza, TecNM/ITSTII; Julia Patricia Melo Morin, TecNM/ITSP; Noel García Díaz, TecNM/ITC; Rocío Ramos Hernández, TecNM/ITO; Miriam Antonia Chavarría Santibáñez, TecNM/ITZ; Ángel Tlatelpa Becerro, UAEM; Dolores Azucena Salazar Piña, UAEM; Miriam Navarrete Procopio, UAEM; Raquel Sotelo Urueta, UAEM; Alejandra Guadalupe Silva Trujillo, UASLP; Pedro Cruz Alcantar, UASLP; Felipe Noh Pat, UAC; Julia Yazmín Arana Llanes, UAH; María Del Rocío Ochoa Montiel, UAT; Jorge Aurelio Brizuela Mendoza, UdeG; María Susana Ávila García, UdeG; Irving Osiris Hernández López, UdeS; Manuel De Jesús Palacios Gallegos, UPC; María De Los Ángeles Cosío León, UPP; Miguel Ángel Ruiz Jaimes, UPEM; Luis Guillermo Carreto Hernández, UTGZ; Esther Lugo González, UTM; Miguel Alberto Domínguez Gurría, UTM; Roberto Alvarado Juárez, UTCV; César Humberto Guzmán Valdivia, UPA; Cristina Blanco González, UPC; Yadira Toledo Navarro, UPEM; Yessenia Olazo Gómez, USon; Eduardo Javier Cortés Zetina, UTCV; Iván Alcalá Borjas, UTEZ; Camilo Leonardo Sandoval Rodríguez, UTS; Carlos Manuel Lara Barrios; Juan José Acevedo Fernández; Leonor Zavaleta Ave



TECNOLOGÍA Y CIENCIA APLICADA

Mensaje Editorial

La revista digital “Tecnología y ciencia aplicada” del Tecnológico Nacional de México/CENIDET comprometida con la divulgación científica de los conocimientos más actuales generados en diversas áreas científicas y aplicados en beneficio de la sociedad, presenta en el volumen 9, número 1, artículos enfocados a la investigación y desarrollo tecnológico. En la revista se presentan un total de 68 artículos divididos en áreas como Ciencias de la Ingeniería, Ingeniería Electrónica, Ingeniería Mecánica, Ciencias Computacionales, Mecatrónica y otras Ingenierías. En estos trabajos de investigación se proporcionan soluciones innovadoras a problemas enmarcados en los Programas Nacionales Estratégicos. Los artículos fueron evaluados por comités especializados en cada área del conocimiento, con el acompañamiento del comité editorial de la revista, lo que aseguró la calidad de cada uno de ellos.

“La revista digital “Tecnología y ciencia aplicada” mantiene su compromiso en promover que la ciencia sea para todos y un puente de comunicación entre la comunidad científica y la sociedad.”

Dr. Ricardo Fabricio Escobar Jiménez
Editor en jefe

CONTENIDO

Ciencias computacionales



001

MARCO: Methodology of Agentic Reporting and Composition Orchestration for Academic Documents

009

Optimizing Nutrient Solutions Using Swarm Intelligence

017

Theoretical Model for Detecting Chemical Components Using Machine Learning Approach

025

Web system for the process of managing climatological information of the state of Zacatecas

033

Validation of an NLP Model for Thematic Organization of Study Programs: Teacher Diagnosis, Gold Standard and Extrinsic Evaluation



040

Fine Grained Classification of Pollinating Insects with EfficientNet B4 for Monitoring in Urban Gardens

047

Exploring the Evolution of the Mercury Beating Heart with the Use of AI

053

E-learning and the "instant generation" in technological higher education: challenges and opportunities for institutional management toward intelligent social learning models.

061

ROS 2 Localization for Indoor Mobile Robots: A Systematic Review and Comparative Analysis

069

Web platform for human-computer interaction for the detection and vocational guidance referral of applicants to higher education

077

Water Turbidity Prediction Using Mathematical Models and Machine Learning on an Open Dataset

CONTENIDO

Ciencias computacionales



085

Quantum Software Engineering Education: Current Landscape, Challenges, and Emerging Directions

093

Systematic Study of Visual and Interactive Tools and Sequences for Refactoring

101

Edge AI and Human-in-the-Loop Architecture for Automated Voice Presentation Generation

107

EvoNHI: Constraint-Aware Remediation Planning for Non-Human Identity Attack Paths in Kubernetes Using Multi-Objective Evolutionary Optimization

113

Abstract Syntax Tree vs. Generalist Ecosystems: Performance Analysis in Markup Language Parsing

119

Corn Crop Analysis Using Computer Vision For Pest And Disease Detection Via Unmanned Aerial Vehicles

125

SARA Desktop: Governed Memory and Safe Local Actions for a Persistent Desktop Companion

133

Local-First Semantic Mapping with Constrained LLMs: A Concurrent Framework for Multimodal Knowledge Extraction

139

FlxPCG: Procedural Controlled Generation of Levels and Entities Implementation Methodology in 2D Using the HaxeFlixel Framework

147

Automated Conversion of Datasets for Panoptic Segmentation Models

153

Asynchronous Monitoring System for Thermal Anomalies Associated with Breast Tumor Formation



CONTENIDO

Ciencias de la Ingeniería



160

Comparative performance of free water surface and horizontal subsurface flow constructed wetlands planted with native macrophytes under tropical conditions

168

Study of the availability of resources in Mexico for the production of synthetic fuels and analysis of states with the greatest potential

176

Electromyographic Assessment of Quadriceps Femoris Muscles Under Different Exertion Conditions

184

Valorization of rice by-products, a sustainable approach for a circular economy: Literature review

192

Early Forest Fire Detection using the Internet of Things and Artificial Intelligence



202

Artificial Intelligence in Companies: PRISMA Bibliometric Review in Scopus and Web of Science

208

Energy efficiency of power supply cables for 1 MW temporary generators under the NOM-001-SEDE standard and international standards and thermo-electrical optimization under real operating conditions

216

A comparative analysis of PEM and alkaline electrolyzers for H₂ production in areas with high solar

223

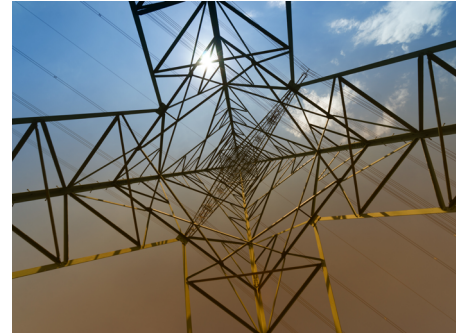
Thermodynamic simulation of sustainable syngas production with Reverse Water-Gas Shift using captured carbon dioxide and green hydrogen

231

The Importance of Applying an Adequate Distribution in Wind Speeds

CONTENIDO

Ingeniería Electrónica



239

A Review of Electrical Stimulation Therapy (EST) as an Adjunct to Myofascial Release Therapy (MFRT) for Posterior Lower-Leg Muscle Pain.

247

Dynamic Analysis of a High-Order Buck-Boost Converter with Reduced Power Processing

254

Observer-Based Fault Diagnosis From Sampled Data

260

Design of a Buck-Boost Converter Based on the Concept of Reduced Redundant Power Processing

267

Use of Comparative LPWAN Studies as a Didactic Resource in Engineering Education

275

Performance Evaluation of Fermat Spiral Distributions for Grating Lobe Suppression in Sparse UAV Networks

282

Spatiotemporal Modeling of Cloudiness Using ConvLSTM Networks: Applications to Photovoltaic Energy

289

Energy Management Strategies in Water Pumping Systems Supplied by Renewable Energy: A Review

297

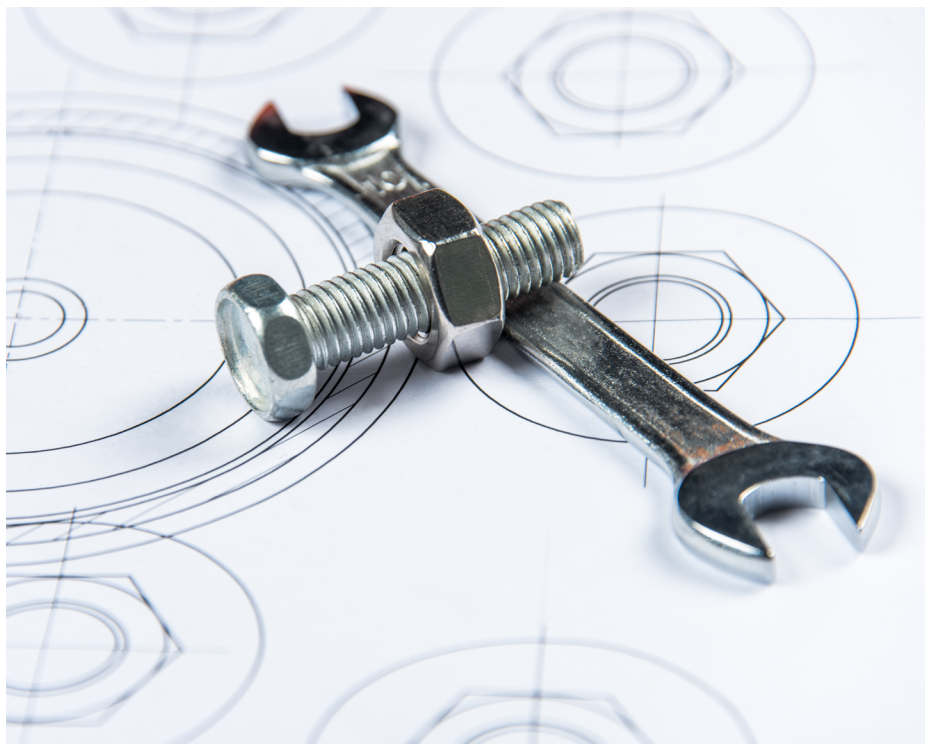
Contribution to the Design Methodology of DC-DC LLC Resonant Converter for Electric Vehicles

305

A Single Source Nine Levels Inverter for Electromobility Applications

CONTENIDO

Ingeniería Mecánica



312

Effect of the Modeling Approach on Porous Hip Implants: Structured vs. Homogenized

320

Bioclimatic Evaluation of Thermal Comfort in Morelos Using the Givoni Chart: Comparison Between Historical Data and In-Situ Monitoring

328

Numerical Analysis of Mixing in Lug-Bolt Burners in Transition from Subcritical to Trans-critical Pressure

336

Assessment of Thermal Comfort in Social Housing Through Occupant Perception And Environmental Monitoring In A Warm Subhumid Climate

344

Parametric Numerical Analysis of an Earth-Air Heat Exchanger under Different Operating Conditions

352

Effect of WAAM-CMT Process Parameters on Microstructure and Electrical Conductivity of Pure Aluminum

360

Thermal Matching and Annual Performance of PCM Retrofit in Commercial Opaque Envelopes: A Whole-Building Parametric Study

370

CFD analysis of the effect of a wind catcher on cross ventilation and thermal comfort in a room of a house in the state of Morelos

CONTENIDO

Ingeniería Mecatrónica

377

From Static Navigation to Dynamic Risk Assessment: Analysis of Assistive Technologies for Visual Impairment

384

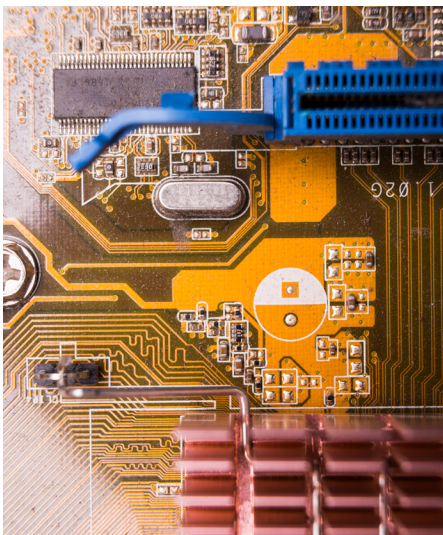
Analysis and Design of an Didactic Platform for Monitoring and Fault Detection in Hydraulic Systems

390

Meshing Contour Detection-based Triangulation for 2D Topology Optimization

398

Mechanical Design of an ASRS Prototype with Artificial Vision as a Learning Resource for Industrial Engineering



406

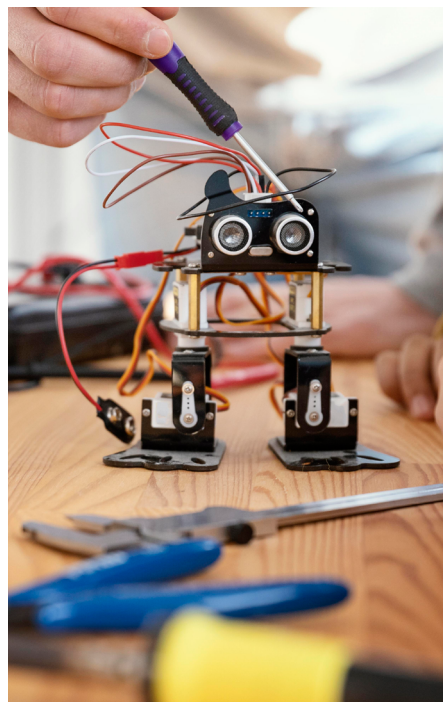
Low-Cost Open-Source Virtual Instrumentation Architecture for Distributed Industrial Monitoring and Control

414

Technical Proposal for an Automated Liquid Filling System with Integrated Computer Vision and Water Quality Monitoring

422

Inverse Kinematics Verification and Validation of a 3-RPS Parallel Robot for a Cervical Orthosis Prototype for Rehabilitation



CONTENIDO

Otras Ingenierías

429

Multiscale Drought Prediction using SVR in Mexico City

437

Analysis and Impact of Student Failure Factors in Basic Sciences at the Delicias Technological Institute

445

TiO₂ nanotubes with CuO quantum dots supported on thin films and their application in the photocatalytic production of hydrogen

452

Systematic PID tuning for ethanol dehydration by extractive distillation using literature-based transfer functions

460

Design of a Production and Marketing Model for Tamarind Derivates with a Focus on Education for Sustainability



468

FlexAI Gen: Hybrid Postural Evaluation System with Local LLM and Adaptive State Machines

476

Experimental Research and Development of the Validation and Verification (V&V) of a Miniaturized, Optimized MicroRheological Device for Point-of-Care Testing (POCT) Based Anemia Monitoring in Clinical Settings

484

Design and Development (D&D) of a Vortex-Based Bead-Assisted Mechanical Disruption System in a Confined Chamber

493

A Fiber-Optic pH Sensor for Water Quality Monitoring Based on a Michelson Interferometer With PMMA/ Graphene Coating

499

Systematic PID tuning for ethanol dehydration by extractive distillation using literature-based transfer functions

507

Evaluation of the Type of Fertilization and Application of Abiotic Stress on the Bioactive Compounds of Chiltepín (*Capsicum annum* var. *aviculare*)



MARCO: Methodology of Agentic Reporting and Composition Orchestration for Academic Documents

Walter A. Mata-López*

*Facultad de Ingeniería Mecánica y Eléctrica/Universidad de Colima. Carretera Colima-Coquimatlán

km. 9, Coquimatlán, Colima; e-mail: wmata@ucol.mx

Abstract: An agentic Human-in-the-Loop (HITL) framework for orchestrating scientific publications is proposed. The system integrates master workflows, modular skills, and governance rules: they extract parametric constraints from the target editorial template, guide large-scale literature curation, orchestrate sequential section drafting, and homologate references, keeping the researcher as epistemic arbiter at every phase. The framework was validated autopoietically: the methodology was used to draft this manuscript under IFAC guidelines (CITCA 2026). Sixteen articles were processed; the researcher acted simultaneously as system architect, direct co-author, and epistemic arbiter, with contributions ranging from the prior design of skills and rules to lexical editing and authorship of foundational passages. Results confirm that rules-based governance prevents factual hallucination, that human contribution is structurally broader than any partial log can capture, and that the framework is agnostic to the target editorial standard.

Keywords: agentic artificial intelligence, human-in-the-loop, academic writing, workflows, large language models.

1. INTRODUCCIÓN

La integración de los Modelos de Lenguaje de Gran Escala (LLM, por sus siglas en inglés) en la redacción científica, la automatización de las revisiones bibliográficas y la publicación académica está transformando rápidamente los flujos de trabajo de la investigación. El enfoque *Human-In-The-Loop* (HITL), en el que los LLM prestan asistencia, pero no sustituyen por completo la experiencia humana, son cada vez más habituales y ofrecen una mayor eficiencia en tareas como la selección de bibliografía, la extracción de datos, la redacción de manuscritos y la revisión por pares. Los flujos de trabajo que coordinan múltiples agentes de IA Generativa (IAG) o LLM modulares para automatizar procesos complejos como las revisiones sistemáticas o incluso ciclos de investigación completos, están surgiendo como un nuevo paradigma. Si bien estas tecnologías pueden reducir drásticamente la carga de trabajo manual y acelerar la síntesis de la evidencia, también plantean retos relacionados con la precisión, la transparencia, el uso ético y el riesgo de una dependencia excesiva del contenido generado por la IAG. La bibliografía destaca tanto el potencial como los riesgos de estas innovaciones: los LLM pueden igualar o superar el rendimiento humano en determinadas tareas de revisión bajo condiciones específicas, pero siguen requiriendo una supervisión humana rigurosa para garantizar la calidad y la integridad. A medida que su adopción crece en todas las disciplinas, es preponderante establecer directrices claras para el uso responsable y la evaluación continua de los sistemas agénticos a fin de mantener la confianza en la comunicación científica.

Esta tecnología ha derivado recientemente en la búsqueda de la automatización total del proceso de descubrimiento e iteración científica. Sistemas multi-agente contemporáneos,

como The AI Scientist (Lu et al., 2024), CycleResearcher (Weng et al., 2025) o el framework de la Universidad de Tampere (Sami et al., 2025), demuestran la viabilidad de delegar con tecnología de IAG, procesos que abarcan desde el planteamiento de una hipótesis hasta la evaluación simulada por pares y redacción de manuscritos a un costo mínimo. Paralelamente, en el dominio de las Revisiones Sistemáticas de Literatura (SLR, por sus siglas en inglés), la aparición de infraestructuras interconectadas como LatteReview (Rouzrokh et al., 2024) y PROMPHEUS (Torres et al., 2025) prometen mitigar el severo cuello de botella provocado por el crecimiento exponencial de la literatura científica, asumiendo tareas operativas críticas como la extracción de datos y el escrutinio automatizado de títulos y resúmenes (Ye et al., 2024; Dennstädt et al., 2024).

Sin embargo, someter la producción intelectual enteramente a una delegación no supervisada provoca una peligrosa homogeneización del discurso e imprecisión fenomenológica. La literatura advierte que los modelos, a pesar de sus impresionantes tasas de recuperación y abstracción, fracasan al momento de discriminar sesgos metodológicos complejos en estudios contradictorios, el cual es un razonamiento que sigue requiriendo el criterio cognitivo humano (Khraisha et al., 2024). Confiar a ciegas en las capacidades narrativas de un agente no solo desvincula el análisis humano de la proximidad con el objeto de estudio, sino que induce el llamado *automation bias* o sesgo de automatización (Drori & Te'eni, 2024; Silva et al., 2025). Ante este problema, se hace imperativo formalizar un ecosistema donde la orquestación del conocimiento combine el vasto poder computacional de la parametrización de bases de datos con el rigor epistemológico que solo el experto humano provee (Malik & Terzidis, 2025).



Para que este balance sistémico HITL rinda frutos, el orquestador artificial debe ser disociado de respuestas creativas “crudas” e integrarse bajo infraestructuras estrictas de validación. Métodos de ajuste iterativo propuestos mediante Ajuste Fino con Eficiencia en los Parámetros (PEFT, por sus siglas en inglés) por Susnjak et al. (2024), o controles de versiones asimétricos sobre redacción asistida (Pividori & Greene, 2024), sugieren que la IA rinde mejor cuando opera bajo flujos de trabajo agénticos que restringen o limitan la acción del modelo a tareas discretas y auditables.

Como respuesta a este desafío técnico y conceptual, el presente artículo propone y documenta MARCO (Metodología Agéntica para la Redacción y Composición Orquestada): un framework de orquestación HITL de aplicación universal para la producción de documentos académicos. MARCO no está concebido exclusivamente para artículos de congreso: su flujo de cuatro fases: 1) parametrización del destino, 2) curaduría literaria, 3) redacción modular y 4) ensamblaje; es estructuralmente invariante para tesis, ensayos, resúmenes ejecutivos, reportes técnicos, propuestas de investigación o investigaciones experimentales. Lo que varía entre tipos de documento es únicamente la configuración de las habilidades (*skills*) y las restricciones paramétricas del destino, no la arquitectura del proceso. Yendo un paso más allá de una revisión teórica, este manuscrito opera como un experimento metacognitivo vivo: la metodología formulada fue utilizada (de manera autopoietica) para planificar, esquematizar, curar literatura y redactar este mismo documento, adaptándose a la parametrización de formato de la conferencia CITCA 2026. La Sección 2 fundamenta los conceptos de agentes, *skills*, reglas (*rules*) y flujos de trabajo (*workflows*); la Sección 3 detalla cada paso de MARCO como guía replicable, incluyendo recomendaciones de implementación; la Sección 4 reporta los resultados del experimento metacognitivo; y la Sección 5 formula las conclusiones y limitaciones.

2. MARCO TEÓRICO

2.1 De la automatización total a la orquestación HITL

La adopción de la Inteligencia Artificial en la producción académica se ha polarizado recientemente en dos extremos conceptuales, dejando un vacío metódico que justifica la arquitectura propuesta. En un extremo se encuentra la automatización total o delegación no supervisada: sistemas que intentan marginar al investigador de la ecuación epistémica. Plataformas de frontera como The AI Scientist (Lu et al., 2024) aspiran a ejecutar el ciclo completo de descubrimiento, experimentación y redacción de artículos científicos de forma autónoma. Si bien suponen un hito de ingeniería, delegar la abstracción fenomenológica en algoritmos estadísticos fomenta la alucinación, carece de anclaje contextual tácito y compromete severamente la integridad y responsabilidad autoral.

En el extremo opuesto se ubica el paradigma del prompting conversacional: el uso de chatbots operando interacciones aisladas (Ahn, 2024). Aunque este enfoque devuelve el control al usuario, adolece de memoria sistemática, es incapaz de ejecutar instrucciones paramétricas complejas y somete al texto al sesgo de “estilo IA” genérico. Paralelamente en el

mundo de los frameworks de desarrollo, librerías multipropósito como AutoGen (Wu et al., 2023) o sistemas enciclopédicos como STORM (Shao et al., 2024) resuelven la orquestación técnica, pero están diseñados para despliegues de programación o resúmenes tipo Wikipedia, careciendo de formalidad metodológica institucional.

Es precisamente en medio de esta dicotomía donde se hace imperativa la orquestación metodológica HITL (Human-in-the-Loop). Al dotar a los modelos base (Claude, Gemini) de una gobernanza rígida, el agente no actúa como “creador” ni como un simple “chat”, sino como un ecosistema de instrumentación (Baek et al., 2024). Plataformas instrumentales como Antigraity (Google DeepMind) permiten instanciar esta separación entre el motor cognitivo (el LLM) y la gobernanza del entorno, asegurando que sea el investigador quien dicte el ensamblaje científico paso a paso bajo control auditable.

2.2 Flujos de trabajo agénticos (*workflow*)

Un *workflow* es una especificación formal y secuencial de las etapas que un agente debe ejecutar para alcanzar un objetivo de producción complejo. En el contexto de la producción académica, un *workflow* maestro actúa como una metaestructura que orquesta subsecuentemente la ingesta de literatura, la revisión por secciones, el formateo bibliográfico y el ensamblaje editorial final. Su diferencia con una lista de instrucciones vaga reside en su adherencia a puntos de control explícitos (*checkpoints*) que requieren confirmación del humano antes de avanzar a la siguiente fase, garantizando que se mantenga el arbitraje epistémico en todo momento. MARCO implementa este principio mediante un *workflow* maestro único (*/write-article*) que actúa como punto de entrada invariante para cualquier tipo de documento académico, ya sea este artículo, tesis, ensayo, reporte técnico o investigaciones experimentales realizadas en un laboratorio, adaptando las fases de ejecución a las restricciones paramétricas del destino sin alterar la arquitectura de orquestación.

La importancia de este diseño ha sido constatada empíricamente: flujos de trabajo semi-automatizados que mantienen al humano como punto de control en etapas críticas de inclusión/exclusión logran una precisión superior a la de los modelos operando en modo totalmente autónomo (Ye et al., 2024). El *workflow* no solo disciplina o regula al agente, sino que estructura el conocimiento al obligar a declarar explícitamente los criterios antes de cada paso de automatización.

2.3 *Skills* (habilidades) y *rules* (reglas): modularidad y gobernanza del sistema

Los *skills* (habilidades modulares especializadas) son unidades funcionales reutilizables que encapsulan una tarea específica del proceso editorial: revisar un artículo en especial, procesar una carpeta de literatura en lote, redactar una sección del manuscrito u homologar referencias bibliográficas. Cada *skill* opera bajo instrucciones propias que lo especializan, pero hereda las restricciones globales del sistema. La Figura 1 ilustra la relación jerárquica entre estas capas.

3. METODOLOGÍA

El *framework* propuesto estructura el ciclo de producción académica en cuatro fases secuenciales y un eje transversal de gobernanza. La secuencia es invariante: ninguna fase de escritura puede iniciarse sin que haya concluido la precedente, y se establece un punto de control humano explícito en cada transición. La Figura 2 sintetiza el flujo general del *pipeline*.



Figura 1. Arquitectura del ecosistema agéntico.

Estas restricciones globales se articulan mediante *rules* (reglas de gobernanza del sistema): directivas persistentes que rigen el comportamiento del agente en todas sus interacciones, con independencia del *skill* activo. En el contexto académico, las *rules* son el mecanismo que previene las patologías más frecuentes de la IA generativa: la alucinación de datos, el lenguaje críptico o artificialmente exagerado, y la invención de referencias (Scherbakov et al., 2025). Una *rule* puede especificar, por ejemplo, que el agente debe basar exclusivamente sus afirmaciones en los documentos suministrados como contexto, abstenerse de inferir datos no presentes en el corpus, y respetar el estilo de citación definido por la plantilla activa. La combinación de *skills* modulares y *rules* globales produce un sistema que es, simultáneamente, flexible en su alcance y rígido en su integridad epistémica.

2.4 Plantillas editoriales como restricciones paramétricas

Las normativas de publicación de revistas y congresos (IEEE, APA, IFAC, Vancouver) pueden conceptualizarse, desde la perspectiva del sistema agéntico, como restricciones paramétricas: límites operativos que definen el espacio de soluciones válidas para el manuscrito. Estas restricciones abarcan la extensión máxima (número de páginas o palabras), la jerarquía de secciones obligatorias, el formato de la lista de referencias, el estilo tipográfico para tablas y figuras, y los requerimientos del resumen (palabras clave, estructura IMRaD o similar).

La ventaja del enfoque agéntico es que este universo de restricciones puede ser extraído, analizado e incorporado como contexto activo por el agente desde el inicio del proceso, mediante la lectura directa del archivo de plantilla, y mantenido durante toda la sesión de redacción. Este paso, que en flujos de trabajo convencionales implica relecturas manuales repetidas y errores de formato costosos en tiempo, queda automatizado como la Fase Inicial del *workflow* maestro, antes de haber escrito una sola palabra de forma manual. Ello garantiza coherencia paramétrica desde el primer archivo generado hasta el ensamblaje del manuscrito final.

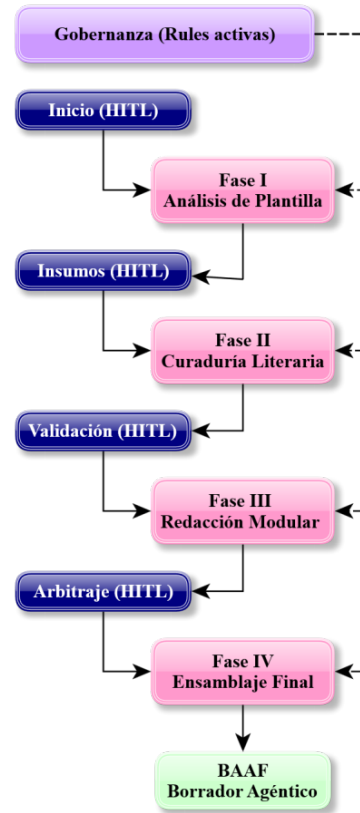


Figura 2. Pipeline MARCO: de la plantilla editorial al manuscrito final.

El *framework* descrito en este artículo no es dependiente de un modelo o plataforma específicos; no obstante, el experimento reportado en la Sección 4 se ejecutó sobre un entorno tecnológico concreto que conviene declarar para garantizar la reproducibilidad. El entorno agéntico utilizado fue Antigravity (Google DeepMind), un entorno de desarrollo integrado (IDE) que provee la capa de orquestación sobre modelos base. Los modelos subyacentes empleados durante la sesión incluyeron Gemini 3.1 Pro (High) (Google DeepMind), así como Claude Sonnet y Claude Pro (Anthropic), seleccionados de forma híbrida en distintas fases según la demanda cognitiva o técnica de cada tarea. El sistema operativo fue Windows 11, y el espacio de trabajo fue gestionado como un repositorio local de archivos Markdown estructurados.

Los *workflows*, *skills* y *rules* se definieron como archivos de configuración propios del entorno de trabajo, independientes del modelo subyacente. Esto confirma la portabilidad del *framework*, pues la arquitectura de orquestación puede reimplementarse sobre entornos similares como Cursor,

Windsurf o cualquier IDE con capacidades agénticas y ejecución de herramientas.

3.1 Fase I: Análisis de la plantilla y generación del esqueleto

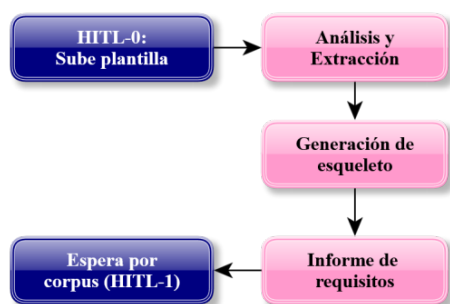


Figura 3. Diagrama de flujo de la Fase I (parametrización).

El proceso no comienza con la redacción, sino con un acto de intervención explícita (véase la Figura 3): la invocación del *workflow* maestro (/write-article). Este comando, diseñado previamente como un archivo de configuración del entorno agéntico, es el punto cero del ciclo: su activación carga las *rules* globales, establece la secuencia de fases y solicita los dos insumos iniciales imprescindibles: la ruta del archivo de plantilla editorial y la carpeta del corpus de literatura. Sin estos dos datos provistos por el humano, el sistema no puede avanzar (primer punto de control HITL, previo a cualquier generación de texto). A partir de esa activación, el agente procede a la parametrización del entorno editorial destino mediante las siguientes operaciones:

- **Extracción de restricciones paramétricas.** El agente lee el documento y extrae las directrices editoriales críticas: extensión máxima del manuscrito, sistema de citación requerido (APA, IEEE, IFAC, etc.), estructura obligatoria de secciones, formato del resumen (límite de palabras, campos requeridos) y normativas para tablas y figuras.
- **Generación del esqueleto modular.** A partir de las secciones identificadas, el agente genera automáticamente los archivos .md correspondientes al manuscrito (uno por sección), insertando en la cabecera de cada archivo las restricciones extraídas como recordatorio activo durante la redacción posterior.
- **Instrucción para recopilación.** Una vez creado el esqueleto, el agente emite un informe de requisitos, indicando el corpus de literatura que debe recopilarse (tipología de fuentes, idiomas aceptados, rango de años) y cualquier dato empírico primario que deberá proveer para las secciones de método y resultados. Este paso materializa el segundo punto de control HITL del proceso: el agente espera la entrega activa de los insumos antes de avanzar.

3.2 Eje transversal: Gobernanza por rules

De forma paralela a todas las fases, el comportamiento del agente está gobernado por un archivo de *rules* (reglas de gobernanza del sistema) que actúa como sistema normativo

persistente. Estas directrices cumplen tres funciones críticas: 1) control de veracidad, 2) control de estilo y 3) control de formato.

Las *rules*, *skills* y *workflows* no fueron generados por la IA, fueron diseñados, redactados y refinados con anterioridad a la sesión de escritura, como producto acumulado de la experiencia metodológica. Este trabajo de arquitectura del sistema constituye una inversión intelectual previa determinante para la calidad del proceso.

3.2.1 Tipología de la Contribución Humana en el Ciclo HITL

La participación humana en la metodología no se limita a revisar borradores generados por el agente. El ciclo HITL comprende cinco modalidades de contribución cualitativamente distintas: 1) el diseño del sistema (exclusivo humano) donde antes de iniciar cualquier sesión, se definen los *workflows*, *skills* y *rules* del entorno agéntico a partir de su conocimiento disciplinar; 2) la redacción autónoma donde se escriben directamente fragmentos del manuscrito que el agente posteriormente expande, referencia o complementa; 3) la orquestación de prompts donde se formulan instrucciones precisas al agente para cada tarea, ajustando el nivel de detalle, el alcance temático y las fuentes autorizadas por sección; 4) la revisión iterativa profunda donde en el ciclo de revisión se lee párrafo a párrafo, se detectan términos fuera del registro académico del artículo, se identifica inconsistencias de voz, solo por citar algunas, y el proceso puede acumular horas de trabajo en una sesión de redacción; y el arbitraje epistémico en el que las decisiones sobre qué incluir, qué omitir, cómo jerarquizar los argumentos y qué posición adoptar ante la literatura son responsabilidad exclusiva del factor humano.

3.3 Fase II: Curaduría literaria masiva

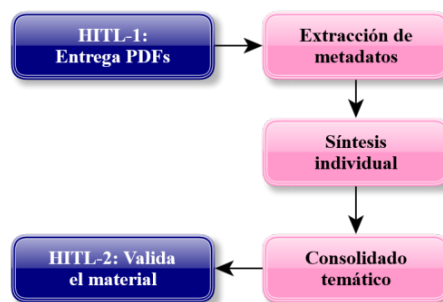


Figura 4. Diagrama de flujo de la Fase II (curaduría literaria).

Una vez que se deposita el corpus de documentos fuente, como se ilustra en la Figura 4 (artículos recopilados en formato PDF), se activa el *skill* de revisión masiva (*Batch Review Papers*). Este *skill* opera de forma sistemática sobre cada documento del directorio. Posteriormente se revisan y validan los resúmenes generados, corrigiendo posibles inexactitudes de extracción antes de que el contenido sea incorporado como base argumentativa del texto.

3.4 Fase III: Redacción modular secuencial

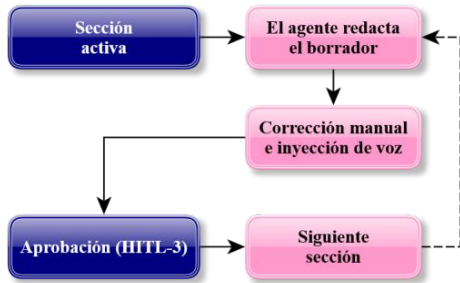


Figura 5. Diagrama de flujo de la Fase III (redacción modular).

Con el esqueleto, las *rules* activas y el consolidado literario disponibles, el agente procede a la redacción sección por sección (véase la Figura 5). Cada sección se trabaja mediante un *skill* especializado que conoce su estructura requerida y sus restricciones paramétricas específicas: Introducción y Marco Teórico, Metodología, Resultados, Discusión y Conclusiones. En cada sección, se ejerce el rol de árbitro final: revisa, corrige terminología, añade matices y aprueba el contenido antes de avanzar a la siguiente fase (tercer punto de control HITL permanente a lo largo de toda la fase).

3.5 Fase IV: Ensamblaje, resumen y homologación bibliográfica



Figura 6. Diagrama de flujo de la Fase IV (ensamblaje final).

Una vez redactadas todas las secciones, como se puede observar en la Figura 6, se ejecutan tres operaciones de cierre: 1) la redacción del borrador del resumen donde el agente redacta el resumen ajustándose al límite de palabras y a los campos obligatorios extraídos en la Fase I; 2) la homologación bibliográfica (Format References) donde el agente mapea todas las citas presentes en el texto y genera la lista de referencias en el formato requerido por la plantilla, verificando coherencia entre las citas en texto y las entradas bibliográficas finales; y 3) la revisión de consistencia paramétrica donde el agente realiza una verificación final de congruencia entre el manuscrito ensamblado y las restricciones editoriales de la Fase I: extensión total, número de figuras y tablas, formato del título y la afiliación institucional.

El producto resultante de esta Fase IV es el Borrador Agéntico de Alta Fidelidad (BAAF), el cual consisten en un manuscrito que cumple prácticamente con todas las restricciones paramétricas establecidas en las *rules*, pero que por ningún motivo constituye el documento final. A partir de este documento, la metodología da paso a la última Fase: la curaduría final manual, llevada a cabo de manera íntegra fuera del entorno agéntico, dedicada a afinar el tono discursivo, ajustar figuras o valores de tablas manualmente, y consolidar la voz definitiva del manuscrito.

La parametrización de variables específicas de entorno, como la gestión de la ventana de contexto, la asignación de tokens por regla, o la extensibilidad a diseños colaborativos multiautor, escapan al enfoque empírico del presente caso de estudio y constituyen el objeto de futuras ampliaciones teóricas de la metodología.

4. RESULTADOS

Esta sección reporta los resultados de la aplicación de la metodología descrita en la Sección 3 a un caso concreto: la redacción del presente manuscrito, adaptado a las directrices de la conferencia CITCA 2026 (plantilla IFAC). El proceso fue simultáneamente el objeto de estudio y el instrumento de producción (lo que constituye su carácter autopoietico), requiriendo una documentación paralela y sistemática en tiempo real de cada iteración significativa entre el agente y la dirección humana.

Para la producción del presente manuscrito empírico, se instanció un ecosistema adaptado específicamente a este proyecto sobre el IDE Antigravity. Una fortaleza de MARCO es que se pueden inyectar *skills* y *rules* genéricos o especializados para ayudar en la estructuración de los datos del proyecto particular. La arquitectura empleada se desglosa en la Tabla 1.

Tabla 1. Componentes arquitectónicos instanciados para la redacción del proyecto.

Componente	Archivo / Nombre	Función Genérica en el Proyecto
<i>Workflow</i>	<i>/write-article</i>	Metarutina de orquestación central de 4 fases para la elaboración del artículo.
<i>Rules</i> (Gobernanza)	<i>reglas_escritura.md</i>	Bloqueo activo de superlativos empalagosos, prevención de alucinación bibliográfica y control del formato "modern print".
<i>Skills</i> (Curaduría)	<i>Batch Review / Review Paper</i>	Extracción y síntesis masiva de los 16 PDFs

Componente	Archivo / Nombre	Función Genérica en el Proyecto
		de literatura consumidos por el sistema.
<i>Skills</i> (Redacción)	Draft Introduction... (x5)	Segmentación narrativa de cada capítulo adaptándose a la plantilla (IFAC).
<i>Skills</i> (Edición)	Format References	Homologación bibliográfica final de todas las citas extraídas.

Como producto de esta arquitectura iterativa, y tras un ciclo de evaluación continua documentado cualitativamente mediante observación participante, que engloba las múltiples iteraciones, ajustes de formato y refinamientos epistémicos ejecutados de forma manual a lo largo del proceso, el sistema generó el Borrador Agéntico de Alta Fidelidad (BAAF). Este manuscrito sirvió como base estructural estable antes de transitar hacia la curaduría humana definitiva.

4.1 Resultados de la Fase I: Parametrización del entorno

El experimento se ejecutó sobre el entorno agéntico Antigravity (Google DeepMind) utilizando Gemini 3.1 Pro (High) en conjunto con Claude Sonnet / Claude Pro (Anthropic) como modelos base, aplicando el framework MARCO (Metodología Agéntica para la Redacción y Composición Orquestada) en su primera validación empírica documentada. El proceso se inició con la invocación explícita del *workflow* maestro */write-article* por vía manual (el acto fundacional que activó la carga de las *rules* globales y estableció la secuencia de fases). Los *workflows*, *skills* y *rules* fueron configurados como archivos locales del directorio *.agents/* del espacio de trabajo, habiendo sido diseñados en sesiones previas. La plantilla editorial de CITCA 2026 (formato IFAC, archivo *ifacconf.docx*) fue procesada al inicio de la sesión.

El esqueleto modular fue generado automáticamente en 6 archivos .md independientes, creando el entorno de trabajo que estructuró toda la sesión subsecuente. Cabe precisar que el producto de las fases agénticas no es un documento final compilado en un procesador de texto (ej. Word). El ensamblaje visual final en formato .docx y su maquetación a doble columna se ejecutaron íntegramente fuera del sistema durante la Fase de curaduría humana.

4.2 Resultados de la Fase II: Curaduría literaria

Se proveyó un corpus de 18 artículos científicos en formato PDF, recuperados de bases de datos como arXiv, PubMed, Semantic Scholar y repositorios de acceso abierto. La ejecución del *skill* de revisión masiva procesó la totalidad del corpus mediante un script de extracción de texto, generando 18 síntesis individuales estructuradas y un consolidado temático agrupado en 4 ejes de debate. La intervención HITL fue determinante: se curó y aprovisionó manualmente el corpus, lo que garantizó su pertinencia y calidad temática.

4.3 Resultados de la Fase III: Redacción modular y puntos de control HITL

La documentación de la intervención humana durante esta fase adoptó un enfoque cualitativo de observación participante: se registraron los episodios significativos de toma de decisión, corrección o redirección durante la sesión de escritura. Este enfoque (análogo al registro etnográfico aplicado a procesos cognitivos situados) permite capturar la textura de la intervención humana sin reducirla a métricas cuantitativas que, en el estado actual de la investigación, carecen de instrumentos de medición consensuados. Los ejemplos reportados en la Tabla 3 son instancias representativas seleccionadas por su valor metodológico ilustrativo, no un inventario exhaustivo de todas las microintervenciones producidas durante la sesión. La redacción de las secciones se ejecutó de forma secuencial desde la introducción hasta la curaduría del manuscrito.

Los puntos de control HITL no fueron obstáculos al proceso: en todos los casos, la corrección inmediata propagó el ajuste al resto del texto de forma consistente. La estimación de tiempos (Tabla 2) desglosa los minutos invertidos y una muestra cualitativa de las intervenciones registradas hasta alcanzar el Borrador Agéntico de Alta Fidelidad (BAAF). Es importante interpretar este tiempo no como el fin del proceso editorial, sino como el esfuerzo requerido para estructurar la base del documento con un alineamiento paramétrico perfecto. A partir de la generación del BAAF, el manuscrito transitó a la última Fase (fuera del entorno agéntico), donde la curaduría humana requirió múltiples horas de trabajo para redibujar figuras especializadas, inyectar semántica de tablas y afinar la estructura del manuscrito final.

Tabla 2. Intervenciones HITL registradas durante la redacción modular del manuscrito.

Fase / Sección	Producto Resultante	Tiempo estimado	Intervenciones HITL
Fase I: Parametrización	Esqueleto modular	~10 min	1
Fase II: Curaduría masiva	Síntesis del corpus	~15 min	2
Fase III: Introducción	Borrador §1	~20 min	3
Fase III: Marco Teórico	Borrador §2	~15 min	4
Fase III: Metodología	Borrador §3	~25 min	2
Fase IV: Ensamblaje	Referencias y Resumen	~15 min	1
Subtotal (MARCO)	Borrador Agéntico (BAAF)	~100 min	Múltiples
Fase V (Humana)	Manuscrito definitivo	Variable (hrs – días)	Fuera del Sistema agéntico

El tiempo estimado engloba la revisión cognitiva humana integral de cada salida (lectura profunda, verificación léxica y



de formato), no únicamente el tiempo de latencia computacional del agente generativo.

4.4 Eficacia de la gobernanza por rules

La capa de *rules* demostró su efectividad en la prevención de las patologías más documentadas de la IAG. En ningún momento el agente incorporó referencias bibliográficas no presentes en el corpus proporcionado, ni adoptó un tono laudatorio o autorreferencial. Las correcciones se concentraron exclusivamente en ajustes de nivel léxico y arquitectónico, lo que sugiere que el control de veracidad de las *rules* operó de forma efectiva.

5. CONCLUSIONES

El presente artículo propuso, describió y demostró empíricamente MARCO (Metodología Agéntica para la Redacción y Composición Orquestada): una metodología HITL de aplicación universal para la orquestación de documentos académicos. El experimento metacognitivo, utilizar el propio MARCO para redactar el artículo que lo documenta, aportó evidencia funcional sobre su viabilidad y sus límites. De los resultados se derivan las siguientes conclusiones.

El paradigma agéntico supera cualitativamente al *prompting* conversacional en la producción académica. La combinación de *workflows* maestros, *skills* modulares y *rules* de gobernanza produce un sistema capaz de mantener coherencia paramétrica, trazabilidad bibliográfica y consistencia estilística a lo largo de un manuscrito completo, capacidades que un LLM en modo conversacional estándar no puede garantizar.

Las *rules* son el componente crítico del sistema, no un accesorio. La única desviación grave registrada durante el experimento ocurrió cuando el agente operó sin consultar su configuración de *rules* y *workflow* locales, generando una metodología genérica en lugar de la específica del entorno. Esto confirma que la gobernanza explícita no es opcional: sin ella, el modelo reproduce patrones estadísticos generales en lugar de razonamiento situado.

El HITL no es un mecanismo de corrección, sino de dirección epistémica. La muestra de intervenciones documentadas formalmente no respondió a errores de veracidad factual, sino a decisiones de nivel arquitectónico, terminológico y léxico, sumadas a múltiples microiteraciones que requieren de un criterio situado continuo por parte de la dirección del manuscrito. Esto distingue al HITL agéntico del simple “revisar lo que genera la IA”: el humano dirige la estructura conceptual del conocimiento; el agente ejecuta su instrumentación.

MARCO es generalizable a cualquier normativa editorial, tipo de documento académico o incluso investigación experimental. La separación entre la capa de restricciones paramétricas y la capa de redacción garantizan que el sistema sea agnóstico al congreso, revista o tipo de documento destino. La Tabla 2 documenta su aplicabilidad a artículos, tesis, ensayos, reportes técnicos y propuestas de investigación.

La recolección de literatura primaria permanece como dominio inalienable de nivel supervisor. El intento de automatizar la descarga del corpus vía API fue bloqueado sistemáticamente por las políticas de los repositorios académicos, confirmando que la curaduría de literatura requiere intervención humana deliberada. El agente puede procesar y sintetizar; no puede sustituir el juicio de pertinencia al seleccionar el corpus base.

5.1 Limitaciones

El experimento se basó en una sola sesión de redacción con dirección humana unitaria, lo que limita la generalización de los tiempos reportados en la Tabla 4. Asimismo, la estimación de tiempos es aproximada y no contempla el tiempo de configuración previa del entorno agéntico (instalación, definición de *rules* y *skills*), que puede ser significativo para investigadores que adopten el sistema por primera vez. La ausencia de un instrumento automatizado de registro de las intervenciones humanas constituye una de las limitaciones del experimento: la documentación se basó en el registro cualitativo de observación participante, por lo que los tiempos y conteos de intervención reportados son estimaciones estructuradas y no métricas extraídas paramétricamente del sistema.

Adicionalmente, se documentó una fricción entre la restricción paramétrica inicial y las iteraciones humanas (Fase V): aunque el agente asimiló la directriz de “extensión máxima de 6 páginas”, la curaduría cualitativa y la inyección de argumentación manual expandieron el borrador final hasta alcanzar las 10 páginas de longitud. Lejos de ejecutar un recorte metodológico que mutilara la evidencia, se tomó la decisión epistémica explícita de mantener la extensión de 10 páginas, asumiendo la superación del límite original con el fin de preservar y documentar todo el proceso orquestal de la metodología con rigor. Esta tendencia documentada (la expansión del corpus provocada por el refinamiento HITL que rompe los límites del manuscrito) constituye no solo una limitante paramétrica, sino una prueba del peso cognitivo humano en la fase V. Finalmente, la evaluación de la calidad del manuscrito producido queda pendiente de revisión por pares externa, que constituye el paso de validación definitivo del proceso.

5.2 Trabajo futuro

La arquitectura propuesta puede extenderse en al menos cinco direcciones: 1) la incorporación de un *skill* de evaluación automatizada que simule criterios de revisión por pares, siguiendo el modelo de CycleReviewer (Weng et al., 2025); 2) el desarrollo de métricas objetivas de calidad para los resúmenes generados por el *Batch Review*, más allá de la validación manual; 3) la adaptación del framework a flujos de trabajo colaborativos multiautor, donde los puntos de control HITL se distribuyen entre varios investigadores con roles diferenciados; 4) el diseño de una taxonomía operacionalizable de niveles de agencia humana en documentos co-producidos con IAG, que permita a investigadores, editores y comités de ética declarar con precisión el grado de intervención humana en cada sección de un manuscrito, superando la declaración binaria (uso / no uso)



vigente en IEEE, Nature o Elsevier; y 5) el diseño e integración de una Bitácora de Trazabilidad HITL sistemática en futuras iteraciones de la metodología en cuestión: un instrumento de registro parametrizado que capture automáticamente el tiempo real de cada fase, el tipo de intervención y la naturaleza de la decisión humana. Esto permitiría transformar la observación participante cualitativa de este primer estudio exploratorio en datos duros cuantificables y comparables entre proyectos, dotando al proceso de un nivel de audibilidad aún mayor

REFERENCIAS

- Ahn, S. (2024). The transformative impact of large language models on medical writing and publishing: current applications, challenges and future directions. *The Korean Journal of Physiology & Pharmacology: Official Journal of the Korean Physiological Society and the Korean Society of Pharmacology*, 28, 393 - 401. <https://doi.org/10.4196/kjpp.2024.28.5.393>.
- Back, J., Jauhar, S., Cucerzan, S., & Hwang, S. (2024). ResearchAgent: Iterative Research Idea Generation over Scientific Literature with Large Language Models. *ArXiv*, abs/2404.07738. <https://doi.org/10.48550/arxiv.2404.07738>.
- Dennstädt, F., Zink, J., Putora, P., Hastings, J., & Cihoric, N. (2024). Title and abstract screening for literature reviews using large language models: an exploratory study in the biomedical domain. *Systematic Reviews*, 13. <https://doi.org/10.1186/s13643-024-02575-4>.
- Drori, I., & Te'eni, D. (2024). Human-in-the-Loop AI Reviewing: Feasibility, Opportunities, and Risks. *J. Assoc. Inf. Syst.*, 25, 7. <https://doi.org/10.17705/1jais.00867>.
- Khraisha, Q., Put, S., Kappenberg, J., Warraitch, A., & Hadfield, K. (2023). Can large language models replace humans in systematic reviews? Evaluating GPT-4's efficacy in screening and extracting data from peer-reviewed and grey literature in multiple languages. *Research Synthesis Methods*, 15, 616 - 626. <https://doi.org/10.1002/jrsm.1715>.
- Lu, C., Lu, C., Lange, R., Foerster, J., Clune, J., & Ha, D. (2024). The AI Scientist: Towards Fully Automated Open-Ended Scientific Discovery. *ArXiv*, abs/2408.06292. <https://doi.org/10.48550/arXiv.2408.06292>.
- Malik, F., & Terzidis, O. (2025). A hybrid framework for creating artificial intelligence-augmented systematic literature reviews. *Management Review Quarterly*. <https://doi.org/10.1007/s11301-025-00522-8>.
- Pividori, M., & Greene, C. (2023). A publishing infrastructure for AI-assisted academic authoring. *bioRxiv*. <https://doi.org/10.1101/2023.01.21.525030>.
- Rouzrokh, P., & Shariatnia, M. (2025). LatteReview: A Multi-Agent Framework for Systematic Review Automation Using Large Language Models. *ArXiv*, abs/2501.05468. <https://doi.org/10.48550/arxiv.2501.05468>.
- Sami, M., Rasheed, Z., Kemell, K., Waseem, M., Kilamo, T., Saari, M., Nguyen-Duc, A., Systä, K., & Abrahamsson, P. (2024). System for systematic literature review using multiple AI agents: Concept and an empirical evaluation. *ArXiv*, abs/2403.08399. <https://doi.org/10.48550/arxiv.2403.08399>.
- Scherbakov, D., Hubig, N., Jansari, V., Bakumenko, A., & Lenert, L. (2024). The emergence of large language models as tools in literature reviews: a large language model-assisted systematic review. *Journal of the American Medical Informatics Association: JAMIA*, 32, 1071 - 1086. <https://doi.org/10.1093/jamia/ocaf063>.
- Shao, Y., Jiang Y., Kanell T. A., XU, P., Khattab, O., Lam, M. S. (2024). Assisting in Writing Wikipedia-like Articles From Scratch with Large Language Models. *arXiv preprint arXiv:2402.14207*. <https://doi.org/10.48550/arXiv.2402.14207>.
- Silva, J., Gouveia, R., Zielinski, K., Oliveira, M., Amancio, D., Bruno, O., & Oliveira, O. (2025). AI-Assisted Tools for Scientific Review Writing: Opportunities and Cautions. *ACS Applied Materials & Interfaces*, 17, 47795 - 47805. <https://doi.org/10.1021/acsami.5c08837>.
- Sušnjak, T., Hwang, P., Reyes, N., Barczak, A., McIntosh, T., & Ranathunga, S. (2024). Automating Research Synthesis with Domain-Specific Large Language Model Fine-Tuning. *ACM Transactions on Knowledge Discovery from Data*, 19, 1 - 39. <https://doi.org/10.1145/3715964>.
- Torres, J., Mulligan, C., Jorge, J., & Moreira, C. (2025). PROMPTHEUS: A Human-Centered Pipeline to Streamline Systematic Literature Reviews with Large Language Models. *Inf.*, 16, 420. <https://doi.org/10.3390/info16050420>.
- Weng, Y., Zhu, M., Bao, G., Zhang, H., Wang, J., Zhang, Y., & Yang, L. (2024). CycleResearcher: Improving Automated Research via Automated Review. *ArXiv*, abs/2411.00816. <https://doi.org/10.48550/arxiv.2411.00816>.
- Wu, Q., Bansal, G., Zhang, J., Wu, Y., Li, B., Zhu, E., Jiang, L., Zhang, X., Zhang, S., Liu, J., Awadallah, A. H., White, R., Burger, D., & Wang, C. (2023). AutoGen: Enabling Next-Gen LLM Applications via Multi-Agent Conversation. *arXiv preprint arXiv:2308.08155*. <https://doi.org/10.48550/arXiv.2308.08155>.
- Ye, A., Maiti, A., Schmidt, M., & Pedersen, S. (2024). A Hybrid Semi-Automated Workflow for Systematic and Literature Review Processes with Large Language Model Analysis. *Future Internet*, 16, 167. <https://doi.org/10.3390/fi16050167>.

Optimizing Nutrient Solutions Using Swarm Intelligence

Guerra Ibarra Juan Pablo* Segundo Sevilla Ana Celia*
Rocha Rocha Aaron Jr.*

* *Tecnológico Nacional de México, Campus Zamora, Zamora, Michoacán 59720 México (e-mail: juan.gi@zamora.tecnm.mx, ana.ss@zamora.tecnm.mx, aaron.rr@zamora.tecnm.mx).*

Abstract: The reduction in arable land and population growth require the optimization of agricultural processes. The application of modern technology to the production of plant-based foods has given rise to precision agriculture, with hydroponic systems being a key example. In these systems, plant nutrition management is crucial to crop yield. This study reports the results of an adaptation of the Swarm Intelligence algorithm to minimise the relative error in macronutrient dosing. The model optimizes the nutrient solutions composition by adjusting quantities of different fertilizers in grams. Over ten iterations, an average relative error of 0.01459% was obtained for Nitrogen, Phosphorus, Potassium and Chlorine.

Keywords: Metaheuristics, Optimization, Plant nutrition, Precision agriculture, Relative Error.

1. INTRODUCTION

Food from various agricultural crops has been, is, and will continue to be fundamental to the development of human society (Awasthi (2020); Ray (2017)). The growing global population, expected to reach nearly 10 billion by 2050, and the reduction of arable land, according to The Food and Agriculture Organization (FAO (2023)) present a challenge to modern agriculture. In response to the current challenges associated with food production from cultivating the land, what is known as Precision Agriculture (PA) has been developed. PA consists of applying different technological advancements to agriculture, with the aim of increasing its productivity and minimizing the environmental impact from water consumption or overuse of chemicals such as fertilizers and pesticides (Cisternas et al. (2020)).

In PA, the development and implementation of hydroponic farming systems are used to combat the reduction in arable land. Hydroponic Agriculture (HA) essentially eliminates soil and replaces it with an inert substrate, which may be simply sand or combined with a range of natural fibres such as those found in coconut husks (Kannan et al. (2022)). Fertilisation processes are fundamental to HA, as without a natural source of nutrients such as soil, nutrition depends entirely on Nutrient Solutions (NS) (Fathidarehniheh et al. (2024)).

Nutrient solutions (NS) are essentially a mixture of varying quantities of fertilizers dissolved in water; this mixture is delivered to the plants in the crops via irrigation systems, primarily through drip irrigation (Ali Al Meselmani (2023)). Fertilizers consist of various plant nutrients, which are classified as macronutrients and micronutrients; this classification is based on the quantities required by

the crops. The main macronutrients are Nitrogen (N), Phosphorus (P) and Potassium (K). Determining the exact quantities of the various fertilizers to be dissolved in the nutrient solution to ensure that the crops receive the nutrients they require is essentially a combinatorial problem that can be addressed using various tools, including Metaheuristic Optimization Algorithms (MOA), such algorithms imitate nature's successful adaptation to the environment. There are MOAs based on evolutionary theory, on the behaviour of insects and animals, and on trajectory-based approaches, to name but a few.

The Particle Swarm Optimization (PSO) algorithm is a population-based stochastic optimization metaheuristic (Eberhart and Kennedy (1995)); it is classified within the Swarm Intelligence paradigm. PSO is based on the social and collective behaviour of groups of animals, primarily flocks of birds or schools of fish. The behaviour that the algorithm seeks to mimic is the successful identification of food sources (the solution to the problem), which depends on the interaction between individual learning and the group transmission of information (Engelbrecht (2007)). The PSO is notable for its computational simplicity, as it requires only a few tuning parameters compared to other types of meta-heuristic algorithms, and for its high efficiency in solving non-linear and continuous optimization problems. It is ideal for optimising complex functions or those with multiple discontinuities.

This study describes the adaptation of the PSO algorithm to minimise the Relative Error (RE) in the concentrations of the macronutrients N, P and K within a NS. The model optimises the quantities of various fertilizers to be dissolved in water, in order to produce a NS with the values defined by agricultural experts in the algorithm's objective function. Given its versatility, the proposed methodology is applicable to the formulation of NS for various crops, allowing the necessary adjustments to be made to the

* This research was funded by *Tecnológico Nacional de México, Campus Zamora*.



concentration in Parts Per Million (ppm) according to the requirements of each phenological stage of the crops.

The organization of the paper is as follows: Section 2 presents a series of works related to land cultivation processes for food production using PA methods. The third Section describes the fundamentals of crop nutrition, NS formulation, and the computational architecture and procedures developed to adapt the PSO algorithm, with the aim of automating and optimizing NS formulation in accordance with the PPM values established in the PSO objective function. The experimental design, including the target objectives and the algorithmic control settings for PSO NS optimization, is established in Section 4. The results obtained with the PSO adaptation for the formulation of NS are presented in the Section 5, the Section 6 provides the concluding remarks, summarizes the key findings of this research and practical implications of the study are synthesized. The final Section outlines future initiatives to be implemented to improve the results generated by the methodology presented.

2. RELATED WORK

One of the main objectives of Precision Agriculture is to provide crops with the inputs they require, such as water, nutrients and plant protection products, at the precise doses and during the appropriate phenological stages (Awasthi (2020)). This strategic approach aims to promote optimal plant growth, maximise agricultural food production and reduce the negative environmental impacts of farming. To achieve these goals, PA employs advanced technological solutions, from specialized hardware systems for monitoring crop conditions to software systems for data processing, automated decision-making and process optimization.

Within the scope of precision agriculture, plant nutrition management is a critical factor in the success of hydroponic systems. Consequently, the current state of the art includes various methods aimed at accurately determining the concentrations of essential macronutrients through leaf analysis, with particular emphasis on *N*, *P* and *K*, given their importance in crop development. Lu et al. (2020) used groups of images with different *K* contents to establish the relationship between reflectance spectra and leaf *K* content. Qi et al. (2021) proposed a method for determining the *N* content in images of wheat crops acquired with mobile devices using histogram manipulation algorithms. *P* is another of the main macronutrients for crop development. Meiqing et al. (2016) presented a method for measuring how *P* concentration affects electrical impedance in tomato crops. Sun et al. (2018) used dynamic leaf capture of rice leaves by scanning to investigate the changing regulation of leaf characteristics under nutritional stress from *N*, *P*, and *K*. Tran et al. (2019) implemented convolutional neural network model and Autoencoder to classify and predict macronutrients deficiencies.

Once the nutritional status of the crops has been assessed and classified, the next stage in the PA is the formulation of NS that maintain macronutrient and micronutrient levels at the ideal range for the crop's stage of development. Various methods for generating efficient fertilization

processes are reported in the literature, Xu et al. (2017) implemented an expert system to develop N, P and K fertilization regimes for use in rice cultivation. Cropper and Comerford (2005) developed a genetic algorithm to determine the minimum amount of fertilizer required to maintain the health of crops. Ahmed et al. (2021) implemented a genetic algorithm that uses data from different types of sensors to make adjustments to fertilization processes, using the neighborhood-based strategy for the exploration and exploitation stages. Chen et al. (2021) proposed an optimization method based on an optimization algorithm inspired by the behavior of the gray wolf with multiple strategies to make fertilization suggestions for the main macronutrients. Guerra Ibarra et al. (2026) implemented a simulated annealing metaheuristic algorithm to formulate NS for tomato crops, demonstrating high efficiency in achieving the algorithm's objective function.

3. MATERIALS AND METHODS

3.1 An overview of plant nutrition

Nutrition in the plant kingdom occurs through the absorption of various nutrients. These nutrients can be classified into two main groups: macronutrients and micronutrients, both groups are necessary for proper plants development (Dimkpa and Bindraban (2016)). A nutrient is classified into one group or the other based on the quantity required, rather than its importance. The concentration range varies from 30 to 400 PPM for macronutrients and 0.01 to 5 PPM for micronutrients (Resh (2022)). Table 1 shows the classification of the 15 elements necessary for the proper development of crops.

Table 1. Nutrient classification into macronutrients or micronutrients.

Nutrients essential for healthy plant growth		
Category	Element	Symbol
Macronutrients	Nitrogen	N
	Phosphorus	P
	Potassium	K
	Calcium	Ca
	Magnesium	Mg
	Sulfur	S
Micronutrients	Iron	Fe
	Manganese	Mn
	Copper	Cu
	Zinc	Zn
	Boron	B
	Molybdenum	Mo
	Chlorine	Cl
	Nickel	Ni
	Sodium	Na

In hydroponic agriculture, as there is no soil to provide some or all of the elements listed in Table 1, crop nutrition is provided entirely through the formulation of nutrient solutions (NS). Essentially, a nutrient solution is a mixture of varying quantities of commercial fertilizers dissolved in water, which is then distributed to the crops via irrigation systems. Table 2 shows a list of commonly used fertilizers and their respective chemical formulas

Table 2. List of commercial fertilizers.

Fertilizer	Chemical formula	Molecular Weight g/mol
Calcium nitrate	$Ca(NO_3)_2$	164.06
Potassium nitrate	KNO_3	101.10
Manganese sulfate	$MnSO_4$	151.00
Diammonium phosphate	$(NH_4)_2HPO_4$	132.06
Ammonium nitrate	NH_4NO_3	80.04
Monopotassium phosphate	KH_2PO_4	136.09
Potassium sulfate	K_2SO_4	174.26
Ammonium sulfate	$(NH_4)_2SO_4$	132.14
Urea	$CO(NH_2)_2$	60.06
Magnesium nitrate	$Mg(NO_3)_2 \cdot 4H_2O$	222.37
Calcium potassium nitrate	$KCa(NO_3)_3$	263.21
Monoammonium phosphate	$NH_4H_2PO_4$	115.03
Potassium chloride	KCl	74.55
Calcium chloride	$CaCl_2$	110.98
Anhydrous ammonia	NH_3	17.03
Ammonium chloride	NH_4Cl	53.49
Phosphoric rock	$Ca_5(PO_4)_3F$	504.30
Superphosphate	$Ca(H_2PO_4)_2 \cdot H_2O$	252.07
Phosphoric acid	H_3PO_4	97.99
Phosphorite	$Ca_5(PO_4)_3F$	504.30
Potassium carbonate	K_2CO_3	138.21
Sodium nitrate	$NaNO_3$	84.99
Kieserite	$MgSO_4 \cdot H_2O$	138.38
Magnesium nitrate	$Mg(NO_3)_2$	148.31
Sodium chloride	$NaCl$	58.43

3.2 PPM calculation for a given amount of fertilizer

By dissolving a specific amount of fertilizer in water to prepare NS, it is necessary to calculate the PPM of each nutrient of interest present in the fertilizer's chemical composition. The PPM of a nutrient present in a fertilizer is calculated using the following equation:

$$PPM_e^f = \frac{1000 \times q_f \times a_e \times n_e^f}{mw^f \times V}, \quad (1)$$

where f refers to a particular fertilizer, and e is the nutrient that forms part of the fertilizer, 1000 is a constant conversion factor to convert grams of fertilizer to milligrams, q_f is the amount in grams of fertilizer, a_e is the atomic weight of the element selected for its quantification, n_e^f is the number of atoms of the chemical element present in the chemical formula of one fertilizer particularly, aw^f is the fertilizer molecular weight and V is total volume of the NS in liters.

A NS consists of the elements contained in each of the dissolved fertilizers; it is therefore necessary to quantify the total concentration in PPM of each macronutrient or micronutrient present in the nutrient solution. This is achieved by applying the following equation:

$$PPM_e^{NS} = \sum_{f=0}^n PPM_e^f, \quad (2)$$

where PPM_e^{NS} represents the total concentration of the nutrient element e present in the NS, while PPM_e^f corresponds to the amount of PPM contributed by a specific

amount of fertilizer f for nutrient element e , calculated by applying Equation (1).

3.3 Features of the selected crop

The cucumber (*Cucumis sativus*) is a herbaceous plant of the Cucurbitaceae family; its origins date back to the tropical regions of India over three millennia ago, and it spread through Mesopotamia before reaching Greece and Rome. It was introduced to the Americas by Christopher Columbus in 1494. Figure 1 shows a cucumber crop grown using protected cultivation techniques



Fig. 1. Cucumber cultivation in protected agriculture in the Zamora Valley, Michoacán, Mexico.

Cucumber is one of the most popular vegetables worldwide Sonneveld and Voogt (2009). The plant is a climber, using tendrils and angular, hairy stems to grow.

Like any other crop, the growth of cucumbers involves different stages, each with its own specific nutritional requirements. Table 3 shows the stages of phenological development and their N , P , K and Cl requirements (Resh (2022)).

Table 3. N , P , K and Cl requirements of cucumbers in ppm during their developmental stages

Phenological stage	N PMM		P PMM		K PMM		Cl PMM	
	Min	Max	Min	Max	Min	Max	Min	Max
Seedling	60	80	25	35	50	70	1.2	1.5
Vegetative growth	120	150	35	45	150	200	1.8	2.2
Flowering	160	180	40	50	200	250	2.2	2.5
Harvest	180	220	45	55	280	350	2.5	3.0

3.4 Adaptation of the PSO algorithm for NS formulation.

The following is a detailed description of the Particle Swarm Optimization Algorithm for Formulating Nutrient Solutions (PSOAFNS), the aim of which is to minimise the average relative error in the PPM of the nutrients specified in the target function and those contained in the nutrient solution created by mixing various fertilizers. PSOAFNS employs a multi-target function, which is



formally defined as a vector of dimension e , where each component represents the concentration in PPM of the required nutrients. This formulation allows for the simultaneous evaluation of the convergence of multiple objectives within the algorithm's search space. The structure of the objective function is presented in Equation:

$$PPM_e = \{PPM_1, PPM_2, \dots, PPM_{e-1}, PPM_e\}, \quad (3)$$

where subscript e refers to the nutrient element established as the target to be achieved.

The PSO algorithm is based on a set of stochastic elements, where each particle is defined by a position vector representing a potential solution within the multivariate search space of the problem domain. In the context of PSOAFNS, the position of each particle is encoded as an n dimensional vector, whose components correspond to the quantities, expressed in grams, of the fertilizers available for the formulation of the NS. Each particle constitutes a configuration whose feasibility is evaluated by the objective function. The structure of these particles is defined by the following Equation:

$$P_i = \{q_1, q_2, \dots, q_{f-1}, q_f\}, \quad (4)$$

where f is the number of fertilizers available for the formulation of NS and i ranges from 1 to n , where n is the total number of particles in the swarm.

Each initial q_f is initialized randomly by applying next Equation:

$$q_f = \text{random}(ll, ul), \quad (5)$$

where the quantity in grams for fertilizer f is between the lower limit (ll) and the upper limit (ul), which are used for all fertilizers available for the formulation of NS.

To evaluate the performance of the PSOAFNS algorithm in converging toward the PPMs for each nutrient defined in Equation 2, the Absolute Relative Error is used. This metric is employed to standardize the disparity in magnitudes between macronutrients and micronutrients, ensuring that they are comparable regardless of the concentration scale. The absolute relative error is defined by the following Equation:

$$RE_e = \frac{|PPM_e - PPM_e^{NS}|}{PPM_e}, \quad (6)$$

where RE_e is the absolute relative error of nutrient e , in relation to the PPM_e^{NS} defined in Equation (2) with respect to PPM_e .

The total relative error in the NS is determined by applying the Equation:

$$RE_{total} = \sum_{e=1}^O RE_e, \quad (7)$$

subindex e takes values from 1 to the total number of nutrients defined in the Equation 3.

Applying Equation 7 allows us to calculate the fitness for each particle in the swarm. It is important to remember that, since this is a minimization problem, a fitness value

tending toward zero indicates better performance, which represents a more accurate solution in the NS formulation.

The search space is explored to identify optimal fertilizer combinations for formulating the NS that meets the PPM concentrations specified in Equation 3, using the following Equation:

$$v_f(t+1) = w \times v_f(t) + c_1 \times r_1 \times (pbest.q_f - q_f(t)) + c_2 \times r_2 \times (gbest.q_f - q_f(t)), \quad (8)$$

where $v_f(t+1)$ is the velocity that will change the particle's position relative to the fertilizer quantity f , w is the coefficient of inertia, which determines the effect of the previous speed, c_1 and c_2 are the acceleration coefficients for the constants $pbest$ and $gbest$, respectively, r_1, r_2 are random values uniformly distributed over the interval $[0, 1]$, which introduce randomness into the search, finally, $pbest$ is the best position of particle i , and $gbest$ is the best position of the swarm; remember that both are composed of f numbers of fertilizers quantities.

Adjusting the formulation by increasing or decreasing the amounts of fertilizers available for NS formulation using the following Equation:

$$P_i = \{q_1 \pm v_1, q_2 \pm v_2, \dots, q_{f-1} \pm v_{f-1}, q_f \pm v_f\}. \quad (9)$$

The quantities of each fertilizer q_f are increased or decreased randomly, with the same probability of occurrence.

The algorithm's behavior is regulated by two fundamental parameters: the cognitive coefficient c_1 , linked to the individual memory of each particle $pbest$, and the social coefficient (c_2), associated with the collective knowledge of the swarm's best individual $gbest$. The former acts as a factor for local exploration, while the latter promotes the exploitation of the search space. To optimize convergence, a dynamic tuning strategy is implemented: during the initial stages, a higher value is assigned to c_1 to favor global exploration of the search space; as the iterative process progresses, c_1 decreases and c_2 gradually increases until they converge to predefined values, thereby prioritizing the refinement of the optimal solution.

The reduction factor for c_1 is calculated using the following Equation:

$$\Delta c_1 = -1 \times \frac{ivc_1 - fvc_1}{i}, \quad (10)$$

where ivc_1 and fvc_1 are the values of c_1 during the first and last iterations, respectively, while i is the total number of iterations of the PSOAFNS.

In order to progressively increase the value of c_2 and ensure that, during the final iterations, the PSOAFNS intensifies its exploration of the search space, the mechanism described in Equation 11 is implemented.

$$\Delta c_2 = \frac{fvc_2 - ivc_2}{i}, \quad (11)$$

where fvc_2 and ivc_2 are the values of c_2 during the last and first iterations, respectively.

At the end of each iteration, the coefficients c_1 and c_2 are dynamically adjusted to ensure an optimal balance between the exploration (pbest) and exploitation (gbest) phases, a critical balance for the performance of meta-heuristic algorithms. These acceleration parameters are updated by implementing Equations 12 and 13, respectively, allowing the swarm to converge.

$$c_1 = c_1 + \Delta c_1 \quad (12)$$

$$c_2 = c_2 + \Delta c_2 \quad (13)$$

Figure 2 shows the general adaptation of the PSO algorithm for the NS formulation.

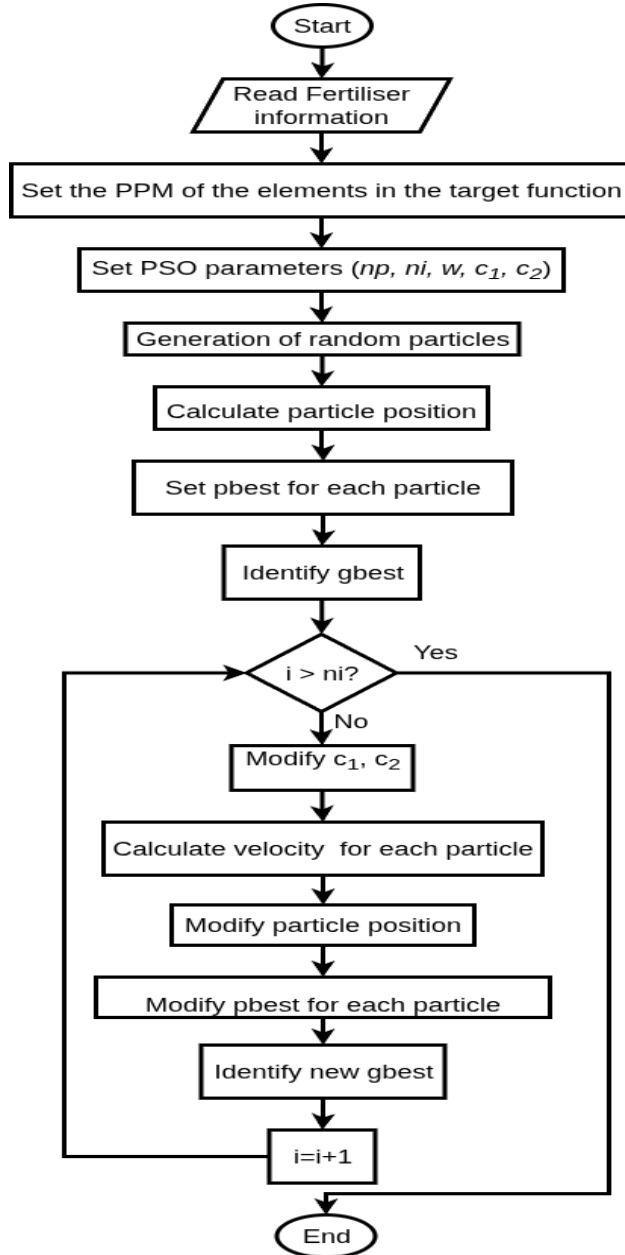


Fig. 2. Adaptation of the PSO for the formulation of NS.

4. EXPERIMENTATION

In order to evaluate the effectiveness of the PSOAFNS, a target function was designed based on the concentration of the three primary macronutrients and one essential micronutrient. The target values, expressed in PPM, were established based on the midpoint between the minimum and maximum nutritional requirements of the cucumber crop during its vegetative growth stage, as specified in Table 3. Table 4 details the target concentrations for the NS, the formulation of which is based on the combination of fertilizers specified in Table 2.

Table 4. Target PPM concentrations of nutrient elements.

No	Element	Symbol	Target PPM
1	Nitrogen	N	135
2	Phosphorus	P	40
3	Potassium	K	175
4	Chlorine	Cl	2

The experimental phase to evaluate the performance and robustness of the PSOAFNS consisted of iterative runs until ten instances of satisfactory convergence were achieved that met the nutritional requirements specified in Table 4. The parameter settings used for the PSOAFNS are shown in Table 5.

Table 5. Parameter values used by PSOAFNS during the experimental stage.

PSOAFNS parameters used		
No.	Parameter	Value
1	w	0.73
2	c_1	3.05
3	c_2	0.5
4	r_1	Random number [0,1]
5	r_2	Random number [0,1]
6	Number of particles	20
7	i	300
8	ivc_1	3.05
9	fv_1	0.5
10	fv_2	2.5
11	ivc_2	0.5
12	ll	10
13	ul	150
114	V	1,000

5. RESULTS

Table 6 presents the results obtained from ten independent runs of the PSOAFNS. It can be seen that the RE_{total} remained consistently at values less than or equal to 0.03. This threshold was established as a criterion for satisfactory convergence, ensuring that the difference between the obtained nutrient concentrations and the target values set by the algorithm is near 0.

The statistical analysis of the ten runs of the PSOAFNS, detailed in Table 6, shows an avg. RE_{total} of 0.01459 and a std of 0.01480, indicating minimal variation among the experimental trials. These indicators confirm the consistency

Table 6. Results of the implementation of the PSOAFNS to achieve the targets in Table 4.

Results of the implementation of the PSOAFNS to achieve the targets in Table 4								
PPM	N	PPM	P	PPM	K	PPM	CI	ER_{total}
Target	in NS	Target	in NS	Target	in NS	Target	in NS	
135.0	139.15212	40.0	39.97237	175.0	175.15577	2.0	2.00000	0.03234
135.0	135.00386	40.0	40.00000	175.0	174.94051	2.0	2.00001	0.00037
135.0	135.00044	40.0	40.00012	175.0	173.16477	2.0	2.00000	0.01049
135.0	133.03546	40.0	39.96681	175.0	174.99620	2.0	2.00000	0.01540
135.0	135.02373	40.0	40.00002	175.0	174.99984	2.0	2.00000	0.00018
135.0	134.99937	40.0	40.00000	175.0	172.92344	2.0	2.00000	0.01187
135.0	135.00000	40.0	39.87958	175.0	175.53766	2.0	2.00000	0.00608
135.0	134.80929	40.0	40.00000	175.0	182.95226	2.0	2.00000	0.04685
135.0	135.00553	40.0	39.90877	175.0	177.68596	2.0	2.00000	0.01767
135.0	134.97718	40.0	40.00000	175.0	175.77760	2.0	2.00000	0.00461

of PSOAFNS and its ability to converge reproducibly toward optimal NS.

Figure 3 illustrates, using a logarithmic scale, the magnitude of the total relative error (RE_{total}) corresponding to the ten independent runs of the PSOAFNS.

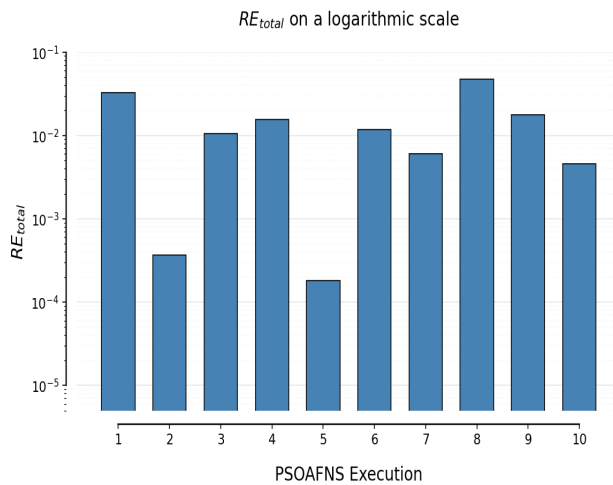


Fig. 3. Logarithmic scale comparison of RE_{total} for the ten runs of the PSOAFNS.

The comparative analysis reveals that the ninth execution exhibits the most efficient performance, achieving convergence close to the target values set forth in Table 4. It is important to note that, in general, all ten runs demonstrate excellent performance.

Table 7 presents a comparative analysis between the best NS and the statistics calculated using the amounts of fertilizer applied in the ten runs of the PSOAFNS. A high standard deviation is observed for the 25 fertilizers used, indicating the existence of a multimodal search space, where various combinations of fertilizers allow convergence toward the nutritional requirements specified in Table 4. This phenomenon confirms the algorithm's flexibility in finding viable solutions.

Figure 4 illustrates the convergence curve of the optimal particles in each of the ten runs of the PSOAFNS. Trajectory analysis reveals an initial divergence phase, characteristic of an intensive exploration process. In contrast, during the final stages of the iterative process, a stabilization of particle fitness is observed, indicating a marked exploitation of the search space aimed at refining the NS.

Table 7. Comparison of fertilizer amounts in the Best NS and Statistics from the 10 NS.

Fertilizer amounts in the Best and the 10 NS in Table 6.			
Fertilizer	Best gr.	Avg. gr	Std. gr
1	0.00000	77.98715	74.63284
2	69.09891	78.07013	65.91351
3	206.90453	113.10392	66.30600
4	0.00000	53.97313	55.42053
5	123.68846	70.57751	66.42933
6	0.50837	44.32177	44.16609
7	132.50930	74.97227	57.39992
8	0.00000	18.88855	36.56701
9	0.00000	71.80593	79.25943
10	152.25628	76.12854	69.08793
11	46.37598	45.7863	47.26491
12	135.58543	26.17878	43.46184
13	0.00000	14.06872	19.48852
14	0.00000	23.45216	32.03456
15	13.38296	17.40660	22.18038
16	62.83433	11.42958	20.61567
17	12.39889	82.39511	81.72789
18	0.00000	39.28600	45.55442
19	0.00000	27.34670	25.79050
20	68.98888	32.51759	37.64669
21	170.07912	185.05964	54.34677
22	9.21624	74.6653	52.30675
23	102.58750	114.67514	34.0844
24	72.17263	64.41675	61.16361
25	0.00000	17.02272	27.73467

Convergence of the Best Particle in the Ten Execs of PSOAFNS

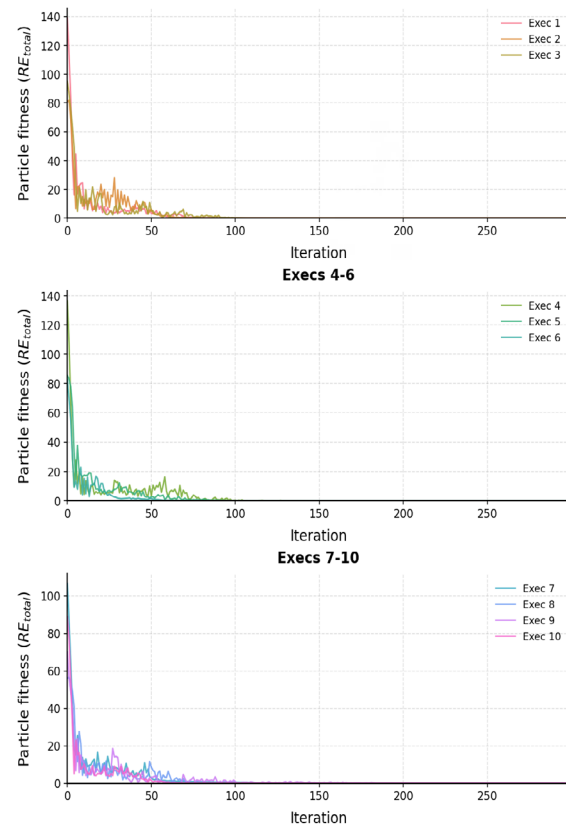


Fig. 4. Fitness performance of the best particles 10 Exec.



6. CONCLUSION

The experimental results detailed in Section 5 validate the efficacy of the PSOAFNS in achieving precise target concentrations for both macronutrients and micronutrients. These findings hold substantial scientific value, particularly given the limited literature regarding automated NS formulations that integrate micronutrients essential for physiological development. A key contribution of this methodology is its capacity to simultaneously optimize micronutrient levels that are orders of magnitude lower than macronutrients within a single solution. Furthermore, the architecture offers significant operational flexibility, enabling seamless adaptation to diverse crops through specific PPM adjustments. Statistically, the algorithm proves both accurate and robust when navigating complex, multivariate search spaces

7. FUTURE WORK

To enhance the performance of the PSOAFNS, future research will focus on redefining the particle fitness function to integrate multi-objective variables, such as economic costs and phytotoxicity thresholds. Additionally, the model will be expanded toward a comprehensive nutritional analysis by increasing the library of available fertilizers, enabling the precise management of micronutrient concentrations at PPM levels significantly lower than those of macronutrients. This refinement is expected to increase the algorithm's robustness and applicability in highly complex NS. Crucially, a preliminary physical validation phase is planned to bridge the gap between computational optimization and practical application. This will involve preparing optimized NS formulations and conducting laboratory chemical analyses to verify that theoretical concentrations align with physical measurements.

ACKNOWLEDGEMENTS

The PSOAFNS research was conducted with support from Tecnológico Nacional de México, Campus Zamora, and the Instituto de Ciencia, Tecnología e Innovación de Michoacán.

REFERENCES

- Ahmed, U., Lin, J.C.W., Srivastava, G., and Djenouri, Y. (2021). A nutrient recommendation system for soil fertilization based on evolutionary computation. *Computers and Electronics in Agriculture*, 189(August), 106407. doi:10.1016/j.compag.2021.106407.
- Ali Al Meselmani, M. (2023). Nutrient Solution for Hydroponics. In *Recent Research and Advances in Soilless Culture*. IntechOpen. doi:10.5772/intechopen.101604.
- Awasthi, Y. (2020). Press "A" for Artificial Intelligence in Agriculture: A Review. *International Journal on Informatics Visualization*, 4(3), 112–116. doi:10.30630/ijov.4.3.387.
- Chen, C., Wang, X., Chen, H., Wu, C., Mafarja, M., and Turabieh, H. (2021). Towards precision fertilization: Multi-strategy grey wolf optimizer based model evaluation and yield estimation. *Electronics (Switzerland)*, 10(18), 1–26. doi:10.3390/electronics10182183.
- Cisternas, I., Velásquez, I., Caro, A., and Rodríguez, A. (2020). Systematic literature review of implementations of precision agriculture. *Computers and Electronics in Agriculture*, 176(May), 105626. doi:10.1016/j.compag.2020.105626.
- Cropper, W.P. and Comerford, N.B. (2005). Optimizing simulated fertilizer additions using a genetic algorithm with a nutrient uptake model. *Ecological Modelling*, 185, 271–281. doi:10.1016/j.ecolmodel.2004.12.010.
- Dimkpa, C.O. and Bindraban, P.S. (2016). Fortification of micronutrients for efficient agronomic production: a review. *Agronomy for Sustainable Development*, 36(1), 1–26. doi:10.1007/s13593-015-0346-6.
- Eberhart, R. and Kennedy, J. (1995). A new optimizer using particle swarm theory. In *MHS'95. Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, 39–43. IEEE. doi:10.1109/MHS.1995.494215.
- Engelbrecht, A. (2007). *Computational Intelligence An Introduction*. John Wiley & Sons, Ltd., second edition.
- FAO (2023). Our approach — Food Systems — Food and Agriculture Organization of the United Nations Retrieved from <http://www.fao.org/food-systems/our-approach/en/>. URL <http://www.fao.org/food-systems/our-approach/en/>.
- Fathidarehnijeh, E., Nadeem, M., Cheema, M., Thomas, R., Krishnapillai, M., and Galagedara, L. (2024). Current perspective on nutrient solution management strategies to improve the nutrient and water use efficiency in hydroponic systems. *Canadian Journal of Plant Science*, 104(2), 88–102. doi:10.1139/cjps-2023-0034.
- Guerra Ibarra, J.P., Cuevas de la Rosa, F.J., and Rocha Rocha, A.J. (2026). Formulation of Nutrient Solutions Using Simulated Annealing. *Agriculture*, 16(4), 449. doi:10.3390/agriculture16040449.
- Kannan, M., Elavarasan, G., Balamurugan, A., Dhanusiya, B., and Freedom, D. (2022). Hydroponic farming – A state of art for the future agriculture. *Materials Today: Proceedings*, 68, 2163–2166. doi:10.1016/j.matpr.2022.08.416.
- Lu, J., Yang, T., Su, X., Qi, H., Yao, X., Cheng, T., Zhu, Y., Cao, W., and Tian, Y. (2020). Monitoring leaf potassium content using hyperspectral vegetation indices in rice leaves. *Precision Agriculture*, 21(2), 324–348. doi:10.1007/s11119-019-09670-w.
- Meiqing, L., Jinyang, L., Hanping, M., and Yanyou, W. (2016). Diagnosis and detection of phosphorus nutrition level for *Solanum lycopersicum* based on electrical impedance spectroscopy. *Biosystems Engineering*, 143, 108–118. doi:10.1016/j.biosystemseng.2016.01.005.
- Qi, X., Zhao, Y., Huang, Y., Wang, Y., Qin, W., Fu, W., Guo, Y., and Ye, Y. (2021). A novel approach for nitrogen diagnosis of wheat canopies digital images by mobile phones based on histogram. *Scientific Reports*, 11(1), 1–15. doi:10.1038/s41598-021-92431-5.
- Ray, P.P. (2017). Internet of things for smart agriculture: Technologies, practices and future direction. *Journal of Ambient Intelligence and Smart Environments*, 9(4), 395–420. doi:10.3233/AIS-170440.
- Resh, H.M. (2022). *Hydroponic Food Production : A Definitive Guidebook for the Advanced Home Gardener*



- and the Commercial Hydroponic Grower. *Hydroponic Food Production*. doi:10.1201/9781003133254.
- Sonneveld, C. and Voogt, W. (2009). *Plant nutrition of greenhouse crops*. Springer Netherlands, Dordrecht. doi: 10.1007/978-90-481-2532-6.
- Sun, Y., Tong, C., He, S., Wang, K., and Chen, L. (2018). Identification of Nitrogen, Phosphorus, and Potassium Deficiencies Based on Temporal Dynamics of Leaf Morphology and Color. *Sustainability*, 10(3), 1–14. doi:10.3390/su10030762. URL <http://www.mdpi.com/2071-1050/10/3/762>.
- Tran, T.T., Choi, J.W., Le, T.T.H., and Kim, J.W. (2019). A comparative study of deep CNN in forecasting and classifying the macronutrient deficiencies on development of tomato plant. *Applied Sciences (Switzerland)*, 9(8). doi:10.3390/app9081601.
- Xu, X., He, P., Yang, F., Ma, J., Pampolino, M.F., Johnston, A.M., and Zhou, W. (2017). Methodology of fertilizer recommendation based on yield response and agronomic efficiency for rice in China. *Field Crops Research*, 206, 33–42. doi:10.1016/j.fcr.2017.02.011. URL <http://dx.doi.org/10.1016/j.fcr.2017.02.011>.

Theoretical Model for Detecting Chemical Components Using Machine Learning Approach

Luis Arturo Tapia-Alarcón* Braulio Misael Villegas-Martínez*, ** Gabriela Hernández-Luna*, ***
Roberto Alan Beltran-Vargas* José Antonio Marban-Salgado* Omar Palillero-Sandoval*

* Centro de Investigación en Ingeniería y Ciencias Aplicadas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Mor 62209
MX (e-mail: luis.tapiaa@uaem.edu.mx).

**Loyola University Chicago, Chicago, IL, USA

*** Escuela de Técnicos Laboratoristas, Universidad Autónoma del Estado de Morelos, Cuernavaca, Mor 62209

Abstract: Accurate detection of inorganic contaminants in water remains a significant challenge. In this work, we use a theoretical slot waveguide approach together with machine learning to identify dissolved chemical species in aqueous media. Variations in the concentration of copper sulfate, lead sulfate, potassium nitrate and calcium carbonate are analyzed through their effect on field confinement in the slot and the propagation constant of the fundamental Transversal Magnetic mode (TM mode). Synthetic data were generated and used to train and evaluate Decision Tree, Random Forest, XGBoost and multilayer perceptron algorithms. Results demonstrate that the Multilayer Perceptron model achieves exceptional classification accuracy and stability under both deterministic and noisy conditions, providing clear discrimination between chemical components based on their optical response.

Keywords: water quality, waveguides, photonics sensors, contaminant detection, machine learning.

1. INTRODUCTION

Water contamination by dissolved inorganic compounds remains a major challenge for environmental monitoring and public health, particularly in regions affected by industrial discharge, agricultural runoff, and groundwater salinization (Srivastav & Ranjan, 2020). In countries such as Mexico, the presence of contaminants including arsenic, fluoride, nitrates, and high salinity levels has been associated with geological conditions and unsustainable water-management practices, posing significant long-term health risks (Camacho et al., 2011).

Conventional water-quality monitoring commonly relies on chemical or electronic sensors; however, these approaches often present limitations related to sensitivity, response time, and operational cost (Pérez-Beltrán et al., 2024). In this context, photonic platforms, particularly slot waveguides, have emerged as promising alternatives due to their strong electromagnetic-field confinement and high sensitivity to refractive-index variations, enabling label-free detection of ionic species (Mamtmin, 2020).

The complex and multidimensional nature of the optical responses generated by these systems motivates the use of machine-learning (ML) techniques capable of identifying nonlinear correlations within the electromagnetic descriptors, particularly under noisy or uncertain conditions (Pérez-Beltrán et al., 2024). The integration of computational photonics and machine learning, therefore, provides a promising framework

for intelligent optical sensing and data-driven chemical discrimination.

In this work, we present a computational framework combining slot-waveguide simulations with ML classification for the analysis of dissolved chemical species in water. Synthetic electromagnetic descriptors derived from the TM-mode response under varying concentrations and operating wavelengths were used to construct a multidimensional optical feature space. Decision Tree (DT), Random Forest (RF), Extreme Gradient Boosting (XGBoost), and Multilayer Perceptron (MLP) models were comparatively evaluated under both clean and noisy conditions to assess classification robustness against perturbations in the simulated optical variables.

2. Methodology

2.1 Chemical-Optical Input Modeling

Chemical species identification was modeled through the relationship between analyte concentration and refractive-index variation in aqueous solutions. Four compounds were chosen based on their environmental significance and compliance with the Mexican drinking water regulation NOM-127-SSA1-2021 (Secretaría de Salud, 2022): Lead Sulfate (PbSO_4), Copper Sulfate (CuSO_4), Nitrate (KNO_3) and Calcium Carbonate (CaCO_3) (Djam, S. et al., 2022; Santucci and Scully, 2020).

The modeling framework proceeds in two integrated stages. First, the chemical concentration is transduced into an optical

parameter for the waveguide system. In this case, the optical response of each solute was modeled from its molar mass M_{solute} and molar refractivity R_{solute} , which characterize the concentration-dependent polarizability of the aqueous solution (Atkins and de Paula, 2018). The molar mass is derived from the stoichiometric sum of the constituent atomic weights. For instance, in the case of PbSO_4 is $M_{\text{solute}} = 303.27 \text{ g/mol}$. Complementarily, the molar refractivity values were obtained from standard physicochemical databases and reference handbooks (Atkins and de Paula, 2018). For PbSO_4 , the R_{solute} is taken as $38.8 \text{ cm}^3/\text{mol}$, CuSO_4 is $33.72 \text{ cm}^3/\text{mol}$, KNO_3 $12.4 \text{ cm}^3/\text{mol}$ is and CaCO_3 $13.1 \text{ cm}^3/\text{mol}$. The analyte concentration expressed in mg/L can be converted into molar concentration through

$$C_{\text{mol/L}} = \frac{C_{\text{mg/L}}}{M_{\text{solute}} \cdot 1000}, \quad (1)$$

where the factor 10^3 converts milligrams into grams, ensuring dimensional consistency in mol/L . Under dilute conditions, the molar concentration determines the corresponding mole fraction of the solute, χ_{solute} , within the aqueous mixture

$$\chi_{\text{solute}} = \frac{C_{\text{mol/L}}}{C_{\text{mol/L}} + \frac{1000 \text{ g/L}}{M_{\text{H}_2\text{O}}}}, \quad (2)$$

where $M_{\text{H}_2\text{O}} = 18.015 \text{ g/mol}$ is the molar mass of water. The mole fraction acts as a weighting factor in the molar refractivity of the mixture R_{mix} as

$$R_{\text{mix}} = (1 - \chi_{\text{solute}}) \cdot R_{\text{H}_2\text{O}} + \chi_{\text{solute}} \cdot R_{\text{solute}}, \quad (3)$$

being the molar refractivity of water $R_{\text{H}_2\text{O}} = 3.71 \text{ cm}^3/\text{mol}$ and R_{solute} act as the fundamental optical signatures of each component (Atkins and de Paula, 2018). Subsequently, the refractive index of the mixture n is derived using the Lorentz-Lorenz equation (Talebian and Talebian, 2013), linking macroscopic polarizability to the optical density of the medium

$$n = \sqrt{\frac{1 + 2(R_{\text{mix}}/V_{\text{mol}})}{1 - (R_{\text{mix}}/V_{\text{mol}})}}. \quad (4)$$

Within this expression, V_{mol} represents the molar volume, defining the space occupied by one mole of the solution. For trace-level concentrations in aqueous media, it is assumed that no significant volumetric changes occur; therefore, V_{mol} is approximated to the molar volume of pure water, $V_{\text{mol}} = 18.07 \text{ cm}^3 \text{ mol}^{-1}$. This parameter effectively scales the microscopic electronic polarizability of the molecules to the macroscopic refractive index. Under dilute conditions, the refractive-index variation of aqueous solutions can be locally approximated through a first-order dependence on analyte concentration. This approximation follows from the Lorentz-Lorenz relation, where small perturbations in molecular polarizability produce correspondingly small variations in the effective refractive index of the medium (Saleh, B. E. A, 2019). Accordingly, the refractive index was parameterized as

$$n_{\text{linear}} = n_{\text{H}_2\text{O}} + k \cdot C_{\text{mg/L}} \quad (5)$$

where $n_{\text{H}_2\text{O}} = 1.333$ is the refractive index of water, $k = 10^3 R_{\text{solute}}/M_{\text{solute}}$ is an effective optical sensitivity coefficient expressed in L/mg . while the factor 10^3 ensures dimensional consistency between molar refractivity and concentration expressed in mg/L . The resulting parametrization preserves the relative optical contrast among analytes and enables concentration-dependent refractive-index discrimination. As summarized in Table 1, the proposed model predicts a progressive increase in refractive index with increasing contaminant concentration. For instance, PbSO_4 , starting from a pure water, $n_{\text{H}_2\text{O}} = 1.333$, the model records a progressive increase in optical density until reaching a value of $n_{\text{linear}} = 1.3978$ at a concentration of 505.05 mg/L .

Table 1. Concentration-dependent refractive-index values n_{linear} , for PbSO_4 obtained from the first-order optical parametrization using pure water as the reference baseline.

Solute	Concentration mg/L	n_{linear}
PbSO_4	0	1.333
	101.010101	1.345962
	202.020202	1.35892399
	505.050505	1.39780999

2.2 Slot waveguide model and solution

Once the concentration-dependent refractive index, n_{linear} , is obtained from the chemical-optical model described in subsection 2.1., it is incorporated as the refractive index of the analyte region in the sensing structure. Variations in analyte concentration modify the dielectric environment surrounding the guided mode, producing measurable changes in the optical response of the waveguide. Then, in the second stage of the model, the optical behavior of the sensor is studied using the theoretical slot-waveguide configuration proposed by ALMEIDA ET AL. (2004). The structure consists of two parallel high-refractive-index dielectric slabs of refractive index n_h and thickness $W = b - a$, separated by a distance $d = 2a$. The region between the slabs is filled with the analyte medium, whose refractive index is defined as $n_s = n_{\text{linear}}$, thereby forming a low-index slot region. The complete structure is embedded within a cladding medium of refractive index n_c .

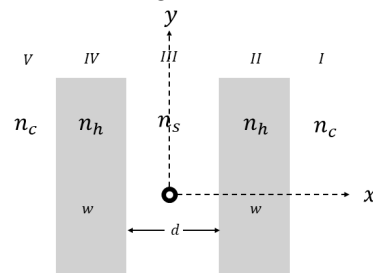


Figure 1. Schematic of the symmetric slot waveguide used for sensing.

As shown in Fig. 1, the system is modeled as a five-layer symmetric structure, enabling strong electromagnetic confinement inside the low-index slot region and enhancing refractive-index sensitivity. Assuming propagation along the z -direction and invariance along y , the problem reduces to a one-dimensional transverse analysis along x . The optical response is described using Maxwell's equations under transverse magnetic (TM) polarization, selected due to its enhanced field confinement within the slot region (Agrawal, 2005). The magnetic field distribution associated with the guided TM mode is expressed as $H(x, y, z, t) = \phi_y(x)e^{i(\omega t - \beta z)}$, where β is the propagation constant and $\omega = \frac{2\pi c}{\lambda}$ is the angular frequency related to the operating wavelength λ , with c denoting the speed of light in vacuum. The transverse field profile $\phi_y(x)$ is evaluated across the five regions of the structure illustrated in Fig. 1. Following the analytical model by Almeida et al. (2004), the profile is defined as

$$\phi(x) = A \begin{cases} \frac{1}{n_c^2} \left[\cosh(\gamma_s a) \cos(k_h W) + \frac{\gamma_s}{k_h} \left(\frac{n_h}{n_s} \right)^2 \sinh(\gamma_s a) \sin(k_h W) \right] e^{-\gamma_c(x-b)}, & x > b \\ \frac{1}{n_h^2} \cosh(\gamma_s a) \cos[k_h(x-a)] + \frac{\gamma_s}{n_s^2 k_h} \sinh(\gamma_s a) \sin[k_h(x-a)], & -a < x < b \\ \frac{1}{n_s^2} \cosh(\gamma_s x), & -a < x < a \\ \frac{1}{n_h^2} \left[\cosh(\gamma_s a) \cos(k_h(x+a)) + \frac{\gamma_s}{k_h} \left(\frac{n_h}{n_s} \right)^2 \sinh(\gamma_s a) \sin(k_h(x+a)) \right], & -b < x < a \\ \frac{1}{n_c^2} \left[\cosh(\gamma_s a) \cos(k_h W) + \frac{\gamma_s}{k_h} \left(\frac{n_h}{n_s} \right)^2 \sinh(\gamma_s a) \sin(k_h W) \right] e^{\gamma_c(x+b)}, & x < -b \end{cases} \quad (6)$$

where A is a normalization constant, while $\gamma_c^2 = \beta^2 - \left(\frac{\omega}{c} n_c\right)^2$, $\gamma_s^2 = \beta^2 - \left(\frac{\omega}{c} n_s\right)^2$ and $k_h^2 = \left(\frac{\omega}{c} n_h\right)^2 - \beta^2$, correspond to the transverse field parameters in the cladding, slot, and high-index slabs, respectively. Guided TM modes satisfy $\beta > \omega n_s/c$, $\omega n_c/c$, ensuring evanescent decay outside the guiding region. By imposing continuity of the transverse magnetic field and its weighted derivative at each interface interfaces between the j -th and k -th layer, $\phi_{j,y}(x) = \phi_{k,y}(x)$ and $\frac{d}{dx} \phi_{j,y}(x) = \left(\frac{n_j}{n_k}\right)^2 \frac{d}{dx} \phi_{k,y}(x)$, the dispersion relation governing the propagation constant β of the guided modes is obtained as (Liu et al., 2015)

$$\arctan \left[\frac{n_h^2 \gamma_c}{n_c^2 k_h} \right] + \arctan \left[\frac{n_h^2 \gamma_s}{n_s^2 k_h} \tanh(\gamma_s a) \right] + m\pi = k_h(b-a), \quad (7)$$

where m denotes the mode order. Only the fundamental TM mode ($m = 0$) was considered to ensure single-mode operation. Eq. (7) enables the calculation of the propagation constant β for each optical wavelength λ . By systematically sweeping both the operating wavelength λ in a range from 200 nm and 1400 nm and the slot refractive index n_s , the modal response of the structure was evaluated through the propagation constant β , which characterizes the effective electromagnetic behavior of the guided mode. The structural parameters (W, d, n_h, n_c) were kept fixed, with $n_h = 3.48$, $n_c = 2$, $a = 50$ nm, $b = 500$ nm and considering a concentration from 0 to 1000 mg/L according to NOM-127-SSA1-2021.

Once β is determined, the electromagnetic field distribution is reconstructed from Eq. (6) to evaluate the relevant sensing quantities. In particular, the normalized electromagnetic

energy confined within the slot region is quantified through the confinement factor η (Liu et al., 2015):

$$\eta = \frac{P_{slot}}{P_{total}} = \frac{n_s^2 \int_{-a}^a |\phi(x)|^2 dx}{\sum_j n_j^2 \int_{region j} |\phi(x)|^2 dx}, \quad (8)$$

where P_{slot} and P_{total} represent the electromagnetic energy confined within the slot region and the entire waveguide structure, respectively (Liu et al., 2015). The group velocity v_g , describing energy transport through the waveguide, is obtained from the dispersion relation as (Agrawal, 2005)

$$v_g = \left(\frac{d\beta}{d\omega} \right)^{-1}, \quad (9)$$

where the derivative is evaluated numerically from the wavelength-dependent data using a Savitzky-Golay filter to reduce discretization noise. Assuming steady-state guided mode propagation, the optical power confined within the slot region, $P_{slot} = P_{in} \eta$, where P_{in} denotes the input optical power associated with the guided mode. The corresponding electromagnetic energy density within the slot region is then estimated as (Han et al., 2022)

$$u_{slot} = \frac{P_{in} \cdot \eta}{v_g}. \quad (10)$$

The conversion of the optical response into an effective electrical signal was subsequently modeled from the electromagnetic energy density. Considering the cross-sectional dimensions of the sensing region, the effective electric field within the slot region, E_{eff} , was estimated as (AGRAWAL, 2005)

$$E_{eff} = \sqrt{\frac{2u_{slot}}{\epsilon_0 n_s^2}}, \quad (11)$$

where ϵ_0 is the vacuum permittivity. Once E_{eff} is determined, the dissipated electrical power within the ionic medium was estimated through

$$P_{elec} = \sigma |E_{eff}|^2 V_{vol}. \quad (12)$$

The electrical conductivity of the analyte solution is defined as $\sigma = \Lambda \cdot C_{mol/L}$, where Λ denotes the molar conductivity of the specific ionic species. For the analytes under evaluation, the nominal values of Λ at 25°C are: 150.02 S · cm²/mol for PbSO₄, 133.62 S · cm²/mol CuSO₄, 144.90 S · cm²/mol KNO₃ and 128.80 S · cm²/mol to CaCO₃ (Atkins and de Paula, 2018) and V_{vol} corresponds to the effective interaction volume of the sensing region (Saleh. B., 2019). This quantity provides an additional descriptor associated with the electrical response of the analyte-loaded waveguide. Complementarily, standard performance metrics for integrated photonic sensing were extracted from the simulated electromagnetic response. The spectral sensitivity S_λ was evaluated as the variation of the effective modal index $n_{eff} = \beta\lambda/2\pi$ with respect to analyte concentration (Khonina et al., 2022)



$$S_{\lambda} = \frac{\partial n_{eff}}{\partial C}. \quad (13)$$

In addition, the electromagnetic interaction factor ξ was calculated as (Saleh and Teich, 2019; Han et al., 2022)

$$\xi = \frac{\int_{slot} |E|^2 dA}{\int_{total} |E|^2 dA}, \quad (14)$$

From these physically derived quantities, a simulation-driven electromagnetic dataset was generated using MATLAB by systematically varying both analyte concentration and operating wavelength. The resulting descriptors, including β , η , v_g , E_{eff} , P_{elec} , S_{λ} , and ξ , define a multidimensional electromagnetic signature associated with each chemical species. These descriptors were subsequently processed in Python to construct the feature space employed in the machine-learning classification stage.

3. Classification models

The theoretical optical responses obtained from the slot waveguide simulations serve as the input for a comparative classification study. We implemented four supervised learning architectures—DT, RF, XGBoost and MLP—to evaluate their capability to discriminate between chemical species based on multidimensional electromagnetic descriptors.

3.1 DT model

The DT classifier was used as an interpretable baseline to evaluate the discriminatory capacity of the electromagnetic descriptors. For a dataset $D = \{(x_i, y_i)\}_{i=1}^n$, the model recursively partitions the feature space using threshold rules $x_j \leq \tau$, selecting splits that minimize the Gini impurity, $G = 1 - \sum_{i=1}^K p_i^2$, where p_i denotes the class probability (Quinto, 2020). Model complexity was controlled by limiting tree depth and the minimum number of samples per split to reduce overfitting.

3.2 RF model

The RF was selected to improve robustness and reduce variance through bootstrap aggregation and random feature selection (Quinto, 2020). The model constructs an ensemble of decision trees trained on resampled subsets of the data, with each split considering a subset of features. Predictions are obtained via majority voting, $\hat{H}(x) = \arg \max_k \sum_{b=1}^B I(\hat{y}_b(x) = k)$, where B is the number of trees and $I(\cdot)$ the indicator function. Model complexity was controlled by tuning the number of estimators and maximum tree depth to balance accuracy and generalization.

3.3 XGBoost

The XGBoost was considered as a sequential ensemble-learning strategy based on gradient-boosted decision trees (Quinto, 2020). The model iteratively fits decision trees to minimize a regularized objective function combining a differentiable loss and a complexity penalty. At iteration t , the prediction is updated as $\hat{y}_i^{(t)} = \hat{y}_i^{(t-1)} + f_t(x_i)$, where f_t denotes the newly added tree. Regularization terms controlling

tree complexity were included to prevent overfitting. Key hyperparameters, including learning rate, maximum depth, and number of estimators, were optimized to balance bias-variance trade-offs.

3.4 MLP model

The MLP was employed to model nonlinear relationships in the electromagnetic feature space. For an input vector $x = [x_1, \dots, x_n]$, each neuron computes a linear transformation followed by a nonlinear activation, $z = w^T x + b$. The network parameters were optimized via backpropagation using stochastic gradient-based methods. To improve generalization, dropout regularization and validation-based early stopping were applied during training (Quinto, 2020).

All the above models are evaluated using standard classification metrics, including Accuracy, Recall, and the F1-score (Hastie, T., 2009). Complementarily, we used confusion matrices, learning curves and analysis to evaluate class separability and overall model discrimination performance.

4. Results

The classification performance of the DT, RF, XGBoost, and MLP models was evaluated using the electromagnetic dataset generated from the slot-waveguide simulations described in Section 2. The dataset consisted of 47,600 synthetic samples obtained under different ionic concentrations and operating wavelengths. The feature space was defined by 12 optical and physical descriptors derived from the effective refractive index, propagation constant, TM-mode response, confinement factor, group velocity, spectral sensitivity, slot voltage, electrical power density, interaction factor, and wavelength-dependent sensitivity variation. The classification task considered four analytes: CaCO_3 , CuSO_4 , KNO_3 and PbSO_4 . To assess robustness under realistic sensing conditions, all classifiers were evaluated using both clean and noisy datasets, where perturbations were introduced into the optical descriptors to emulate measurement uncertainty and fluctuations in the electromagnetic response of the sensor. Performance was quantified using accuracy, precision, recall, and F1-score metrics under a stratified partition consisting of 38,400 training samples and 9,600 testing samples.

The DT classifier was optimized through grid-search tuning using a maximum depth of 12, a minimum split size of 7 samples, 4 samples per leaf node, and cost-complexity pruning with $\alpha = 2.8 \times 10^{-4}$. Split selection was based on the log-loss criterion, while class imbalance was compensated using `class_weight = balanced`. Under clean conditions, the model achieved a validation accuracy of approximately 61.5%. As shown in Fig.2a, the learning curves exhibit limited convergence between training and validation accuracy, indicating restricted generalization despite the applied regularization strategy. The metrics summarized in Table 2a reveal heterogeneous performance across analytes, with KNO_3 reaching the highest F1-score of 0.75. Consistently, the confusion matrix in Fig.3a shows partial overlap between chemically similar optical signatures. Under noisy conditions, the validation accuracy shown in Fig. 2b remained substantially lower and weakly coupled to the training curve,

indicating limited robustness under perturbations in the electromagnetic feature space. This degradation is further reflected in Table 2b and Fig. 3b, where the off-diagonal overlap increases considerably relative to the clean-condition case. These results indicate that the threshold-based structure of a single DT is highly sensitive to fluctuations in the optical sensing variables.

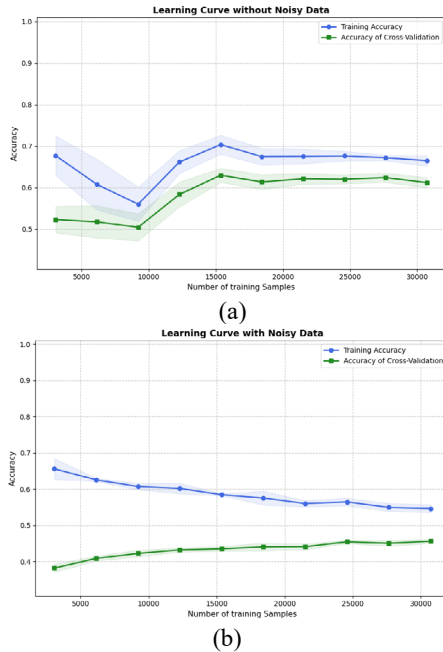


Figure 2. Learning curve with clean data (a) and with noisy data (b) of the DT classifier.

Table 2. Evaluation metrics by component of the DT model, using the clean data (a) and noisy data (b).

Chemical component	precision	recall	F1-score
$CaCO_3$	0.54	0.44	0.49
$CuSO_4$	0.66	0.71	0.69
KNO_3	0.75	0.76	0.75
$PbSO_4$	0.50	0.54	0.52

(a)

Chemical component	precision	recall	F1-score
$CaCO_3$	0.25	0.25	0.29
$CuSO_4$	0.61	0.61	0.56
KNO_3	0.58	0.68	0.63
$PbSO_4$	0.35	0.33	0.34

(b)

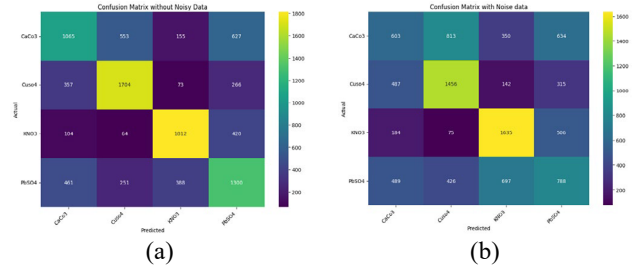


Figure 1. DT Confusion matrix, clean data(a), and noisy data (b).

For the noiseless dataset, the RF model achieved a validation accuracy of approximately 98%, substantially outperforming the DT classifier. As shown in Fig. 4a, the training and validation curves converge smoothly, indicating enhanced generalization through ensemble averaging. The RF classifier was configured as an ensemble of 200 decision trees with a maximum depth of 20 and a minimum of 2 samples per leaf node. Feature sampling at each split was restricted through $\text{max_features} = \text{sqrt}$, reducing inter-tree correlation and improving ensemble generalization across partially overlapping electromagnetic signatures.

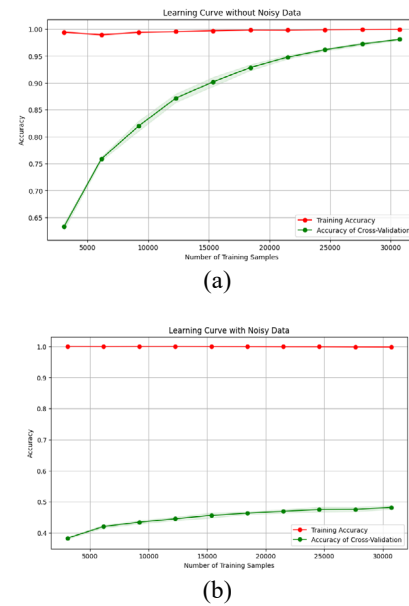


Figure 4. Learning curves of the RF classifier under clean (a) and noisy (b) sensing conditions, showing the degradation in generalization caused by perturbations in the electromagnetic descriptors.

Table 3. Precision, recall, and F1-score obtained with the Random Forest classifier using clean (a) and noisy (b) datasets.

Chemical component	precision	recall	F1-score
$CaCO_3$	0.48	0.47	0.47
$CuSO_4$	0.74	0.80	0.77
KNO_3	0.84	0.92	0.88
$PbSO_4$	0.58	0.50	0.54

(a)

Chemical component	precision	recall	F1-score
$CaCO_3$	0.34	0.27	0.30
$CuSO_4$	0.55	0.67	0.60
KNO_3	0.62	0.73	0.67
$PbSO_4$	0.36	0.29	0.32

(b)

The metrics summarized in Table 3a demonstrate strong classification performance across all analytes, with KNO_3 reaching the highest F1-score of 0.88.

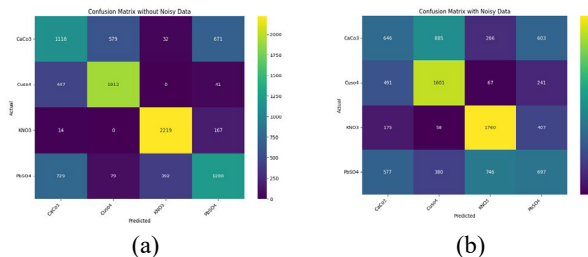


Figure 5. Confusion matrix with clean data (a) and Noisy data (b) for RF.

This behavior is consistent with the confusion matrix in Fig. 5a, where the dominant diagonal structure indicates high separability among the refractometric signatures. When noise was introduced, the validation accuracy shown in Figure 4b stabilized near 48% despite the training accuracy remaining close to unity, revealing reduced robustness under perturbed sensing conditions. The degradation observed in Table 3b and Figure 5b indicates that fluctuations in the electromagnetic descriptors increase overlap between chemically similar analytes, weakening the stability of the ensemble decision boundaries. The introduced noise particularly affects field-interaction variables, such as the ξ factor, weakening the correlation between ion concentration and the electromagnetic energy confined within the slot region. As the optical boundaries between analytes become less distinguishable, the classifier increasingly confuses noisy refractometric signatures despite the robustness provided by the ensemble architecture. These results suggest that fluctuations in the optical channel significantly affect the stability of the sensing response.

For the XGBoost evaluation, the classifier was optimized using a learning rate of 0.0996, maximum depth of 12, subsample = 0.6138, colsample_bytree = 0.9780, minimum samples per leaf parameter of 0.0010, min_child_weight = 1, and L1/L2 regularization coefficients of 0.0282 and 0.00028, respectively. Using the deterministic dataset, XGBoost achieved the highest discrimination capability among the tree-based models, reaching near-perfect classification performance. As shown in Fig. 6a, the training and validation curves rapidly converge, indicating efficient sequential residual optimization and strong nonlinear separability within the electromagnetic feature space.

The metrics summarized in Table 4a confirm high classification performance across all analytes, while the confusion matrix in Fig. 7a exhibits minimal off-diagonal overlap, demonstrating that the boosted ensemble effectively captures the refractometric differences between chemical species. After introducing perturbations into the electromagnetic descriptors, the boosted architecture exhibited reduced stability, with the validation accuracy in Fig. 6b saturating near 50.8% despite the continued increase in training performance. This degradation is further reflected in Table 4b and Fig. 7b, where the overlap between chemically similar analytes increases considerably relative to the deterministic case. The metrics reported in Table 4a further indicate that $CuSO_4$ and KNO_3 exhibit the highest separability within the electromagnetic feature space, which is consistent with the dominant diagonal structure observed in Figure 7a.

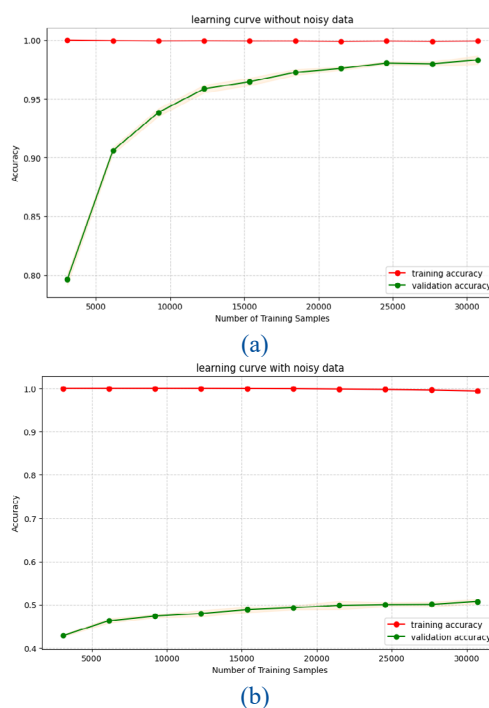


Figure 6. Learning curves without clean data (a) and Noisy data (b) for XGBoost.

Table 4. Evaluation metrics by component of the XGBoost model, using the clean data (a) and noisy data (b).

Chemical component	precision	recall	F1-score
$CaCO_3$	0.97	0.98	0.98
$CuSO_4$	1.0	1.0	1.0
KNO_3	1.0	1.0	1.0
$PbSO_4$	0.98	0.98	0.98

(a)

Chemical component	precision	recall	F1-score
--------------------	-----------	--------	----------



$CaCO_3$	0.36	0.31	0.33
$CuSO_4$	0.59	0.66	0.62
KNO_3	0.67	0.75	0.71
$PbSO_4$	0.37	0.33	0.35

(b)

Under noisy conditions, the overlap between these refractometric signatures increases, reducing the classification stability of the boosted ensemble.

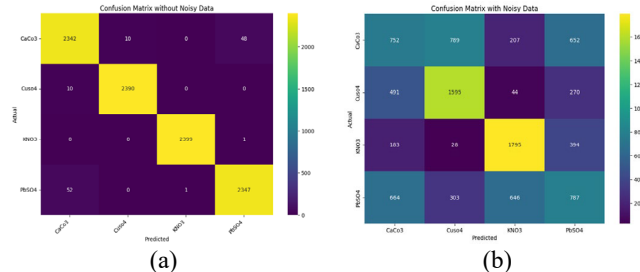


Figure 7. Confusion matrix with clean data (a) and with Noisy data (b) for XGBoost.

The MLP was implemented as a fully connected architecture with three hidden layers with 100, 50, and 25 neurons using ReLU activation functions, followed by a four-neuron Softmax output layer for multiclass classification. Training was conducted via the Adam optimizer with a categorical initial learning rate of 0.001, cross-entropy loss, L2 regularization, and early stopping to improve generalization under noisy sensing conditions.

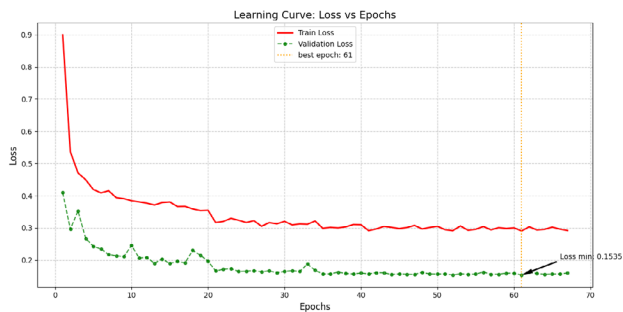


Figure 8. Learning curve of MLP.

The training dynamics of the network are illustrated in Figure 8, where both training and validation losses decrease monotonically and stabilize after approximately 60 epochs without significant divergence between the curves. This behavior indicates stable optimization and limited overfitting during the learning process. Unlike the tree-based models, the MLP preserved high classification stability under both deterministic and noisy conditions. As shown in Figure 9a, the training and validation accuracy curves rapidly converge toward near-unity values, indicating efficient learning of the nonlinear relationships embedded within the electromagnetic feature space. The metrics summarized in Table 5a confirm strong classification performance across all analytes, while the confusion matrix in Figure 10a exhibits minimal off-diagonal

overlap, demonstrating that the neural-network architecture effectively preserves the separability of the refractometric signatures under clean sensing conditions. Under noisy conditions, the MLP maintained substantially higher validation accuracy than DT, RF, and XGBoost models, as observed in Figure 9b. Although perturbations in the electromagnetic descriptors introduce a moderate reduction in classification performance, the metrics reported in Table 5b remain consistently high across all analytes, particularly for $CuSO_4$ and $KN,,,O_3$ which preserve the highest F1-scores within the evaluated dataset.

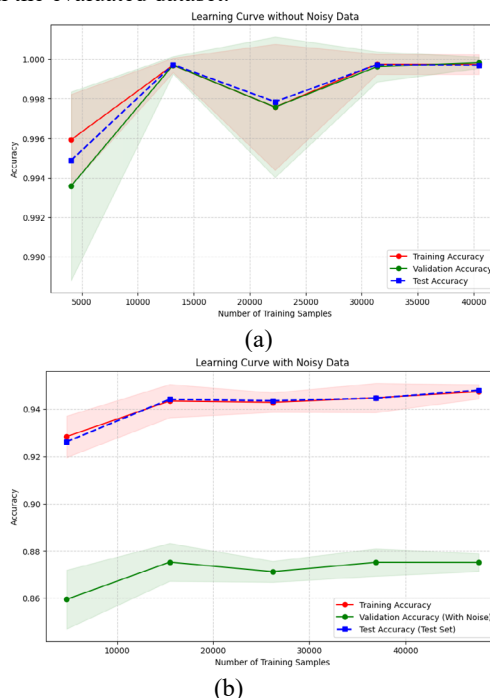


Figure 9. Learning Curves with clean data (a) and with Noisy data (b) for MLP.

Table 5. Evaluation metrics by component of the MLP model – Validation Set, using the clean dataset (a) and Noisy dataset (b).

Chemical component	precision	recall	F1-score
$CaCO_3$	1.0	1.0	1.0
$CuSO_4$	1.0	1.0	1.0
KNO_3	1.0	1.0	1.0
$PbSO_4$	1.0	1.0	1.0

(a)

Chemical component	precision	recall	F1-score
$CaCO_3$	0.93	0.90	0.92
$CuSO_4$	0.98	0.98	0.98
KNO_3	0.99	1.0	1.0
$PbSO_4$	0.92	0.95	0.93

(b)

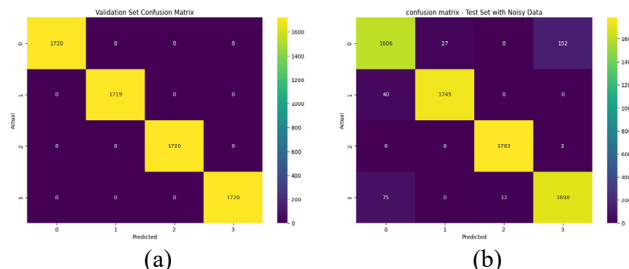


Figure 10. Confusion matrix without Noisy (a) and with Noisy (b) for MLP.

Consistently, the confusion matrix in Figure 10b shows only limited off-diagonal overlap, indicating that the smoother nonlinear decision boundaries learned by the neural network improve tolerance to fluctuations in the optical sensing variables and preserve the discrimination capability of the electromagnetic feature space under uncertainty.

6. CONCLUSIONS

We presented a theoretical framework combining slot-waveguide electromagnetic modeling with machine-learning classification for the analysis of dissolved chemical species in water. Synthetic electromagnetic descriptors derived from the TM-mode response enabled the construction of a multidimensional optical feature space suitable for data-driven chemical discrimination. The results indicate that tree-based models, DT, RF and XGBoost, preserve strong discrimination under deterministic conditions but exhibit limited robustness once perturbations distort the optical descriptors. In contrast, the MLP maintained higher classification stability under noisy conditions, suggesting improved nonlinear generalization within the simulated electromagnetic feature space. Although the present study is based on synthetic data generated from theoretical electromagnetic simulations, the proposed framework demonstrates the feasibility of integrating computational photonics and machine learning for intelligent optical sensing applications. Future work will incorporate experimentally acquired signals and more realistic perturbation models to evaluate the transferability of the approach under practical sensing conditions.

REFERENCES

- Agrawal, G. P. (2005) *Lightwave technology: components and devices*. Hoboken, NJ: Wiley-Interscience.
- Almeida, V. R., et al. (2004) 'Guiding and confining light in void nanostructure', *Optics Letters*, 29(11), pp. 1209–1211. doi: 10.1364/OL.29.001209.
- Atkins, P. W. and de Paula, J. (2018) *Physical Chemistry*. 11^a ed. Oxford: Oxford University Press.
- Camacho, L. M. et al. (2011) 'Occurrence and treatment of arsenic in groundwater and soil in northern Mexico and southwestern USA', *Chemosphere*, 83(3), pp. 211–225. doi: 10.1016/j.chemosphere.2010.12.067.
- Djam, S. et al. (2022). Assessment of significant anions of nitrite, nitrate, fluoride, chloride, sulfate and phosphate in mineral and drinking bottled waters and their roles in contamination. *Nutr. Food Sci. Res*, 9(2), 17–21.
- Han, J. et al. (2022) 'Highly Sensitive Liquid M-Z Waveguide Sensor Based on Polymer Suspended Slot Waveguide Structure', *Polymers*, 14(19), p. 3967. doi: 10.3390/polym14193967.
- Hastie, T., Tibshirani, R. y Friedman, J. (2009) *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. 2^a ed. Nueva York: Springer.
- Khonina, S. et al. (2022) 'Environmental Monitoring: A Comprehensive Review on Optical Waveguide and Fiber-Based Sensors', *Biosensors*, 12(11), p. 1038. doi: 10.3390/bios12111038.
- Liu, Y., Kong, M. y Jiang, Y. (2015) 'Transverse magnetic modes in planar slot waveguides', *Journal of the Optical Society of America B*, 32(10), pp. 2052–2060. doi: 10.1364/JOSAB.32.002052.
- Mamtmin, G. et al. (2020) '5, 10, 15, 20-tetrakis-(4-methoxyphenyl) porphyrin film/K⁺ ion-exchanged optical waveguide gas sensor', *Optics and Laser Technology*, 128, p. 106260. doi: 10.1016/j.optlastec.2020.106260.
- Pérez-Beltrán, C. H. et al. (2024) 'Artificial intelligence and water quality: From drinking water to wastewater', *TrAC Trends in Analytical Chemistry*, 172, p. 117597. doi: 10.1016/j.trac.2024.117597.
- Quinto, B. (2020). Introduction to machine learning. In *Next-Generation Machine Learning with Spark: Covers XGBoost, LightGBM, Spark NLP, Distributed Deep Learning with Keras, and More* (pp. 1-27). Berkeley, CA: Apress.
- Saleh, B. E. A. y Teich, M. C. (2019). *Fundamentals of Photonics*. 3^a ed. Hoboken, NJ: Wiley.
- Santucci, R. J. y Scully, J. R. (2020) 'The pervasive threat of lead (Pb) in drinking water: Unmasking and pursuing scientific factors that govern lead release', *Proceedings of the National Academy of Sciences*, 117(38), pp. 23211–23218. doi: 10.1073/pnas.1913749117.
- SECRETARÍA DE SALUD. (2022). Agua para uso y consumo humano. Límites permisibles de la calidad del agua. NOM-127-SSA1-2021. Diario Oficial de la Federación, 2 de mayo de 2022.
- Srivastav, A. L., & Ranjan, M. (2020). Inorganic water pollutants. In *Inorganic pollutants in water* (pp. 1-15). Elsevier.
- Talebian, E. y Talebian, M. (2013) 'A general review on the derivation of Clausius-Mossotti relation', *Optik*, 124(16), pp. 2324–2326. doi: 10.1016/j.ijleo.2012.06.075.

Web system for the process of managing climatological information of the state of Zacatecas

Núñez Adame Jose de Jesús, Torres Hernández Luis Eduardo, López Veyna Jaime Iván, Márquez Martínez Luis Enrique

Área de participación: Sistemas Computacionales

**Tecnológico Nacional de México / Instituto Tecnológico de Zacatecas, Zacatecas, 98000, México.*

(m19450853@itz.edu.mx, luis.torres@zacatecas.tecnm.mx, ivanlopezveyna@itz.edu.mx, luisse.mm@itz.edu.mx)

Abstract: This paper describes the design and implementation of a web-based system for the management of meteorological data generated by conventional weather stations of the Meteorological Forecast Center of the National Water Commission (CONAGUA), Zacatecas Local Directorate. The system aims to improve the efficiency of data capture, storage, and visualization processes. The development was carried out using the Scrum agile methodology, enabling an iterative and incremental approach. The Laravel framework was used under the Model-View-Controller (MVC) architectural pattern to ensure proper system organization and scalability. To evaluate its effectiveness, a quantitative experimental study was conducted comparing the traditional data capture method with the proposed web system. The evaluation considered data entry time and the number of errors generated. The results were analyzed using Student's t-test for independent samples, obtaining statistically significant improvements ($p < 0.05$), reducing both data entry time and the number of errors generated. These findings demonstrate that the web system improves the efficiency of meteorological data management, supporting better decision-making.

Keywords: Web system; meteorological data; data management; Scrum; Laravel; experimental evaluation.

1. INTRODUCCION

La medición de datos meteorológicos es importante para predecir y comprender las condiciones del clima, esto es considerable para diferentes actividades como la agricultura, aviación y respuesta ante desastres naturales. Ahí es donde las estaciones climatológicas terrestres han sido la principal fuente de datos (Acosta y Meléndez, 2024). En la figura 1 se muestra una fotografía de una estación meteorológica.



Figura 1. Estación Meteorológica (fuente propia)

Actualmente, la gestión de datos meteorológicos enfrenta dificultades relacionadas con la organización, almacenamiento y consulta eficiente de información. Estas limitaciones afectan

la disponibilidad de datos necesarios para el análisis meteorológico y la toma de decisiones.

Ante esta problemática, surge la necesidad de implementar una herramienta que permita sistematizar los procesos de captura, almacenamiento, procesamiento y visualización de la información.

La presente investigación se desarrolla en el área de Tecnologías de la Información, específicamente en el campo de la Ingeniería de Software aplicada a la gestión de datos meteorológicos. Este trabajo está enfocado en el diseño, desarrollo e implementación de un sistema web orientado a optimizar la administración de la información generada por las estaciones climatológicas convencionales del Centro de Previsión Meteorológica de la Comisión Nacional del Agua (CONAGUA), Dirección Local Zacatecas.

1.1 Objetivo General

Desarrollar e implementar un sistema web para la gestión de datos meteorológicos que permita optimizar el proceso de captura de información y mejorar la calidad, organización y precisión de los datos climatológicos en el Centro de Previsión Meteorológica de la Comisión Nacional del Agua en el estado de Zacatecas.

1.2 Justificación



Las estaciones meteorológicas constituyen una de las principales fuentes para la obtención de información atmosférica relevante, como la temperatura, humedad, presión y velocidad del viento, datos fundamentales para el análisis del comportamiento climático y la toma de decisiones en diversos sectores (Acosta y Meléndez, 2024).

En este contexto, la Comisión Nacional del Agua (CONAGUA), Dirección Local Zacatecas, cuenta con una red de 136 estaciones climatológicas distribuidas en el estado, las cuales generan diariamente un volumen considerable de información. Esta situación implica la necesidad de contar con mecanismos eficientes para la captura, almacenamiento y gestión de los datos meteorológicos.

Sin embargo, el proceso actual de registro presenta limitaciones asociadas al uso de métodos tradicionales, lo que puede generar errores, inconsistencias y tiempos elevados en la captura de información, afectando la calidad y disponibilidad de los datos para su análisis, como se muestra en la figura 2.



Figura 2. Flujo de captura convencional (fuente propia)

La captura de datos climatológicos se realiza en hojas de cálculo, como se observa en la figura 3, presentando diversas vulnerabilidades que comprometen la calidad y confiabilidad de la información meteorológica. Entre ellas destacan los errores de validación, como la presencia de valores atípicos o inconsistencias entre variables, los cuales pueden derivar en resultados incorrectos durante el análisis y la generación de pronósticos.

TERMINOMETRO AL ABRIGO			LLUVIA EN mm		MICRÓMETRO		ESTADO DEL TIEMPO A LA HORA DE LA OBSERVACIÓN							ESTADO DEL TIEMPO HS. ANTERIORES A OBSERVACIÓN			
AMBIENTE	MAXIMA	MINIMA	RELATIVIDAD %	LECTURA	LECTURA	EVAP	HELADA	VENTO	VEL	VIS	CIELO	FENOMENOS	TEMPERATURA	GRADOS	HELADA	VELO	
1	4.0	25.0	2.5	0.0	19.11	4.00	0	N	1	7	7	0	0	0	0	0	
2	7.6	23.6	3.6	0.0	15.11	5.10	0	N	1	7	7	0	0	0	0	0	
3	6.5	29.0	3.6	0.0	14.37	149.50	5.08	0	SW	1	7	7	0	0	0	0	
4	7.0	29.5	4.5	0.0	13.75	5.08	0	N	1	7	7	0	0	0	0	0	
5	7.6	29.0	4.5	0.0	13.53	3.41	0	N	1	7	7	0	0	0	0	0	
6	7.0	29.5	4.5	0.0	133.96	3.30	0	N	1	7	7	0	0	0	0	0	
7	7.0	29.5	4.5	0.0	129.96	4.41	0	SW	1	7	7	0	0	0	0	0	
8	7.0	29.5	4.5	0.0	129.06	4.54	0	N	1	7	7	0	0	0	0	0	
9	7.0	29.5	4.5	0.0	121.29	3.29	0	N	1	7	7	0	0	0	0	0	
10	7.0	29.5	4.5	0.0	113.52	4.41	0	N	1	7	7	0	0	0	0	0	
11	7.0	29.5	4.5	0.0	129.79	4.50	0	N	1	7	7	0	0	0	0	0	
12	7.0	29.5	4.5	0.0	121.52	4.35	0	NW	1	7	7	0	0	0	0	0	
13	7.0	29.5	4.5	0.0	113.96	4.59	0	N	1	7	7	0	0	0	0	0	
14	7.0	29.5	4.5	0.0	109.16	4.55	0	N	1	7	7	0	0	0	0	0	
15	7.0	29.5	4.5	0.0	108.82	4.59	0	N	1	7	7	0	0	0	0	0	
16	7.0	29.5	4.5	0.0	101.98	4.55	0	NW	1	7	7	0	0	0	0	0	
17	7.0	29.5	4.5	0.0	109.16	4.35	0	N	1	7	7	0	0	0	0	0	
18	7.0	29.5	4.5	0.0	96.19	4.59	0	N	1	7	7	0	0	0	0	0	
19	7.0	29.5	4.5	0.0	81.52	4.35	0	N	1	7	7	0	0	0	0	0	
20	7.0	29.5	4.5	0.0	85.08	3.33	0	N	1	7	7	0	0	0	0	0	
21	7.0	29.5	4.5	0.0	73.33	4.59	0	N	1	7	7	0	0	0	0	0	
22	7.0	29.5	4.5	0.0	78.29	4.35	0	NW	1	7	1 ROCIO	0	0	0	0	0	
23	7.0	29.5	4.5	0.0	67.32	3.89	0	N	1	7	7	0	0	0	0	0	
24	7.0	29.5	4.5	0.0	61.69	3.41	0	N	1	7	7	0	0	0	0	0	
25	1.5	27.5	4.5	0.0	66.67	3.93	0	N	1	7	7	0	0	0	2	0	
26	1.5	25.9	1.0	0.0	62.63	3.93	0	N	1	7	7	0	0	0	2	1	

Figura 3. Formato de captura convencional (fuente propia)

Asimismo, la naturaleza manual del proceso incrementa el riesgo de errores de transcripción y cálculos incorrectos, afectando la precisión de los datos y, en consecuencia, la toma de decisiones basada en esta información.

Este proyecto contribuye a mejorar la gestión de datos meteorológicos, fortaleciendo la calidad, integridad y confiabilidad de la información, y proporcionando una herramienta que puede ser replicada en instituciones con necesidades similares.

2. REVISION DE LA LITERATURA

Diversos trabajos han abordado la gestión de datos meteorológicos mediante el desarrollo de sistemas orientados a mejorar la organización, procesamiento y calidad de la información. En este contexto, estudios como los de Navarro-Silva et al. (2020) y Pablo Báez-Vásquez y Portilla-Flores (2022) proponen sistemas de gestión enfocados en el manejo de información climatológica y la integración de mecanismos de verificación y homogeneización de datos.

Sin embargo, estos enfoques se centran principalmente en el procesamiento y análisis de datos una vez capturados, dejando en segundo plano la optimización del proceso de captura, que constituye una etapa crítica donde se originan errores e inconsistencias. En contraste, el presente trabajo se enfoca en mejorar dicha etapa mediante la implementación de validaciones en tiempo real durante el registro de la información, reduciendo errores desde su origen.

Por otro lado, se han desarrollado soluciones orientadas a la visualización y análisis de datos meteorológicos utilizando herramientas como Python y la librería Dash, como se observa en el trabajo de Pérez Bellido (2024). Estas plataformas permiten representar información mediante gráficos y paneles interactivos, facilitando su interpretación por parte de los usuarios.

No obstante, estos sistemas priorizan la fase de análisis y presentación de datos, sin abordar de manera directa la eficiencia en la captura de la información.

En el ámbito de la automatización, diversos estudios han explorado el uso de tecnologías como el Internet de las Cosas (IoT) para la recolección de datos meteorológicos en tiempo real. Por ejemplo, Montaña-Blacio et al. (2023) y Pereto Soler (2024) presentan soluciones basadas en sensores que permiten monitorear variables ambientales y transmitir datos de forma automatizada.

Si bien estos enfoques representan avances importantes en la automatización de la captura de datos, su implementación requiere infraestructura especializada, lo cual puede limitar su adopción en entornos donde aún predominan procesos manuales, como en muchas instituciones públicas. En este sentido, la presente investigación propone una solución orientada a optimizar los procesos manuales existentes, ofreciendo una alternativa viable y de bajo costo que mejore la

eficiencia sin requerir cambios significativos en la infraestructura.

En el ámbito del desarrollo de software, el uso de metodologías ágiles como Scrum ha demostrado ser efectivo para la gestión de proyectos tecnológicos. Estudios como el de Bautista-Villegas (2022) destacan su capacidad para adaptarse a cambios en los requerimientos y mejorar la organización del trabajo.

No obstante, estos estudios se enfocan principalmente en aspectos técnicos del desarrollo de software, sin evaluar de manera cuantitativa el impacto de los sistemas implementados en procesos específicos. El presente trabajo no solo aborda el desarrollo del sistema, sino que también incluye una validación experimental que permite medir su impacto en términos de tiempo de captura, reducción de errores y mejora en la productividad.

Finalmente, en sectores como la agricultura, diversos estudios han resaltado la importancia de los datos meteorológicos para la toma de decisiones, como lo señalan Adalberto y Casadiego (2020) y Quintero (2021). Estos trabajos destacan el valor de contar con información confiable y oportuna para optimizar procesos productivos. A continuación, en la tabla 1 se presenta una comparación de la literatura.

Tabla 1. Comparativa literatura (fuente propia)

Trabajo	Enfoque principal	Aportación	Diferencia con el presente trabajo
Navarro-Silva et al. (2020)	Gestión de información meteorológica	Organización y procesamiento de datos	Se enfoca en mejorar la captura mediante validaciones en tiempo real
Báez-Vásquez y Portilla-Flores (2022)	Gestión de datos climáticos	Integración de mecanismos de verificación	Reduce errores desde el origen durante el registro
Pérez Bellido (2024)	Visualización de datos (Dash)	Representación gráfica e interacción	Incrementa productividad y reduce tiempo de registro
Montaño-Blacio et al. (2023)	IoT para monitoreo ambiental	Captura automatizada de datos	Enfocado en procesos manuales existentes sin necesidad de IoT
Pereto Soler (2024)	Monitoreo con sensores y app móvil	Recolección en tiempo real	Solución adaptable a entornos con recursos limitados
Bautista-Villegas (2022)	Metodologías ágiles (Scrum)	Mejora en gestión del desarrollo	Incluye validación experimental del sistema
Presente trabajo	Optimización de captura de datos meteorológicos	Reducción de tiempo, errores y mejora en productividad	Validación experimental en entorno real con impacto operativo

En este contexto, la presente investigación contribuye al fortalecimiento de la calidad de los datos meteorológicos desde su origen, mediante la mejora del proceso de captura, lo que impacta directamente en la confiabilidad de la información utilizada para la toma de decisiones.

3. METODOLOGIA

3.1 Metodología de desarrollo

Para el desarrollo del sistema web se empleó la metodología ágil Scrum. Como se muestra en la figura 4, este enfoque se basa en ciclos iterativos e incrementales que permiten la entrega continua de funcionalidades. Esta metodología permite una mejor adaptación a cambios en los requerimientos, favorece la organización del trabajo y mejora la eficiencia durante el proceso de desarrollo de aplicaciones web (Bautista-Villegas, 2022).

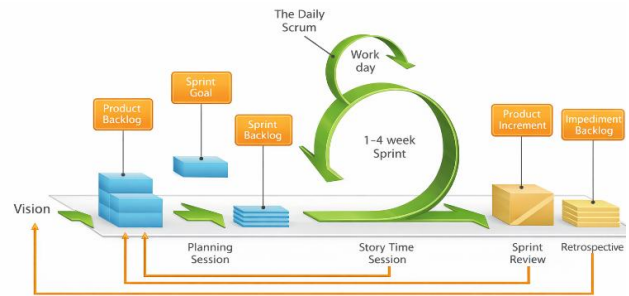


Figura 4. Procesos de Scrum (Avilés Matute et al., 2020)

El desarrollo se estructuró en seis *Sprints* organizados en dos iteraciones. Cada *Sprint* incluyó las fases de análisis de requisitos, diseño, desarrollo, pruebas e implementación, permitiendo la entrega de incrementos funcionales del sistema al final de cada ciclo.

Durante la fase de análisis, los requerimientos del sistema fueron definidos mediante historias de usuario, lo que permitió identificar las funcionalidades principales, entre las que destacan: autenticación de usuarios, gestión de estaciones climatológicas, captura de datos y visualización de información.

La primera iteración se enfocó en el desarrollo de funcionalidades básicas del sistema, incluyendo el acceso, registro de usuarios y gestión inicial de estaciones. La segunda iteración estuvo orientada al desarrollo del módulo de captura de datos climatológicos y su posterior visualización mediante herramientas gráficas. A continuación, en la tabla 2 se presentan las actividades de desarrollo realizadas.

Tabla 2. Actividades de desarrollo (fuente propia)

Actividades			
Iteración	Sprint	HU_ID	DESCRIPCION
1	1	HU01	BIENVENIDA DEL SISTEMA
		HU02	INICIO DE SESIÓN
		HU03	REGISTRO
		HU04	ROLES DE USUARIO
	2	HU05	MÓDULO DE USUARIO
	3	HU06	MÓDULO DE ESTACIONES
2	4	HU07	MÓDULO CAPTURA FRONTEND
	5	HU08	MÓDULO CAPTURA BACKEND
	6	HU09	MÓDULO VISUALIZACIÓN DE DATOS



En cuanto a las tecnologías utilizadas, el sistema fue desarrollado en PHP utilizando el marco de trabajo Laravel, bajo el patrón de arquitectura Modelo-Vista-Controlador (MVC), lo que permitió una adecuada organización del código y facilitó su mantenimiento. Para el control de versiones se utilizó GitLab, mientras que para el diseño de interfaces se emplearon herramientas como AdminLTE y Balsamiq Mockups.

3.2 Diseño experimental y validación

Para evaluar la efectividad del sistema web propuesto, se realizó un estudio cuantitativo con diseño experimental de grupos independientes, enfocado en comparar el método tradicional de captura de datos con el sistema desarrollado.

La población de estudio estuvo conformada por personal operativo del Centro de Previsión Meteorológica de la Comisión Nacional del Agua (CONAGUA), Dirección Local Zacatecas, quienes realizan actividades de captura de datos climatológicos de manera habitual.

A partir de esta población, se seleccionó una muestra de $n = 6$ participantes, todos con experiencia en el registro de información meteorológica, con una experiencia promedio superior a 5 años, lo que garantiza un conocimiento adecuado del proceso de captura.

Los participantes fueron asignados aleatoriamente en dos grupos:

- Grupo de control (GC): utilizó el método tradicional basado en hojas de cálculo.
- Grupo experimental (GE): utilizó el sistema web desarrollado.

El tamaño de la muestra ($n = 6$) corresponde a la totalidad del personal operativo disponible en el Centro de Previsión Meteorológica de la Comisión Nacional del Agua (CONAGUA), Dirección Local Zacatecas, que cuenta con capacitación específica en la captura de datos meteorológicos.

En este sentido, el estudio se desarrolló bajo un enfoque de muestra censal, ya que se incluyó el universo completo de capturistas especializados disponibles en la dependencia al momento de la investigación.

No obstante, se reconoce que el tamaño de la muestra es limitado, por lo que los resultados deben interpretarse dentro del contexto específico del estudio. Sin embargo, la homogeneidad del perfil de los participantes (experiencia superior a 5 años) contribuye a reducir la variabilidad y fortalece la validez interna del experimento.

Adicionalmente, la variabilidad entre usuarios fue minimizada mediante la selección de participantes con experiencia similar en la captura de datos meteorológicos, permitiendo condiciones más homogéneas durante la ejecución del experimento.

Se definieron como variables principales: la variable independiente, correspondiente al uso del sistema web para la captura de datos, y la variable dependiente, asociada a la eficiencia en la captura, medida a través del tiempo de registro y el número de errores cometidos.

3.3 Procedimiento experimental

Para el diseño de la presente investigación se siguieron las fases para la realización de un experimento propuestas por (Vizcaíno Zúñiga, Cedeño Cedeño, y Maldonado Palacios, 2023), adaptándolas al contexto de datos meteorológicos. A continuación, se describen dichas fases:

1. **Selección y asignación de participantes:** Se seleccionaron seis trabajadores del Centro de Previsión Meteorológica de la Comisión Nacional del Agua (CONAGUA), Dirección Local Zacatecas, todos con experiencia en la captura de datos climatológicos (experiencia promedio superior a 5 años).

Los participantes fueron asignados aleatoriamente en dos grupos:

- Grupo de control (método tradicional)
- Grupo experimental (uso del sistema web)

Esta asignación permitió reducir sesgos y garantizar la validez interna del experimento.

El experimento se desarrolló en una sesión controlada durante una jornada de trabajo, en la cual ambos grupos realizaron las tareas de captura bajo las mismas condiciones operativas y utilizando el mismo volumen de información meteorológica.

2. **Manipulación de la variable independiente:** La variable independiente correspondió al método de captura de datos, evaluándose dos condiciones:

- Captura mediante hojas de cálculo (método tradicional)
- Captura mediante el sistema web desarrollado

Previo a la ejecución del experimento, el grupo experimental recibió una capacitación de aproximadamente 4 horas sobre el uso del sistema web, mientras que el grupo de control continuó utilizando el método tradicional con el que ya estaba familiarizado.

Cada participante realizó la captura del mismo conjunto de registros meteorológicos, garantizando igualdad en la carga de trabajo, volumen de información y complejidad de las tareas entre ambos grupos. Asimismo, se utilizaron las mismas condiciones operativas y el mismo tipo de datos para todos los participantes.

3. **Medición de la variable dependiente:** La eficiencia del proceso de captura fue evaluada mediante dos indicadores principales:



Tiempo de captura: medido desde el inicio hasta la finalización de cada tarea utilizando un cronómetro. Los tiempos obtenidos fueron registrados en hojas de control para su posterior análisis estadístico.

Número de errores: identificado mediante validaciones automáticas del sistema en el grupo experimental y mediante la comparación con un conjunto de datos de referencia en ambos grupos. El número de errores obtenidos fue registrado en hojas de control para su posterior análisis estadístico. Se consideraron como errores los valores fuera de rango, inconsistencias entre variables.

4. **Control de variables de confusión:** Para garantizar la validez del experimento, se controlaron factores que pudieran influir en los resultados, tales como:
- Nivel de experiencia de los participantes
 - Condiciones del entorno (equipo de cómputo y software)
 - Complejidad de las tareas
 - Tipo y volumen de datos utilizados

Estas condiciones se mantuvieron constantes para ambos grupos.

5. **Análisis de datos:** Los datos obtenidos fueron analizados mediante la prueba t de *Student* para muestras independientes, con el objetivo de identificar diferencias significativas entre el grupo de control y el grupo experimental, particularmente en términos de tiempo de captura y número de errores.

3.4 Análisis de datos

Para el análisis de los datos se empleó un enfoque estadístico descriptivo e inferencial, con el objetivo de evaluar las diferencias entre el grupo de control y el grupo experimental en términos de eficiencia en la captura de datos meteorológicos.

En primera instancia, se calcularon estadísticas descriptivas para cada variable analizada, incluyendo la media (μ) y la desviación estándar (σ), tanto para el tiempo de captura como para el número de errores. Asimismo, se incorporó la métrica de productividad, definida como la cantidad de registros capturados por unidad de tiempo, con el fin de evaluar la eficiencia operativa del proceso.

Posteriormente, se verificó el supuesto de normalidad de los datos mediante la prueba de Shapiro-Wilk, considerando un nivel de significancia de $\alpha = 0.05$. Los resultados indicaron que no se rechaza la hipótesis de normalidad ($p > 0.05$), lo que permitió el uso de pruebas paramétricas.

Dado que los datos cumplieron con los supuestos de normalidad e independencia de las observaciones, se aplicó la prueba t de *Student* para muestras independientes, con el propósito de comparar las medias entre el grupo de control y

el grupo experimental. El nivel de significancia establecido fue de $\alpha = 0.05$.

El procesamiento y análisis estadístico de los datos se realizó mediante Microsoft Excel y RStudio, utilizando herramientas para estadística descriptiva, pruebas de normalidad, pruebas inferenciales y generación de representaciones gráficas.

Adicionalmente, se calcularon los intervalos de confianza al 95% (IC 95%) para la diferencia de medias, estimándose un rango aproximado de [-12.5, -9.7] minutos para el tiempo de captura y [-3.0, -0.9] errores para la variable de errores, lo que refuerza la consistencia de los resultados observados.

Finalmente, los resultados fueron representados mediante tablas comparativas, gráficas de barras y diagramas de caja (*boxplots*), permitiendo visualizar las diferencias entre ambos grupos y la dispersión de los datos.

3.5 Arquitectura y funcionamiento del sistema

El sistema web desarrollado se basa en una arquitectura de tres capas sustentada en el patrón Modelo-Vista-Controlador (MVC), ampliamente utilizado en el desarrollo de aplicaciones web debido a su capacidad para dividir la funcionalidad del software en componentes bien definidos.

El patrón Modelo-Vista-Controlador (MVC) permite separar la lógica de negocio, la interfaz de usuario y el control de las solicitudes del sistema en tres componentes principales: Modelo, encargado de gestionar la lógica de datos y las interacciones con la base de datos; Vista, responsable de presentar la información al usuario; y Controlador, que actúa como intermediario entre ambos componentes, coordinando las solicitudes y respuestas del sistema (Fonseca Márquez, 2025).

El sistema fue desarrollado utilizando el *framework* Laravel bajo el lenguaje PHP, lo que permitió una adecuada organización del código, facilitando su mantenimiento y escalabilidad. Asimismo, para el control de versiones se empleó GitLab y para el diseño de interfaces se utilizaron herramientas como AdminLTE y Balsamiq Mockups.

La aplicación se compone de diversos módulos que permiten gestionar de manera integral la información meteorológica. En la tabla 3 se muestran módulos desarrollados:

Tabla 3. Módulos del sistema (fuente propia)

Módulo	Descripción
Autenticación	Encargado de gestionar el acceso de usuarios al sistema, garantizando la seguridad de la información.
Gestión de estaciones	Permite registrar, actualizar y consultar información relacionada con las estaciones climatológicas.
Captura de datos	Facilita el registro de datos meteorológicos mediante formularios con validaciones automáticas.
Visualización	Permite consultar y analizar la información mediante tablas y representaciones gráficas.

El funcionamiento del sistema inicia con la autenticación del usuario, quien accede a la plataforma mediante credenciales



previamente registradas. Posteriormente, el usuario selecciona la estación climatológica correspondiente y realiza la captura de datos a través de formularios diseñados para validar la información en tiempo real.

Una vez capturados, los datos son procesados por la lógica del sistema, donde se aplican reglas de validación para evitar inconsistencias. Finalmente, la información es almacenada en la base de datos y puede ser consultada o visualizada mediante los módulos correspondientes.

En la figura 5 se muestra el flujo general de funcionamiento del sistema web.

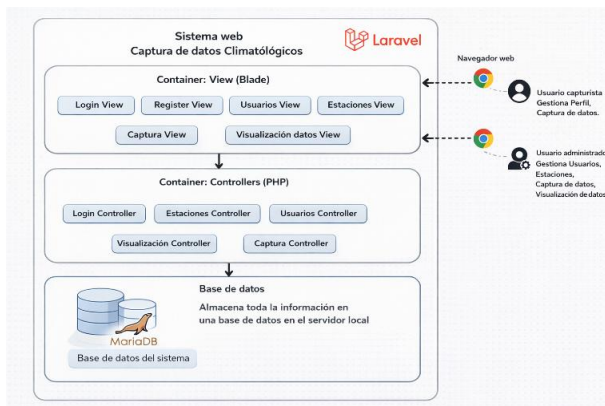


Figura 5. Flujo sistema web (fuente propia)

Este flujo permite reducir errores humanos y mejorar la eficiencia en el proceso de captura de datos meteorológicos.

4. RESULTADOS

Los resultados obtenidos a partir del experimento realizado permitieron evaluar el impacto del sistema web en el proceso de captura de datos meteorológicos. Para ello, se analizaron las variables de tiempo de captura, número de errores y productividad.

La tabla 4 muestra que el grupo experimental obtuvo menores tiempos de captura y ausencia de errores respecto al grupo de control, además de una mayor productividad promedio.

Tabla 4. Estadísticos descriptivos (fuente propia)

Grupo	n	Tiempo (min) $\mu \pm \sigma$	Errores $\mu \pm \sigma$	Productividad (reg/min)
Experimental	12	8.52 $\pm$ 1.73	0.00 $\pm$ 0.00	11.54
Control	12	19.48 $\pm$ 1.62	1.92 $\pm$ 1.44	5.01

Los resultados muestran que el grupo experimental presentó una reducción significativa en el tiempo de captura respecto al grupo de control.

El análisis estadístico mediante la prueba t de *Student* para muestras independientes arrojó un valor de:

- $t = -17.233$
- $p\text{-value} < 0.05$

lo que indica que la diferencia observada es estadísticamente significativa.

Por otra parte, respecto al número de errores, se observó una disminución considerable en el grupo experimental, el cual no presentó errores durante el proceso de captura. La prueba t de *Student* mostró los siguientes resultados:

- $t = -10.34$
- $p\text{-value} < 0.05$

confirmando igualmente diferencias significativas entre ambos grupos.

Adicionalmente, los intervalos de confianza al 95% (IC 95%) mostraron un rango aproximado de [-12.5, -9.7] minutos para el tiempo de captura y [-3.0, -0.9] errores para la variable de errores, reforzando la consistencia de los resultados observados.

En términos de productividad, el sistema web mostró un incremento considerable en la eficiencia operativa, alcanzando una productividad promedio de 11.54 registros por minuto, mientras que el método tradicional obtuvo 5.01 registros por minuto.

En la tabla 5 se presenta una comparación general de los resultados obtenidos.

Tabla 5. Comparación general de resultados (fuente propia)

Variable	Grupo control	Grupo experimental	Mejora
Tiempo de captura	19.48 min	8.52 min	56.3%
Número de errores	1.92	0	100%
Productividad	5.01 reg/min	11.54 reg/min	130%

Para facilitar la interpretación de los resultados, se emplearon representaciones gráficas que permiten visualizar de manera clara las diferencias entre ambos grupos:

En la figura 6 representa una gráfica de barras que compara el tiempo promedio de captura, evidenciando una reducción significativa en el grupo experimental (8.52 minutos) respecto al grupo de control (19.48 minutos).

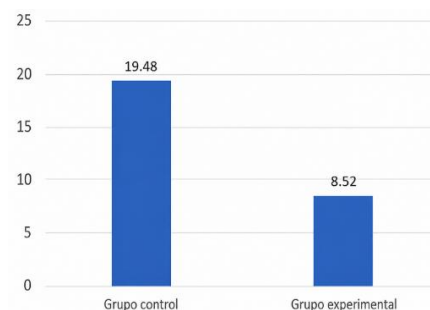


Figura 6. Tiempo promedio de captura (fuente propia)



En la figura 7 se muestra la comparación del número promedio de errores, donde el grupo experimental no presentó errores, en contraste con el grupo de control.

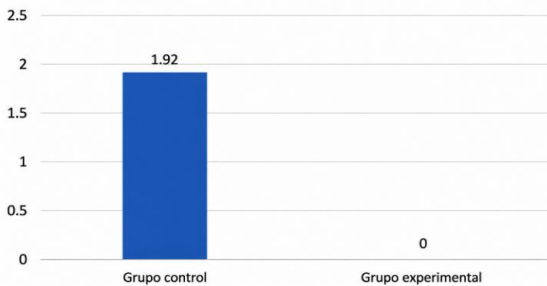


Figura 7. Número promedio de errores (fuente propia)

En la figura 8 se presenta la gráfica de productividad (registros por minuto), donde se observa un incremento considerable en el grupo experimental frente al grupo de control.

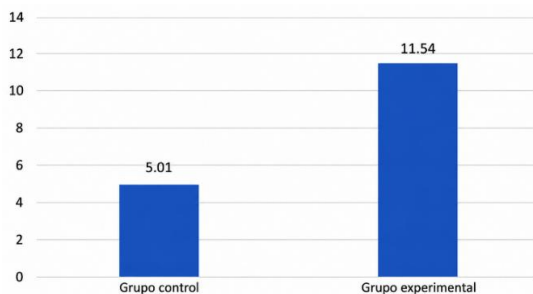


Figura 8. Productividad - registros por minuto (fuente propia)

En la figura 9 se presenta el diagrama de caja correspondiente al tiempo de captura para ambos grupos. Se observa que el grupo experimental presenta una mediana considerablemente menor en comparación con el grupo de control, lo que indica una reducción significativa en el tiempo de registro.

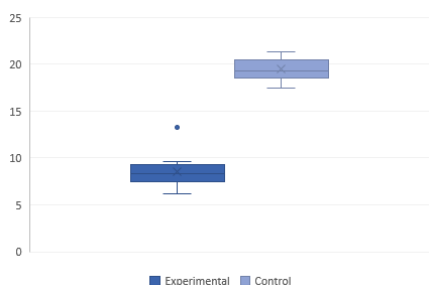


Figura 9. Tiempo de captura ambos grupos (fuente propia)

En la figura 10 se muestra el diagrama de caja correspondiente al número de errores. Se observa que el grupo experimental no presenta variabilidad, manteniendo un valor constante de cero errores, lo que indica alta consistencia en el proceso. En contraste, el grupo de control presenta dispersión en los valores, evidenciando variabilidad en el desempeño y mayor incidencia de errores.

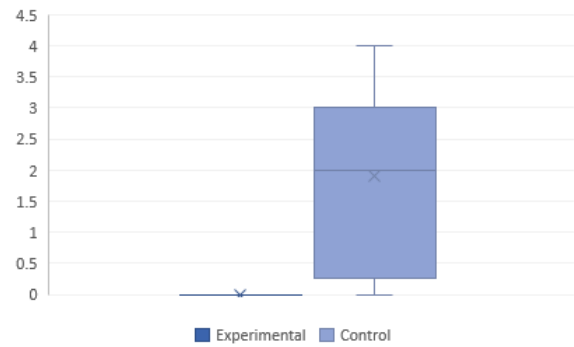


Figura 10. Número de errores ambos grupos (fuente propia)

La ausencia de dispersión en el grupo experimental indica alta consistencia en el proceso de captura debido a las validaciones implementadas en el sistema.

Estas representaciones gráficas complementan el análisis estadístico, proporcionando una visualización clara de las diferencias entre los grupos y reforzando la interpretación de los resultados obtenidos.

5. DISCUSION

Los resultados obtenidos evidencian una mejora significativa en la eficiencia del proceso de captura de datos meteorológicos mediante el uso del sistema web propuesto. En particular, se observó una reducción considerable en el tiempo de captura, así como la eliminación de errores en el grupo experimental, lo que confirma la hipótesis planteada.

La disminución en el tiempo de captura puede atribuirse a la automatización del proceso y a la estructuración de la interfaz, que facilita el registro de información. Asimismo, la ausencia de errores en el grupo experimental se relaciona con la implementación de validaciones automáticas en tiempo real, las cuales previenen inconsistencias desde el origen de los datos.

Los resultados obtenidos son consistentes con estudios previos relacionados con sistemas de gestión y automatización de datos meteorológicos, como los propuestos por Navarro-Silva et al. (2020) y Pérez Bellido (2024), quienes destacan la importancia de mejorar la calidad y disponibilidad de la información meteorológica mediante herramientas tecnológicas. Sin embargo, a diferencia de dichos trabajos, la presente investigación se enfoca específicamente en la etapa de captura de datos, incorporando validaciones en tiempo real que permiten reducir errores desde el origen de la información. Asimismo, los resultados obtenidos coinciden con lo reportado por Bautista-Villegas (2022) respecto a la efectividad del uso de metodologías ágiles para el desarrollo de sistemas web funcionales y adaptables.

Finalmente, el incremento en la productividad confirma que el sistema no solo digitaliza el proceso, sino que lo optimiza de manera integral. No obstante, el tamaño de la muestra limita la generalización de los resultados, aunque las condiciones controladas del experimento fortalecen su validez interna.



6. LIMITACIONES Y TRABAJOS FUTUROS

Una de las principales limitaciones del estudio corresponde al tamaño reducido de la muestra, debido a la disponibilidad limitada de personal especializado en captura de datos meteorológicos.

Como trabajo futuro, se propone ampliar la evaluación del sistema en otras dependencias y considerar la integración de tecnologías IoT para automatizar parcialmente el proceso de captura.

7. CONCLUSIONES

El presente trabajo permitió desarrollar e implementar un sistema web orientado a optimizar el proceso de captura de datos meteorológicos en el Centro de Previsión Meteorológica de la Comisión Nacional del Agua, Dirección Local Zacatecas. Los resultados obtenidos demostraron que el sistema mejora significativamente la eficiencia del proceso, reduciendo el tiempo de captura, eliminando errores e incrementando la productividad en comparación con el método tradicional.

La principal contribución de esta investigación radica en la incorporación de validaciones en tiempo real durante la captura de información, lo que fortalece la calidad y confiabilidad de los datos meteorológicos desde su origen.

Asimismo, el sistema desarrollado representa una solución viable para instituciones públicas que aún dependen de procesos manuales, permitiendo mejorar la gestión de información sin requerir infraestructura especializada.

REFERENCIAS

- Acosta, Luis Gerardo, y Meléndez, Adolfo. (2024, julio 18). Implementación de un Sistema de Medición Meteorológica para el Análisis de Datos | RICT Revista de Investigación Científica, Tecnológica e Innovación. 47-55. Recuperado el 25 de septiembre de 2024 de: <https://revista.ccaitec.com/index.php/ridt/article/view/62>
- Adalberto, Yulian, y Casadiego, Sepúlveda. (2020). Importance of meteorological stations for decision making in agriculture. Importancia de las estaciones meteorológicas para la toma de decisiones en la agricultura. doi:10.13140/RG.2.2.33323.46886
- Avilés Matute, S., Avila-Pesantez, D., y Avila, M. (2020). Desarrollo de sistema Web basado en los frameworks de Laravel y VueJs, para la gestión por procesos: Un estudio de caso. *Revista peruana de computación y sistemas*, 3(2), 3-10. doi:10.15381/RPCS.V3I2.19256
- Bautista-Villegas, Edwin. (2022, enero 25). Vista de Metodologías ágiles XP y Scrum, empleadas para el desarrollo de páginas web, bajo MVC, con lenguaje PHP y framework Laravel. Recuperado el 5 de noviembre de

2024 de:
<https://revistas.unamad.edu.pe/index.php/rad/article/view/168/275>

- Fonseca Márquez, Karla Alejandra. (2025). Vista de Implementación de una arquitectura modular en Laravel para sistemas web mantenibles. doi:<https://doi.org/10.22201/dgtic.ctud.2025.3.1.85>
- Montaño-Blacio, Manuel, González-Escarabay, Jorge, Jiménez-Sarango, Óscar, Mingo-Morocho, Leydi, Carrión-Aguirre, César, Montaño-Blacio, Manuel, ... Carrión-Aguirre, César. (2023). DISEÑO Y DESPLIEGUE DE UN SISTEMA DE MONITOREO BASADO EN IOT PARA CULTIVOS HIDROPÓNICOS. *Ingenius. Revista de Ciencia y Tecnología*, 2023(30), 9-18. doi:10.17163/INGS.N30.2023.01
- Navarro-Silva, Dainelaine, Loredó-Carballo, Néstor, Jiménez-Pearse, Leonardo, Gallardo-Avilés, Yordan, Estrada-Canosa, Eduardo, Bisbé-Sifontes, Luis, ... Nieves-Pérez, Dayan Arturo. (2020). Sistema de Gestión de Información para el pronóstico meteorológico especializado. *Revista Cubana de Meteorología*, 26(2), 2664-0880. Recuperado de: http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S2664-08802020000200002&lng=es&nrm=iso&tlng=es
- Pablo Báez-Vásquez, Juan, y Portilla-Flores, Alberto. (2022). Diseño de un sistema de gestión de datos climáticos bajo una metodología de desarrollo PSP Design of a Climatic Data Management System Under a PSP Development Methodology.
- Pereto Soler, Carla. (2024). Diseño de un sistema de monitorización y control de datos del entorno a través de nodos LoRa accesibles desde una app del teléfono móvil. Recuperado de: <https://riunet.upv.es/handle/10251/201804>
- Pérez Bellido, Rosamaría. (2024). Automatización de la gestión de datos meteorológicos usando Plotly Dash. Recuperado de: <http://repositorio.lamolina.edu.pe/handle/20.500.12996/6650>
- Quintero, Jorge Arcenio. (2021). Sistema para el manejo de datos climáticos de pequeñas producciones agrícolas bajo invernadero: un acercamiento a la agricultura inteligente. Recuperado el 26 de septiembre de 2024 de: <https://repositorio.unal.edu.co/handle/unal/81046>
- Vizcaino Zúñiga, Paulina Iveth, Cedeño Cedeño, Ricardo Javier, y Maldonado Palacios, Israel Alejandro. (2023). Metodología de la investigación científica: guía práctica. *Ciencia Latina Revista Científica Multidisciplinar*, 7(4), 9723-9762. doi:10.37811/el_rcm.v7i4.7658

Validation of an NLP Model for Thematic Organization of Study Programs: Teacher Diagnosis, Gold Standard and Extrinsic Evaluation

César Iván Abraján Barraza*. Andrés Saúl de la Serna Tuya**. Kimberly Lucero Carvajal Estrada*. Ivonne Alejandra Palazuelos Gálvez*

* Instituto Tecnológico de Culiacán, Tecnológico Nacional de México, Culiacán, Sinaloa, México
(Tel.: +52 667 451 1899; e-mail: cesar.ab@culiacan.tecnm.mx; kimcar615@gmail.com;
L22170260@culiacan.tecnm.mx).

** Universidad Popular Autónoma del Estado de Puebla (UPAEP), Puebla, México
(e-mail: asdelaserna@gmail.com).

Abstract: In higher education institutions of the Tecnológico Nacional de México (TecNM), the organization of thematic areas in study programs is performed manually and in a fragmented manner. This work presents the validation of a natural language processing (NLP) model for the automatic thematic organization of the 51 subjects of the Business Management Engineering (IGE) study plan. A diagnostic instrument was designed and applied to 16 teachers at the Instituto Tecnológico de Culiacán, yielding a gold standard of expert classification with 11 effective thematic categories and Fleiss $\kappa = 0.445$ (moderate agreement). Comparison with three model configurations showed that hierarchical Ward clustering ($K = 10$) achieved the highest concordance with experts ($ARI = 0.339$, $V\text{-measure} = 0.706$), followed by NMF $K = 8$ ($V\text{-measure} = 0.615$). The discordance analysis reveals both genuine model limitations in subjects with strong disciplinary identity and complementary distinctions in integrative subjects. The analysis is framed within the TPACK, SAMR, Triple E and RAT technopedagogical models, and proposes five principles of critical contextual appropriation for the integration of AI into curricular management processes.

Keywords: natural language processing, thematic organization, study programs, gold standard, extrinsic validation, educational technology, hierarchical clustering, TecNM.

1. INTRODUCCIÓN

La organización de las áreas temáticas predominantes en los programas de estudio constituye un proceso fundamental para la gestión curricular en instituciones de educación superior. En el contexto del Tecnológico Nacional de México (TecNM), sistema que opera 254 campus y gestiona entre 10 000 y 13 000 programas de estudio activos (TecNM, 2023), este proceso se realiza predominantemente de forma manual mediante la revisión individual de documentos por parte de academias y cuerpos colegiados.

La literatura sobre inteligencia artificial en educación (AIED) ha priorizado aplicaciones orientadas al aprendizaje estudiantil. Zawacki-Richter et al. (2019) identificaron que solo el 8.2 % de los estudios abordan análisis de documentos institucionales o curriculares. Wang et al. (2024) confirman que menos del 5 % de la investigación reciente se relaciona con planificación curricular. García-Peñalvo (2023) observa que las aplicaciones se concentran en interacciones directas con el estudiante, no en la gestión documental institucional. Este vacío es particularmente relevante en contextos latinoamericanos, donde las herramientas de análisis curricular sistemático son escasas (Casillas y Ramírez-Montoya, 2020; Salas-Pilco y Xiao, 2020).

El presente estudio aborda este vacío desde la perspectiva de tecnología educativa. La variable independiente es un modelo de inteligencia artificial basado en procesamiento del lenguaje natural (PLN). La variable dependiente es la organización de las áreas temáticas predominantes en los programas de estudio de educación superior. El objetivo es validar los resultados del modelo de PLN comparándolos contra las clasificaciones de docentes expertos, generando evidencia sobre el grado de concordancia entre la organización automática y el juicio humano.

A diferencia de trabajos previos que reportan únicamente métricas intrínsecas del modelo (Author et al., 2026b), este artículo se centra en la validación extrínseca: la construcción de un gold standard docente, la medición del acuerdo inter-evaluador y la comparación sistemática entre las organizaciones temáticas producidas por el modelo y por los expertos humanos.

2. MARCO TECNOPEDAGÓGICO APLICADO

El modelo educativo del TecNM se sustenta en principios humanistas y constructivistas (TecNM, 2015). Cualquier herramienta tecnológica propuesta debe articularse con estos valores institucionales. Esta sección no repite los fundamentos teóricos desarrollados en Author (2026a), sino que sintetiza su aplicación al análisis de la herramienta de PLN.

2.1 Marcos de integración tecnológica y sus límites institucionales

Los modelos TPACK (Mishra y Koehler, 2006), SAMR (Puentedura, 2006), Triple E (Kolb, 2017) y RAT (Hughes, 2005) ofrecen lenguaje analítico para describir la integración tecnológica en el aula. La herramienta propuesta opera en la intersección TPACK de los dominios tecnológico, de contenido y pedagógico; en SAMR se ubica entre Modificación y Redefinición al permitir el procesamiento simultáneo de 51 documentos, una tarea cualitativamente distinta a la revisión manual; en Triple E satisface las dimensiones Engage, Enhance y Extend al hacer visible una estructura curricular previamente inaccesible; y en RAT funciona como Amplificación en asignaturas de identidad temática clara y como Transformación en asignaturas integradoras donde el juicio humano no logra consenso. Estos cuatro marcos, sin embargo, fueron concebidos para decisiones pedagógicas individuales, no para procesos curriculares colegiados, y por tanto no operacionalizan los valores humanistas que el TecNM (2015) declara fundantes. Esa limitación motiva los principios complementarios que se enuncian a continuación.

2.2 Principios de apropiación crítica contextual

La integración de herramientas de IA en procesos académicos requiere un marco ético (UNESCO, 2023; Gamboa y Presa, 2023). A partir de estas recomendaciones y de la reflexión sobre las condiciones institucionales del TecNM, se proponen cinco principios orientadores: (1) Transparencia interpretativa: los resultados se presentan de forma comprensible para los docentes. (2) Explicabilidad pedagógica: las categorías generadas se interpretan en términos del modelo formativo institucional. (3) Preservación de la agencia académica: el modelo aporta insumos para la deliberación, no decisiones finales. (4) Visibilidad de los límites algorítmicos: las limitaciones del modelo se comunican explícitamente. (5) Horizontalidad epistémica: el conocimiento experto docente y el análisis computacional tienen estatus equivalente. Estos principios no constituyen constructos empíricamente validados sino lineamientos derivados de la convergencia entre los marcos éticos internacionales y la experiencia concreta de este estudio.

3. METODOLOGÍA

3.1 Diseño del instrumento diagnóstico

Se diseñó un instrumento con propósito dual: diagnóstico de las prácticas actuales de organización curricular y generación de un gold standard de clasificación experta. La Tabla 1 presenta la estructura del instrumento.

Tabla 1. Estructura del instrumento diagnóstico Fase 0.

S.	Contenido	Tipo de respuesta	Propósito
I	Datos del participante	Demográficos	Perfil de la muestra

S.	Contenido	Tipo de respuesta	Propósito
II	Percepción organización curricular	Likert 1-5 (10 ítems)	Diagnóstico
III	Prácticas análisis curricular	Opción múltiple (5 preg.)	Diagnóstico
IV	Clasificación 51 asignaturas	Texto libre + confianza 1-5	Gold standard
V	Áreas temáticas globales	Tabla: área → asignaturas	Validación cruzada
VI	Comentarios adicionales	Texto libre	Cualitativo

La Sección IV es el núcleo del instrumento. Cada docente asigna cada asignatura a exactamente un área temática, con denominación libre y restricción de 5 a 12 áreas distintas. Este diseño híbrido se justifica por dos razones. Primero, las opciones predefinidas sesgarían las respuestas hacia una taxonomía particular. Segundo, el formato completamente libre produciría clasificaciones tan fragmentadas que el cálculo de métricas de acuerdo sería inviable. La restricción de 5 a 12 áreas equilibra expresividad con procesabilidad estadística. Cada clasificación se acompaña de una escala de confianza (1-5), que sirve como criterio de desempate cuando la moda no es única y permite identificar respondentes con patrones de aquiescencia.

El instrumento fue sometido a un pilotaje con $n = 3$ docentes. El pilotaje identificó que los tres respondentes asignaron múltiples etiquetas por asignatura (entre 19.6 % y 49.0 %), violando la instrucción de clasificación única. Se implementaron correcciones en Google Forms: instrucciones reforzadas con ejemplo visual, validación de campo y campos obligatorios para datos demográficos.

3.2 Muestra y procedimiento

La muestra se conforma por 16 docentes del Instituto Tecnológico de Culiacán adscritos al Departamento de Ciencias Económico-Administrativas, seleccionados mediante muestreo intencional con tres criterios: antigüedad docente mínima de 2 años, experiencia impartiendo asignaturas del plan IGE y participación activa o reciente en academias del programa. El instrumento se administró mediante Google Forms durante marzo de 2026.

3.3 Construcción del gold standard y evaluación extrínseca

El procedimiento se formaliza en la Tabla 2, que describe las cuatro fases del proceso.

Tabla 2. Procedimiento de construcción del gold standard y evaluación extrínseca.

F.	Etapas	Procedimiento
1	Recolección	Entrada R[16×51] con confianza 1 a 5; extraer primera etiqueta y normalizar.
2	Armonización	138 etiquetas únicas → 12 categorías (11 efectivas) por mapeo manual.



F.	Etapas	Procedimiento
3	Gold standard	Para cada asignatura: GS = moda armonizada. Empate por suma de confianza.
4	Evaluación	κ de Fleiss; ARI y V-measure de los tres modelos vs. GS.

Los criterios de armonización aplicados en la Fase 2 fueron tres, jerarquizados. Primero, sinonimia léxica directa (p. ej., «Finanzas» y «Finanzas y Contabilidad» se fusionan en «Finanzas y Contabilidad»). Segundo, identidad disciplinar (etiquetas que comparten campo académico se consolidan, p. ej., «Estadística», «Probabilidad» y «Métodos Cuantitativos» se mapean a «Ciencias Básicas y Matemáticas»). Tercero, función formativa equivalente en el plan IGE (etiquetas con el mismo rol curricular se agrupan aunque difieran léxicamente, p. ej., «Habilidades Directivas» y «Liderazgo» se mapean a «Humanidades y Formación Integral»). Las decisiones se documentaron en una tabla de mapeo y fueron revisadas por dos investigadores; las discrepancias se resolvieron por discusión hasta consenso. La armonización redujo las 138 etiquetas únicas a 12 categorías candidatas, de las cuales 11 recibieron al menos una asignatura por moda.

La Fase 4 emplea dos métricas extrínsecas complementarias. El Adjusted Rand Index o ARI (Hubert y Arabie, 1985) mide similitud entre particiones ajustada por azar con rango [-1, 1]. La V-measure (Rosenberg y Hirschberg, 2007) es la media armónica de homogeneidad y completitud con rango [0, 1]; es matemáticamente equivalente a la Normalized Mutual Information (NMI) bajo normalización aritmética (Vinh et al., 2010), por lo que se reporta solo V-measure al ofrecer mayor detalle mediante sus dos componentes.

3.4 Implementación técnica del modelo de PLN

El modelo de PLN se implementó en Python sobre Google Colab. La extracción de texto de los 51 programas de estudio (formato PDF) se realizó con pdfplumber. El preprocesamiento aplicó normalización Unicode, eliminación de caracteres no alfabéticos, conversión a minúsculas y stemming con SnowballStemmer (variante española) de NLTK; se omitió lematización con spaCy por restricciones de red en el entorno de ejecución. Se generaron bigramas y trigramas con gensim.Phrases para preservar términos compuestos del dominio.

La representación vectorial empleó TF-IDF de scikit-learn con parámetros $\text{min_df} = 2$, $\text{max_df} = 0.95$ y stopwords personalizadas en español. Para los algoritmos basados en distancia se aplicó reducción de dimensionalidad mediante TruncatedSVD a $d = 10$ componentes. Las tres configuraciones evaluadas son: (i) NMF (sklearn.decomposition.NMF) con $K = 8$, $\text{init} = \text{nndsvd}$, $\text{max_iter} = 500$; (ii) K-Means (sklearn.cluster.KMeans) con $K = 10$, $\text{n_init} = 50$, $\text{random_state} = 42$ sobre la matriz reducida por SVD; (iii) clustering jerárquico

(scipy.cluster.hierarchy.linkage) con $K = 10$, método de Ward sobre la matriz reducida por SVD. La selección de K se basa en el análisis de estabilidad y métricas intrínsecas reportadas en Author et al. (2026b).

4. RESULTADOS

4.1 Diagnóstico de prácticas docentes

Los 16 docentes presentan una antigüedad promedio de 15.4 años ($DE = 8.8$, rango 5-32). El 93.75 % imparte asignaturas de IGE. La distribución de grados incluye 4 licenciaturas, 4 maestrías, 2 doctorados, 1 posdoctorado y 5 con otros grados. El 75 % participa activamente en academias del programa.

La Tabla 3 presenta los estadísticos descriptivos de la escala Likert. El alfa de Cronbach para los 10 ítems fue $\alpha = 0.788$ (aceptable). La media del bloque diagnóstico (ítems 1-7) fue 3.60. El ítem 6 (reuniones de academia como mecanismo de análisis sistemático) obtuvo la media más baja ($\mu = 2.50$). La media del bloque de disposición hacia IA (ítems 8-10) fue 4.21. La brecha de 0.61 puntos entre ambos bloques indica una demanda insatisfecha.

Tabla 3. Estadísticos descriptivos por ítem de la escala Likert (n = 16).

#	Descripción	μ	DE	Bloque
1	Conozco contenidos asignaturas IGE	4.06	1.29	Diag.
2	Identifico áreas temáticas	4.06	0.93	Diag.
3	Existe organización clara	4.19	0.66	Diag.
4	Detecto solapamientos	3.56	1.41	Diag.
5	Identifico vacíos temáticos	3.56	1.21	Diag.
6	Academias permiten análisis	2.50	0.97	Diag.
7	Cuento con herramientas	3.25	1.18	Diag.
8	Herramienta sería útil	4.25	1.06	Disp.
9	IA aporta a gestión curricular	4.31	0.87	Disp.
10	Usaría IA como insumo	4.06	1.12	Disp.

Las prácticas de análisis curricular confirman este panorama: el 62.5 % de los docentes revisa con poca frecuencia los programas de asignaturas que no imparte («rara vez» u «ocasionalmente»), y el 62.6 % analiza relaciones temáticas en reuniones de academia rara vez o nunca. Solo el 18.8 % declara revisar siempre los programas de otras asignaturas.

La satisfacción con los mecanismos actuales para organizar contenidos del plan es predominantemente neutral (8 docentes, 50.0 %) o negativa (7 docentes, 43.7 %). Solo un respondiente se declaró satisfecho. Respecto al uso potencial de una herramienta de IA, las necesidades más mencionadas fueron actualización de contenidos (10/27 menciones, 37.0 %), diseño



de nuevas asignaturas (6 menciones, 22.2 %) y detección de solapamientos (6 menciones, 22.2 %).

4.2 Gold standard y acuerdo inter-evaluador

La armonización consolidó las 138 etiquetas en 12 categorías, de las cuales 11 recibieron al menos una asignatura por moda. La categoría «Innovación y Emprendimiento», presente en etiquetas originales de varios evaluadores, fue absorbida por «Administración y Gestión» en el cálculo de la moda. La Tabla 4 presenta las 11 categorías efectivas.

Tabla 4. Gold standard: 11 áreas temáticas efectivas del plan IGE.

Área temática	n	Asignaturas (ejemplos)
Administración y Gestión	13	Fund. Gestión, Diseño Org.
Ciencias Básicas y Matem.	9	Cálculos, Álgebra, Física
Humanidades y Form. Int.	6	Des. Humano, Ética
Finanzas y Contabilidad	7	Contabilidad, Costos
Producción e Ing. Indust.	4	Gest. Producción, Cadena
Derecho y Marco Legal	3	Legislación Laboral
Investigación	3	Fund. Inv., Talleres I, II
Economía	2	Econ. Empresarial
Tecnología y Sistemas	2	Software Aplic.
Mercadotecnia	1	Mercadotecnia Electrónica
Ambiental y Sustent.	1	Desarrollo Sustentable

El Fleiss kappa (Fleiss, 1971) para 16 evaluadores, 51 asignaturas y 12 categorías fue $\kappa = 0.445$ (acuerdo moderado según la escala de Landis y Koch, 1977). La concordancia observada media fue $\bar{P} = 0.523$ frente a una esperada por azar $P_e = 0.140$. Trece asignaturas alcanzaron acuerdo alto ($P_i \geq 0.7$), de las cuales tres fueron unánimes ($P_i = 1.0$): Desarrollo Humano, Fundamentos de Física y Software de Aplicación Ejecutivo. Las otras 26 asignaturas presentan acuerdo intermedio ($0.3 \leq P_i < 0.7$).

En contraste, 12 asignaturas obtuvieron acuerdo bajo ($P_i < 0.3$): Estrategias Contables y Legales ($P_i = 0.200$), Gestión de Proyectos y Tecnologías Emergentes (0.200), Innovación Tecnológica e Inteligencia de Negocios (0.225), Ingeniería Económica (0.233), El Emprendedor y la Innovación (0.233), Mercadotecnia Electrónica (0.233), Estrategias de Comercio Internacional (0.250), Mercadotecnia (0.250), Economía Empresarial (0.258), Propiedad Intelectual y Transferencia de Tecnología (0.267), Entorno Macroeconómico (0.283) y Administración de la Salud y Seguridad Ocupacional (0.292). Estas asignaturas comparten dos rasgos: pertenecen al segundo y último tercio del plan de estudios y tienen carácter integrador o multidisciplinar.

4.3 Evaluación extrínseca: modelo vs. expertos

La Tabla 5 presenta las métricas extrínsecas de las tres configuraciones del modelo comparadas contra el gold

standard, y la Figura 1 las visualiza para facilitar el análisis comparativo.

Tabla 5. Métricas extrínsecas: modelo PLN vs. gold standard docente (n = 16).

Configuración	ARI	V-meas	Homog	Comp
NMF K = 8	0.281	0.615	0.597	0.635
K-Means K = 10	0.188	0.598	0.637	0.564
Ward K = 10	0.339	0.706	0.758	0.661

El clustering jerárquico Ward (K = 10) obtuvo la mayor concordancia en las cuatro métricas. Su homogeneidad (0.758) indica que dentro de cada clúster del modelo las asignaturas tienden a pertenecer a la misma categoría del gold standard, mientras que su completitud (0.661) refleja que el modelo subdivide algunas categorías humanas en agrupaciones más finas. NMF K = 8 ocupa el segundo lugar (V-measure = 0.615), y K-Means K = 10 obtuvo los valores más bajos, consistente con su menor estabilidad reportada en Author et al. (2026b).

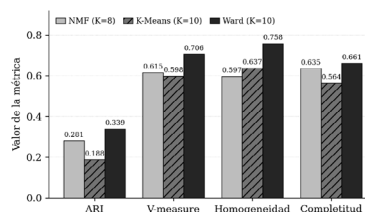


Fig. 1. Comparativa de las cuatro métricas extrínsecas para las tres configuraciones del modelo. Ward K = 10 obtiene la mayor concordancia con el gold standard.

Los valores de ARI son más modestos que los de V-measure porque ARI penaliza las diferencias de granularidad: el gold standard opera con 11 categorías efectivas frente a 8 tópicos (NMF) o 10-12 clústeres. V-measure es más tolerante a esta asimetría, lo que la hace apropiada cuando las particiones tienen diferente número de grupos.

4.4 Análisis de concordancias y discordancias

Para el análisis cualitativo se comparó la asignación NMF (tópico dominante) con el gold standard asignatura por asignatura. Se seleccionó NMF porque, a diferencia del clustering duro, permite examinar pesos de pertenencia múltiple, facilitando la interpretación de discordancias. Dos asignaturas presentaron variaciones de nomenclatura entre archivos del modelo y del instrumento que impidieron alineación automática, por lo que el análisis se reporta sobre las 49 asignaturas con correspondencia verificada.

De las 49 asignaturas alineadas, 35 (71.4 %) presentan concordancia entre el tópico NMF y la categoría del gold standard. Las concordancias se concentran en asignaturas del tronco disciplinar con identidad temática clara: ciencias básicas, investigación, humanidades, marco jurídico y



producción. Las 14 discordancias (28.6 %) se detallan en la Tabla 6.

Tabla 6. Discordancias NMF vs. gold standard (14 de 49 asignaturas alineadas).

Asignatura	Gold standard	Tópico NMF
Ingeniería Económica	Finanzas y Contab.	Economía y sust.
Inv. de Operaciones	Ciencias Básicas	Producción y costos
Gestión Capital Humano	Adm. y Gestión	Marco jurídico
Mercadotecnia	Adm. y Gestión	Mkt. y digital
Adm. Salud y Seguridad	Adm. y Gestión	Marco jurídico
Emprendedor e Innov.	Adm. y Gestión	Des. humano y ética
Sist. Info. Mercadotec.	Adm. y Gestión	Mkt. y digital
Calidad Aplicada Gest.	Adm. y Gestión	Estadística
Plan de Negocios	Finanzas y Contab.	Mkt. y digital
Marca Personal y Com.	Adm. y Gestión	Mkt. y digital
Innov. Tec. e Intel. Neg.	Adm. y Gestión	Mkt. y digital
Comercio Internacional	Adm. y Gestión	Economía y sust.
Estrat. Contables y Leg.	Finanzas y Contab.	Marco jurídico
Gestión Proy. y Tec.	Adm. y Gestión	Mkt. y digital

Las discordancias son heterogéneas y deben interpretarse en dos direcciones, no como evidencia uniforme a favor del modelo. Tres casos representan errores claros del modelo de PLN. Investigación de Operaciones es matemática de optimización y pertenece a Ciencias Básicas; el modelo la asignó a «Producción y costos» por la frecuencia de términos como «producción» y «costo» en los problemas tipo del programa, sin captar la naturaleza metodológica subyacente. Ingeniería Económica corresponde institucionalmente a Finanzas y Contabilidad; el modelo la mapea a «Economía y sustentabilidad» por la presencia léxica del término «económica», pero pierde el contexto financiero-decisional que los expertos reconocen. Plan de Negocios funciona como cierre integrador del programa con énfasis financiero, motivo por el cual los docentes lo ubican en Finanzas y Contabilidad; el modelo capta la dimensión de mercadotecnia presente en el contenido, pero no la función curricular del cierre. Estos tres casos ilustran un límite intrínseco del análisis basado en frecuencia léxica: el modelo no accede al rol formativo que cada asignatura ocupa en la secuencia curricular.

Las restantes 11 discordancias responden a un patrón distinto. Diez de ellas corresponden a asignaturas que el gold standard agrupa bajo «Administración y Gestión» pero que el modelo separa según contenido semántico fino. Mercadotecnia y Sistemas de Información de la Mercadotecnia se asignan al tópico de Mercadotecnia y negocios digitales; Calidad Aplicada a la Gestión aparece en Estadística y calidad; Estrategias Contables y Legales se ubica en Marco jurídico-

laboral. En estos casos no es evidente que el modelo erre: la categoría humana «Administración y Gestión» absorbe 13 asignaturas heterogéneas, lo que sugiere que el modelo aporta distinciones legítimas que el juicio docente no consensuó, posiblemente porque el 18.8 % de los respondientes reconoce clasificar por título sin leer el temario y el 37.5 % reporta dificultad genuina para categorizar asignaturas integradoras.

5. DISCUSIÓN

Los resultados deben interpretarse con cautela. La V-measure de 0.706 indica concordancia estructural fuerte entre el modelo Ward y el gold standard, pero el ARI de 0.339 es moderado y refleja diferencias sustantivas de granularidad. La conclusión defendible no es «alta concordancia» en términos absolutos sino «concordancia heterogénea por tipo de asignatura»: alta en asignaturas con identidad disciplinar clara, moderada-baja en asignaturas integradoras. El Fleiss kappa de 0.445 establece un techo realista del acuerdo humano: los docentes coinciden por encima del azar pero con variabilidad considerable, lo que indica que ningún modelo de PLN podría alcanzar concordancias superiores a las que los expertos logran entre sí.

La capacidad del modelo para procesar simultáneamente los 51 programas e identificar agrupaciones que los docentes no consensúan (12 asignaturas con $P_i < 0.3$) lo sitúa en el nivel de Modificación del modelo SAMR. En la taxonomía RAT, opera como Amplificación para asignaturas de alto consenso y como Transformación para asignaturas integradoras. Sin embargo, el análisis de discordancias muestra que el modelo también comete errores sistemáticos cuando la asignatura combina léxico de un dominio con función formativa de otro (caso Investigación de Operaciones, Ingeniería Económica, Plan de Negocios). La distinción entre fortalezas y limitaciones del modelo debe formar parte de cualquier integración institucional responsable.

La categoría «Administración y Gestión» merece atención especial. Absorbió 13 asignaturas en el gold standard, convirtiéndose en la categoría más heterogénea. El modelo de PLN la fragmenta en componentes específicos (mercadotecnia, calidad, marco legal, economía), lo cual podría reflejar con mayor precisión la estructura de contenidos. González Campos et al. (2024) advierten que los desafíos de la IA en educación superior incluyen la necesidad de interpretar críticamente los resultados computacionales; aquí, la fragmentación propuesta invita a las academias a reconsiderar si la agrupación vigente como categoría única es la más apropiada para la discusión curricular.

Desde una perspectiva operativa, los resultados sugieren un protocolo de uso para las academias del TecNM. El modelo de PLN generaría una propuesta inicial de organización temática que los docentes revisarían en sesión colegiada, concentrando la atención en asignaturas donde el modelo y el consenso previo divergen. Este enfoque reduce el tiempo de análisis y focaliza la discusión académica en los casos donde se requiere juicio experto. Delgado de Frutos et al. (2024) documentan que el profesorado percibe la IA como herramienta de apoyo, no



de sustitución, lo que es coherente con el principio de preservación de agencia académica. Hernández González et al. (2024) reportan que los estudiantes mexicanos demandan marcos éticos claros para el uso de IA generativa, demanda compartida por los docentes de esta muestra (ítems 8-10: $\mu = 4.21$).

5.1 Limitaciones

El gold standard presenta ruido inherente. Cinco de 16 respondientes muestran aquiescencia en la escala de confianza ($\sigma = 0$): asignan confianza máxima sin discriminar. Un sexto respondiente presenta varianza near-zero ($\sigma = 0.14$). Esta proporción (31.25 %) limita la utilidad de la confianza como ponderador, por lo que el gold standard se calculó por moda no ponderada excepto en empates. El 18.8 % de los docentes reconoce clasificar por título sin leer el temario. La muestra procede de un solo campus, lo que limita la generalización a otros contextos del TecNM. El instrumento no captura la dimensión temporal de la organización curricular (secuencia de prerrequisitos), aspecto que el modelo de PLN tampoco incorpora. Estévez et al. (2024) señalan que las percepciones sobre IA varían entre instituciones, lo que refuerza la necesidad de replicar el estudio en otros campus.

6. CONCLUSIONES

El modelo de PLN produce una organización temática con concordancia heterogénea respecto al juicio experto docente. Ward ($K = 10$) alcanzó $V\text{-measure} = 0.706$ (concordancia estructural fuerte) pero $ARI = 0.339$ (concordancia moderada por diferencia de granularidad). El acuerdo entre docentes (Fleiss $\kappa = 0.445$) establece un techo realista: la variabilidad humana es sustantiva, especialmente en asignaturas integradoras.

Las 14 discordancias entre modelo y gold standard son heterogéneas. Tres casos —Investigación de Operaciones, Ingeniería Económica y Plan de Negocios— son errores genuinos del modelo, que captura léxico pero no función formativa. Las restantes once señalan asignaturas donde el modelo aporta distinciones semánticas legítimas que el juicio humano no consensuó, particularmente dentro de la categoría heterogénea «Administración y Gestión». Esta distinción debe integrarse en cualquier protocolo de uso institucional.

El diagnóstico confirma que los docentes carecen de herramientas sistemáticas para analizar la estructura curricular (ítem 6: $\mu = 2.50$) pero muestran alta disposición a utilizar insumos de IA (ítems 8-10: $\mu = 4.21$), algo que debería ser aprovechado por la academia para su implementación. Los cinco principios de apropiación crítica contextual ofrecen un marco orientador para la integración de herramientas de IA en gestión curricular que respeta los valores humanistas y constructivistas del modelo educativo del TecNM. Como trabajo futuro se plantea la validación con docentes de otros campus, la extensión a otros programas de ingeniería y la incorporación de la dimensión temporal del plan en el modelo.

REFERENCIAS

- Abraján Barraza, C. I. y De la Serna Tuya, A. S. (2026a). Tecnología educativa en gestión curricular: tensiones entre automatización algorítmica y autonomía académica humanista en el TecNM. *Arandu UTIC*, 13(1), 1432–1446. <https://doi.org/10.69639/arandu.v13i1.1986>
- Abraján Barraza, C. I., De la Serna Tuya, A. S., Carvajal Estrada, K. L. y Palazuelos Gálvez, I. A. (2026b). Procesamiento de lenguaje natural para la organización temática de programas de estudio de ingeniería: comparativa de modelado de tópicos y agrupamiento. *Research in Computing Science*. Manuscrito aceptado para publicación.
- Casillas, M. A. y Ramírez-Montoya, M. S. (2020). Innovación educativa apoyada en tecnologías emergentes en universidades latinoamericanas. *RIED*, 23(1), 45-67.
- Delgado de Frutos, N., Campo Carrasco, L., Sainz de la Maza, M. y Etxabe-Urbietta, J. M. (2024). Aplicación de la inteligencia artificial IA en educación: los beneficios y limitaciones de la IA percibidos por el profesorado de educación primaria, educación secundaria y educación superior. *Revista Electrónica Interuniversitaria de Formación del Profesorado*, 27(1), 207-224.
- Estévez, J., Gaete-Bravo, F., Valdés, J. y Alarcón, F. (2024). Inteligencia artificial y educación universitaria: percepciones y prácticas emergentes en profesorado español. *Revista CTS*, 19(56), 117-139.
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, 76(5), 378-382.
- Gamboa, M. A. y Presa, D. I. C. (2023). Convivir con inteligencias artificiales en la educación superior: retos y estrategias. *Perfiles Educativos*, 45(Especial), 56-69.
- García-Peñalvo, F. J. (2023). La percepción de la inteligencia artificial en contextos educativos tras el lanzamiento de ChatGPT: disrupción o pánico. *Education in the Knowledge Society*, 24, e31279.
- González Campos, J. A., López Núñez, J. C. y Araya Pérez, C. E. (2024). Educación superior e inteligencia artificial: desafíos para la universidad del siglo XXI. *Aloma*, 42(1), 79-90.
- Hernández González, M., Ramos Quiroz, J. M., Chávez Maciel, F. J. y Trejo Cázares, M. D. C. (2024). Ventajas y riesgos de la inteligencia artificial generativa desde la percepción de los estudiantes de educación superior en México. *European Public and Social Innovation Review*, 9, 1-14.
- Hubert, L. y Arabie, P. (1985). Comparing partitions. *Journal of Classification*, 2(1), 193-218.
- Hughes, J. (2005). The role of teacher knowledge and learning experiences in forming technology-integrated pedagogy. *Journal of Technology and Teacher Education*, 13(2), 277-302.
- Kolb, L. (2017). *Learning first, technology second: The educator's guide to designing authentic lessons*. ISTE, Portland.
- Landis, J. R. y Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174.



- Mishra, P. y Koehler, M. J. (2006). Technological pedagogical content knowledge: a framework for teacher knowledge. *Teachers College Record*, 108(6), 1017-1054.
- Puentedura, R. (2006). Transformation, technology, and education. Disponible en <http://hippasus.com/resources/tte/>
- Rosenberg, A. y Hirschberg, J. (2007). V-Measure: a conditional entropy-based external cluster evaluation measure. En *Proceedings of the 2007 Joint Conference on EMNLP-CoNLL*, pp. 410-420. ACL, Praga.
- Salas-Pilco, S. Z. y Yang, Y. (2022). Artificial intelligence applications in Latin American higher education: a systematic review. *International Journal of Educational Technology in Higher Education*, 19, 21.
- Tecnológico Nacional de México (2015). *Modelo educativo para el siglo XXI: formación y desarrollo de competencias profesionales*. TecNM, Ciudad de México.
- Tecnológico Nacional de México (2023). *Anuario estadístico del Tecnológico Nacional de México 2022-2023*. TecNM, Ciudad de México.
- UNESCO (2023). *Artificial intelligence and education: Guidance for policy-makers* (2.ª ed.). UNESCO, París.
- Vinh, N. X., Epps, J. y Bailey, J. (2010). Information theoretic measures for clusterings comparison: variants, properties, normalization and correction for chance. *Journal of Machine Learning Research*, 11, 2837-2854.
- Wang, S., Christensen, C., Xu, T., Cui, S., Han, R., Zhuo, C. y Xie, J. (2024). Artificial intelligence in education: a systematic literature review. *Journal of Research on Technology in Education*, 1-28.
- Zawacki-Richter, O., Marín, V. I., Bond, M. y Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education. *International Journal of Educational Technology in Higher Education*, 16, 39.

Fine Grained Classification of Pollinating Insects with EfficientNet B4 for Monitoring in Urban Gardens^{*}

González-Vidales, Leonel^{*} Bernal-Beltrán, Carlos-Alberto^{*}
García-Vega, César-Sandino^{*} Cortés Sarabia, Jazmín^{*}
Sánchez Alonso, Itzel^{*}

^{*} TecNM Campus Ciudad Altamirano, Ciudad Altamirano, GRO 40660
MX (e-mail: leonel.gv@cdaltamirano.tecnm.mx,
carlosalberto.bb@cdaltamirano.tecnm.mx,
cesarsandino.gv@cdaltamirano.tecnm.mx,
jazmin.cs@cdaltamirano.tecnm.mx, itzel.sa@cdaltamirano.tecnm.mx).

Abstract: Urban gardens support pollinating insects, but monitoring is still based on manual observation and expert taxonomic validation. In this study, an EfficientNet B4 fine-grained classifier was tested for pollinating insect detection in urban gardens. The dataset consisted of 25,973 images from 42 taxonomic classes and was split into training, validation and test sets. The method consisted of semantic class cleaning, data augmentation, weighted sampling, transfer learning, and full fine-tuning. The accuracy is 97.64%, balanced accuracy 97.44%, Macro F1 0.9724 and top 5 accuracy 99.36% for the model. Error analysis showed that 74.2% of the misclassifications were within the same family or order.

Keywords: Machine Learning, Image processing, Pattern recognition and AI in agriculture, Monitoring, Agricultural robotics, Grading systems and Quality assessment, Modeling and identification of environmental systems

1. INTRODUCCIÓN

Los insectos polinizadores sostienen procesos ecológicos y productivos de alto valor. Abejas, abejorros, avispas, mariposas, moscas y escarabajos visitan flores en busca de alimento o refugio y, durante este proceso, transfieren polen entre estructuras reproductivas florales (Escobés and Vignolo, 2018; García García et al., 2016; Garibaldi et al., 2012; Gob.mx, 2021; Heiblum-Robles, 2019). Esta interacción mantiene la reproducción sexual de gran parte de las plantas con flor. Diversos reportes indican que más del 80 al 90% de estas plantas dependen de vectores animales para completar su ciclo reproductivo (García García et al., 2016; Gob.mx, 2021; Heiblum-Robles, 2019; Nates-Parra et al., 2016; Nava-Bolaños et al., 2022). En agricultura, alrededor del 70 al 75% de los cultivos destinados al consumo humano recibe beneficios de la polinización animal (Garibaldi et al., 2012; Gob.mx, 2021; Heiblum-Robles, 2019).

En las ciudades, los jardines urbanos representan espacios con valor ecológico, ambiental y educativo. Las áreas urbanas concentran pavimento, fragmentación del hábitat y presión térmica. Frente a este contexto, jardines y par-

ques operan como refugios que ofrecen néctar, polen, agua y sitios de nidificación (Botías and Sánchez-Bayo, 2018; Matellano-Villacampa, 2023; sedema.cdmx.mx, 2020; Vergara Briceño et al., 2023). Estos espacios favorecen la agricultura urbana, apoyan flora nativa y fortalecen conectividad biológica dentro de la ciudad (Matellano-Villacampa, 2023; Ochoa-Balboa, 2023). También cumplen una función formativa. Sirven como espacios para educación ambiental, ciencia ciudadana y revalorización de la biodiversidad (Gob.mx, 2021; Gutierrez-Galindo and Renis-Cogollo, 2022; Ochoa-Balboa, 2023). Esta doble función vuelve prioritario su monitoreo.

El monitoreo manual presenta limitaciones. La observación directa enfrenta frecuentes similitudes visuales entre especies, y demás características (Chiranjeevi et al., 2023; Gob.mx, 2021; Vergara Briceño et al., 2023; Wu et al., 2019). Además, la observación visual también dificulta demostrar si la visita de un insecto genera transferencia efectiva de polen o si corresponde a otra interacción (Nava-Bolaños et al., 2022). El monitoreo manual es lento y necesita de especialistas, y para una correcta clasificación es necesario poder diferenciar entre taxones muy similares que comparten características fenotípicas casi idénticas (Amarathunga et al., 2021; Pucci et al., 2023; Xie et al., 2015, 2018).

Una forma concreta de automatizar parte del monitoreo es a través de la visión por computadora. Los algoritmos, sensores y cámaras permiten registrar información visual de forma continua, remota y no invasiva a lo largo de

^{*} Este trabajo se realizó en el marco del proyecto de investigación "Monitoreo automatizado de insectos polinizadores en jardines urbanos utilizando visión por computadora y aprendizaje automático", clave CDAL PYR 2024 19796, registrado ante la Dirección de Posgrado, Investigación e Innovación del Tecnológico Nacional de México. El proyecto fue aprobado y registrado sin apoyo financiero debido a limitaciones presupuestales.

los ciclos diarios y estacionales (De Cesaro Júnior and Rieder, 2020; Høye et al., 2021; Spiesman et al., 2021). Esta estrategia permite reducir el cuello de botella del trabajo manual y preservar especies raras o amenazadas, al evitar su captura o sacrificio (Høye et al., 2021). En este contexto un clasificador fino tiene un aporte de valor directo. Permite separar especies afines y diferenciar organismos benéficos de especies plaga o invasoras, con una resolución taxonómica que resulta útil para estimar riqueza, diversidad y estado de conservación de una comunidad (Bjerge et al., 2023b; Chiranjeevi et al., 2023; Hansen et al., 2020; Pucci et al., 2023; Shamrat et al., 2024; Spiesman et al., 2021).

Se elige EfficientNet por su destacada estrategia de escalado compuesto, que ajusta de forma coordinada la profundidad, el ancho y la resolución de entrada (Doan, 2023; Kundur and Mallikarjuna, 2022; Padhi et al., 2024). El presente diseño permite una relación eficaz entre precisión y coste computacional (Padhi et al., 2024; Ray et al., 2026; Tariku et al., 2023).

El aprendizaje por transferencia se adapta bien a este problema. En entomología es difícil conseguir millones de imágenes bien anotadas de cada especie. El uso de pesos preentrenados en grandes conjuntos de datos permite disminuir la cantidad de datos necesarios, abreviar el tiempo requerido para el entrenamiento y minimizar el riesgo de sobreajuste (Amarathunga et al., 2021; Ayan et al., 2020; Krishna and Kalluri, 2019; Pérez-Aguilar et al., 2021; Thenmozhi and Srinivasulu Reddy, 2019).

Persisten lagunas en cuanto a sistemas para su monitoreo en jardines urbanos. La mayoría de los estudios se centran en plagas agrícolas, cultivos o ambientes controlados (Botías and Sánchez-Bayo, 2018; Thenmozhi and Srinivasulu Reddy, 2019). Sigue un reto abierto de incertidumbre ante especies no vistas durante el entrenamiento y de clasificación jerárquica cuando la red no distingue con seguridad a nivel de especie (Bjerge et al., 2023a,b; Chiranjeevi et al., 2024). A esto se suma una brecha metodológica entre imágenes limpias de colección, y escenas reales con clutter vegetal, iluminación variable y oclusión (Amarathunga et al., 2021; Bjerge et al., 2023a, 2024).

En este contexto, el presente trabajo se centra en la validación del módulo de clasificación fina para insectos polinizadores con interés en jardines urbanos. El estudio usó EfficientNet B4 con pesos preentrenados, aumento de datos, muestreo ponderado por clase y una estrategia de entrenamiento en dos fases sobre un conjunto limpio de 25,973 imágenes distribuidas en 42 clases. La pregunta principal fue directa. ¿Qué tan bien logra una arquitectura convolucional afinada con transferencia de aprendizaje separar 42 clases taxonómicas con alta similitud visual? El artículo responde a esta pregunta mediante una evaluación en conjuntos independientes de entrenamiento, validación y prueba.

2. METODOLOGÍA

2.1 Diseño experimental

Este estudio siguió un diseño supervisado de clasificación fina. El flujo partió de un conjunto de imágenes ya etique-

tadas por clase taxonómica. Después se aplicó una etapa de limpieza semántica sobre los nombres de clase. Luego se construyó un clasificador con aprendizaje por transferencia sobre la arquitectura EfficientNet B4. El entrenamiento se dividió en dos fases. La primera fase ajustó la cabeza clasificadora. La segunda fase ajustó toda la red. La evaluación final se realizó sobre un conjunto de prueba independiente.

2.2 Conjunto de datos

El conjunto final quedó formado por 25,973 imágenes distribuidas en 42 clases taxonómicas. La partición experimental se fijó en tres subconjuntos. El conjunto de entrenamiento reunió 18,161 imágenes. El conjunto de validación reunió 3,879 imágenes. El conjunto de prueba reunió 3,933 imágenes. La estructura final se organizó en carpetas compatibles con ImageFolder. Cada carpeta representó una clase.

La Tabla 1 muestra el tamaño de cada partición. La Tabla 2 muestra la distribución por orden en entrenamiento.

Table 1. Tamaño del conjunto de datos por partición.

Partición	Imágenes	Proporción
Entrenamiento	18,161	69.93%
Validación	3,879	14.94%
Prueba	3,933	15.14%
Total	25,973	100.00%

Table 2. Distribución de imágenes por orden taxonómico en entrenamiento.

Orden taxonómico	Imágenes
Lepidoptera	13,761
Hymenoptera	2,915
Coleoptera	698
Odonata	416
Orthoptera	371

2.3 Extracción de etiquetas taxonómicas

El nombre de cada carpeta codificó la jerarquía taxonómica de la clase. Esta extracción permitió construir diccionarios de apoyo para especie, familia y orden. Ese material se usó después en el análisis por familia, en la matriz de confusión y en el estudio de pares de error.

2.4 Preprocesamiento y aumento de datos

Todas las imágenes se llevaron a una entrada de 380×380 píxeles. Este tamaño coincide con la resolución nativa de EfficientNet B4. El preprocesamiento se definió por partición.

En el entrenamiento se aplicó un esquema de aumento de datos orientado a inducir variación controlada en escala, orientación, color y textura. El detalle se muestra en la Tabla 3.

En validación y prueba se aplicó un flujo determinista. Esta separación entre entrenamiento y evaluación mantuvo un protocolo limpio.



Los parámetros de normalización se fijaron con las estadísticas estándar de ImageNet:

$$\mu = (0.485, 0.456, 0.406), \quad \sigma = (0.229, 0.224, 0.225).$$

2.5 Arquitectura del modelo

El clasificador se construyó con `timm.create_model` sobre la arquitectura `efficientnet_b4` con pesos preentrenados en ImageNet. La red se adaptó a 42 salidas, una por clase. La configuración incluyó `drop_rate = 0.4` en la capa clasificadora y `drop_path_rate = 0.2` para profundidad estocástica. Esta combinación reforzó la regularización durante el ajuste fino.

2.6 Control del desbalance

La distribución por clase no fue uniforme. Para reducir sesgo hacia las clases con mayor número de imágenes, se aplicó un `WeightedRandomSampler`. El peso de cada clase se definió como el inverso de su frecuencia en entrenamiento:

$$w_c = \frac{1}{n_c},$$

donde n_c representa el número de imágenes de la clase c en el conjunto de entrenamiento.

2.7 Estrategia de entrenamiento

El entrenamiento se realizó en dos fases. La primera fase congeló todos los parámetros del backbone. El entrenamiento actualizó únicamente la capa clasificadora final. Esta fase se ejecutó durante 5 épocas. En la segunda fase se cargó el mejor modelo guardado al final de la Fase 1. Luego liberó todos los parámetros de la red. Se continuó entrenando durante 30 épocas más. En esta fase se utilizó un plan de tasa de aprendizaje diferencial.

La configuración principal del entrenamiento se resume en la Tabla 4.

2.8 Evaluación en prueba

La evaluación final se realizó sobre el conjunto de prueba independiente. El flujo de inferencia recorrió todas las imágenes y almacenó las etiquetas reales, las etiquetas predichas, los puntajes y las cinco clases con mayor probabilidad. A partir de esta salida se calcularon las métricas globales y el análisis taxonómico del error.

Table 3. Preprocesamiento y aumento de datos por partición.

Partición	Transformaciones
Entrenamiento	<code>RandomResizedCrop(380, scale=(0.7,1.0), ratio=(0.75,1.33)), RandomHorizontalFlip(p=0.5), RandomVerticalFlip(p=0.1), RandomRotation(20), ColorJitter(brightness=0.3, contrast=0.3, saturation=0.2, hue=0.05), RandomGrayscale(p=0.02), ToTensor(), Normalize(mean, std)</code>
Validación y prueba	<code>Resize(399), CenterCrop(380), ToTensor(), Normalize(mean, std)</code>

2.9 Análisis por clase, familia y orden

El análisis posterior a la inferencia se desarrolló en cuatro niveles.

- (1) Reporte por clase. Se calcularon precisión, recall, F1 y soporte para cada una de las 42 clases.
- (2) Matriz de confusión normalizada.
- (3) Pares de confusión. Se ordenaron los errores por frecuencia y por tasa relativa respecto al total real de la clase.
- (4) Métricas por familia. Se mapearon las etiquetas reales y predichas a su familia taxonómica.

3. RESULTADOS

3.1 Desempeño global del clasificador

La evaluación final en el conjunto de prueba mostró un desempeño alto y estable, como se resume en la Tabla 5. Estos valores muestran que la red fue capaz de resolver de forma consistente la clasificación fina de 42 clases.

El conjunto de prueba reunía 3,933 imágenes. El modelo tuvo 93 errores. Esta cifra representa el 2.36% del total evaluado.

Table 5. Métricas globales en el conjunto de prueba.

Métrica	Valor
Accuracy	97.64%
Balanced accuracy	97.44%
Macro F1	0.9724
Top 5 accuracy	99.36%
Imágenes en prueba	3,933
Errores totales	93

Table 4. Configuración principal del entrenamiento.

Componente	Configuración
Arquitectura	EfficientNet B4
Número de clases	42
Resolución de entrada	380 × 380
Optimizador	AdamW
Decaimiento de peso	1×10^{-4}
Función de pérdida	CrossEntropyLoss con label smoothing = 0.1
Fase 1	5 épocas, backbone congelado
Tasa de aprendizaje Fase 1	1×10^{-3}
Tamaño de lote Fase 1	32
Fase 2	30 épocas, red completa
Tasa de aprendizaje base Fase 2	5×10^{-5}
Tasa de aprendizaje de la cabeza en Fase 2	2.5×10^{-4}
Tamaño de lote Fase 2	24
Programador de tasa	Cosine Annealing
Criterio de selección del mejor modelo	mayor val_acc

3.2 Evolución del entrenamiento

Las curvas de entrenamiento presentaron dos etapas bien definidas, como puede observarse en la Figura 1. En la Tabla 6 se resume estos valores.

Table 6. Resumen de la dinámica de entrenamiento.

Etapas	Época	Train acc	Val acc	Val loss
Inicio de Fase 1	1	0.4952	0.7440	1.8844
Cierre de Fase 1	5	0.7658	0.8100	1.4110
Inicio de Fase 2	6	0.8448	0.9167	1.0441
Mejor tramo final	34	0.9922	0.9781	0.7614
Cierre del proceso	35	0.9922	0.9776	0.7617

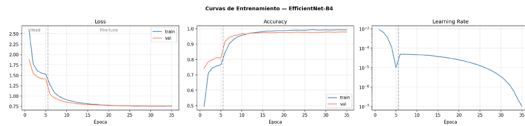


Fig. 1. Curvas de pérdida, accuracy y tasa de aprendizaje durante las dos fases de entrenamiento.

3.3 Resultados por clase

El reporte por clase confirmó que el modelo mantuvo alto rendimiento en casi todas las especies. Las especies con menor F1 se muestran en la Tabla 7.

Este patrón muestra dos hechos. Primero, el error residual quedó concentrado en un subconjunto corto de especies. Segundo, varias de esas especies difíciles tuvieron soporte bajo o similitud morfológica alta con clases cercanas.

Table 7. Especies con menor F1 en prueba.

Especie	Precision	Recall	F1
<i>Glutophrissa drusilla</i>	0.885	0.958	0.920
<i>Urbanus proteus</i>	0.894	0.955	0.923
<i>Hamadryas guatemalena</i>	0.897	0.963	0.929
<i>Ascia monuste</i>	0.932	0.932	0.932
<i>Smyrna blomfieldia</i>	0.982	0.903	0.941

3.4 Matriz de confusión y estructura del error

La matriz de confusión normalizada mostró una diagonal dominante, como se aprecia en la Figura 2. La mayor parte de las especies quedó clasificada con tasas de acierto superiores a 0.95.

Los errores más intensos no aparecieron de forma dispersa. Se concentraron en pares taxonómicos concretos. Estos pares se resumen en la Tabla 8.

Table 8. Pares de confusión con mayor tasa relativa.

Par de confusión	Tasa
<i>Ascia monuste</i> → <i>Glutophrissa drusilla</i>	0.07
<i>Glutophrissa drusilla</i> → <i>Ascia monuste</i>	0.04
<i>Hamadryas guatemalena</i> → <i>Hamadryas februa</i>	0.04
<i>Smyrna blomfieldia</i> → <i>Microtia elva</i>	0.03
<i>Eunica monima</i> → <i>Petrophila jaliscalis</i>	0.03

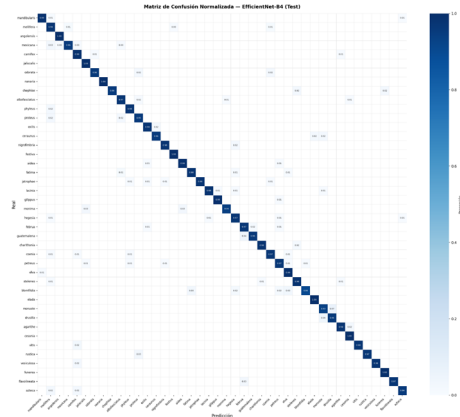


Fig. 2. Matriz de confusión normalizada sobre el conjunto de prueba.

3.5 Resultados por familia

El análisis por familia combinada confirmó que el alto rendimiento persistía al nivel superior taxonómico. Los valores relativos de menor F1 fueron para Formicidae 0.957, Erebidae 0.957 y Pieridae 0.963. Estos valores se resumen en la Tabla 9.

Table 9. Familias con menor desempeño relativo.

Familia	Macro F1	Accuracy
Formicidae	0.957	0.957
Erebidae	0.957	0.957
Pieridae	0.963	0.968
Hesperiidae	0.973	0.974
Apidae	0.975	0.975

3.6 Coherencia taxonómica del error

La distribución del error permitió una clara lectura taxonómica. El mayor número de errores cometidos, el 38.7 por ciento, se produjo dentro de la misma familia. Otro 35.5% ocurrió dentro de la misma orden, pero fuera de la familia correcta. El 25.8% restante procedió de entre distintas órdenes. Los valores sintetizados se presentan en la Figura 3.

Estos resultados muestran que el 74.2% del error se limitó a vecinos taxonómicos cercanos.

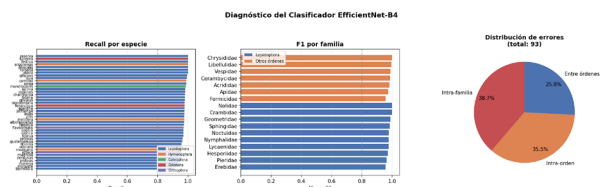


Fig. 3. Diagnóstico final del clasificador. La figura resume recall por especie, F1 por familia y distribución del error.



4. DISCUSIÓN

4.1 Lectura general del desempeño

Los resultados de la Tabla 5 validan el uso de EfficientNet B4 como clasificador fino sobre las 42 clases del conjunto limpio.

Esta combinación de métricas tiene una lectura directa. El conjunto estuvo desequilibrado, con fuerte presencia de Lepidoptera, pero ello no arrastró la decisión final a un sesgo visible. La diferencia entre métricas macro y globales fue pequeña.

4.2 Papel de las dos fases de entrenamiento

La dinámica del entrenamiento mostrada en la Figura 1 y resumida en la Tabla 6 mostró que la mejora fuerte surgió al activar el ajuste fino completo.

Esta secuencia muestra que la cabeza clasificadora resolvió el problema de arranque. El ajuste fino del backbone resolvió la discriminación fina.

4.3 Significado del análisis por clase

El error residual se concentró en pocas especies, como se muestra en la Tabla 7 y en la Figura 3. Esta observación importa por dos motivos. La primera es práctica. El sistema no registró una caída generalizada sobre varias clases. La segunda es de tipo biológico. Las especies difíciles comparten caracteres de color, forma o patrón alar con clases cercanas.

El estudio señala, por lo tanto, los segmentos en los que resulta conveniente reforzar el conjunto de datos o añadir descriptores complementarios.

4.4 Significado del análisis por familia

El análisis por familia dio un resultado útil. Ninguna familia descendió a las clases inferiores. Con esta observación se quiere decir que el problema no está en familias enteras. El reto se sitúa en pares concretos de especies.

4.5 Coherencia taxonómica del error

La distribución del error mostrada en la Figura 3 refuerza la validez del clasificador.

Lo que importa para el objetivo del artículo es esta estructura del error. Para el monitoreo ecológico, un buen clasificador no necesita acertar por casualidad. Debe mantener la lógica taxonómica en caso de error. El modelo satisfizo esta condición.

4.6 Alcance del estudio

El estudio confirma el módulo de la clasificación fina en recortes limpios. Por el momento, el resultado no es todavía un sistema de monitoreo completo en jardín urbano. Todavía faltan pruebas en escenas abiertas, con fondos complicados, variación de luz, cambios de escala, oclusión y muchos insectos en una sola imagen.

Este punto indica el siguiente paso del proyecto. El clasificador ya tiene un rendimiento suficiente para integrarse a un flujo de dos etapas. La primera etapa debe identificar al insecto en la escena. La segunda tiene que clasificar el recorte con EfficientNet B4. Este diseño separa identificación y localización. Permite además el conteo de visitas florales y análisis temporal por cámara o planta hospedera. En situaciones reales de jardines urbanos, se espera que el rendimiento del clasificador sea menor que los valores obtenidos con recortes limpios. Una oclusión parcial del insecto por hojas, flores u otros organismos puede ocultar caracteres morfológicos relevantes. De forma similar, cambios en la iluminación pueden alterar los patrones de color, contraste y textura que la red usa para separar especies cercanas. Sin embargo, la alta precisión top 5 y la coherencia taxonómica del error indican que el modelo puede mantener su utilidad como módulo de apoyo, siempre que se integre con una etapa previa de detección y se valide con imágenes reales de campo.

4.7 Implicaciones para el monitoreo en jardines urbanos

Los resultados proporcionan una base sólida para un sistema automático de apoyo al monitoreo de polinizadores en entornos urbanos.

El estudio deja también una señal práctica. Las zonas de error ya fueron identificadas. En datos futuros, los pares de Nymphalidae y Pieridae merecen un refuerzo específico.

5. CONCLUSIONES

El estudio alcanzó su objetivo central. Con el modelo EfficientNet B4 se logró una clasificación fina de 42 clases de insectos polinizadores con un alto nivel de acierto sobre el conjunto de prueba. La metodología incluyó limpieza taxonómica, aumento de datos, muestreo ponderado y entrenamiento de dos fases.

Los resultados por clase y por familia corroboraron la misma lectura. La mayoría de las especies alcanzaron altos valores de precisión, recall y F1. Las familias con menor rendimiento relativo, Formicidae, Erebidae y Pieridae, permanecieron por encima de 0.95 en F1 o cerca de ese umbral.

Uno de los descubrimientos más útiles del estudio fue la estructura del error. Su lectura resulta clara. El sistema concentró sus dudas en clases con verdadera cercanía morfológica.

Desde una perspectiva aplicada, el artículo deja un resultado concreto. EfficientNet B4 como módulo de clasificación fina para un sistema de monitoreo de polinizadores en jardines urbanos. El clasificador ya resolvió su parte del problema sobre cortes limpios. Esta validación permite llevar el proyecto a una etapa posterior centrada en detección sobre escenas abiertas.

El estudio también define sus límites. El experimento comprobó la clasificación en imágenes ya recortadas y bien etiquetadas. El trabajo no evaluó escenas de campo con fondos complejos, múltiples insectos, oclusión o fuerte variación de luz. El artículo da una base sólida para la parte de identificación y deja bien definida la ruta para el entorno real de jardín urbano.



Como próxima ruta el proyecto debe construir un dataset de escena abierta para jardín urbano. Ese conjunto debe contener flores, fondos naturales, cambios de escala, variación de iluminación y más de un insecto por foto. Luego, el sistema debe incluir un detector para ubicar al visitante floral y pasar el recorte a EfficientNet B4.

ACKNOWLEDGEMENTS

Los autores agradecen al Tecnológico Nacional de México, a través de la Dirección de Posgrado, Investigación e Innovación, por el registro del proyecto de investigación *Monitoreo automatizado de insectos polinizadores en jardines urbanos utilizando visión por computadora y aprendizaje automático*, con clave CDAL-PYR-2024-19796, con vigencia del 1 de enero al 31 de diciembre de 2024.

Los autores agradecen al Instituto Tecnológico de Ciudad Altamirano por el respaldo institucional brindado para el desarrollo, seguimiento y evaluación del proyecto.

Se reconoce el trabajo académico del responsable del proyecto, Leonel González Vidales, por la coordinación general, la conducción técnica del estudio y la articulación de las actividades de análisis, modelado y evaluación. Se agradece la colaboración de César Sandino García Vega, Jazmín Cortes Sarabia e Itzel Sánchez Alonso, quienes aportaron conocimiento especializado en biología y jardines urbanos. Su participación fortaleció la definición del contexto ecológico del proyecto, la interpretación biológica de los resultados y la vinculación del estudio con el monitoreo de polinizadores en entornos urbanos. También se reconoce la contribución de Carlos Alberto Bernal Beltrán como desarrollador de software, por su apoyo en las tareas técnicas del proyecto y por su trabajo en la organización, depuración y preparación del conjunto de datos empleado en el entrenamiento y evaluación del modelo.

Este artículo deriva de un proyecto registrado por su calidad académica y técnica dentro de la Convocatoria 2024 de Proyectos de Investigación Científica, Desarrollo Tecnológico e Innovación del Tecnológico Nacional de México.

REFERENCES

- Amarathunga, D.C., Grundy, J., Parry, H., and Dorin, A. (2021). Methods of insect image capture and classification: A systematic literature review. *Smart Agricultural Technology*, 1, 100023. doi:10.1016/j.atech.2021.100023.
- Ayan, E., Erbay, H., and Varçın, F. (2020). Crop pest classification with a genetic algorithm-based weighted ensemble of deep convolutional neural networks. *Computers and Electronics in Agriculture*, 179, 105809. doi:10.1016/j.compag.2020.105809.
- Bjerge, K., Alison, J., Dyrmann, M., Frigaard, C.E., Mann, H.M.R., and Høye, T.T. (2023a). Accurate detection and identification of insects from camera trap images with deep learning. *PLOS Sustain. Transform*, 2(3), e0000051. doi:10.1371/journal.pstr.0000051.
- Bjerge, K., Geissmann, Q., Alison, J., Mann, H.M., Høye, T.T., Dyrmann, M., and Karstoft, H. (2023b). Hierarchical classification of insects with multitask learning and anomaly detection. *Ecological Informatics*, 77, 102278. doi:https://doi.org/10.1016/j.ecoinf.2023.102278.
- Bjerge, K., Karstoft, H., Mann, H.M., and Høye, T.T. (2024). A deep learning pipeline for time-lapse camera monitoring of insects and their floral environments. *Ecological Informatics*, 84, 102861. doi:10.1016/j.ecoinf.2024.102861.
- Botías, C. and Sánchez-Bayo, F. (2018). Papel de los plaguicidas en la pérdida de polinizadores. *Ecosistemas*, 27(2), 3441. doi:10.7818/ecos.1314.
- Chiranjeevi, S., Saadati, M., Deng, Z.K., Koushik, J., Jubery, T.Z., Mueller, D.S., O'Neal, M., Merchant, N., Singh, A., Singh, A.K., Sarkar, S., Singh, A., and Ganapathysubramanian, B. (2024). Insectnet: Real-time identification of insects using an end-to-end machine learning pipeline. *PNAS Nexus*, 4(1), pgae575. doi:10.1093/pnasnexus/pgae575.
- Chiranjeevi, S., Sadaati, M., Deng, Z.K., Koushik, J., Jubery, T.Z., Mueller, D., Neal, M.E.O., Merchant, N., Singh, A., Singh, A.K., Sarkar, S., Singh, A., and Ganapathysubramanian, B. (2023). Deep learning powered real-time identification of insects using citizen science data.
- De Cesaro Júnior, T. and Rieder, R. (2020). Automatic identification of insects from digital images: A survey. *Computers and Electronics in Agriculture*, 178, 105784. doi:10.1016/j.compag.2020.105784.
- Doan, T.N. (2023). Large-scale insect pest image classification. *JAIT*, 14(2), 328–341. doi:10.12720/jait.14.2.328-341.
- Escobés, R. and Vignolo, C. (2018). Guía de los polinizadores más comunes de las zonas verdes de Madrid. Recuperado el 17 de febrero de 2025, de Real Jardín Botánico en español: <https://rjb.csic.es/educacion/proyectos-educativos/sos-polinizadores/>.
- García García, M., Ríos Osorio, L.A., and Álvarez del Castillo, J. (2016). La polinización en los sistemas de producción agrícola: revisión sistemática de la literatura. *Idesia (Arica)*, 34(3), 5368. doi:10.4067/s0718-34292016000300008.
- Garibaldi, L., Morales, C., Ashworth, L., Chacoff, N., and Aizen, M. (2012). Los polinizadores en la agricultura. *Ciencia Hoy*, 21, 34–43.
- Gob.mx (2021). Estrategia nacional para la conservación y uso sustentable de los polinizadores encusp. Recuperado el 17 de febrero de 2025, de GOB.MX en español: <https://www.gob.mx/agricultura/documentos/estrategia-nacional-para-la-conservacion-y-uso-sustentable-de-los-polinizadores-encusp>.
- Gutierrez-Galindo, K. and Renis-Cogollo, Y.J. (2022). Guía macrofotográfica digital: un camino hacia la enseñanza en ecología y conservación de insectos polinizadores. *Revista Cedotic*, 7(2), 149171. doi:10.15648/cedotic.2.2022.3473.
- Hansen, O.L.P., Svenning, J.C., Olsen, K., Dupont, S., Garner, B.H., Iosifidis, A., Price, B.W., and Høye, T.T. (2020). Species-level image classification with convolutional neural network enables insect identification from habitus images. *Ecology and Evolution*, 10(2), 737–747. doi:https://doi.org/10.1002/ece3.5921.
- Heiblum-Robles, A. (2019). Abejas: insectos polinizadores. Recuperado el 17 de febrero de 2025, de foroconsultivo.mx en español: <https://www.foroconsultivo.org.mx/INCyTU/index.php/notas/1-ambiente-y-energia/134-31-abejas-insectos->

- polinizadores.
- Høye, T.T., Årje, J., Bjerger, K., Hansen, O.L.P., Iosifidis, A., Leese, F., Mann, H.M.R., Meissner, K., Melvad, C., and Raitoharju, J. (2021). Deep learning and computer vision will transform entomology. *Proc. Natl. Acad. Sci. U.S.A.*, 118(2), e2002545117. doi: 10.1073/pnas.2002545117.
- Krishna, S.T. and Kalluri, H.K. (2019). Deep learning and transfer learning approaches for image classification. 7(5).
- Kundur, N.C. and Mallikarjuna, P.B. (2022). Insect pest image detection and classification using deep learning. *IJACSA*, 13(9). doi:10.14569/IJACSA.2022.0130947.
- Matellano-Villacampa, J. (2023). Jardín para polinizadores: estudio del ecosistema urbano e infraestructura verde.
- Nates-Parra, G. et al. (2016). *Iniciativa Colombiana de Polinizadores Capítulo Abejas -ICPA*.
- Nava-Bolaños, A., Osorio-Olvera, L., and Soberón, J. (2022). Estado del arte del conocimiento de biodiversidad de los polinizadores de México. *Revista Mexicana de Biodiversidad*, 93, e933948. doi: 10.22201/ib.20078706e.2022.93.3948.
- Ochoa-Balboa, L.J. (2023). Implementación y desarrollo de jardines polinizadores en el distrito de san isidro, lima.
- Padhi, J., Korada, L., Dash, A., Sethy, P.K., Behera, S.K., and Nanthamornphong, A. (2024). Paddy leaf disease classification using efficientnet b4 with compound scaling and swish activation: A deep learning approach. *IEEE Access*, 12, 126426–126437. doi: 10.1109/ACCESS.2024.3451557.
- Pucci, R., Kalkman, V.J., and Stowell, D. (2023). Comparison between transformers and convolutional models for fine-grained classification of insects.
- Pérez-Aguilar, D.A., Risco-Ramos, R.H., and Casaverde-Pacherrez, L. (2021). Transfer learning en la clasificación binaria de imágenes térmicas. *ings*, (26), 71–86. doi: 10.17163/ings.n26.2021.07.
- Ray, R.K., Chakravarty, S., Dash, S., Ghosh, A., Mohanty, S.N., Reddy Chirra, V.R., Ayouni, S., and Khan, M. (2026). Precision pest management in agriculture using inception v3 and efficientnet b4: A deep learning approach for crop protection. *Information Processing in Agriculture*, 13(1), 142–161. doi: 10.1016/j.inpa.2025.09.005.
- sedema.cdmx.mx (2020). Guía para la creación de jardines polinizadores. Recuperado el 17 de febrero de 2025, de sedema.cdmx.gob.mx en español: <https://www.sedema.cdmx.gob.mx/comunicacion/nota/publica-sedema-guia-para-la-creacion-de-jardines-polinizadores/>.
- Shamrat, F.M.J.M., Idna Idris, M.Y., Zhou, X., Khalid, M., Sharmin, S., Sharmin, Z., Ahmed, K., and Moni, M.A. (2024). PollenNet: A novel architecture for high precision pollen grain classification through deep learning and explainable AI. *Heliyon*, 10(19), e38596. doi: 10.1016/j.heliyon.2024.e38596.
- Spiesman, B.J., Gratton, C., Hatfield, R.G., Hsu, W.H., Jepsen, S., McCornack, B., Patel, K., and Wang, G. (2021). Assessing the potential for deep learning and computer vision to identify bumble bee species from images. *Sci Rep*, 11(1), 7580. doi:10.1038/s41598-021-87210-1.
- Tariku, G., Ghiglieri, I., Gilioli, G., Gentilin, F., Armiraglio, S., and Serina, I. (2023). Automated identification and classification of plant species in heterogeneous plant areas using unmanned aerial vehicle-collected rgb images and transfer learning. *Drones*, 7(10), 599. doi: 10.3390/drones7100599.
- Thenmozhi, K. and Srinivasulu Reddy, U. (2019). Crop pest classification based on deep convolutional neural network and transfer learning. *Computers and Electronics in Agriculture*, 164, 104906. doi: 10.1016/j.compag.2019.104906.
- Vergara Briceño, C.H., Perera Rieder, F., Fierros López, H.E., Jimeno Sevilla, H.D., Sagot, P., Martínez López, Ó.G., Ayala Barajas, R., and Vandame, R. (2023). *Atlas de los Bombus de México*. Ediciones Comunicación Científica. doi:10.52501/cc.123.
- Wu, X., Zhan, C., Lai, Y.K., Cheng, M.M., and Yang, J. (2019). IP102: A Large-Scale Benchmark Dataset for Insect Pest Recognition. In *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 8779–8788. IEEE, Long Beach, CA, USA. doi: 10.1109/CVPR.2019.00899.
- Xie, C., Wang, R., Zhang, J., Chen, P., Dong, W., Li, R., Chen, T., and Chen, H. (2018). Multi-level learning features for automatic classification of field crop pests. *Computers and Electronics in Agriculture*, 152, 233–241. doi:10.1016/j.compag.2018.07.014.
- Xie, C., Zhang, J., Li, R., Li, J., Hong, P., Xia, J., and Chen, P. (2015). Automatic classification for field crop insects via multiple-task sparse representation and multiple-kernel learning. *Computers and Electronics in Agriculture*, 119, 123–132. doi: 10.1016/j.compag.2015.10.015.

Exploring the Evolution of the Mercury Beating Heart with the Use of AI

Valeria Vega López* and Elizeth Ramírez Álvarez*

*Tecnológico Nacional de México, Campus Querétaro, México, 76000 (Tel: +52 744 754 2465 ; e-mail: m23140014@queretaro.tecnm.mx, elizeth.ra@queretaro.tecnm.mx).

Abstract: Modeling the behavior of active particles remains a challenging task due to their complex and non-linear dynamics. This work presents a data-driven approach for analyzing the temporal evolution of a mercury droplet using computer vision and deep learning techniques. Video frames are processed to segment the droplet and extract its contour, from which spatial points are obtained to construct a time-series dataset. This dataset is used to train a Long Short-Term Memory (LSTM) model to predict the droplet's dynamic behavior over time. The results show that the model is capable of capturing general deformation patterns, achieving an average prediction error of approximately 14.85% under optimized conditions. These findings demonstrate the potential of combining image processing and sequence modeling techniques for the study of deformable active systems. This approach opens the possibility of applying data-driven methods to a wide range of complex dynamical systems.

Keywords: Active Matter, Time Series Analysis, Computer Vision, Deep Learning, Particle Dynamics.

1. INTRODUCTION

The modeling of active particle systems has become an area of growing interest due to its relevance in disciplines such as physics, biology, and materials science. These systems exhibit complex and often non-linear dynamic behaviors, particularly when deformable structures are involved, which poses significant challenges for analysis and prediction.

Active matter systems have been extensively studied due to their ability to exhibit complex collective behaviors driven by energy consumption at the microscopic scale (Marchetti et al., 2013).

The term active matter describes both natural and synthetic systems composed of particles capable of extracting energy from their local environment to perform mechanical work. In the case of artificial systems, active matter encompasses a broad range of scales, from molecular-level structures to self-propelled microparticles currently under development (Cichos et al., 2020).

The interest in determining the fundamental principles governing the spatiotemporal dynamics of active matter is a frontier problem because these are non-equilibrium systems.

Furthermore, to construct a fundamental model, it is necessary to understand the origin of the motion and deformation of the active particles or matter, and this information is not always available.

Biological systems across multiple scales (as shown in Fig. 1), including molecular motors, cells, bacteria, and collective organisms, operate far from equilibrium through the conversion of chemical energy into mechanical processes.

Active matter across length scales

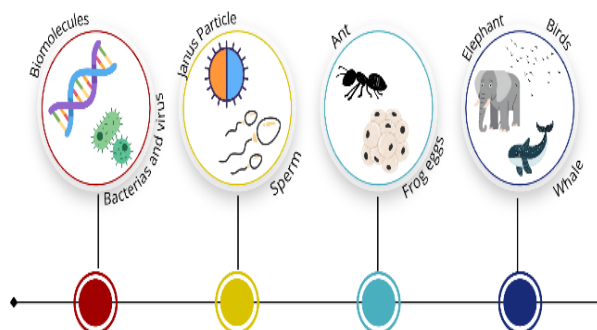


Fig. 1. Active matter across length scales (adapted from Cichos et al., 2020)

The interaction between biological networks and their environment governs the evolution of their dynamic behavior, enabling the execution of specific functions. In particular, bacteria can exhibit deformable particle-like behavior, which facilitates their mobility and interaction with cellular environments (Sengupta, Carrara and Stocker, 2017).

Recent advances in computer vision and machine learning have opened new possibilities for analyzing dynamic systems from visual data. Video analysis involves treating the evolution of shape as a time series. Models such as RNN, LSTMs, and ConvLSTM have been used to predict the evolution of visual sequences.

Specifically, (Shi et al., 2015) demonstrated the efficacy of ConvLSTM in predicting dynamic phenomena in images, which suggests potential applications to the study of droplets in fluids.

Other studies (Lu and Yang, 2022) have highlighted the growing importance of deep learning approaches for modeling complex temporal and nonlinear systems in scientific applications. These approaches have demonstrated strong capabilities in feature extraction, pattern recognition, and predictive modeling from time-series data.

This work presents preliminary results of ongoing study and aims to contribute to the understanding of deformable active systems through data-driven approaches for analyzing the dynamic behavior of a mercury droplet using computer vision techniques.

Video frames are processed to segment the droplet and extract its contour, from which a set of spatial points is obtained to describe its shape over time.

These data are used to construct a time-series dataset that serves as input for a Long Short-Term Memory (LSTM) neural network, enabling the modeling of temporal dynamics.

2. METHODOLOGY

This section describes the methodology applied to the experiment known as the Mercury Beating Heart (MBH), with the aim of developing this proposal using data science tools.

After several trials and adjustments, these steps were identified as an essential part of the methodology applied to provide a system that describes the dynamics of mercury droplet.

The methodological framework consists of five parts as it shows in *Figure 2*.

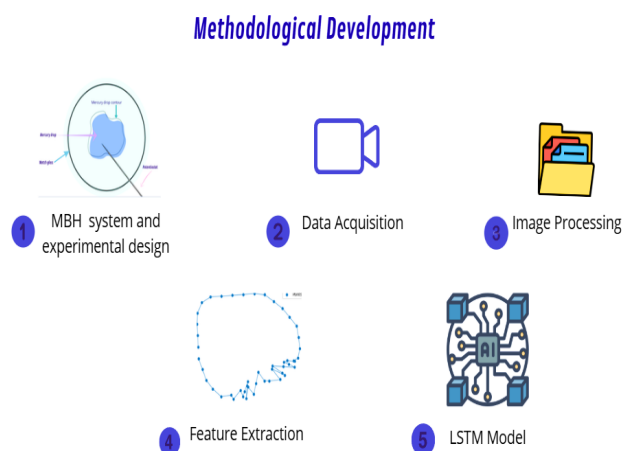


Fig. 2. Methodological Development applied to MBH system.

Each section below provides a detailed description of the methodological steps.

2.1 MBH system and experimental design.

The “mercury beating heart” is a system consisting of a volume of mercury that due to surface electrochemical reactions behaves as an active particle that upon deformation moves.

The proposed system consists of the construction of an artificial intelligence algorithm using videos of the behavior of the mercury active matter dynamics which is dependent on the applied potential (Ocampo-Espindola et al., 2015). Fig. 3. shows the procedure.

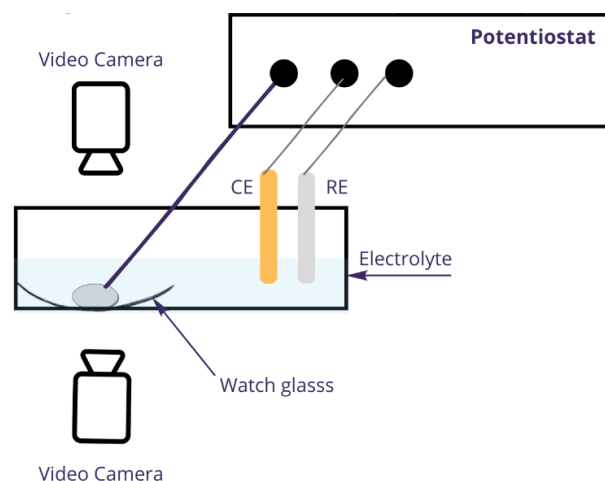


Fig. 3. Schematic of the experimental setup consisting of a three-electrode electrochemical cell; one in direct contact with the droplet of mercury and the other two simply immersed in the electrolyte. Adapted from (Ocampo-Espindola et al., 2015).

The MBH system consists in a mercury droplet (WE) connected to a platinum wire, also a copper rod (CE), and a saturated calomel electrode (RE) as shown in *Fig. 3*. The dynamics can be monitored with a video camera placed at the top or the bottom of the electrochemical cell. Adapted from (Ocampo-Espindola et al., 2015).

For a clearer picture of what we are working on, *Figure 4* below shows the MBH from a different perspective, allowing us to see each of its components.

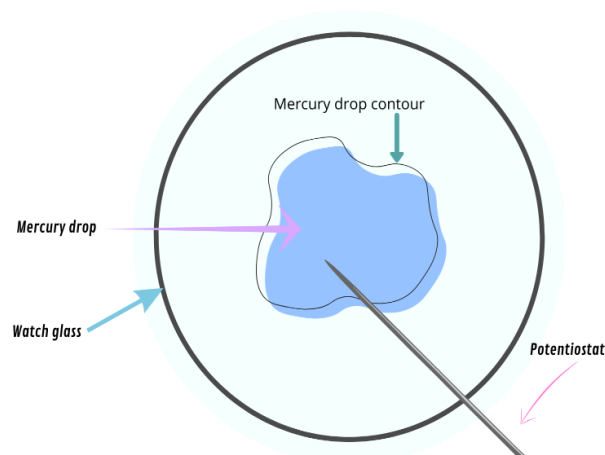


Figure. 4. Front view of the mercury drop, with labels indicating the components of the experimental setup with detail.

2.2 Data Acquisition

The dataset used in this study was obtained from video recordings of a mercury droplet exhibiting oscillatory behavior. During the mercury tests, the applied potential

ranged from -1800 mV to -2400 mV, with precise control over the potential.

Throughout the experiments, the generation of shape patterns in the droplet was observed, indicating a dynamic response to the applied potential. The videos capture the temporal evolution of the droplet under controlled experimental conditions as shown in Fig. 5.

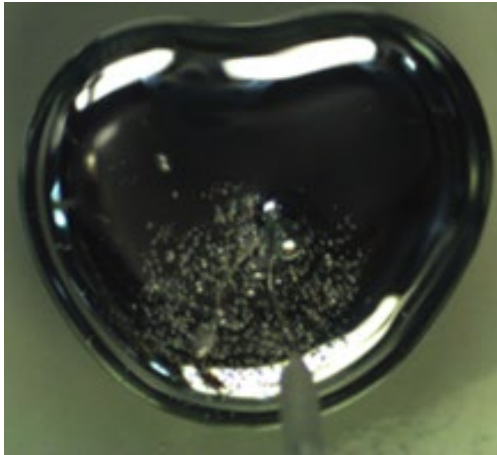


Fig. 5. Frame without treatments, initial part.

A library of information on this system is available, consisting of 15 videos; in this experimental phase, a single video with controlled motion was analyzed to subsequently replicate the same processes in at least two other videos and train the neural network to obtain a more accurate prediction in this study.

2.3 Image Processing

To extract meaningful information from the videos, a computer vision pipeline was implemented. Each frame was processed to segment the mercury droplet from the background. This segmentation step enables the isolation of the region of interest and facilitates the analysis of its geometric properties.

Binarization, segmentation, and morphological operations were applied to improve image quality and extract the relevant information—namely, the droplet's outline.

Image segmentation partitions an image into homogeneous regions, while contour extraction focuses on describing the geometry of the object. Active contour methods (snakes) have shown strong performance in accurately detecting object boundaries. In the case of droplets, the outer contour is particularly important for characterizing their temporal evolution (Kass, Witkin and Terzopoulos, 1988).

The original video is 1 minute and 23 seconds long, with 83.25 frames per second. It should be noted that for this case study, 19.27 seconds were analyzed, resulting in 1,613 processed frames; that is, they underwent binarization and edge detection. This analysis is shown in Fig. 6.



Fig. 6. Binarized frame with low blur values to appreciate droplet reflections and their contrast with the contour.

2.4 Feature Extraction

From the extracted contour, a fixed number of spatial points was sampled to represent the droplet shape. In this work, 50 points were selected for each frame, capturing the geometry of the droplet in a compact form.

These points were defined in terms of their spatial coordinates (x, y), forming a structured representation of the droplet at each time step. By stacking these representations across consecutive frames, a time-series dataset was constructed to describe the temporal evolution of the system Fig. 7 shows the result of this step.

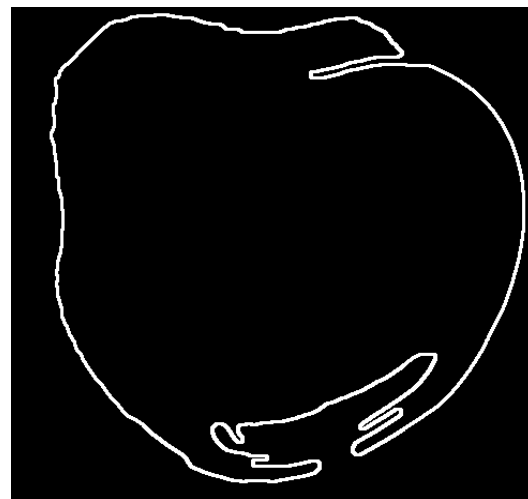


Fig. 7. Shows the segmentation and contour extraction of the mercury droplet from the video frames. Ideal frame for obtaining coordinates (X, Y).

2.5 LSTM Model

Long Short-Term Memory (LSTM) networks were developed to overcome the vanishing gradient problem present in traditional RNNs, enabling the learning of long-term dependencies in sequential data (Shi et al., 2015; Goodfellow et al., 2016). By incorporating memory cells and gating

mechanisms—namely the forget, input, and output gates—LSTMs regulate the flow of information across time steps, allowing relevant information to be retained while irrelevant data is discarded.

Recurrent neural networks, particularly LSTM architectures, have demonstrated strong performance in modeling sequential data due to their ability to capture long-term dependencies (Goodfellow et al., 2016).

Due to these characteristics, LSTMs have been widely used in time-series modeling and prediction tasks, where capturing temporal dependencies is essential. Their architecture, which includes both hidden and cell states, allows the model to effectively represent complex dynamic behaviors.

In this work, the input to the LSTM model consists of sequences of spatial coordinates (x,y) , representing the contour of the mercury droplet over time. Each sequence captures the temporal evolution of the droplet shape across consecutive frames.

The decision to use 50 (x,y) coordinate points stems from the importance of creating a relatively small yet functional neural network, given the nature of the MBH system and the datasets, which allow us to obtain equally functional sequences regardless of the number of data inputs.

The model is trained to learn the underlying temporal patterns of these sequences, enabling it to predict the future state of the droplet based on previous observations. This approach allows the representation of the dynamic behavior of the system as a time-dependent process.

By leveraging the sequential nature of the data, the LSTM model provides a suitable framework for modeling the evolution of deformable structures, where past configurations influence future states.

The model was trained using a supervised learning approach, where input sequences of spatial coordinates were mapped to future states of the system. The training process involved minimizing a regression-based loss function to reduce the difference between predicted and actual values.

3. RESULTS

3.1 Droplet dynamics analysis

The analysis of the processed video frames reveals that the mercury droplet exhibits dynamic and oscillatory deformation patterns over time depending on the injected potential. The extracted contour and the corresponding spatial points provide a consistent representation of the droplet shape across consecutive frames.

The temporal evolution of these points reflects the non-linear behavior of the system, capturing variations in shape and position that characterize the droplet dynamics.

3.2 Model Performance

The Long Short-Term Memory (LSTM) model was trained using the constructed time-series dataset derived from the spatial coordinates of the droplet. The model demonstrates the ability to capture the general temporal behavior of the system, producing predictions that follow the overall trend of the droplet deformation.

While the predicted outputs do not perfectly match the real data, they preserve key dynamic patterns, indicating that the model successfully learns relevant temporal dependencies.

Fig. 8 below shows the architecture used for this system, which includes the system's specific features.

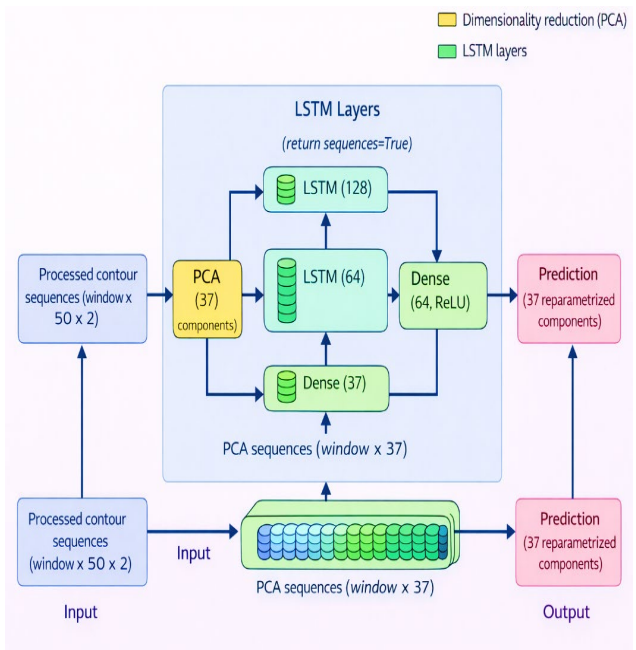


Fig. 8 Architecture of the LSTM model used in this study.

3.3 Visualization of results

After segmentation, the contour of the droplet was extracted for each frame. The contour provides a precise representation of the droplet shape, allowing the study of its deformation across time.

As a result in the image processing, a robust segmentation of the external contour of the droplet was obtained, eliminating shadows, reflections and internal bubbles.

Figure 9 was created to illustrate what the computer analyzes. It shows the outline of the droplet represented by a solid white line, with cyan dots distributed along it and connected by blue lines, resulting in a visual representation of the concept of the 50 (x,y) points.

The results were exported in visual (video) and numerical (CSV) formats, for temporal analysis and predictive modeling.

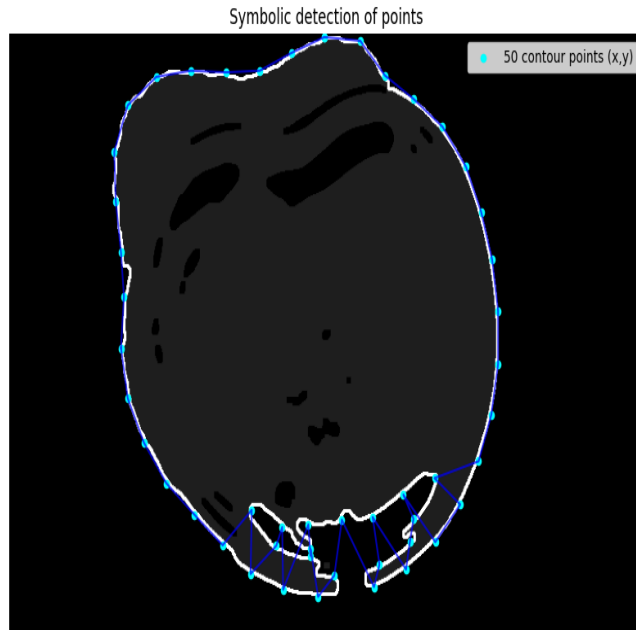


Fig. 9. Illustrates the spatial points sampled from the contour, which are used to construct the dataset.

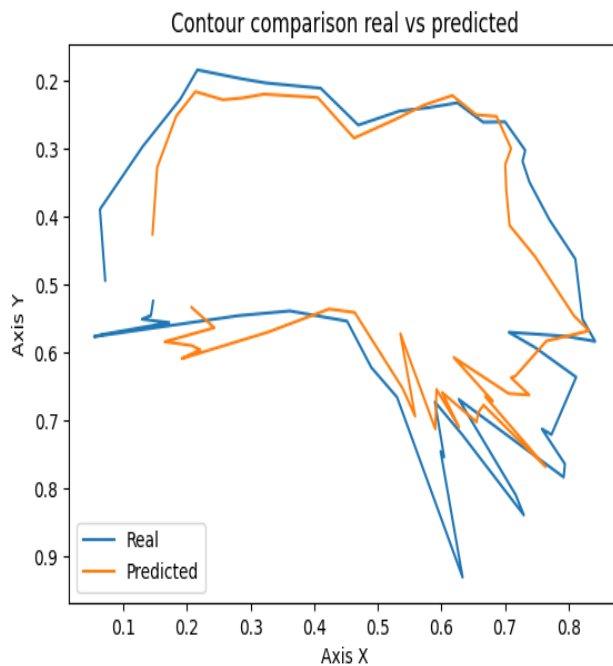


Fig. 10. Shows a comparison between the real and predicted values generated by the LSTM model, highlighting the model's ability to approximate the temporal evolution of the system with average error per point of 14.85% with 50 epochs and PCA resampling = 50.

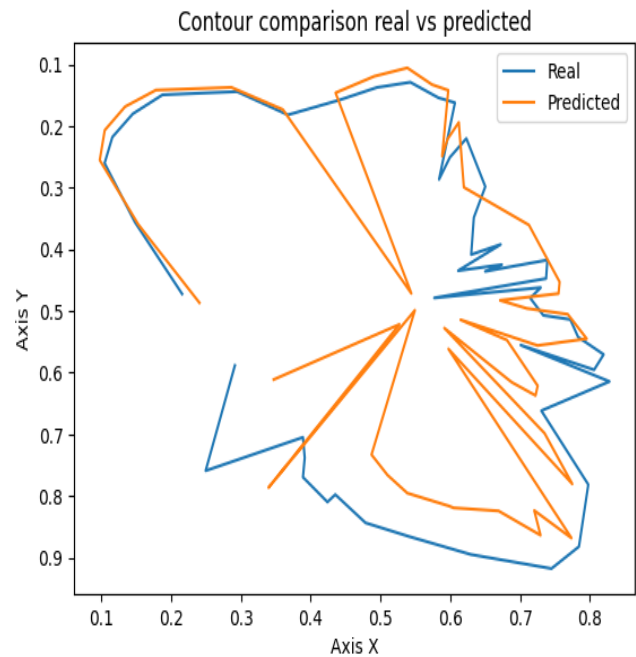


Fig. 11. Presents a comparison between the real and predicted values generated by the LSTM model, highlighting the model's ability to approximate the temporal evolution of the system with data average error per point of 14.71% with 100 epochs and resampling PCA = 37.

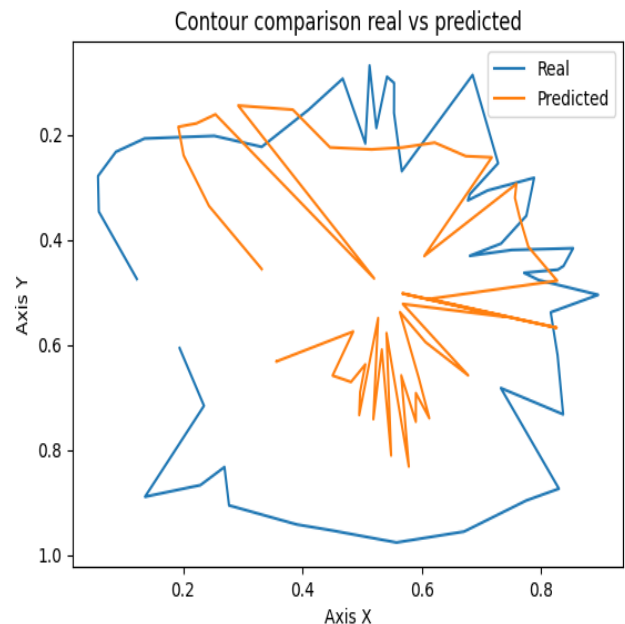


Fig. 12. Presents a comparison between the real and predicted values generated by the LSTM model, highlighting the model's ability to approximate the temporal evolution of the system with data average error per point of 25.77% with 100 epochs, Windows=20 and PCA + deltas resampling = 37.

These results (shown in Fig. 10, Fig. 11 & Fig. 12) provide further insight into the behavior of active particles and support the applicability of data-driven approaches in related fields.



It is observed that applying deltas to the model worsens the error rate, increasing it from 14.85% to 25.77%; therefore, the model performs well when Principal Component Analysis (PCA) is applied and the number of epochs is increased.

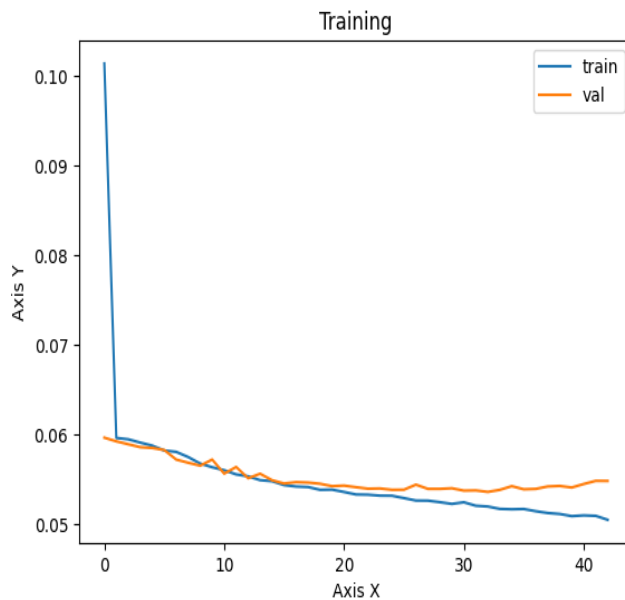


Fig. 13. Training and validation loss curves of the LSTM model during training. The divergence between both curves indicates the onset of overfitting.

Fig. 13 shows the training and validation loss curves during the learning process. Both curves initially decrease, indicating that the model effectively learns the underlying temporal patterns of the system. However, after a certain number of epochs, the validation loss begins to increase while the training loss continues to decrease, suggesting the presence of overfitting.

This behavior indicates that, although the model captures the training data effectively, its generalization capability is limited under the current configuration.

These results suggest that additional strategies such as early stopping or regularization could further improve model performance.

4. CONCLUSIONS

We have decided to continue experimenting with PCA values and the number of epochs until we find the “ideal” setting or one that reduces the error, as well as to implement a change in the loss function. Future work will focus on optimizing hyperparameters, improving the loss function, and refining the temporal window selection to enhance model performance.

The proposed network has not yet achieved the optimal values for the expected representation of the mercury drop; however, this is an ongoing project, and future improvements are expected to reduce the prediction error and improve model generalization, which would be acceptable for our predictive model.

This work presented an approach for analyzing the dynamic behavior of a mercury droplet using computer vision and machine learning techniques. By processing video frames, segmenting the droplet, and extracting spatial features, a time-series dataset was constructed to model the temporal evolution of the system.

The use of a Long Short-Term Memory (LSTM) neural network enabled the identification of temporal patterns associated with the deformation of the droplet.

Although the results are preliminary, the model demonstrated the ability to capture the general dynamics of the system, highlighting the potential of data-driven approaches for studying complex deformable structures.

In summary, although certain obstacles have been presented due to the noise caused by refraction in the droplet, the initial findings indicate that the system responds dynamically to the applied stimuli, which opens new avenues to further explore this phenomenon in detail.

Future work will focus on improving the accuracy of the model, refining the feature extraction process, and exploring alternative architectures to better represent the system dynamics.

REFERENCES

- Cichos, F., Gustavsson, K., Mehlig, B. and Volpe, G. (2020) ‘Machine learning for active matter’, *Nature Machine Intelligence*, 2, pp. 94–103.
- Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep Learning*. Cambridge, MA: MIT Press.
- Kass, M., Witkin, A. and Terzopoulos, D. (1988) ‘Snakes: Active contour models’, *International Journal of Computer Vision*, 1, pp. 321–331.
- Lu, D. and Yang, S. (2022) ‘A survey of the analysis of complex systems based on complex network theory and deep learning’, *International Journal of Performability Engineering*, 18(4), pp. 241–250.
- Marchetti, M.C., Joanny, J.F., Ramaswamy, S., Liverpool, T.B., Prost, J., Rao, M. and Simha, R.A. (2013) ‘Hydrodynamics of soft active matter’, *Reviews of Modern Physics*, 85(3), pp. 1143–1189.
- Ocampo-Espindola, J.L., Ramirez-Álvarez, E., Montoya, F., Parmananda, P. and Rivera, M. (2015) ‘Inducing rotational motion in the mercury beating heart system’, *Journal of Solid State Electrochemistry*, 19, pp. 3297–3303.
- Sengupta, A., Carrara, F. and Stocker, R. (2017) ‘Phytoplankton can actively diversify their migration strategy in response to turbulent cues’, *Nature*, 543, pp. 555–558.
- Shi, X., Chen, Z., Wang, H., Yeung, D.Y., Wong, W.K. and Woo, W.C. (2015) ‘Convolutional LSTM network: A machine learning approach for precipitation nowcasting’, in *Advances in Neural Information Processing Systems (NeurIPS)*.

E-learning and the “instant generation” in technological higher education: challenges and opportunities for institutional management toward intelligent social learning models.

Areli Soledad Ruiz Martínez*, Ricardo Gutiérrez Feria, Elizabeth Cabrera Mendoza***, Víctor Josué Ruiz Martínez******

**TecNM/ Centro Interdisciplinario de Investigación y Docencia en Educación Técnica, Santiago de Querétaro, Querétaro (Tel: 01 (442) 216 - 3744; (e-mail: areli.rm@ciidet.tecnm.mx) elizabeth.cm@ciidet.tecnm.mx).*

***TecNM/ Instituto Tecnológico de Álvaro Obregón, Ciudad de México, México (Tel: 01 (55) 5602-1984; e-mail: ricardo.gf@aobregon.tecnm.mx).*

***** TecNM/ Centro Nacional de Investigación y Desarrollo Tecnológico, Cuernavaca, Morelos (Tel: (777) 362 - 7770;., (e-mail: josue.rm@cenidet.tecnm.mx).*

Abstract: The incorporation of e-learning and Artificial Intelligence (AI) in technological higher education has generated a substantial transformation in learning processes, shaping a new student profile characterized by immediacy, hyperconnectivity, and a preference for brief content. This study aims to analyze the usage dynamics of these technologies among students of the Tecnológico Nacional de México, within the context of the so-called “on-demand generation” or culture of digital immediacy. A quantitative descriptive approach was employed with a sample of 500 students through the application of a structured questionnaire. The results show a high prevalence of mobile learning (88.1%), a significant preference for microcontent (55.5%), and a strong demand for immediate feedback (94.5%). Additionally, 51% of students use AI tools as academic support. The findings demonstrate a transition toward learning models centered on cognitive efficiency, which raises tensions between the speed of access to information and the depth of learning. It is concluded that the main institutional challenge lies in redesigning educational environments toward intelligent social learning models that integrate the critical use of AI, source verification, and the development of critical thinking.

Keywords: Artificial Intelligence, E-learning, Microlearning, Social Learning, Technological Higher Education

1. INTRODUCCIÓN

La educación superior tecnológica se encuentra en un proceso de transformación profunda impulsado por la digitalización y la incorporación de tecnologías emergentes. En este contexto, el e-learning y la inteligencia artificial (IA) han dejado de ser herramientas complementarias para convertirse en elementos estructurales del ecosistema educativo contemporáneo. Esta transformación no solo ha modificado los medios de acceso al conocimiento, sino también las formas en que los estudiantes interactúan con la información procesan contenidos y construyen aprendizaje.

Diversos estudios recientes señalan que la digitalización ha generado cambios significativos en los estilos cognitivos de los estudiantes, particularmente en términos de atención, procesamiento de información y expectativas de inmediatez (Bond et al., 2023). En este sentido, emerge el concepto de la “generación del ya o cultura de la inmediatez digital”, que describe a estudiantes que se encuentran en un rango de edad de 18 a 23 años (Tirocchi, 2024), que demandan respuestas rápidas, accesibilidad permanente y experiencias de aprendizaje ágiles.

En el contexto del Tecnológico Nacional de México, esta transformación adquiere especial relevancia debido a la naturaleza tecnológica de sus programas educativos, donde el uso de plataformas digitales y herramientas de IA se ha intensificado. Sin embargo, esta adopción también plantea desafíos importantes, como la superficialidad del aprendizaje, la dependencia tecnológica y la necesidad de desarrollar competencias críticas para la validación de información.

El presente estudio tiene como propósito analizar las prácticas de e-learning y el uso de inteligencia artificial en estudiantes de educación superior tecnológica, identificando sus implicaciones para el diseño de modelos de aprendizaje social inteligente.

2. MARCO TEÓRICO

2.1 Conectivismo y aprendizaje en red

El conectivismo, propuesto por Siemens (2005), plantea que el aprendizaje en la era digital se produce a través de redes de información interconectadas, donde el conocimiento no reside exclusivamente en el individuo, sino en la capacidad de acceder, filtrar y utilizar información distribuida. Este enfoque resulta relevante en entornos digitales, donde los estudiantes



interactúan constantemente con múltiples fuentes de información.

En este sentido, el aprendizaje se concibe como un proceso dinámico y continuo, mediado por tecnologías digitales que amplían las posibilidades de acceso al conocimiento, colocando a la IA como un nodo activo capaz de generar, procesar y distribuir información.

2.2 Espiral del conocimiento y construcción del aprendizaje

La teoría de la espiral del conocimiento de Nonaka y Takeuchi (1995), retomada y validada en estudios recientes (Han et al., 2026; Žatuchin, 2024), ofrece un marco conceptual para comprender cómo se transforma el conocimiento en entornos organizacionales y educativos. Este modelo describe la conversión del conocimiento tácito en explícito mediante procesos de socialización, externalización, combinación e internalización.

En el contexto del e-learning y la IA, este modelo permite analizar cómo los estudiantes convierten la información obtenida de herramientas digitales en conocimiento significativo. Sin embargo, cuando el proceso se limita a la obtención rápida de respuestas, sin reflexión crítica, se corre el riesgo de interrumpir la espiral del conocimiento, afectando la profundidad del aprendizaje.

2.3 Inteligencia artificial en educación superior

La inteligencia artificial ha emergido como una herramienta clave en la educación, facilitando la personalización del aprendizaje, la automatización de procesos y el acceso inmediato a la información (Luckin et al., 2016). No obstante, investigaciones recientes advierten que su uso no crítico puede generar dependencia cognitiva y limitar el desarrollo de habilidades analíticas (Zawacki-Richter et al., 2023).

Asimismo, Holmes et al. (2022) destacan que la integración efectiva de la IA en educación requiere el desarrollo de competencias como la alfabetización digital, el pensamiento crítico y la verificación de fuentes, elementos esenciales para garantizar un aprendizaje significativo.

2.4 E-learning y transformación digital educativa

El e-learning ha evolucionado hacia modelos flexibles, interactivos y centrados en el estudiante. De acuerdo con Bond et al. (2023), la transformación digital en la educación superior implica no solo el uso de tecnologías, sino también la redefinición de los procesos pedagógicos, la evaluación y la interacción docente-estudiante.

En este contexto, el microaprendizaje (microlearning) ha ganado relevancia como estrategia didáctica, al ofrecer contenidos breves, específicos y accesibles, alineados con las nuevas dinámicas de atención y consumo de información.

2.5 Generación del “ya” y cultura de la inmediatez en el aprendizaje

La denominada generación del “ya” o cultura de la inmediatez hace referencia a un fenómeno sociotecnológico caracterizado por la expectativa de acceso inmediato a la información, la reducción de los tiempos de espera y la preferencia por

respuestas rápidas y soluciones instantáneas, impulsado por el uso intensivo de tecnologías digitales, redes sociales y dispositivos móviles. En los entornos contemporáneos, esta dinámica se ha intensificado debido a la hiperconectividad y al consumo de contenidos digitales fragmentados, lo que ha modificado los patrones de atención, interacción y aprendizaje de los estudiantes (Pedrero-Esteban & Barrios-Rubio, 2024; Rathnasekara et al., 2025; Li et al., 2025)

Desde una perspectiva educativa, esta generación presenta características específicas:

- Disminución del tiempo de atención sostenida
- Preferencia por contenidos breves y visuales
- Búsqueda de retroalimentación inmediata
- Tendencia a la multitarea digital

De acuerdo con estudios recientes en educación digital, la inmediatez no solo modifica las prácticas de acceso a la información, sino también los procesos cognitivos, favoreciendo aprendizajes rápidos, pero potencialmente superficiales si no se acompañan de procesos reflexivos (Carr, 2010; Zawacki-Richter et al., 2023).

En este sentido, la cultura de la inmediatez representa un desafío para los modelos educativos tradicionales, ya que exige la incorporación de estrategias pedagógicas que integren:

- Interactividad en tiempo real
- Retroalimentación inmediata
- Aprendizaje basado en la experimentación
- Uso de simulaciones y entornos dinámicos

Así, la generación del “ya” no debe entenderse únicamente como una limitación, sino como un cambio paradigmático que redefine la manera en que los estudiantes aprenden, interactúan y construyen conocimiento en la educación superior tecnológica.

3. METODOLOGÍA

3.1 Enfoque metodológico y diseño experimental

El presente estudio se desarrolló bajo un enfoque cuantitativo con diseño no experimental de tipo transversal, con alcances descriptivos y analíticos, orientado a identificar patrones de comportamiento en el uso del e-learning y la inteligencia artificial en estudiantes de educación superior tecnológica cuya edad oscila entre los 18 y 23 años.

El estudio se estructuró como un proceso de caracterización de variables educativas, estableciendo relaciones entre: uso de dispositivos, preferencias de contenido, uso de inteligencia artificial, nivel de inmediatez percibida y comportamiento de abandono



Este enfoque permite aproximarse a un análisis de tipo sistema socio-tecnológico, donde el estudiante es entendido como un nodo dentro de un ecosistema digital de aprendizaje.

3.2 Definición de variables de estudio

En congruencia con el enfoque de estudio, se operacionalizaron las variables en tres categorías:

- Variables independientes: tipo de dispositivo de acceso, plataforma de aprendizaje y herramientas de inteligencia artificial utilizadas.
- Variables dependientes: nivel de satisfacción del aprendizaje, percepción de inmediatez y abandono de actividades.
- Variables intervinientes: Tipo de contenido (microcontenido vs tradicional), frecuencia de uso y nivel académico.

Permitiendo con ello identificar elementos clave acordes al objeto de estudio del presente artículo.

3.3 Población, muestra y contexto de aplicación

La población estuvo conformada por estudiantes del Instituto Tecnológico de Cuautla, Instituto Tecnológico de Álvaro Obregón del Tecnológico Nacional de México y por la Universidad Aztlán del Estado de Morelos, en donde se seleccionó una muestra de 500 participantes mediante un muestreo no probabilístico por conveniencia, por ser un criterio común en estudios aplicados exploratorios, investigaciones cualitativas y situaciones de difícil acceso.

El contexto de aplicación corresponde a entornos de aprendizaje digital, caracterizados por, uso intensivo de plataformas virtuales, acceso móvil predominante e Integración creciente de inteligencia artificial.

3.4 Instrumento de medición

Se utilizó un cuestionario estructurado de 17 reactivos, diseñado bajo criterios de medición de variables tecnológicas y educativas.

El instrumento fue organizado en tres dimensiones:

- Acceso y uso de e-learning
- Interacción con inteligencia artificial
- Percepción del aprendizaje e inmediatez

Cada reactivo fue construido en escala categórica, permitiendo su análisis mediante frecuencias y porcentajes.

3.5 Procedimiento experimental

El procedimiento se desarrolló en tres fases:

Fase 1: Diseño del instrumento: Definición de variables, construcción de reactivos y validación preliminar.

Fase 2: Recolección de datos: Aplicación digital del cuestionario, registro automatizado y filtrado de respuestas inconsistentes.

Fase 3: Análisis de datos: Procesamiento en entorno estadístico, generación de distribuciones e identificación de patrones.



Figura 1. Fases de la metodología utilizada, 2026. (fuente propia)

3.6 Técnicas de análisis de datos

El análisis se realizó mediante:

- Estadística descriptiva: Frecuencias, porcentajes y distribuciones
- Análisis comparativo: Relación entre variables (ej. IA vs abandono), Identificación de tendencias, análisis interpretativo y categorización de respuestas abiertas.

Este enfoque es consistente con estudios tecnológicos donde se analizan indicadores de desempeño del sistema, como en redes IoT, pero aplicado a comportamiento educativo.

4. RESULTADOS

A continuación, se expone el análisis de los datos obtenidos a partir de la aplicación del instrumento diseñado, cuyo propósito fue caracterizar las dinámicas de aprendizaje digital y el uso de inteligencia artificial en estudiantes de educación superior tecnológica de entre 18 y 23 años. Los resultados se derivan del procesamiento estadístico de la información recabada mediante un cuestionario digital, lo que permitió identificar patrones de comportamiento, tendencias de uso tecnológico y percepciones asociadas al aprendizaje en entornos mediados por tecnologías emergentes.

El análisis se estructuró en tres niveles: descriptivo, inferencial e interpretativo. En el primer nivel, se presentan las distribuciones de frecuencia y proporciones que permiten caracterizar el perfil sociodemográfico y tecnológico de la muestra. Posteriormente, se identifican relaciones relevantes entre variables clave, tales como el uso de dispositivos móviles, la preferencia por formatos de contenido, la necesidad de inmediatez y la adopción de herramientas de inteligencia artificial. Finalmente, se realiza una interpretación integral de los hallazgos, orientada a comprender las implicaciones de estos comportamientos en los procesos de enseñanza-aprendizaje.



Los resultados evidencian la consolidación de un perfil de estudiante digital caracterizado por la movilidad, la inmediatez y la interacción constante con tecnologías inteligentes. Observándose una creciente integración de herramientas de inteligencia artificial como apoyo en actividades académicas, lo cual plantea tanto oportunidades para la optimización del aprendizaje como desafíos relacionados con la profundidad del conocimiento y el desarrollo del pensamiento crítico.

En conjunto, los hallazgos permiten no solo describir el comportamiento de los estudiantes en entornos digitales, sino también aportar elementos para la reflexión sobre la transformación de los modelos educativos en el contexto de la educación superior tecnológica, particularmente en instituciones del Tecnológico Nacional de México.

4.1 Perfil sociodemográfico del estudiante digital

Género

Se evidencia una ligera mayor participación femenina (58.06% mujeres y 41.94% hombres), lo cual puede influir en:

- Preferencias de aprendizaje
- Uso de tecnología
- Interacción en entornos digitales

Participación por edad

La muestra considero jóvenes estudiantes de educación superior tecnológica de entre 18 y 23 años, lo cual:

- Refuerza la validez del concepto de “generación del ya”
- Explica la alta adaptación tecnológica
- Justifica la preferencia por inmediatez y formatos dinámicos

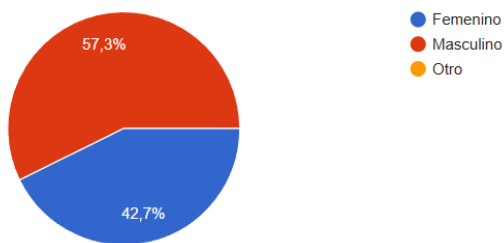


Figura 2. Participación por edad, 2026. (fuente propia)

4.2 Perfil de acceso y comportamiento digital

Dispositivo de acceso

- 88.06% móvil
- 10% computadora
- <2% otros

Las respuestas emitidas confirman que el aprendizaje ocurre en un ecosistema mobile-first, lo que implica:

- Aprendizaje fragmentado
- Acceso en contextos no formales
- Reducción del tiempo de atención prolongado

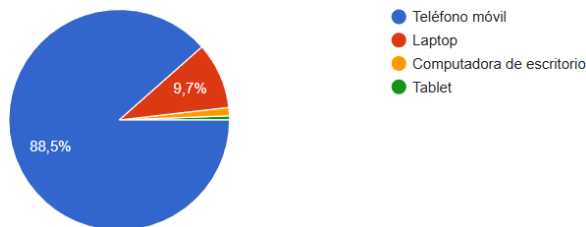


Figura 3. Tipos de dispositivos de acceso, 2026. (fuente propia)

Frecuencia de uso de plataformas

- 47.74% uso constante
- 25.16% frecuente
- 20% moderado
- 7.1% bajo

Más del 70% presenta alta interacción digital, lo que indica:

- Dependencia funcional del entorno digital
- Alta exposición a contenidos educativos y no educativos

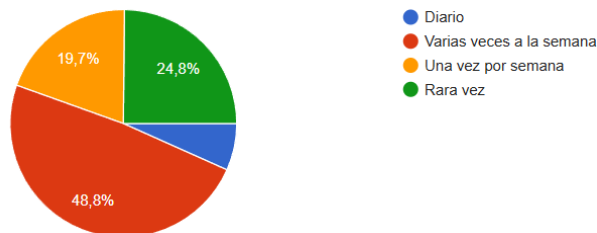


Figura 4. Frecuencia de uso de plataformas, 2026. (fuente propia)

Contenido preferido

- 55.48% videos cortos
- 20.65% textos
- 15.81% recursos mixtos
- 8.06% otros

Prevalce un modelo de microaprendizaje visual, asociado a:

- Plataformas tipo TikTok, Reels, Shorts
- Consumo rápido de información
- Reducción del esfuerzo cognitivo

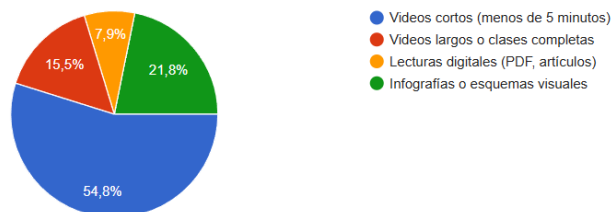


Figura 5. Tipo de contenido preferido, 2026. (fuente propia)

4.3 Inmediatez y comportamiento cognitivo

El estudiante no solo prefiere la inmediatez, sino que condiciona su permanencia en el proceso de aprendizaje a ella, implicando: Baja tolerancia a la espera, necesidad de estímulo constante y dependencia de retroalimentación inmediata.

Desde un enfoque cognitivo, se fortalece el aprendizaje reactivo y se debilita el aprendizaje reflexivo profundo

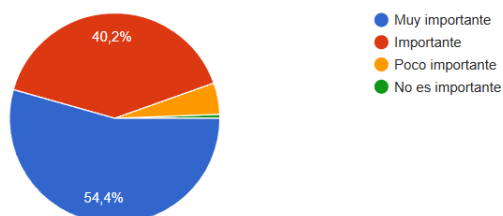


Figura 6. Importancia de obtención de respuestas rápidas, 2026. (fuente propia)

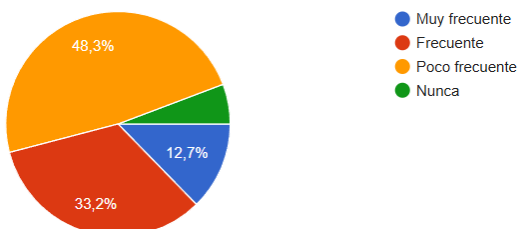


Figura 7. Frecuencia de abandono de actividades en línea por respuestas tardías, 2026. (fuente propia)

4.4 Interacción y aprendizaje social

Herramientas de interacción

- Digitalización.
- Trabajo colaborativo (pero no es predominante).
- Motivación mediada por: Tecnología, Inmediatez y Retroalimentación
- 92.58% de los estudiantes utilizan plataformas digitales para interactuar y generar trabajo colaborativo (principalmente mensajería/IA)

4.5 Uso de inteligencia artificial

Herramientas más utilizadas

- ChatGPT: 44.84%
- Gemini: 29.35%

- Meta AI: 4.52%
- Copilot: 3.55%
- Otras (DeepSeek, Claude, Perplexity): <3%

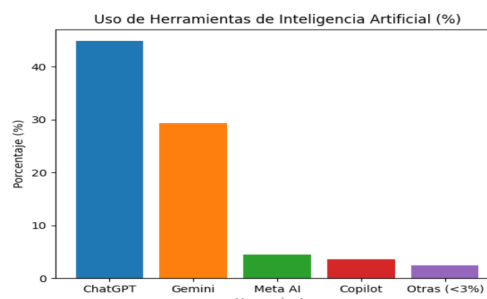


Figura 8. Porcentaje de uso de herramientas de Inteligencia Artificial, 2026. (fuente propia)

Existe dominación del uso de Chatgpt, posicionándose como la herramienta principal de aprendizaje, funcionando como un tutor digital, sin embargo, existe la apertura al uso de nuevas tecnologías.

4.6 Interacción y aprendizaje social

El 91.9% de los estudiantes utiliza redes sociales para actividades académicas, destacando el uso de WhatsApp como principal medio de comunicación.

Este hallazgo evidencia una transición hacia modelos de aprendizaje social digital informal, que coexisten con los entornos institucionales.

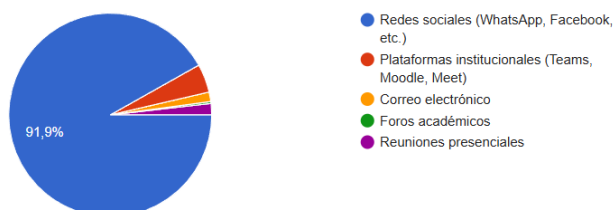


Figura 9. Principales medios utilizados para interactuar y realizar actividades académicas, 2026. (fuente propia)

4.7 Análisis cualitativo

El análisis de los datos evidencia una relación dual en la percepción del uso de herramientas digitales, particularmente aquellas asociadas a la inteligencia artificial. Por un lado, los estudiantes reconocen beneficios claramente alineados con las características de la denominada “generación de la inmediatez”, donde la rapidez en las respuestas, el apoyo en la resolución de tareas y el acceso inmediato a la información se posicionan como elementos altamente valorados. Estas ventajas reflejan un cambio en las expectativas del aprendizaje, orientado hacia la eficiencia, la optimización del tiempo y la disponibilidad constante de recursos, lo cual fortalece procesos de aprendizaje autónomo y flexible.

La identificación de una falta de profundidad en las respuestas sugiere que, si bien la información es accesible de manera rápida, no siempre cumple con los niveles de análisis requeridos en contextos académicos de mayor rigor. Las dudas

sobre la veracidad de la información evidencian una brecha en las competencias críticas de los estudiantes para evaluar fuentes digitales, lo que puede derivar en la reproducción de conocimientos poco confiables o superficiales.

Adicionalmente, la dependencia tecnológica emerge como un riesgo significativo, ya que puede limitar el desarrollo de habilidades cognitivas fundamentales como el pensamiento crítico, la reflexión analítica y la construcción autónoma del conocimiento. En este sentido, los resultados ponen de manifiesto la necesidad de replantear las estrategias pedagógicas en la educación superior tecnológica, incorporando enfoques que no solo integren el uso de la inteligencia artificial, sino que también fortalezcan la formación en habilidades metacognitivas, el análisis crítico de la información, el uso ético y responsable de la tecnología.

5. DISCUSIÓN

El análisis de la información permite identificar un patrón estructural claramente definido en el comportamiento de los estudiantes, el cual puede sintetizarse en la secuencia mostrada en la figura 10, que a continuación se muestra:

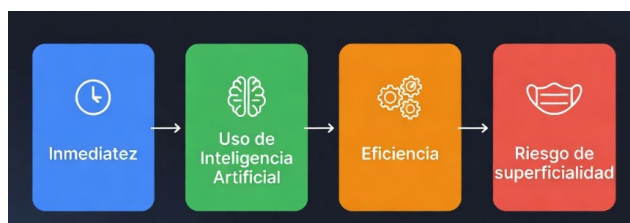


Figura 10. Patrón de comportamiento de los estudiantes, 2026.
(fuente propia)

Este modelo evidencia que la lógica dominante en el aprendizaje actual no es únicamente tecnológica, sino también cultural, marcada por la búsqueda de respuestas rápidas y la optimización del tiempo. En este sentido, la inteligencia artificial se inserta como un catalizador que potencia la eficiencia, pero que simultáneamente introduce desafíos relacionados con la profundidad y calidad del conocimiento construido.

Del análisis de la relación entre variables, se observa que la convergencia del uso de dispositivos móviles, el consumo de videos cortos y la utilización de herramientas de IA configura un tipo de aprendizaje caracterizado por su rapidez, pero también por su fragmentación. Este fenómeno sugiere que el conocimiento ya no se construye de manera lineal o profunda, sino en unidades breves y dispersas, lo cual puede afectar la integración conceptual y el desarrollo de estructuras cognitivas complejas.

La relación entre inmediatez y abandono revela un riesgo significativo de deserción digital. La necesidad constante de retroalimentación inmediata genera una baja tolerancia a la espera o a procesos formativos prolongados, lo que puede derivar en la interrupción de actividades cuando no se satisfacen dichas expectativas. Este hallazgo resulta especialmente relevante para el diseño de experiencias

educativas, ya que implica que la permanencia del estudiante está estrechamente vinculada con la dinámica de interacción y respuesta del entorno de aprendizaje.

Por otro lado, la asociación entre el uso de inteligencia artificial y las dudas sobre la veracidad de la información pone en evidencia la urgente necesidad de fortalecer la alfabetización crítica. Los estudiantes no solo requieren acceder a la información, sino desarrollar la capacidad de evaluarla, validarla y contextualizarla, particularmente en entornos donde la generación automatizada de contenidos puede incluir sesgos, errores o simplificaciones excesivas.

A nivel científico, los hallazgos permiten caracterizar al estudiante de educación superior tecnológica como un sujeto digital, inmediato, visual y altamente dependiente de la retroalimentación, lo cual redefine las bases tradicionales del proceso de enseñanza-aprendizaje. Por lo que, la inteligencia artificial deja de ser una herramienta complementaria para convertirse en un componente estructural del aprendizaje contemporáneo, influyendo directamente en la forma en que se accede, procesa y utiliza el conocimiento.

Sin embargo, emerge una dicotomía entre la eficiencia del aprendizaje y la profundidad del conocimiento. Mientras que las tecnologías digitales permiten aprender más rápido y con mayor accesibilidad, también pueden propiciar aprendizajes superficiales si no se integran bajo un enfoque pedagógico sólido.

Se hace necesario la incorporación de microcontenidos estructurados en el proceso de enseñanza-aprendizaje, capaces de responder a las dinámicas de atención actuales, pero sin sacrificar la coherencia y profundidad del aprendizaje. Por lo que, resulta indispensable implementar esquemas de evaluación continua con retroalimentación inmediata, que mantengan el compromiso del estudiante y reduzcan el riesgo de abandono.

Finalmente, la integración formal de la inteligencia artificial en los procesos educativos debe realizarse bajo un enfoque guiado, orientado al desarrollo de competencias críticas.

6. RETOS Y OPORTUNIDADES

La integración del e-learning y la inteligencia artificial en la educación superior representa una de las transformaciones más significativas en los modelos de enseñanza-aprendizaje. Estas tecnologías han ampliado las posibilidades de acceso al conocimiento, favoreciendo entornos flexibles, personalizados y centrados en el estudiante.

6.1 Retos en la integración del e-learning y la inteligencia artificial

a) Superficialidad del aprendizaje y fragmentación del conocimiento

La superficialidad del aprendizaje, derivada del predominio del microaprendizaje, el consumo de contenidos breves y la utilización de herramientas de inteligencia artificial orientadas a la obtención inmediata de respuestas, se relaciona con una transformación en los procesos cognitivos de los estudiantes,



caracterizada reducción del tiempo de atención y la preferencia por la información sintetizada.

b) Dependencia tecnológica y reducción de la autonomía cognitiva

El uso recurrente de herramientas de inteligencia artificial como apoyo en actividades académicas puede derivar en una dependencia excesiva, donde el estudiante delega procesos fundamentales como la búsqueda, el análisis y la síntesis del conocimiento.

Desde el enfoque del Humanismo Mexicano, este reto adquiere una dimensión ética, ya que la educación no solo debe formar individuos técnicamente competentes, sino sujetos capaces de pensar de manera autónoma, reflexiva y responsable. (Secretaría de Educación Pública [SEP], 2022; UNESCO, 2023).

c) Déficit en la verificación de información y pensamiento crítico

A pesar del uso extendido de herramientas de inteligencia artificial, persisten dudas significativas sobre la veracidad y confiabilidad de la información obtenida. Necesitándose reconfigurar los procesos educativos para incorporar el análisis crítico y uso ético de la información.

d) Deserción asociada a la falta de inmediatez

Uno de los hallazgos más relevantes del estudio es la fuerte relación entre la necesidad de inmediatez y el abandono de actividades académicas. Evidenciando que los estudiantes condicionan su permanencia en el proceso de aprendizaje a la disponibilidad de retroalimentación inmediata.

e) Desarticulación entre entornos formales e informales de aprendizaje

El uso predominante de plataformas como WhatsApp y otras redes sociales para fines académicos evidencia una migración hacia entornos informales de aprendizaje, lo cual genera una desarticulación con los sistemas institucionales tradicionales.

El reto consiste en reconocer estos entornos como espacios legítimos de aprendizaje y desarrollar estrategias que permitan su integración. Finalmente, en la gestión de las instituciones educativas el verdadero desafío no consiste en adaptarse a la era digital, sino a liderar su integración con sentido pedagógico, ético y estratégico.

6.2 Oportunidades para el desarrollo de aprendizaje social inteligente

a) Rediseño pedagógico basado en microaprendizaje estructurado

La preferencia por contenidos breves representa una oportunidad para rediseñar los procesos educativos mediante estrategias de microaprendizaje estructurado. Permitiendo organizar el conocimiento en unidades progresivas, facilitando su comprensión y retención.

b) Integración guiada de la inteligencia artificial en el aprendizaje

La adopción de herramientas de IA ofrece la oportunidad de incorporarlas de manera formal en los procesos educativos, como un recurso pedagógico que potencie la comprensión y la interacción, Desarrollando competencias específicas en los estudiantes, como el diseño de *prompts* efectivos, la evaluación de respuestas generadas por IA y el uso ético de la tecnología.

c) Integración de recursos educativos de retroalimentación inmediata

La necesidad de inmediatez puede ser aprovechada mediante la elaboración e implementación recursos educativos capaces de proporcionar retroalimentación automática (plataformas interactivas y herramientas de evaluación continua).

7. CONCLUSIONES

En síntesis, los hallazgos confirman que la denominada “generación del ya o cultura de la inmediatez digital” no constituye únicamente una tendencia emergente, sino un cambio estructural en los procesos de enseñanza-aprendizaje, caracterizado por la inmediatez, la movilidad y la orientación a la eficiencia. En este contexto, el e-learning y la inteligencia artificial se consolidan como componentes centrales del ecosistema educativo contemporáneo, redefiniendo las formas de acceso, procesamiento y construcción del conocimiento.

En este marco, los modelos de aprendizaje social inteligente emergen como una alternativa pertinente y viable, al articular la interacción colaborativa, el uso estratégico de la tecnología y el desarrollo de competencias críticas. Su implementación permite no solo responder a las demandas del contexto educativo actual, sino también avanzar hacia un enfoque formativo más integral.

En cuanto al futuro de los estudiantes como profesionistas, este nuevo perfil formativo tendrá un impacto directo en su desempeño en el ámbito laboral. Se prevé que los egresados desarrollen un estilo de trabajo caracterizado por la adaptabilidad, la autonomía, la capacidad de aprendizaje continuo y el uso intensivo de herramientas digitales e inteligencia artificial.

8. RECOMENDACIONES

- Implementar modelos pedagógicos híbridos que integren de manera equilibrada el e-learning y la inteligencia artificial para el aprendizaje profundo.
- Promover el uso guiado y crítico de herramientas de inteligencia artificial en el proceso formativo
- Desarrollar estudios longitudinales sobre el impacto del uso de la inteligencia artificial en el aprendizaje a lo largo del tiempo en relación con el desarrollo del pensamiento crítico y la autonomía cognitiva.
- Investigar el impacto del nuevo perfil de egresado en el ámbito laboral, evaluando su adaptabilidad, el aprendizaje continuo y el desempeño profesional.



9. REFERENCIAS

- Anderson, T. and Dron, J. (2011) 'Three generations of distance education pedagogy', *The International Review of Research in Open and Distributed Learning*, 12(3), pp. 80–97. Available at: <https://doi.org/10.19173/irrodl.v12i3.890>
- Bozkurt, A. and Sharma, R.C. (2020) 'Emergency remote teaching in a time of global crisis due to COVID-19 pandemic', *Asian Journal of Distance Education*, 15(1), pp. i–vi.
- Castells, M. (2010) *La sociedad red: Una visión global*. Madrid: Alianza Editorial.
- García-Peñalvo, F.J. (2023) 'Digital transformation in higher education: Challenges and opportunities in the age of artificial intelligence', *Education in the Knowledge Society*, 24, e31271. Available at: <https://doi.org/10.14201/eks.31271>
- Luckin, R., Holmes, W., Griffiths, M. and Forcier, L.B. (2016) *Intelligence unleashed: An argument for AI in education*. London: Pearson.
- Nonaka, I. and Takeuchi, H. (1995) *The knowledge-creating company: How Japanese companies create the dynamics of innovation*. Oxford: Oxford University Press.
- OECD (2021) *Artificial intelligence in education: Challenges and opportunities for sustainable development*. Paris: OECD Publishing. Available at: <https://doi.org/10.1787/0d7c9a1f-en>
- Selwyn, N. (2016) *Education and technology: Key issues and debates*. 2nd edn. London: Bloomsbury Academic.
- Siemens, G. (2005) 'Connectivism: A learning theory for the digital age', *International Journal of Instructional Technology and Distance Learning*, 2(1), pp. 3–10.
- UNESCO (2023) *Guidance for generative AI in education and research*. Paris: UNESCO Publishing.
- Han, L., Zhang, Y. and Liu, X. (2026) 'The SECI model of knowledge creation as an enabler of students' creative behavior: The mediating role of absorptive capacity', *Frontiers in Psychology*, 17, 1708055. Available at: <https://doi.org/10.3389/fpsyg.2025.1708055>
- Żatuchin, D. (2024) 'An analysis of the SECI model's application in course design in digital learning environments', *Education and Information Technologies*. Available at: <https://doi.org/10.1007/s44217-024-00229-0>
- Alavi, M. and Leidner, D. (2024) 'Knowledge management and artificial intelligence: Extending the SECI model in digital contexts', *Journal of Knowledge Management*.
- Tirotchi, S. (2024) 'Generation Z, values, and media: From influencers to BeReal', *Social Media + Society*. Available at: <https://doi.org/10.1177/20563051241234567>
- Secretaría de Educación Pública (SEP) (2022) *Plan de estudio para la educación básica 2022*. Ciudad de México: Gobierno de México. Available at: <https://www.gob.mx/sep>

ROS 2 Localization for Indoor Mobile Robots: A Systematic Review and Comparative Analysis

Damian Macedo-Garcia*, Jonathan Villanueva-Tavira, * Andrea Magadán-Salazar *, Jorge Fuentes-Pacheco*, Raúl Pinto-Elías*

**Tecnológico Nacional de México, CENIDET, C.P. 62490, Cuernavaca, Morelos, México. (e-mail: {m24ce057, jonathan.vt, andrea.ms, jorge.fp, raul.pe} @cenidet.tecnm.mx).*

Abstract: This paper presents a comprehensive analysis of the main ROS 2 localization tools used in indoor mobile robotics. Localization in indoor environments is one of the most important challenges for autonomous navigation. In this work, the principal tools reported in the literature which resolve this problem are systematically reviewed and compared. Each method is analyzed from four key perspectives: its algorithmic foundation, computational cost, required sensor inputs, and supported localization modality. The results show that Extended Kalman Filter-based approaches provide the best trade-off between accuracy and computational efficiency for IMU + odometry platforms (RMSE = 0.045 m), whereas factor graph-based methods achieve state-of-the-art precision for LiDAR + IMU systems (ATE = 0.08 m), although at a considerably higher computational cost.

Keywords: ROS 2, Mobile Robot Localization, Sensor Fusion, SLAM, Autonomous Navigation.

1. INTRODUCCIÓN

Los robots móviles han adquirido una importancia creciente en los últimos años debido a su capacidad para operar sin la supervisión de un humano. Su uso se ha expandido en numerosas aplicaciones, como la logística, la inspección, la agricultura, la manufactura, la vigilancia así como los robots de asistencia y servicio.

La navegación autónoma se refiere a la capacidad de un robot para desplazarse desde una posición inicial hasta una ubicación objetivo tomando decisiones de movimiento en base a lo que conoce. Esta capacidad requiere la integración de múltiples funciones. En primer lugar, el robot debe percibir su entorno mediante datos sensoriales. Después, debe estimar su propia posición dentro del entorno, determinar una trayectoria óptima para alcanzar el destino y ejecutar los comandos de movimiento correspondientes para seguir dicha trayectoria de forma segura y eficaz. Por esta razón, la navegación autónoma se entiende generalmente como un proceso multidisciplinario que involucra percepción, localización, cognición y control del movimiento.

Entre los distintos componentes que conforman la navegación autónoma, la localización ocupa un papel central, ya que proporciona una estimación de la posición y orientación del robot comúnmente denominada *pose*, la cual es esencial para garantizar un desplazamiento preciso y eficiente en el entorno. En este sentido, la localización es ampliamente reconocida como uno de los pilares fundamentales de la navegación autónoma.

En respuesta a esta relevancia, se han desarrollado múltiples herramientas y enfoques que facilitan su implementación. En particular, el uso de ROS 2 (del inglés *Robot Operating System*) ha permitido estandarizar y simplificar el despliegue de diversos algoritmos de localización. Esta plataforma

proporciona una arquitectura modular que favorece la integración de métodos heterogéneos, al tiempo que permite su adaptación a distintas configuraciones de hardware, reduciendo así la complejidad asociada a su implementación en sistemas robóticos reales.

1.1 Motivación y contribución

Este trabajo tiene como objetivo realizar una revisión sistemática de la literatura con el fin de analizar y comparar los principales enfoques de localización de robots móviles en entornos interiores implementados y de acceso libre en ROS 2. Las preguntas de investigación que se determinaron fueron las siguientes:

Q1: ¿Qué algoritmos de localización en ROS 2 ofrecen el mejor balance entre precisión y costo computacional para entornos interiores?

Q2 ¿Cuáles son las métricas de evaluación más apropiadas para determinar la superioridad entre algoritmos?

Q3 ¿Cómo ROS2 mejora las capacidades de localización respecto a ROS1?

Asimismo, se abordan la clasificación y los conceptos fundamentales de los algoritmos de localización en robótica móvil. Los criterios de selección priorizaron artículos de revisión y trabajos de investigación que presentaran resultados sólidos, ya sea mediante simulaciones o validaciones experimentales, con un enfoque en publicaciones de la última década, a fin de garantizar la inclusión de técnicas de vanguardia.

2. METODOLOGÍA

Esta sección detalla el procedimiento sistemático empleado para la selección y análisis de las herramientas de localización en el ecosistema ROS 2, siguiendo una adaptación del

protocolo PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) para asegurar la transparencia y reproducibilidad del estudio.

2.1 Estrategia de Búsqueda y Fuentes de Información

Para garantizar una recopilación exhaustiva de literatura científica y técnica, se consultaron las siguientes bases de datos académicas y repositorios especializados: IEEE Xplore, ScienceDirect (Elsevier), Scopus, Web of Science y Google Académico. Adicionalmente, se revisó la documentación oficial de ROS 2 y repositorios de código abierto en GitHub para verificar el estado de mantenimiento de los paquetes analizados.

2.2 Palabras Clave y Cadena de Búsqueda

La localización de información relevante se realizó mediante el uso de términos representativos del área de estudio. Se emplearon combinaciones de palabras clave en inglés y español: "ROS 2", "localization", "mobile robot", "sensor fusion", "SLAM", "Factor Graphs", "EKF" y "AMCL". Un ejemplo de la cadena de búsqueda utilizada en bases indexadas fue:

("ROS 2" AND "localization" AND ("SLAM" OR "sensor fusion" OR "factor graphs"))

2.3 Criterios de Inclusión y Exclusión

Con el propósito de centrar de manera efectiva el análisis, se definieron criterios específicos para la selección de las herramientas y estudios:

Criterios de Inclusión:

- Publicaciones y herramientas desarrolladas o portadas a ROS 2 (distribuciones Humble, Iron o Jazzy) entre 2020 y 2025.
- Software de código abierto con disponibilidad de repositorio público.
- Estudios que reporten métricas de precisión cuantificables ATE (del inglés Absolute Trajectory Error) o RMSE (del inglés Root Mean Square Error) o análisis de desempeño computacional

Criterios de Exclusión:

- Herramientas exclusivas de ROS 1 sin soporte o migración activa a ROS 2.
- Documentos de carácter no académico, como cartas al editor o manuales de usuario sin validación técnica.
- Software con documentación incompleta o que requiera licencias comerciales restrictivas.

2.4 Fases de Inclusión y Exclusión

El proceso se estructuró en cuatro fases críticas para filtrar el volumen inicial de registros hasta identificar las 8 herramientas clave analizadas en este trabajo:

- 1) Identificación: Localización de registros en bases de datos y repositorios mediante las cadenas de búsqueda mencionadas ($n \approx 84$ registros).
- 2) Cribado: Eliminación de duplicados y descarte de herramientas diseñadas únicamente para entornos de simulación sin potencial de aplicación real ($n \approx 77$).
- 3) Elegibilidad: Revisión a texto completo para asegurar que las herramientas cumplieran con los estándares de la arquitectura Navigation2 (Nav2) y la convención de marcos de coordenadas REP-105 ($n \approx 35$).
- 4) Inclusión: Selección final de herramientas que cubren los principales paradigmas de localización: estimado, filtrado, probabilístico, optimización por grafos y visual ($n = 8$ herramientas sustentadas por 17 estudios principales)

Las herramientas seleccionadas son: (1) robot_localization (EKF/UKF), (2) nav2_amcl, (3) slam_toolbox, (4) Google Cartographer, (5) rtabmap_ros, (6) IRIS LaMa, (7) ZED ROS 2 Wrapper y (8) Isaac ROS Map Localization

3. FUNDAMENTOS TEÓRICOS

De acuerdo con Siegwart et al. (2011), la navegación autónoma se estructura en percepción, localización, cognición y control. En este contexto, la localización aborda la estimación de la pose del robot, respondiendo a la pregunta fundamental: ¿dónde estoy?.

La localización robótica se define como el proceso de estimar la pose del robot posición y orientación dentro de un entorno. Esta estimación, típicamente representada por un vector de estado (x, y, z, θ) , es esencial para la interpretación de datos sensoriales, la planificación de trayectorias y la toma de decisiones. El esquema de la figura 1 funciona como la base general de cualquier algoritmo de localización.

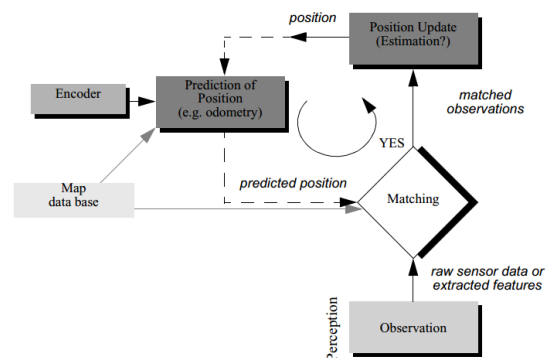


Figura 1. Diagrama general del algoritmo de un sistema de localización para un robot móvil (Siegwart et al., 2011).

Según Siegwart et al. (2011), existen dos enfoques principales:

Local (relativa): Se enfoca en estimar el movimiento del robot respecto a un estado previo, utilizando generalmente sensores propioceptivos como los codificadores de las ruedas. Es un

proceso iterativo de navegación estimada (*dead reckoning*) que proporciona datos a alta frecuencia, pero es vulnerable a la acumulación ilimitada de errores o deriva (*drift*)

Global (absoluta): Busca determinar la posición y orientación del robot con respecto a un marco de referencia fijo o un mapa previo del entorno. Su función principal es corregir el error acumulado por los métodos locales mediante la observación de características externas.

3.1 ROS 2 como Middleware en Robótica

El Robot Operating System 2 (ROS 2) es un middleware de código abierto diseñado para el desarrollo de sistemas robóticos distribuidos y escalables. A diferencia de un sistema operativo convencional, ROS 2 no gestiona directamente los recursos de hardware, sino que proporciona una infraestructura de comunicación modular en la que las aplicaciones se estructuran como un grafo de nodos interconectados.

Una de las principales características de ROS 2 es su adopción del estándar *Data Distribution Service* (DDS) como capa de comunicación subyacente (Macenski et al., 2022). Este enfoque permite el descubrimiento automático de nodos y una comunicación descentralizada, eliminando la necesidad de un nodo maestro central y reduciendo los puntos únicos de fallo. Asimismo, la interacción entre los distintos componentes del sistema se lleva a cabo mediante mecanismos de comunicación bien definidos, como *topics* para el intercambio de datos continuos, *services* para interacciones síncronas, *actions* para la ejecución de tareas de larga duración con retroalimentación, y parámetros para la configuración dinámica de los nodos, la Figura 2 muestra la arquitectura de comunicación interno que usa ROS2.

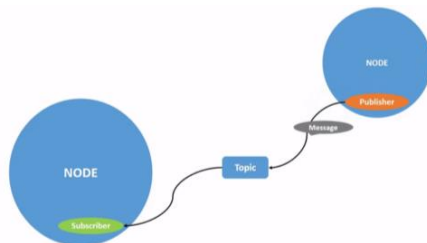


Figura 2. Arquitectura de comunicación entre nodos que utiliza ROS2.

3.2 Clasificación de algoritmos de localización.

Desde una perspectiva de algoritmos, las herramientas se subclasifican en las siguientes clases

3.3 Navegación Estimada

La odometría constituye la base de la localización en robots móviles, permitiendo estimar la pose a partir de sensores internos (Siegwart et al., 2011).

Desde una perspectiva matemática, en robots de tracción diferencial, la pose del sistema se representa mediante el vector de estado $x_k = [x, y, \theta]^T$, donde x e y describen la posición del robot en el plano cartesiano, y θ corresponde a su orientación con respecto a un marco de referencia, el cual se actualiza mediante la integración incremental de los

desplazamientos de las ruedas. En particular, el cambio de orientación se obtiene a partir de la diferencia de desplazamiento entre las ruedas, normalizada por la distancia entre ellas, como se aprecia en la figura 3.

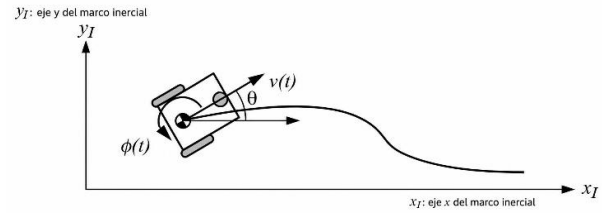


Figura 3. Representación de la cinemática planar (2D) del robot (Siegwart et al., 2011).

En términos de incertidumbre, los errores asociados se clasifican en dos categorías: deterministas, derivados de imperfecciones de calibración (por ejemplo, diferencias en el diámetro de las ruedas), y no deterministas, relacionados con fenómenos estocásticos como el deslizamiento o la interacción con el terreno. Estos errores se propagan de forma no lineal en el tiempo, generando un crecimiento no acotado de la incertidumbre, fenómeno conocido como drift.

3.4 Filtrado No Lineal y Fusión Sensorial

Este tipo de sistemas buscan fusionar la información de diferentes sensores en diferentes niveles de información con el objetivo de aumentar el contenido útil de las lecturas y mitigar el efecto del ruido sensorial, es decir, aumentar la información de varios sensores ayuda a tener un conocimiento del entorno mucho más amplio. Típicamente se puede encontrar de la siguiente forma.

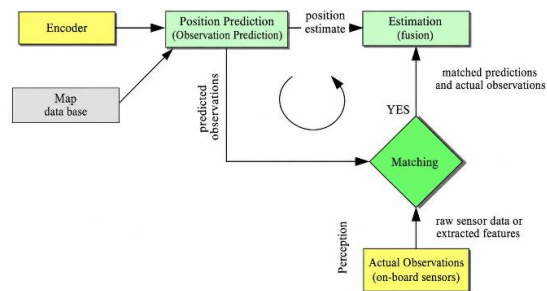


Figura 4. Esquema general de fusión de datos multisensor para la estimación de la pose (Siegwart et al., 2011).

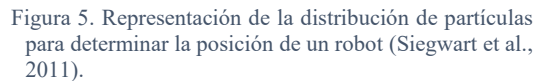
De acuerdo al algoritmo seleccionado, este proceso puede realizarse, por ejemplo, mediante la extracción de características del entorno (como líneas, puertas o marcas visuales) o a través del procesamiento directo de datos crudos mediante arquitecturas fuertemente acopladas.

Entre los métodos más conocidos y utilizados para la estimación recursiva de estados en robótica encontramos el Filtro de Kalman convencional, diseñado para sistemas lineales, y el Filtro de Kalman Extendido (EKF), el cual es la opción preferida para sistemas no lineales, al linealizar los modelos de movimiento y observación mediante expansiones de Taylor de primer orden y Jacobianos. Asimismo, destaca



3.5 Localización Probabilística basada en Filtros de Partículas

A diferencia de los métodos de hipótesis única, este enfoque mantiene una distribución de creencias sobre las posibles posiciones del robot, permitiendo representar múltiples estados simultáneamente como se ve en la Figura 5.



3.6 Mediante grafos de factores

```

graph LR
    subgraph Input_Sensor_Data [Input Sensor Data]
        RangeData[Range Data  
Laser/range/  
Point clouds]
        OdometryPose[Odometry  
Pose]
        IMUData[IMU Data  
(linear  
acceleration/  
Angular  
velocity)]
    end

    subgraph Local_SLAM [Local SLAM]
        VoxelFilter[Voxel Filter  
(fixed size)]
        PoseEstimator[PoseEstimator]
        AdaptiveVoxelFilter[Adaptive  
Voxel Filter]
        ScanMatching[Scan Matching  
(score)]
        PoseObservation[PoseObservation]
        VoxelFilterSize[Voxel Filter Size  
(thresholding for  
motion or time)]
        Dropped[Dropped]
    end

    subgraph Global_SLAM [Global SLAM (background threads)]
        GlobalSLAM[Global SLAM (background threads)]
        ComputeConstraints[Compute Constraints  
(on the order of 10^4  
INTER. loop closures)]
        SystemPoseAdjustment[System Pose  
Adjustment]
        ExtrapolatePoses[Extrapolate all  
poses that were  
added later]
        Submap[Submap]
        VoxelGridUpdate[Voxel Grid  
Update  
(batch)]
        MovementOrOld[Movement or Old]
        SearchRebuild[Search/Rebuild  
(time, pose,  
range data,  
submap)]
    end

    RangeData --> VoxelFilter
    OdometryPose --> PoseEstimator
    IMUData --> PoseEstimator
    VoxelFilter --> PoseEstimator
    PoseEstimator --> AdaptiveVoxelFilter
    AdaptiveVoxelFilter --> ScanMatching
    AdaptiveVoxelFilter --> PoseObservation
    PoseEstimator --> PoseEstimate[PoseEstimate]
    PoseEstimate --> GlobalSLAM
    PoseEstimate --> VoxelFilterSize
    VoxelFilterSize --> Dropped
    PoseEstimate --> Submap
    Submap --> VoxelGridUpdate
    Submap --> MovementOrOld
    VoxelGridUpdate --> SearchRebuild
    GlobalSLAM --> ComputeConstraints
    ComputeConstraints --> SystemPoseAdjustment
    SystemPoseAdjustment --> ExtrapolatePoses

```

Nodos (Vértices): Representan el estado del robot (su pose) en distintos instantes de tiempo a lo largo de su recorrido.

Aristas (Restricciones): Corresponden a los vínculos espaciales o restricciones derivadas de la odometría y de las observaciones sensoriales. Estas se dividen principalmente en dos tipos:

Restricciones Locales: Generadas por sensores propioceptivos (como la odometría de ruedas o IMU), que vinculan nodos consecutivos en el tiempo para definir el movimiento relativo.

Restricciones Globales (Cierres de Bucle): Se originan cuando el sistema, mediante sensores exteroactivos (como LiDAR o visión), reconoce un entorno visitado previamente. En este momento, se añade una arista que vincula nodos no consecutivos, permitiendo corregir la deriva acumulada.

El objetivo matemático de este paradigma es resolver un problema de optimización no lineal para minimizar el error global de todo el grafo simultáneamente. Al detectar un cierre de bucle, se emplean “resolvedores” numéricos avanzados para ajustar la posición de todos los nodos de manera que las restricciones impuestas por los sensores sean consistentes entre sí. Dando un sistema que en conjunto se ve de la siguiente forma.

3.7 Localización basada en visión

La localización basada en visión se articula principalmente a través de dos metodologías: la Odometría Visual (VO) y el SLAM Visual (VSLAM). Estos enfoques utilizan sensores ópticos para estimar el movimiento del robot mediante el seguimiento de características o rasgos distintivos en el entorno a través de secuencias de imágenes.

Odometría Visual Inercial (VIO): Para obtener una estimación de estado 3D robusta, las fuentes destacan la combinación de



cámaras (monoculares, estéreo o RGB-D) con unidades de medición inercial (IMU).

Localización basada en Apariencia: Este método permite al robot reconocer lugares visitados previamente mediante un mecanismo conocido como "Bolsa de Palabras" (Bag of Words).

En lugar de depender exclusivamente de la geometría, el sistema identifica patrones visuales únicos para detectar "cierres de bucle" (loop closures), lo que permite corregir de forma retroactiva el error acumulado en la trayectoria.

4. COMPARATIVA DE LAS HERRAMIENTAS

Esta sección presenta y compara las ocho herramientas de localización seleccionadas conforme a tres dimensiones complementarias: paradigma algorítmico, eficiencia computacional y precisión posicional. Las Tablas 1, 2 y 3 sintetizan, respectivamente, la clasificación por familia algorítmica, el consumo de recursos y las métricas de exactitud reportadas en la literatura revisada. A partir de este análisis cuantitativo se responden, en la Sección 4.4, las preguntas de investigación planteadas en la Sección 1.

4.1 Clasificación por paradigma algorítmico

La Tabla 1 clasifica las herramientas según su familia algorítmica y los sensores que requieren. Esta categorización permite comprender el fundamento técnico de cada solución y orienta la selección en función del hardware disponible en la plataforma robótica.

Tabla 1. Herramientas clasificadas por algoritmo

Herramienta	Familia Algorítmica	Sensores
slam_toolbox	Grafo de poses	LiDAR 2D, IMU
Google Cartographer	Grafo de poses	LiDAR 2D/3D, IMU, Odometría
rtabmap_ros	Grafo de poses	Cámaras RGB-D, Estéreo, LiDAR 2D
zed-ros2-wrapper	Visual SLAM / VIO	Cámaras ZED GNSS
iris_lama	Filtro de partículas	LiDAR 2D
nav2_amcl	Filtro de partículas	LiDAR 2D, Odometría
robot_localization	Fusión de sensores y Navegación estimada	Odometría de ruedas, IMU, GPS
Isaac ROS Map Localization	Filtro de partículas	LiDAR 2D

Como se aprecia en la Tabla 1, tres de las ocho herramientas analizadas se apoyan en grafos de poses (slam_toolbox, Google Cartographer y rtabmap_ros), lo que refleja la consolidación de este paradigma como el enfoque de mayor precisión para SLAM en interiores. Las herramientas basadas en filtros de partículas (nav2_amcl, iris_lama e Isaac ROS Map Localization) requieren exclusivamente LiDAR 2D y destacan por su bajo consumo de recursos; el manejo del mapa de ocupación subyacente se apoya en bibliotecas de representación genérica de entornos como la descrita por Fankhauser and Hutter (2016). Rtabmap_ros y zed-ros2-wrapper soportan cámaras RGB-D, estéreo y LiDAR, convirtiéndolas en opciones flexibles para entornos con requisitos de percepción multimodal (Labbé and Michaud, 2019). Finalmente, robot_localization ocupa una categoría diferenciada como sistema que combina odometría, IMU y GPS para proporcionar una estimación de pose continua y de baja latencia (Moore and Stouch, 2014).

4.2 Desempeño computacional

La Tabla 2 presenta el consumo de CPU y RAM, la tasa de procesamiento en fotogramas por segundo (FPS) y el hardware mínimo requerido para cada herramienta. Los valores provienen de evaluaciones reportadas en la literatura y son indicativos: pueden variar según la configuración del entorno, los parámetros del sistema y la carga del resto del stack de navegación.

Tabla 2. Consumo de recursos computacionales por herramienta.

Algoritmo	Uso CPU	RAM	FPS	Hardware
iris_lama	Bajo (~20-30%)	100-200 MB	20-30 FPS	CPU estándar (i7)
Google Cartographer	Alto (60-90%)	500 MB/1 GB	5-10 FPS	i7/RPi4
nav2_amcl	Muy bajo (~10%)	50-100 MB	50+ FPS	Cualquier CPU
robot_localization	Muy bajo (~5%)	<50 MB	Real-time	Cualquier CPU
slam_toolbox	Moderado (30-50%)	200-400 MB	15-25 FPS	RPi4/i7
rtabmap_ros	Alto (50-80%)	1-4 GB	5-15 FPS	i7+ RAM alta
zed-ros2-wrapper	Moderado (25-40%)	300 MB	15-30 FPS	Jetson/RPi
Isaac ROS Map Loc.	CPU mínimo (<10%)	500 MB/1B	30+ FPS	Jetson/RTX GPU

Los datos de la Tabla 2 evidencian una clara correlación entre la complejidad algorítmica y la demanda de recursos. `robot_localization` y `nav2_amcl` representan los extremos de máxima eficiencia, operando con menos del 10 % de CPU y menos de 100 MB de RAM, lo que los hace aptos para hardware de gama baja. En el extremo opuesto, `rtabmap_ros` exige entre 1 y 4 GB de RAM y una utilización de CPU de hasta el 80 %, lo que restringe su uso a plataformas de alto rendimiento. `slam_toolbox` y Google Cartographer se ubican en un rango intermedio, posicionándose como soluciones balanceadas para hardware de clase Raspberry Pi 4 o Intel i7. Isaac ROS Map Localization constituye un caso particular: aunque requiere una GPU NVIDIA (Jetson o RTX), su carga de CPU es mínima (<10 %), lo que ilustra el potencial de la aceleración por hardware especializado para reducir la latencia de procesamiento.

4.3 Precisión posicional

La Tabla 3 resume el error cuadrático medio (RMSE, del inglés Root Mean Square Error) reportado para cada herramienta en los conjuntos de datos de referencia disponibles en la literatura. Las comparaciones directas entre herramientas deben realizarse con cautela, ya que los resultados se obtuvieron sobre conjuntos de datos y condiciones experimentales heterogéneas.

Tabla 3. Precisión de las herramientas en términos de RMSE.

Algoritmo	RMSE(m)	Entorno/Dataset
<code>iris_lama</code>	0.02	Dataset Intel (Online SLAM)
Google Cartographer	0.037	Dataset KITTI/ACES
<code>nav2_amcl</code>	0.038	Dataset industrial.
<code>robot_localization</code>	0.045	Dataset indoor industrial
<code>slam_toolbox</code>	0.13	Dataset TUM-like
<code>rtabmap_ros</code>	0.004-0.14	Dataset TUM RGB-D
<code>zed-ros2-wrapper</code>	0.05-0.20	Dataset visual handheld
Isaac ROS Map Loc.	No reportado	Dataset no especificado

Los resultados de la Tabla 3 revelan que `iris_lama` obtiene el menor RMSE (m) reportado (0.02 m) sobre el dataset Intel de referencia para SLAM en línea (eupedrosa, 2019), seguida de `rtabmap_ros` (0.004–0.14 m) (Labbé and Michaud, 2019), cuya amplia varianza refleja su sensibilidad a la calidad del sensor y la textura del entorno. Google Cartographer y `nav2_amcl` ofrecen precisiones comparables (~0.037–0.038 m) sobre datasets industriales de interiores, consolidándose como referencias para entornos con geometría bien definida (Ínce et al., 2025; Martín Rico et al., 2024). `Robot_localization` reporta un RMSE de 0.045 m, resultado notable dado su

mínimo consumo computacional (Moore and Stouch, 2014). `Slam_toolbox` presenta el mayor error (0.13 m) en configuración de simulación tipo TUM (Ínce et al., 2025), lo que puede estar relacionado con la parametrización del dataset más que con las capacidades reales del algoritmo. Isaac ROS Map Localization no reporta métricas cuantitativas en los estudios revisados, lo que limita su comparabilidad directa con el resto de las herramientas.

4.4 Discusión: respuestas a las preguntas de investigación

El análisis cuantitativo de las Tablas 1–3 permite responder de forma fundamentada las preguntas de investigación planteadas en la Sección 1.

En respuesta a la **Q1**, los trabajos revisados identifican al Filtro de Kalman Extendido (EKF) el cual es implementado en `robot_localization` (Moore and Stouch, 2014), como la solución más equilibrada para plataformas embebidas de bajo costo, reportando un RMSE de 0.045 m con una latencia de procesamiento de 4.8 ms. Esta característica lo posiciona como la línea base de referencia en aplicaciones donde la ejecución determinista en tiempo real constituye un requisito estricto. En el extremo superior del espectro de rendimiento, los métodos basados en grafos de factores, particularmente `slam_toolbox` y Cartographer alcanzan precisión del estado del arte con un Error de Trayectoria Absoluta (ATE) de 0.13 m en simulación y 0.17 m en entornos físicos interiores; sin embargo, esta precisión conlleva una utilización de CPU de hasta el 81.9 %, lo que puede comprometer su viabilidad en hardware embebido con recursos limitados. RTAB-Map, por su parte, se distingue por su capacidad de fusión multimodal combinando LiDAR, cámara RGB-D y odometría (Labbé and Michaud, 2019) y por soportar cierre de bucle (*loop closure*) en entornos interiores de geometría compleja, a costa de una mayor huella de memoria.

Respecto a la pregunta **Q2**, la literatura revisada evidencia que los criterios de evaluación han evolucionado sustancialmente más allá de las métricas puramente geométricas. Si bien el RMSE y el ATE continúan siendo indicadores fundamentales, la eficiencia computacional cuantificada mediante perfiles de utilización de CPU y RAM y la consistencia estructural del mapa en sesiones prolongadas han emergido como factores determinantes para la viabilidad operativa en plataformas embebidas. Este desplazamiento refleja la creciente demanda de sistemas de localización desplegables en tiempo real, en contraste con soluciones de alta fidelidad concebidas para procesamiento fuera de línea. Para entornos interiores con mapa previo disponible, AMCL implementado sobre Nav2 se consolida como el estándar de referencia por su alta precisión posicional (RMSE: 0.038 m) y su capacidad de relocalización global ante pérdida de seguimiento (Martín Rico et al., 2024). Por su parte, `iris_lama` ofrece un rendimiento hasta cinco veces superior al de AMCL en términos de tiempo de procesamiento (eupedrosa, 2019), consolidándose como alternativa eficiente en entornos con recursos limitados.

En respuesta a la **Q3**, las ventajas arquitectónicas de ROS 2 sobre su predecesor tienen como fundamento la adopción del middleware Data Distribution Service (DDS) (Open Robotics, 2024). Al eliminar el nodo centralizado rosmaster, ROS 2

suprime el punto único de fallo inherente a ROS 1 (Macenski et al., 2022), mientras que las políticas de Calidad de Servicio (QoS) habilitan garantías configurables de confiabilidad, durabilidad y latencia. Estas características resultan especialmente relevantes en entornos interiores con obstrucciones físicas que degradan la comunicación inalámbrica entre nodos distribuidos.

5. CONCLUSIONES

Del análisis de las ocho herramientas se desprende que no existe una solución universal: el rendimiento depende del paradigma algorítmico, del hardware disponible y del contexto operativo. `robot_localization` con EKF (RMSE = 0.045 m, < 5% CPU) constituye la referencia para plataformas embebidas con requisitos de tiempo real (Moore and Stouch, 2014), mientras que `iris_lama` ofrece la mayor precisión reportada en la comparativa (RMSE = 0.02 m) con un consumo moderado de recursos (eupedrosa, 2019). Para entornos que requieren construcción simultánea de mapa y localización, `slam_toolbox` y Google Cartographer alcanzan mayor precisión de trayectoria con un ATE de 0.17 m en entornos físicos interiores, aunque con una demanda de CPU de hasta el 90% (Ínce et al., 2025). RTAB-Map se posiciona como la opción más versátil para plataformas con sensores RGB-D y necesidad de cierre de bucle robusto (Labbé and Michaud, 2019), a costa de una mayor huella de memoria. La integración de IA para el ajuste dinámico de covarianza del EKF abre una vía prometedora para superar estas limitaciones (Abaza, 2025).

En cuanto a las métricas, el RMSE y el ATE continúan siendo referencias fundamentales, pero la eficiencia computacional y la consistencia estructural del mapa en sesiones prolongadas han emergido como criterios igualmente determinantes para la viabilidad en hardware embebido. La heterogeneidad de los datasets empleados en la literatura limita la comparabilidad directa entre herramientas, lo que evidencia la necesidad de benchmarks estandarizados para esta área. Desde el punto de vista arquitectónico, la adopción del middleware DDS en ROS 2 elimina el punto único de fallo del nodo maestro y habilita políticas QoS configurables, ventajas especialmente relevantes en entornos interiores con obstrucciones físicas que degradan la comunicación inalámbrica. El fin de vida de ROS Noetic en mayo de 2025 refuerza además la pertinencia de consolidar el ecosistema sobre ROS 2.

Finalmente, los resultados reportados en la literatura sugieren que la integración de técnicas de aprendizaje automático para la sintonización adaptativa de parámetros como el ajuste dinámico de matrices de covarianza del EKF mediante modelos de Random Forest produce mejoras medibles sobre las implementaciones base, permitiendo reducir el MAE (del inglés Mean Absolute Error) en sistemas de localización con ROS 2 (Abaza, 2025).

5.1 Trabajo futuro

Se identifican tres líneas de investigación prioritarias. Primero, evaluar las herramientas sobre hardware embebido real bajo condiciones dinámicas, considerando plataformas con distintos esquemas de locomoción (Mamani et al., 2025) y arquitecturas de adquisición de señales de bajo costo (Tinoco-

Ramírez et al., 2026), así como la integración de sensores RGB-D para detección de objetos en tiempo real (Sernaque-Julca and Quiroz-Compeán, 2026). Segundo, estudiar sistemáticamente el ajuste dinámico de covarianza del EKF mediante aprendizaje automático (Abaza, 2025) y su impacto sobre precisión y latencia. Tercero, explorar estrategias de mapeo móvil con sensores heterogéneos (González-Santamarta et al., 2025) y localización híbrida interior-exterior (Cao et al., 2025) para superar las limitaciones individuales de cada paradigma.

REFERENCIAS

- Abaza, B.F. (2025). AI-Driven Dynamic Covariance for ROS 2 Mobile Robot Localization. *Sensors*, 25 (10), 3026. MDPI, Basel.
- Cao, Y., Usai, A., Gu, W. and Rizzo, A. (2025). Localization Handover for Mobile Robots: A Seamless Indoor-Outdoor Approach. *Journal of Intelligent & Robotic Systems*, 111 (3). Springer Nature, Cham.
- Pedrosa, E., Pereira, A., & Lau, N. (2018). A Sparse-Dense Approach for Efficient Grid Mapping. En 2018 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC).
- Fankhauser, P. and Hutter, M. (2016). A Universal Grid Map Library: Implementation and Use Case for Rough Terrain Navigation. In A. Koubaa (ed.), *Robot Operating System (ROS) – The Complete Reference (Volume 1)*, chapter 5. Springer, Berlin.
- González-Santamarta, M.Á., Rodríguez-Lera, F.J. and Matellan-Olivera, V. (2025). VAULT: A Mobile Mapping System for ROS 2-based Autonomous Robots. *arXiv preprint*, 2506.09583v1. Cornell University, Ithaca.
- Ínce, Í., Yiltas-Kaplan, D. and Keleş, F. (2025). From Simulation to Reality: Comparative Performance Analysis of SLAM Toolbox and Cartographer in ROS 2. *Electronics*, 14 (24), 4822. MDPI, Basel.
- Labbé, M. and Michaud, F. (2019). RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *Journal of Field Robotics*, 36 (2), 416-446. Wiley, Hoboken.
- Macenski, S., Foote, T., Gerkey, B., Lalancette, C. and Woodall, W. (2022). *Robot Operating System 2: Design, architecture, and uses in the wild*. Science Robotics, 7 (66). AAAS, Washington.
- Mamani, M., Uchamaco, J., Meneses, D., Silva, Y. and Apaza, J. (2025). Implementación y evaluación de un robot con sistema de locomoción wheeled-legged para la optimización del protocolo de inspección. *Ingenius, Revista de Ciencia y Tecnología*, 34, 126-139.
- Martín Rico, F., Guerrero, J.M., Pérez-Rodríguez, R., Díaz-Pascual, J. and García-Juan, A. (2024). Open Source



- Robot Localization for Non-Planar Environments. arXiv preprint, 2309.12744v2. Cornell University, Ithaca.
- Meeussen, W. (2023). Coordinate Frames for Mobile Platforms (REP-105). Open Robotics. Available from: <https://ros.org/reps/rep-0105.html>.
- Moore, T. and Stouch, D.W. (2014). A Generalized Extended Kalman Filter Implementation for the Robot Operating System. In Proceedings of the IEEE International Conference on Robotics and Automation (ICRA). IEEE, Hong Kong.
- Open Robotics (2024). ROS 2 Documentation: Humble Hawksbill. Available from: <https://docs.ros.org/en/humble/index.html>.
- Sernaque-Julca, J.C. and Quiroz-Compeán, G. (2026). Detección en tiempo real de las dimensiones de objetos usando cámaras RGB-D para navegación móvil. Padi Boletín Científico de Ciencias Básicas e Ingenierías, 14 (Especial), 75-83.
- Siegwart, R., Nourbakhsh, I.R. and Scaramuzza, D. (2011). Introduction to Autonomous Mobile Robots. MIT Press, Cambridge, MA.
- Tinoco-Ramírez, I., Sernaque-Julca, J.C. and Quiroz-Compeán, G. (2026). Diseño e implementación de un sistema de adquisición de señales para motores CD con ESP32 y Raspberry Pi. Padi Boletín Científico de Ciencias Básicas e Ingenierías, 14 (Especial), 39-45.
- Vizzo, I., Guadagnino, T., Mersch, B., Wiesmann, L., Behley, J. and Stachniss, C. (2023). KISS-ICP: In defense of point-to-point ICP – simple, accurate, and robust registration if done the right way. IEEE Robotics and Automation Letters, 8 (2), 1029–1036. IEEE, Piscataway.

Web platform for human-computer interaction for the detection and vocational guidance referral of applicants to higher education

Elena Elsa Bricio-Barrios *
Martha Cecilia Ramírez-Guzmán ** Santiago Arceo-Díaz *

* *Tecnológico Nacional de México, Instituto Tecnológico de Colima,
Departamento de Ciencias Básicas, e-mail:*

elena.bricio@colima.tecnm.mx, santiago.arceo@colima.tecnm.mx

** *Tecnológico Nacional de México, Instituto Tecnológico de Colima,
Departamento de Desarrollo Académico, e-mail:
cecilia.ramirez@colima.tecnm.mx*

Abstract: It reports the design and implementation of a web application based on the CHASIDE scale to identify the compatibility between the chosen degree program and the vocational profile of applicants to the Instituto Tecnológico de Colima. This platform, using a human-computer interaction approach, integrates data processing with interface design oriented toward user experience, facilitating interaction among applicants, coordinators, and tutors through automated reports and personalized recommendations.

Keywords: Classification algorithm; Data processing; Information system; Educational system; User interface; Decision support systems.

1. INTRODUCCIÓN

En México, de cada 100 niñas y niños que inician la educación primaria, aproximadamente 40 acceden a la educación superior y, de estos, solo 20 concluyen su formación universitaria (SEP, 2024). Aun entre quienes egresan, solo 3 están satisfechos con su elección de vocación (Gallup, 2024). Este déficit es significativo durante el primer año de la licenciatura, cuando se ha documentado la mayor deserción escolar (UNAD, 2024).

La elección vocacional representa un proceso complejo y multidimensional que involucra la integración de variables individuales como intereses, aptitudes, personalidad y valores, en interacción con factores como la oferta educativa, las condiciones socioeconómicas y las demandas del mercado laboral (Super, 1957) la presión por continuar con una profesión asociada a la tradición familiar, la elección de carreras universitarias influida por las tendencias sociales, la percepción de que ciertas áreas laborales ofrecen mejores oportunidades económicas (Rochin, 2021).

Este proceso adquiere especial relevancia durante la adolescencia y la juventud temprana, etapas en las cuales los individuos consolidan su identidad y proyectan su desarrollo profesional (SEP, 2024).

Si bien estos factores pueden inclinar la balanza en la elección de una carrera, una decisión vocacional más sólida requiere considerar la relación entre intereses y aptitudes (Dillon et al. 2023). Los intereses reflejan aquello que atrae, motiva o despierta curiosidad en la persona, mientras que las aptitudes hacen referencia a su potencial para

desenvolverse en determinadas áreas de conocimiento o acción (Pérez y Rodríguez, 2023).

Ante la dificultad de diferenciar con precisión entre aptitudes e intereses al elegir una carrera, las escalas vocacionales surgen como herramientas que permiten posicionar al participante en un perfil vocacional específico (Madrigal, 2024).

Este desafío ha sido atendido, a través de la orientación vocacional, cuyo objetivo facilita la toma de decisiones informadas y congruentes con las características personales del participante. Para ello, se apoya en el uso de instrumentos psicométricos que permiten evaluar de manera sistemática los intereses y preferencias ocupacionales, contribuyendo al autoconocimiento y a la reducción de la incertidumbre vocacional (Quishpe, 2022).

Estos instrumentos están diseñados para explorar de manera estructurada distintas dimensiones personales y reducir, en la medida de lo posible, sesgos de respuesta que podrían distorsionar la interpretación del resultado (Brown y Lent, 2013).

De las escalas disponibles, CHASIDE es una herramienta de orientación vocacional. Su nombre es un acrónimo de las áreas que mide: Administrativas, Humanísticas, Artísticas, Salud, Ingeniería, Seguridad y Ciencias Exactas (César et al. 2024). Esta herramienta de cribado primario se caracteriza por su practicidad, accesibilidad y utilidad.

A diferencia de otras escalas CHASIDE, ha mostrado alta eficiencia en su aplicación donde se requiere optimizar tiempos sin comprometer la calidad del proceso de ori-



entación (Anastasi y Urbina). El test proporciona información estructurada que facilita la toma de decisiones, disminuyendo la ansiedad asociada a la elección de carrera y aumentando la probabilidad de congruencia entre el perfil del estudiante y su elección profesional (Aragónés y Américo, 2010).

Esta escala está compuesta por 98 afirmaciones dicotómicas (sí/no), diseñada para identificar áreas de afinidad profesional a partir del análisis conjunto de intereses y aptitudes. Su estructura permite puntuar dimensiones asociadas a los perfiles vocacionales con el objetivo de reconocer fortalezas relativas que apoyen en la elección de la carrera universitaria (Holland, 1997).

Este instrumento está organizado de manera que en dos componentes complementarios se evalúan los intereses y las aptitudes. El conteo de respuestas positivas en cada grupo permite obtener puntajes parciales por dimensión, que posteriormente se integran para estimar el perfil vocacional predominante (Aragónés y Américo, 2010).

La interpretación del CHASIDE no se limita a qué área salió más alta la puntuación, sino que implica un análisis más fino en tres niveles: 1) En la identificación de áreas dominantes: Se identifican puntuaciones altas que representan intereses consolidados, tendencias vocacionales iniciales y posibles campos de estudio y/o carreras. Ejemplo Alto en S (Social), Inclinación hacia ayuda, interacción y servicio. 2) En el análisis de combinaciones se permite perfilar mejor al sujeto y evita interpretaciones rígidas y abre posibilidades interdisciplinarias. Ejemplo Alto en A (Arte) más H (Humanidades), Inclinación en la expresión artística con base humanista. 3) Detección de discrepancias. Ejemplos Interés alto más baja aptitud percibida, existe un desarrollo. Aptitud alta más bajo interés, posible talento no motivado. Las puntuaciones homogéneas hay una indecisión vocacional y puntuaciones muy bajas generales hay desmotivación o confusión.

La escala CHASIDE ha sido extensamente empleada en países latinoamericanos. Por ejemplo, en Ecuador se logró la evaluación de 350 estudiantes de nivel bachiller, reportando que el 49% de ellos coincidía la carrera elegida con su vocación profesional (Vacasela, 2019). En Colombia participaron 261 estudiantes y se cuantificó que el 58.3% del estudiantado mostraba intereses requeridos en la carrera universitaria seleccionada, mientras que el 44.4% coincidió en sus aptitudes con el perfil vocacional (Lopera et al. 2018). En México, la Universidad Autónoma de Tamaulipas y la Universidad Autónoma de Chalco utilizan esta escala para concientizar al estudiantado en la elección de carrera (Moreno, 2016).

La escala CHASIDE está disponible para imprimir en papel y ser respondida por la persona interesada y, a continuación, una persona con capacitación básica realiza el cruce de aptitudes e intereses para identificar las áreas vocacionales más fuertes. También se puede acceder a esta encuesta y contestarla en línea de forma automatizada y obtener de manera inmediata los resultados.

Esta propuesta es viable para la atención individualizada; no obstante, se convierte en un desafío cuando se desea evaluar grandes concentraciones de aspirantes para ingresar a una licenciatura, así como la identificación y canal-

ización oportuna del estudiantado para orientarlos hacia la carrera que se adecua más a su vocación profesional. Aunado a esto, la extensión del instrumento puede provocar cansancio en los participantes, favoreciendo respuestas repetitivas o poco diferenciadas, reportando puntajes elevados en todas las áreas vocacionales y, en consecuencia, perfiles difíciles de interpretar con confiabilidad.

El desarrollo de sistemas de interacción humano-máquina ha hecho posible la automatización de estas tareas, tanto para la obtención de resultados individuales como para el análisis integral de los elementos contenidos en la base de datos. Asimismo, estos sistemas permiten establecer categorizaciones automáticas que facilitan la localización de cada participante dentro del universo analizado y la generación inmediata de reportes específicos, útiles para procesos de seguimiento, orientación y toma de decisiones.

Los sistemas de interacción humano-computadora han comenzado a aplicarse en el análisis de perfiles profesionales y en la orientación vocacional. Estudios recientes de Ayay et al. (2024), diseñaron una plataforma web de consejería con interacción con el usuario que analiza la información académica. Massi et al. (2025) evaluaron la utilidad del uso de estas plataformas web y cómo cambió la decisión de la persona participante al obtener el resultado de su análisis. Seesukong et al. (2024) propusieron un sistema personalizado de recomendación de carrera basado en competencias y habilidades a través de información disponible en la red.

De manera complementaria, también se han desarrollado aplicaciones móviles que fortalecen la autoeficacia en la toma de decisiones vocacionales y sistemas de recomendación personalizados que relacionan competencias, intereses o habilidades con trayectorias posibles. En conjunto, estos antecedentes sugieren que las herramientas digitales pueden contribuir a la identificación temprana de desajustes vocacionales y a la generación de recomendaciones antes de que estos se traduzcan en rezago o abandono académico.

Se centra en la evaluación de intereses predominantes a partir de la afinidad del sujeto hacia diversas actividades. Su lógica se alinea con modelos de congruencia persona-ambiente, especialmente con la teoría de Holland (1997), donde se asume que las elecciones vocacionales exitosas dependen del ajuste entre intereses personales y contextos ocupacionales.

Por tanto, se reporta un sistema de recomendación vocacional basado en la escala CHASIDE, diseñado para aspirantes al Instituto Tecnológico de Colima. A diferencia de otras aplicaciones web de orientación vocacional, la plataforma desarrollada considera el contexto real del estudiantado de la institución académica e incorpora un sistema para la detección de patrones de respuesta homogéneos, así como la generación de alertas ("red flags") ante posibles inconsistencias en los resultados. Además, integra las necesidades de coordinadores y tutores mediante la emisión de reportes automatizados que facilitan la toma de decisiones académicas y la intervención oportuna.

2. PROCEDIMIENTO

2.1 Base de datos

Este estudio, de carácter transversal, se llevó a cabo durante el periodo comprendido entre el 30 de junio y el 3 de julio de 2025 en el Instituto Tecnológico de Colima. La recolección de datos se realizó mediante un muestreo no probabilístico de tipo bola de nieve, donde el personal del Departamento de Desarrollo Académico actuó como principal nodo de difusión, extendiendo la invitación por correo electrónico a todos los aspirantes interesados en ingresar a alguna de las diez carreras de nivel licenciatura ofertadas por la institución.

El instrumento (escala CHASIDE) fue aplicado mediante un formulario electrónico en Google Forms y estuvo compuesto por dos módulos:

- **Módulo introductorio:** se presenta el objetivo general del cuestionario y se explica el uso académico y confidencial de los datos, incluyendo la posible integración en un expediente electrónico de seguimiento académico y orientación. En esta fase se incorporó un proceso de consentimiento informado explícito, en el que el usuario podía aceptar o rechazar participar. En caso de no aceptar, el sistema agradecía su participación y finalizaba automáticamente la sesión sin registrar información adicional, garantizando así el cumplimiento de criterios éticos de participación voluntaria.
- **Módulo de aplicación:** si el participante confirmaba el consentimiento informado, se solicitaba el nombre completo de la persona participante y la carrera de interés a ingresar. Posteriormente, las afirmaciones de la escala CHASIDE se despliegan en bloques de siete afirmaciones por pantalla donde cada afirmación se alineaba con una dimensión de la escala CHASIDE y el usuario seleccionaba una respuesta dicotómica (“Sí” o “No”).

2.2 Identificación del perfil académico del aspirante

La Figura 1 muestra la conceptualización general del flujo de trabajo implementado en la aplicación web.

La Fig. 1 muestra un diagrama distribuido en tres módulos:

- (1) **Presentación:** Se incorporó un menú lateral izquierdo donde:
 - El usuario ingresa el vínculo asociado a las respuestas en formato .xls/.xlsx o en formato tabular exportable desde Google Forms, lo que permite reutilizar o clonar periódicamente la escala (<https://forms.gle/5Qviu3UJHcAMEowb6>) y, una vez concluida la participación de los aspirantes, cargar el vínculo para el procesamiento automatizado de los datos. Esta estrategia permite la actualización continua de la base de datos sin requerir modificaciones estructurales en la lógica de la aplicación.
 - Se ajusta la relación de ponderación entre intereses y aptitudes, la cual puede ser seleccionada a discreción de los especialistas responsables del

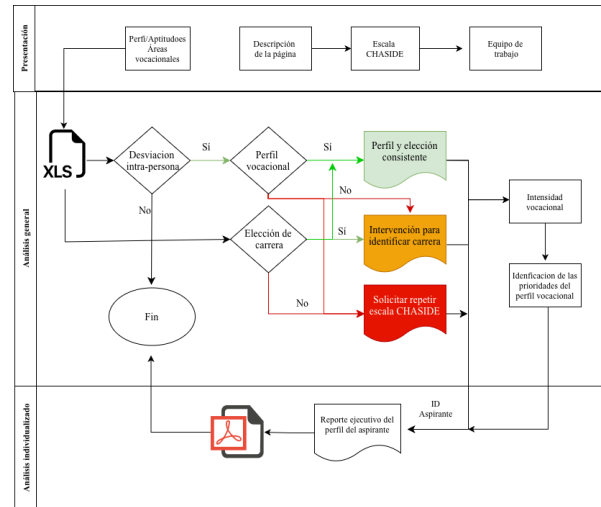


Fig. 1. Descripción de los módulos de la aplicación web.
Fuente: Elaboración propia.

análisis. Este ajuste permite modificar la sensibilidad del modelo respecto a la contribución relativa de ambos componentes, facilitando la exploración de distintos escenarios de decisión vocacional.

- Se define el perfil vocacional esperado para cada una de las carreras, de manera que el coordinador de carrera y el jefe de departamento puedan seleccionar cuáles son las “letras” de la escala con mayor afinidad con la ingeniería o licenciatura que representan. Este componente convierte a la aplicación en una herramienta paramétrica, adaptable a criterios disciplinares y a la experiencia contextual del personal académico.
- Una vez seleccionados estos parámetros, ya sea mediante los valores predeterminados o mediante configuraciones introducidas por personal calificado, se habilita el proceso de identificación automatizada del perfil vocacional del aspirante y la propuesta de acciones de mediación u orientación académica.
- (2) **Análisis general:** A partir del vínculo que contiene las respuestas de los participantes, la aplicación realiza una conversión de las respuestas dicotómicas a valores numéricos, asignando 0 a la opción “No” y 1 a “Sí”. Esta transformación permitió estructurar la información en forma matricial, facilitando su procesamiento computacional y análisis estadístico posterior.

Cada afirmación del instrumento fue asociada a una de las siete áreas vocacionales de la escala CHASIDE y se contabilizó el número total de respuestas afirmativas por cada área. Adicionalmente, se estimó la proporción de respuestas positivas respecto al total de reactivos por dimensión, lo que permitió estandarizar la comparación entre áreas con diferente número de ítems.

Posteriormente, se analizó cada fila, correspondiente a cada aspirante, mediante una prueba de desviación intrapersonal con el objetivo de identificar patrones de respuesta homogéneos o potencialmente aleatorios (Fernández et al. 2002). Aquellos



participantes que se ubicaron en el percentil inferior (10%) de la distribución de la desviación intrapersonal fueron etiquetados con la categoría “Respondió siempre igual” (color gris). Estos registros fueron excluidos del análisis subsecuente, al considerarse de baja confiabilidad para la inferencia del perfil vocacional.

Para los participantes que superaron este umbral, se realizó el cruce de las respuestas afirmativas con la matriz de intereses y aptitudes para obtener el perfil vocacional predominante. Este perfil se definió a partir de la dimensión con mayor puntuación combinada, considerando una ponderación ajustable entre intereses y aptitudes.

El perfil obtenido se comparó con la carrera seleccionada por el aspirante, generando una clasificación categórica:

- Si el perfil vocacional coincide con la carrera elegida, se asigna la categoría “Perfil corresponde con la carrera elegida” (color verde).
- Si el perfil vocacional no coincide con la elección, se asignó la categoría “Perfil NO va acorde con la carrera elegida” (color amarillo).
- Si el participante presentó una distribución homogénea de puntuaciones en todas las áreas, indicando ausencia de dominancia vocacional, se asignó la categoría “No se observa perfil prioritario” (color rojo).

Cada uno de los perfiles identificados fue almacenado en una base de datos interna para su análisis estadístico descriptivo. Se construyeron visualizaciones como diagramas de pastel para estimar la distribución porcentual de las categorías vocacionales y evaluar la necesidad de implementar estrategias de intervención durante el curso propedéutico. Asimismo, se generaron diagramas de barras apiladas para analizar la distribución por carrera y facilitar la identificación de grupos de riesgo o perfiles inconsistentes a nivel departamental.

A partir de los registros por carrera, se construyó un diagrama de Sankey que permite visualizar las transiciones entre la carrera elegida y las opciones vocacionales compatibles. Este recurso facilita la identificación de trayectorias académicas alternativas y promueve la articulación entre coordinaciones de distintas carreras.

Con base en los registros clasificados como “Perfil corresponde con la carrera elegida”, se realizó una estratificación por percentiles: los valores inferiores fueron etiquetados como “Perfil en transición”, mientras que los superiores fueron clasificados como “Joven promesa”. De manera análoga, para los registros “Perfil NO acorde con la carrera elegida”, se definieron las categorías “Sin perfil” y “Perfil en riesgo”. Estas categorías fueron organizadas por carrera con el fin de segmentar a los aspirantes según su nivel de ajuste vocacional.

Finalmente, se realizó un análisis comparativo entre grupos mediante el cálculo del error porcentual y su representación a través de un diagrama de Pareto. Este análisis permitió identificar las áreas vocacionales con mayor brecha entre perfiles de bajo y alto desempeño, orientando así la priorización de con-

tenidos y estrategias dentro del curso propedéutico. De esta manera, el curso no solo cumple una función de nivelación académica en habilidades básicas, sino que se configura como un espacio de intervención vocacional dirigido al fortalecimiento de competencias específicas alineadas al perfil del estudiante.

- (3) Información individual. A partir de los registros generados en el análisis general, este módulo permite consultar el diagnóstico particular de cada estudiante. En esta etapa se integra la información individual derivada del procesamiento automatizado, permitiendo evaluar si la carrera seleccionada presenta correspondencia con el perfil vocacional identificado. Asimismo, el sistema reporta las áreas en las que el participante muestra mayor fortaleza relativa respecto a otros aspirantes que eligieron la misma carrera, así como aquellas dimensiones que requieren reforzamiento.

En los casos en que se detecta un desajuste significativo entre el perfil vocacional del estudiante y el perfil esperado de la carrera, el módulo sugiere opciones académicas afines con base en la compatibilidad observada entre dimensiones CHASIDE y programas educativos. Este nivel de análisis permite pasar de una interpretación grupal a una intervención focalizada, útil para procesos de orientación vocacional y toma de decisiones académicas.

3. RESULTADOS

El código fuente fue desarrollado en Python y posteriormente integrado en un repositorio de GitHub para su control de versiones y mantenimiento, desde donde se migró a Streamlit para la visualización e interacción de la plataforma web disponible y gratuita en <https://chaside-v8e9yahvs2sszwulqfx3n.streamlit.app/>. La implementación final permitió desplegar un entorno de análisis accesible, orientado tanto al procesamiento automatizado de datos como a la consulta de resultados globales e individuales.

Previo a la visualización de los resultados generales y particulares de los participantes, se realizaron tres configuraciones iniciales:

- El ingreso del vínculo web generado por el cuestionario en Google Forms, el cual contenía los datos crudos de los participantes.
- En la sección “Ajustes del algoritmo”, personal del Departamento de Desarrollo Académico recomendó una ponderación de 80% para intereses y 20% para aptitudes, al considerar que, en la etapa inicial de ingreso a licenciatura, la afinidad vocacional declarada por el aspirante constituye un indicador más estable de elección profesional que la autopercepción de habilidades específicas.
- Para el apartado “Perfil esperado de la carrera” ofertada por el Instituto Tecnológico de Colima, se seleccionaron las dimensiones de la escala CHASIDE que reportan mayor correspondencia con el perfil de egreso del profesionista, así como con las áreas laborales y de investigación reportadas por los egresados de cada programa académico.

Identificados los parámetros del modelo, se ingresó al módulo “Análisis general”, donde se reportó la partici-



pación de 486 aspirantes. En el módulo introductorio, 6 participantes indicaron no desear colaborar, por lo que fueron descartados del presente estudio, mientras que el resto continuó con el módulo de aplicación.

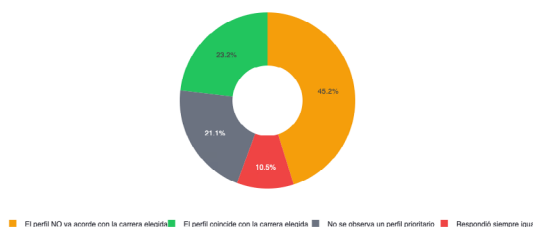


Fig. 2. Distribución de respuestas de la escala CHASIDE global. Fuente: Elaboración propia.

En la Fig. 2, se observa que el 23.2% presentó correspondencia entre el perfil vocacional identificado y la carrera elegida. En contraste, 45.2% mostró falta de correspondencia entre su perfil vocacional y la carrera seleccionada, lo que sugiere que sus intereses y aptitudes no se alinean de manera consistente con su elección académica. Este resultado podría ser un indicador temprano de riesgo de desajuste vocacional y eventual dificultad de adaptación al plan de estudios de la carrera elegida.

Asimismo, 21.1% de los aspirantes fueron clasificados en la categoría “No se observa un perfil prioritario”, lo que indica ausencia de una dominancia vocacional clara dentro de las dimensiones evaluadas. En términos operativos, este grupo requiere atención prioritaria mediante procesos de orientación vocacional, con el fin de profundizar en el conocimiento del plan de estudios, las áreas de formación y los posibles campos de desempeño profesional asociados a cada carrera.

Finalmente, 10.5% de los participantes presentaron una desviación intrapersonal ubicada en el criterio de baja confiabilidad, lo que llevó a clasificarlos en la categoría “Respondió siempre igual”. Estos registros fueron excluidos del resto del análisis, debido a que su patrón de respuesta no proporcionó evidencia suficiente para una identificación vocacional confiable y, por tanto, limitó la posibilidad de emitir recomendaciones de canalización académica.

La Fig. 3 tiene por objetivo identificar, por carrera, el listado nominal de los participantes, con el fin de que cada jefe de departamento, en coordinación con el coordinador de carrera, localice a los aspirantes clasificados en las categorías “No se observa un perfil prioritario” y “Respondió siempre igual”. Esta segmentación permitió establecer una ruta de seguimiento focalizada, orientada a la reaplicación de la escala CHASIDE y a la posterior definición de acciones de acompañamiento u orientación académica.

El objetivo operativo de esta fase fue intervenir de manera oportuna antes de la conclusión del curso propedéutico, incrementando la probabilidad de canalización temprana en los casos con mayor incertidumbre vocacional o baja confiabilidad en la respuesta.

Por otro lado, la Fig. 4 permite al coordinador de carrera estimar el nivel de heterogeneidad vocacional de la cohorte

de ingreso, así como identificar la proporción de aspirantes con perfiles de alto ajuste y aquellos que podrían representar mayores desafíos de acompañamiento académico. Esta información puede ser socializada con la academia del departamento y con el cuerpo de tutores, a fin de implementar estrategias de seguimiento más cercanas con el estudiantado ubicado en las secciones amarilla y roja, antes de que avance el semestre y comiencen a manifestarse dificultades en la acreditación de las asignaturas.

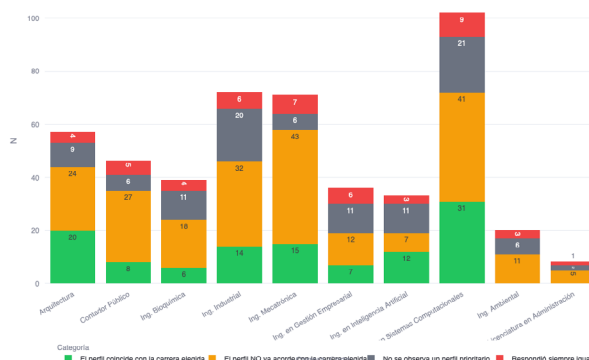


Fig. 3. Distribución de respuestas de la escala CHASIDE por carrera. Fuente: Elaboración propia.

Una vez identificados los participantes que no presentaban una correspondencia clara entre su elección de carrera y su perfil vocacional, así como aquellos clasificados en la categoría “Respondió siempre igual”, estos registros fueron separados del análisis de intensidad vocacional. A partir de los casos restantes, se construyó un diagrama de barras apiladas en el que se clasificó a los aspirantes según su nivel de ajuste vocacional en categorías de alto desempeño, desempeño intermedio y necesidad de atención individualizada.

Esta estratificación permite distinguir a los estudiantes con mayor congruencia entre perfil y elección profesional, de aquellos que podrían requerir intervención preventiva al inicio de su trayectoria universitaria. En particular, los aspirantes ubicados en la sección roja fueron considerados como el grupo de mayor prioridad, debido a su potencial riesgo de rezago académico y desajuste vocacional durante las primeras etapas de formación.

Es importante señalar que el estudiantado ubicado en las categorías roja y naranja no debe interpretarse, de manera directa, como alumnado de bajo desempeño académico. Estas categorías indican, más bien, una baja correspondencia entre la carrera elegida y el perfil vocacional estimado a partir de las habilidades e intereses recabados mediante la escala CHASIDE. En este sentido, la clasificación obtenida refleja un posible desajuste vocacional, no una medida de capacidad académica global.

Tomando en cuenta lo anterior, se construyó un diagrama de Sankey (Fig. 5) en el que se establece como punto de partida la carrera elegida por el aspirante y como punto de llegada el perfil vocacional compatible identificado por la aplicación. Esta representación permite visualizar de manera intuitiva las posibles transiciones entre la elección inicial del estudiante y aquellas carreras con mayor afinidad

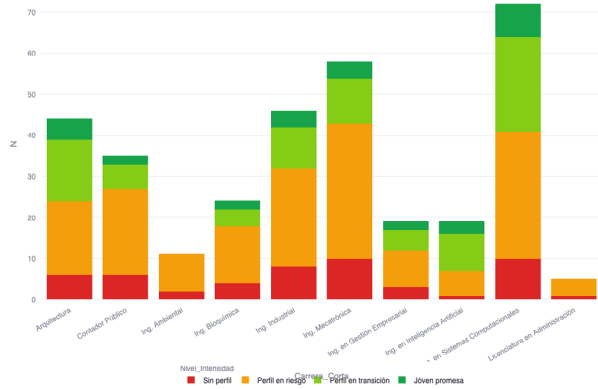


Fig. 4. Intensidad de perfil vocacional por carrera. Fuente: Elaboración propia.

vocacional, facilitando así la canalización hacia el programa académico para el cual el participante se encuentra mejor perfilado. A partir de esta información, el aspirante puede valorar, de manera informada, si desea mantener su elección original o considerar un posible cambio de carrera.



Fig. 5. Transición vocacional por carrera. Fuente: Elaboración propia.

Como última etapa, se construyó un diagrama de Pareto que permite identificar las áreas vocacionales prioritarias de intervención para cada carrera. Esta herramienta hace posible reconocer qué dimensiones concentran la mayor parte de la brecha entre aspirantes con menor y mayor ajuste vocacional, aportando un criterio cuantitativo para la toma de decisiones pedagógicas.

Con base en esta información, el curso propedéutico puede trascender su función tradicional de nivelación y convertirse en un espacio de fortalecimiento vocacional y académico, en el que el profesorado diseñe actividades específicas para atender las áreas de oportunidad detectadas. De igual forma, estos hallazgos pueden ser aprovechados por docentes de Ciencias Básicas y de asignaturas disciplinares para estructurar estrategias de acompañamiento acordes al perfil real del alumnado de nuevo ingreso.

En la Fig. 6 se observa que, para Ingeniería Mecatrónica, las áreas que requieren mayor fortalecimiento corresponden a Ciencias Experimentales (E), donde se sugiere la necesidad de reforzar habilidades asociadas con la observación sistemática, la experimentación, la interpretación de datos, el método y el pensamiento crítico. En segundo lugar, Enseñanzas Técnicas (I) indican la importancia de fortalecer el pensamiento lógico, el modelado, el cálculo,

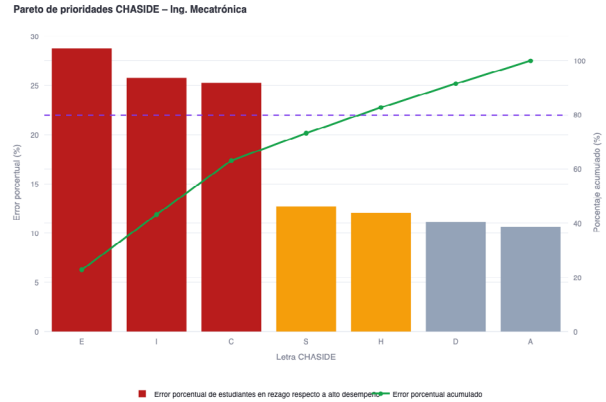


Fig. 6. Prioridades vocacionales por carrera. Fuente: Elaboración propia.

el uso de herramientas, las prácticas guiadas y la resolución técnica de problemas. Administrativo (C) reporta la conveniencia de incorporar actividades orientadas a la organización, planeación, seguimiento de instrucciones, gestión del tiempo y resolución estructurada de problemas. Ciencias de la Salud (S) mostró la necesidad de favorecer la observación, la precisión, el estudio de casos, la empatía profesional y las actividades con orientación al servicio. Humanidades y Sociales (H) promueve comunicación oral y escrita, argumentación, comprensión de textos, análisis de casos y trabajo colaborativo.

Finalmente, en el módulo de información individual, el usuario tiene acceso a dos niveles de selección. En el primero, se elige la carrera seleccionada por el aspirante y, a partir de esta, se despliega automáticamente la lista de participantes asociados a dicho programa académico. Esta organización permite una consulta focalizada y facilita la revisión individual por parte de coordinadores, tutores o del propio estudiante.

Una vez seleccionado el participante, el sistema genera de manera inmediata un reporte ejecutivo individual, en el que se informa el nombre del estudiante, el grado de correspondencia entre la carrera elegida y el perfil vocacional identificado, y, en caso de detectarse una falta de coincidencia, la carrera hacia la cual el aspirante presenta mayor afinidad vocacional. De esta manera, el módulo traslada los resultados del análisis general a una escala individual, permitiendo apoyar procesos de orientación y toma de decisiones académicas con base en evidencia.

Como validación indirecta de la eficiencia de esta propuesta exploratoria, se rastrearon las calificaciones parciales de dos estudiantes que participaron en este estudio. En la Fig. 7 y 8 se muestran las calificaciones de una estudiante que no tiene el perfil. En cambio, la Fig. 9 y 10 muestran a un estudiante que está en transición, es decir, su elección de carrera coincide con su vocación profesional, pero esta es algo débil.



Reporte individual CHASIDE

Estudiante: yulemi
Carrera: Ingeniería Bioquímica
Perfil identificado: El perfil NO va acorde con la carrera elegida
Intensidad vocacional: Perfil en riesgo

Conclusión y recomendación

El estudiante presenta coincidencia mínima entre su perfil vocacional y la carrera elegida. El análisis compatible sugiere mejor ajuste hacia Ingeniería en Inteligencia Artificial. Se recomienda seguimiento tutorial temprano y orientación vocacional.

Fig. 7. Reporte ejecutivo del perfil vocacional de un participante cuyo perfil no coincide con su vocación profesional. Fuente: Elaboración propia.

	Estudiante: XXXXXXXXXXXXXXXXX ARTURO No. control: XXXXXXXXXX									
Carrera: INGENIERIA EN INTELIGENCIA ARTIFICIAL Modalidad: ESCOLARIZADA										
Asignatura	Temas									Promedio
	1	2	3	4	5	6	7	8	9	
CALCULO INTEGRAL ELENA ELSA BRICIO BARRIOS	94	93	98	82						91
PROBABILIDAD Y ESTADISTICA SANTIAGO ARCEO DIAZ	92	94	100	90						93
COMUNICACION Y LIDERAZGO MA. MAGDALENA TORRES PACHECO	96	88	94	100						94
PROGRAMACION ORIENTADA A OBJETOS PATRICIA ELIZABETH FIGUEROA MILLAN	82	98	95	88	88					89
BASE DE DATOS JORGE ESTEBAN GONZALEZ VALLADARES	94	82	74	100						87
TALLER DE ETICA (2023) MARIA GUADALUPE VENEGAS MOCTEZUMA	100	100	100	0						75
TUTORIA II PATRICIA ELIZABETH FIGUEROA MILLAN	100	100	0							66

Fig. 10. Reporte de calificaciones parciales de segundo semestre. La calificación de cero indica que el docente no ha realizado la evaluación de ese tema. Fuente: Elaboración propia.



Estudiante:

XXXXXXXXXXXXXXXXXXXX

YULEMI

No. control:

XXXXXX

Carrera:

INGENIERIA BIOQUIMICA

Modalidad:

ESCOLARIZADA

Asignatura	Temas									Promedio
	1	2	3	4	5	6	7	8	9	
ALGEBRA LINEAL ALICIA OLIVERA MONTEJANO	98	75	70	70	70					76
QUIMICA ANALITICA MARIA FELIPA ANDRADE URZUA	79	83	78	76						78
BIOLOGIA JUAN ENRIQUE CORTES VALLE	70	99	90	93	92					88
QUIMICA ORGANICA I FRANCISCO JAVIER DELGADO VIRGEN	51	59	60	49						55
ADMINISTRACION Y LEGISLACION DE EMPRESAS ROCIO ABRIL MORALES LOYA	93	99	99	99	98	97				97
TUTORIA II MARIA AURORA ANDRADE URZUA	100	80	80							86
CALCULO INTEGRAL ELENA ELSA BRICIO BARRIOS	88	80	100	100						92

Fig. 8. Reporte de calificaciones parciales de segundo semestre. La calificación de cero indica que el docente no ha realizado la evaluación de ese tema. Fuente: Elaboración propia.

SIITEC

Usuario:

Seleccionar carrera

CONVOCATORIA 2026 PARA EL INGRESO A LICENCIATURA

Como parte del criterio 2 de selección, en esta convocatoria vigente, La evaluación sumativa de admisión, que comprende los resultados y la evaluación del desarrollo de habilidades y saberes en los diversos aspectos del curso propedéutico (60%) se le solicita responder la siguiente escala vocacional que encontrará en este hipervínculo: como parte del seguimiento, se le requerirá la evidencia durante su curso propedéutico.

☐ ARQUITECTURA

☐ INGENIERIA INDUSTRIAL

☐ INGENIERIA EN SISTEMAS COMPUTACIONALES

☐ INGENIERIA BIOQUIMICA

☐ LICENCIATURA EN ADMINISTRACION

☐ INGENIERIA MECATRONICA

☐ INGENIERIA AMBIENTAL

☐ INGENIERIA EN GESTION EMPRESARIAL

☐ CONTADOR PUBLICO

☐ INGENIERIA EN INTELIGENCIA ARTIFICIAL

Fig. 11. Impresión de pantalla de la plataforma SIITEC donde se realiza el proceso de reclutamiento de aspirantes edición 2026. Elaboración propia.

Reporte individual CHASIDE

Estudiante: Arturo
Carrera: Ingeniería en Inteligencia Artificial
Perfil identificado: El perfil coincide con la carrera elegida
Intensidad vocacional: Perfil en transición

Conclusión y recomendación

El estudiante muestra una congruencia vocacional funcional con la carrera elegida. Se recomienda acompañamiento académico preventivo y seguimiento durante el primer semestre.

Fig. 9. Reporte ejecutivo del perfil vocacional de un participante cuyo perfil sí coincide con su vocación profesional. Fuente: Elaboración propia.

4. CONCLUSIONES

La aplicación web desarrollada a partir de la escala CHASIDE permitió automatizar el procesamiento de respuestas de aspirantes de nuevo ingreso al Instituto Tecnológico de Colima e identificar el nivel de compatibilidad entre la carrera elegida y el perfil vocacional de cada participante. De esta manera, el presente estudio sienta las bases para consolidar un sistema institucional de orientación y prevención del rezago académico apoyado en analítica vocacional.

Más allá de su valor como sistema digital de análisis, la principal aportación de la plataforma radica en su capacidad para traducir los resultados de una escala voca-



cional; además de clasificar a los aspirantes según su ajuste vocacional, la plataforma generó sugerencias de canalización hacia carreras afines, reportes ejecutivos individuales y áreas prioritarias de fortalecimiento por programa académico.

Estos resultados muestran que la herramienta desarrollada constituye un apoyo útil para la toma de decisiones preventivas durante el curso propedéutico y para la implementación de acciones tempranas de orientación académica.

Agradecimientos. Los autores agradecen a los aspirantes que participaron en este estudio por su disposición y colaboración en la respuesta del instrumento aplicado.

REFERENCES

- Acuña, C., Acosta, C., Peñaralda, E., Gatica, G. and Córdova, A. (2024). A CHASIDE test-based analysis for identifying adolescents characteristics impacting their vocational orientation. *Procedia Computer Science*, 231, pp.654–659.
- Anastasi, A., and Urbina, S. (1997). *Psychological Testing*. 7th ed. Prentice Hall, Upper Saddle River, NJ, USA.
- Aragón, G., and Américo, J. C. (2010). *Psicología Ambiental*. Pirámide, Madrid, España.
- Atay, S., Yildirim, M. A., and Demir, A. K. (2024). A web-based career information system for supporting career decision-making. *Education and Information Technologies*.
- Brown, S. D., and Lent, R. W. (2013). *Career Development and Counseling: Putting Theory and Research to Work*. 2nd ed. Wiley, Hoboken, NJ, USA.
- Dillon-Pérez, F., Rojas-Londoño, D., Lara-Ramos, E., and Freire-Muñoz, I. (2023). Orientación vocacional y profesional como alternativa en la elección de carreras universitarias. *Revista Cátedra*, 6(1), pp.78–91.
- Fernández, S. F., Sánchez, J. M. C., Córdoba, A., and Largo, A. (2002). *Estadística descriptiva*. Esic Editorial.
- Gallup (2024). La mayoría de egresados universitarios no le encuentra sentido a su empleo, afirma encuesta de Gallup. [En línea] Available at: <https://observatorio.tec.mx/edu-news/la-mayoria-de-egresados-universitarios-no-le-encuentra-sentido-a-su-empleo-afirma-encuesta-de-gallup/>
- Holland, J. L. (1997). *Making Vocational Choices: A Theory of Vocational Personalities and Work Environments*. Psychological Assessment Resources.
- Lopera, D. A. T., Correa, A. M. C., Restrepo, S. L., and Torres, M. (2018). Correspondencia entre la carrera cursada y resultados del test CHASIDE. *Psicoespacios*, 12(21), pp.76–93.
- Madrigal Hernández, M. (2024). *Factores que influyen en la elección de carrera profesional: el papel de la vocación y la autoeficacia*.
- Masi, A., Ricci, L., and Gori, F. (2025). Mobile career counseling applications and their impact on career decision self-efficacy. *International Journal for Educational and Vocational Guidance*.
- Moreno Aguirre, S. E. (2016). *Mejora de la determinación del grado aptitudinal y orientación profesional mediante un sistema experto*.
- Pérez, D. E. P., and Rodríguez-Alberto, M. (2023). Interés y aptitud vocacional como predictores del rendimiento académico universitario. *Revista de Psicología y Ciencias del Comportamiento*, 14(2), pp.70–87.
- Quishpe, J. A. P. (2022). Diseño de un instrumento de orientación vocacional para estudiantes de educación tecnológica. *Polo del Conocimiento*, 7(6), pp.1169–1186.
- Rochin Berumen, F. L. (2021). Deserción escolar en la educación superior en México: revisión de literatura. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 11(22).
- Seesukong, P., Khamnuan, T., and Phanniphong, S. (2024). CARIA: A personalized career recommendation system based on individual competencies. *Journal of Computer Assisted Learning*.
- Secretaría de Educación Pública (2024). *Principales cifras del Sistema Educativo Nacional 2023–2024*. SEP, Ciudad de México.
- Super, D. (1957). *The Psychology of Careers*. Harper & Row, New York, NY, USA.
- Universidad Abierta y a Distancia de México (2026). Orientación vocacional y profesional. [En línea] Available at: <https://www.unadmexico.mx/tramites-y-servicios/orientacion-vocacional-y-profesional>
- Vacasela Lema, N. B. (2019). *Vocación profesional en estudiantes de Psicopedagogía*. Bachelor's thesis, Riobamba.

Water Turbidity Prediction Using Mathematical Models and Machine Learning on an Open Dataset

AZ Aguilar-Sántiz* GF Pérez-García* VI Ruiz-Pérez*
D López-Cortés* A de Jesús Flores-García**

* *Facultad de Ciencias en Física y Matemáticas, Universidad Autónoma de Chiapas, Tuxtla Gutiérrez, Chiapas, México (e-mail: gilbert.perez@unach.mx).*

** *Tecnológico Nacional de México / Instituto Tecnológico de Tuxtla Gutiérrez, Tuxtla Gutiérrez, Chiapas, México.*

Abstract: This work presents a supervised modeling study for estimating water turbidity from physicochemical variables obtained from an open online water quality dataset. The research was conducted using a reproducible sample of 220 observations extracted from the original dataset, considering turbidity as the target variable and six relevant physicochemical descriptors as predictors. Based on this dataset, three modeling approaches were implemented: a linear solution based on least squares, an iterative optimization scheme using gradient descent, and a simple nonlinear approach based on sigmoid activation. The study focuses on the methodological comparison and interpretation of model behavior within a supervised regression framework, evaluating their fitting capability and their limitations in representing relationships among water quality variables. The results establish an analytical basis for discussing the scope of linear approaches in comparison with simple nonlinear formulations for turbidity prediction from open environmental data.

Keywords: water turbidity, water quality, supervised regression, least squares, gradient descent, nonlinear modeling, sigmoid function, open dataset.

1. INTRODUCCIÓN

La calidad del agua es un tema de interés prioritario en el ámbito ambiental, sanitario e industrial, debido a que su deterioro afecta tanto la salud pública como la estabilidad de los ecosistemas y la confiabilidad de distintos procesos productivos. En este contexto, el monitoreo de variables fisicoquímicas constituye una herramienta esencial para caracterizar el estado del recurso hídrico y detectar condiciones que puedan asociarse con contaminación, arrastre de sedimentos o alteraciones en la composición del agua.

Entre estas variables, la turbidez representa un indicador particularmente relevante, ya que expresa el grado en que el agua pierde transparencia por la presencia de partículas suspendidas. Su comportamiento suele relacionarse con fenómenos físicos y químicos de interés práctico, por lo que su estimación resulta útil no sólo desde una perspectiva descriptiva, sino también como parte de esquemas de análisis predictivo orientados al monitoreo ambiental.

Tradicionalmente, el análisis de variables de calidad del agua se ha apoyado en mediciones directas y en relaciones empíricas entre parámetros fisicoquímicos. Sin embargo, cuando se dispone de conjuntos de datos multivariados,

este problema también puede formularse como una tarea de regresión supervisada, en la que una variable continua de interés se estima a partir de un conjunto de predictores observables. Esta perspectiva resulta especialmente pertinente en inteligencia artificial aplicada, porque permite estudiar la capacidad de distintos modelos para aproximar relaciones entre variables ambientales sin depender exclusivamente de formulaciones analíticas cerradas.

En años recientes, la disponibilidad de datasets abiertos ha favorecido el desarrollo de estudios reproducibles en problemas de monitoreo y predicción ambiental. No obstante, una parte importante de estos trabajos se concentra en modelos cada vez más complejos, mientras que con menor frecuencia se presenta una comparación metodológica clara entre enfoques lineales clásicos, esquemas iterativos de optimización y formulaciones no lineales simples sobre un mismo problema de estimación. Desde una perspectiva formativa y científica, esta comparación sigue siendo valiosa, ya que permite entender con mayor precisión qué aporta cada estrategia de modelado, cuáles son sus alcances reales y bajo qué condiciones una mayor complejidad puede o no traducirse en una mejora efectiva del ajuste.

Con base en ello, en este trabajo se estudia la predicción de la turbidez del agua a partir de variables fisicoquímicas obtenidas de un dataset abierto de calidad del agua disponible en línea. Para el análisis se empleó una muestra

* Este trabajo forma parte de una línea de investigación sobre modelado matemático y aprendizaje automático aplicado al análisis de variables de calidad del agua.



reproducibles de 220 observaciones del conjunto original, considerando como variable objetivo la turbidez (*Turbidity (NTU)*) y como variables predictoras seis descriptores físico-químicos seleccionados. Sobre esta base se implementaron tres enfoques de modelado: una solución lineal por mínimos cuadrados, un procedimiento iterativo mediante descenso de gradiente y una aproximación no lineal simple con activación sigmoide.

El objetivo de esta investigación es comparar el comportamiento de estos tres enfoques en la estimación de la turbidez, analizando sus diferencias desde el punto de vista del ajuste, la formulación matemática y la capacidad de representación del problema. De manera específica, se busca establecer una base metodológica clara para discutir el alcance de los modelos lineales frente a una formulación no lineal simple en un escenario de datos ambientales abiertos.

La hipótesis de trabajo plantea que los modelos lineales constituyen una referencia sólida para capturar una parte importante de la estructura presente en los datos, mientras que una aproximación no lineal elemental puede ofrecer información adicional sobre la posible presencia de relaciones no estrictamente lineales entre los predictores físico-químicos y la turbidez. En este sentido, la contribución principal del artículo no radica en proponer una arquitectura compleja, sino en presentar una comparación rigurosa y reproducible entre tres estrategias de modelado supervisado aplicadas a un problema real de calidad del agua.

2. ESTADO DEL ARTE

El análisis predictivo de variables de calidad del agua ha evolucionado desde enfoques empíricos y estadísticos clásicos hacia esquemas de aprendizaje automático capaces de capturar relaciones no lineales entre múltiples parámetros ambientales. En términos generales, esta evolución ha estado impulsada por dos necesidades complementarias: por un lado, mejorar la capacidad de estimación de variables de interés a partir de datos observables y, por otro, construir herramientas de apoyo a la decisión en contextos donde el monitoreo directo puede resultar costoso, limitado o insuficiente. En este sentido, los modelos lineales y de regresión continuaban ocupando un lugar relevante, no sólo por su simplicidad, sino también porque proporcionan una referencia interpretable frente a métodos de mayor complejidad (Haghiabi et al., 2018; Zhu et al., 2022).

Dentro de este campo, la literatura ha mostrado que la predicción de calidad del agua puede abordarse exitosamente mediante técnicas de aprendizaje automático y redes neuronales artificiales. Revisiones recientes han documentado el uso de modelos como regresión múltiple, máquinas de soporte vectorial, árboles de decisión, redes neuronales y ensamblados para distintos parámetros físico-químicos y escenarios de monitoreo. En conjunto, estos trabajos coinciden en que los enfoques basados en datos ofrecen ventajas importantes cuando las relaciones entre variables son complejas, aunque también advierten que la ganancia en precisión suele venir acompañada de una menor interpretabilidad o de mayores exigencias de preprocesamiento y ajuste (Chen et al., 2020; Zhu et al., 2022).

En el caso específico de la turbidez, varios estudios recientes confirman esta transición desde modelos estadísticos hacia estrategias más flexibles. Stevenson y Bravo desarrollaron una herramienta de soporte para la predicción operativa de eventos de turbidez superiores a 1 NTU hasta con siete días de anticipación, comparando un modelo lineal generalizado con Random Forest. Sus resultados mostraron que el modelo basado en bosque aleatorio podía mejorar el desempeño al capturar interacciones no lineales, aunque el problema se mantuvo fuertemente condicionado por las particularidades del sitio y del esquema de monitoreo (Stevenson and Bravo, 2019). De forma similar, Zhang et al. propusieron un modelo Random Forest para predecir turbidez en agua cruda tipo lago usando variables meteorológicas preprocesadas, mostrando que el uso de entradas derivadas y etapas de limpieza de datos puede fortalecer la predicción, pero también hace que el modelo dependa de una configuración muy específica del contexto estudiado (Zhang et al., 2021).

Otros trabajos han explorado formulaciones aún más especializadas. Rele et al. compararon regresión dinámica, redes LSTM y modelos aditivos generalizados para la predicción de turbidez en series temporales de corrientes, incorporando además un meta-modelo para seleccionar el mejor predictor en cada instante. Este planteamiento resulta valioso para escenarios temporales complejos, pero también desplaza el problema hacia arquitecturas híbridas cuya interpretación es menos directa (Rele et al., 2023). En una línea distinta, Devi y Mamatha estudiaron la predicción de turbidez en el lago Saki a partir de variables derivadas de percepción remota, comparando regresión lineal, k-NN y árboles de decisión. Sus resultados mostraron ventajas para el árbol de decisión ajustado, aunque las variables de entrada provinieron de reflectancias espectrales y no de un conjunto físico-químico tabular convencional (Devi and Mamatha, 2024). Más recientemente, McKelvey et al. compararon regresión lineal, k-NN, regresión polinomial y redes neuronales artificiales para predecir turbidez de agua sedimentada en una planta de tratamiento. Un aspecto especialmente relevante de este estudio es que plantea explícitamente una secuencia de modelos de complejidad creciente para equilibrar interpretabilidad y capacidad predictiva; sin embargo, el mejor desempeño correspondió al modelo ANN y el análisis se mantuvo dentro de un entorno operacional de planta y con datos altamente específicos del proceso (McKelvey et al., 2024).

En conjunto, esta literatura permite identificar varias tendencias. Primero, existe evidencia consistente de que los modelos no lineales y los ensamblados pueden superar a los modelos lineales cuando el fenómeno estudiado incorpora interacciones complejas o comportamiento temporal marcado (Stevenson and Bravo, 2019; Zhang et al., 2021; Rele et al., 2023). Segundo, una parte importante de los estudios de turbidez se ha desarrollado en contextos muy particulares, como series temporales hidrológicas, sistemas de tratamiento o esquemas de percepción remota, lo que dificulta comparar directamente los resultados entre trabajos (Devi and Mamatha, 2024; McKelvey et al., 2024). Tercero, aunque varios autores incluyen modelos lineales como referencia, con frecuencia estos aparecen sólo como línea base frente a arquitecturas más sofisticadas y no como parte de una discusión metodológica central sobre

su papel analítico e interpretativo (Chen et al., 2020; Zhu et al., 2022; McKelvey et al., 2024).

A partir de estas observaciones, se identifica una oportunidad clara para estudios comparativos más transparentes y reproducibles en los que la complejidad de los modelos aumente de forma controlada sobre un mismo problema y un mismo conjunto de datos. Por esta razón, la presente investigación se diferencia de buena parte de la literatura reciente en tres aspectos. En primer lugar, se formula la estimación de turbidez a partir de un *dataset* abierto de calidad del agua y de un conjunto reducido de variables fisicoquímicas con significado directo. En segundo lugar, se comparan tres enfoques con niveles de complejidad claramente diferenciados: una solución cerrada por mínimos cuadrados, un esquema iterativo de descenso de gradiente y una aproximación no lineal simple con activación sigmoide. En tercer lugar, el aporte del estudio no se centra en proponer una arquitectura de alto desempeño, sino en ofrecer una comparación metodológica clara que permita discutir qué tanto aportan los modelos lineales como referencia y en qué medida una no linealidad elemental modifica la capacidad de ajuste del problema. Bajo esta perspectiva, el valor principal del trabajo radica en su trazabilidad analítica, su reproducibilidad y su utilidad como base rigurosa para posteriores desarrollos de modelado predictivo en calidad del agua.

3. METODOLOGÍA

En esta sección se describe la metodología empleada para la estimación de la turbidez del agua a partir de variables fisicoquímicas contenidas en un *dataset* abierto. El estudio se planteó como un flujo secuencial que inicia con la adquisición y depuración de datos, continúa con la formulación del problema de regresión y culmina con la comparación entre tres enfoques de modelado supervisado: mínimos cuadrados, descenso de gradiente y una aproximación no lineal simple con activación sigmoide. La secuencia general del procedimiento se resume en la Figura 1.

3.1 Fuente de datos y conformación de la muestra

El estudio se desarrolló a partir de un conjunto de datos abierto de calidad del agua disponible en línea, empleado como fuente primaria de observaciones para el análisis, lo que permitió formular un problema de predicción reproducible bajo un entorno de datos reales y trazables.

El archivo fuente contiene 1963 registros. A partir de este conjunto se seleccionó una muestra aleatoria reproducible de 220 observaciones, fijando una semilla de aleatoriedad para garantizar la replicabilidad del procedimiento. Previamente a la selección final, se realizó una depuración básica consistente en conservar únicamente las columnas de interés y eliminar registros con valores faltantes en las variables utilizadas por el modelo. Esta decisión permitió trabajar con una base compacta, suficiente para un análisis comparativo inicial y coherente con el alcance metodológico del estudio.

La configuración general de los tres enfoques evaluados se resume en la Tabla 1. Esta tabla concentra el tamaño de muestra, el número de variables empleadas, los modelos

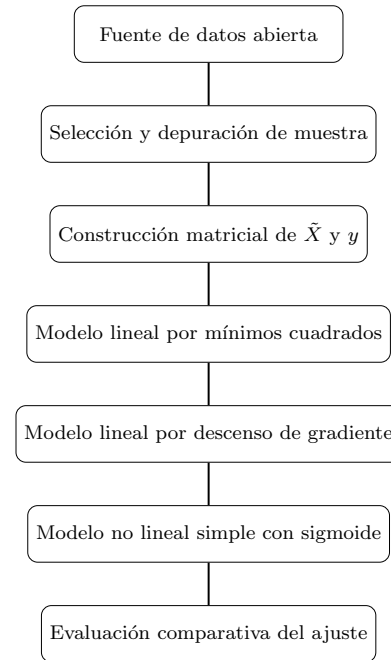


Figura 1. Flujo metodológico general del estudio.

considerados y los principales parámetros de entrenamiento utilizados en los métodos iterativos.

Tabla 1. Configuración general de los modelos evaluados.

Elemento	Configuración
Tamaño de muestra	220 observaciones
Variables de entrada	6 variables fisicoquímicas
Variable objetivo	Turbidity (NTU)
Modelo 1	Regresión lineal por mínimos cuadrados
Modelo 2	Regresión lineal por descenso de gradiente
Modelo 3	Aproximación no lineal simple con sigmoide
Inicialización de pesos (iterativos)	Vector nulo
Tasa de aprendizaje lineal	0,05
Épocas del descenso de gradiente	3000
Tasa de aprendizaje no lineal	0,05
Épocas del modelo sigmoidal	4000
Estandarización	Aplicada a modelos iterativos
Normalización de la salida	Aplicada al modelo sigmoidal

3.2 Definición de variables

La variable objetivo del trabajo fue la turbidez del agua, expresada en unidades NTU. Como variables predictoras se seleccionaron seis descriptores fisicoquímicos presentes en el *dataset*, elegidos por su relación potencial con la condición general del agua y por su disponibilidad completa dentro de la muestra empleada. La Tabla 2 resume las variables utilizadas.



Tabla 2. Variables empleadas en el modelo de predicción de turbidez.

Variable	Tipo	Descripción
Potassium (mg/l)	Entrada	Concentración de potasio en la muestra.
Dissolved Oxygen mg/L	Entrada	Oxígeno disuelto en el agua.
Chloride (mg/l)	Entrada	Concentración de cloruros.
Total Organic Carbon (mg/l)	Entrada	Contenido de carbono orgánico total reportado en el conjunto de datos.
Conductivity (uS/cm)	Entrada	Conductividad eléctrica del agua.
Nitrate (mg/l)	Entrada	Concentración de nitratos.
Turbidity (NTU)	Salida	Variable objetivo asociada a la presencia de partículas suspendidas.

3.3 Formulación matricial del problema

Sea n el número de observaciones y $p = 6$ el número de variables predictoras. El problema se formuló como una tarea de regresión supervisada a partir de una matriz de características

$$X \in R^{n \times p},$$

y un vector objetivo

$$y \in R^n.$$

Cada fila de X representa una muestra de agua y cada componente de y contiene el valor observado de turbidez correspondiente. Para incorporar el término independiente del modelo lineal, se construyó una matriz aumentada

$$\tilde{X} = [\mathbf{1} \ X],$$

de modo que la estimación lineal puede expresarse como

$$\hat{y} = \tilde{X}\tilde{w},$$

donde $\tilde{w}$ contiene el sesgo y los coeficientes asociados a las seis variables de entrada.

Previo a los procedimientos iterativos, se generó también una versión estandarizada de las variables predictoras, con media cero y desviación estándar unitaria. Esta transformación no fue necesaria para la solución cerrada por mínimos cuadrados en escala original, pero sí resultó conveniente para el descenso de gradiente y para la formulación no lineal sigmoideal, debido a que ambos métodos son sensibles a diferencias de escala entre variables.

3.4 Modelo lineal por mínimos cuadrados

Como primera referencia se empleó un modelo lineal ajustado mediante solución cerrada por mínimos cuadrados. Este enfoque permite obtener directamente el vector de parámetros que minimiza el error cuadrático medio sobre el conjunto analizado. La estimación se formuló mediante la ecuación normal

$$\tilde{w} = (\tilde{X}^T \tilde{X})^{-1} \tilde{X}^T y.$$

Con el propósito de mantener estabilidad numérica y verificar la consistencia del cálculo, también se consideró la resolución equivalente del sistema lineal

$$(\tilde{X}^T \tilde{X})\tilde{w} = \tilde{X}^T y,$$

mediante un solucionador matricial estándar. Este modelo cumple una doble función dentro del estudio: por un lado, constituye la base analítica más interpretable; por otro, actúa como referencia para comparar el comportamiento de los enfoques iterativo y no lineal.

3.5 Modelo lineal mediante descenso de gradiente

Como segunda estrategia se implementó el mismo problema lineal, pero ahora resuelto mediante un procedimiento iterativo de optimización. La función de costo utilizada fue

$$J(\tilde{w}) = \frac{1}{n} \|\tilde{X}_s \tilde{w} - y\|^2,$$

donde $\tilde{X}_s$ representa la matriz aumentada construida a partir de las variables estandarizadas.

El gradiente de la función de costo está dado por

$$\nabla J(\tilde{w}) = \frac{2}{n} \tilde{X}_s^T (\tilde{X}_s \tilde{w} - y),$$

y la regla de actualización adoptada fue

$$\tilde{w}^{(t+1)} = \tilde{w}^{(t)} - \eta \nabla J(\tilde{w}^{(t)}),$$

donde η es la tasa de aprendizaje. El vector de pesos se inicializó en cero y el proceso se iteró durante un número fijo de épocas. La finalidad de este bloque no fue únicamente reproducir la solución lineal por una vía distinta, sino observar la convergencia del método y contrastar una solución analítica exacta con un esquema iterativo clásico de optimización.

3.6 Modelo no lineal simple con activación sigmoide

Como tercera estrategia se incorporó una formulación no lineal simple basada en activación sigmoide. En este caso, sobre la matriz estandarizada aumentada se definió

$$z = \tilde{X}_s \tilde{w}, \quad \hat{y}_{\text{norm}} = \sigma(z) = \frac{1}{1 + e^{-z}}.$$

Debido a que la función sigmoide produce salidas en el intervalo $(0, 1)$, la variable objetivo fue normalizada previamente al entrenamiento. Una vez obtenido el modelo, las predicciones fueron reescaladas al dominio original de turbidez para permitir su comparación directa con los otros enfoques.

El entrenamiento se realizó minimizando el error cuadrático medio sobre la salida normalizada. La actualización de parámetros se obtuvo mediante la regla de la cadena, incorporando la derivada de la sigmoide

$$\sigma'(z) = \sigma(z)(1 - \sigma(z)).$$

Este modelo no se introdujo como una red neuronal profunda ni como una propuesta de alta complejidad, sino como una extensión elemental del esquema lineal que permitiera examinar si una no linealidad explícita modificaba la capacidad de ajuste del problema.



3.7 Criterios de evaluación

La comparación entre modelos se realizó mediante tres métricas cuantitativas:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2,$$

$$RMSE = \sqrt{MSE},$$

y

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}.$$

El error cuadrático medio permitió evaluar la discrepancia promedio entre valores observados y estimados; la raíz del error cuadrático medio facilitó la interpretación en las mismas unidades de la variable objetivo; y el coeficiente de determinación permitió valorar la proporción de variabilidad explicada por cada modelo. De forma complementaria, en la sección de resultados se recurrió a representaciones gráficas de valores observados frente a predichos y a curvas de convergencia de la función de costo para los métodos iterativos.

4. RESULTADOS

En esta sección se presentan los principales resultados obtenidos a partir de la muestra reproducible de 220 observaciones empleada en el estudio. Siguiendo una lógica de análisis progresiva, primero se resume el comportamiento general del conjunto de datos, después se reportan los resultados del modelo lineal por mínimos cuadrados, posteriormente se compara dicho ajuste con el obtenido por descenso de gradiente y, finalmente, se analiza el desempeño de la aproximación no lineal simple con activación sigmoide.

La Tabla 3 resume la estadística descriptiva de las seis variables predictoras y de la variable objetivo. Se observa que las variables presentan escalas significativamente distintas, lo que justifica el uso de estandarización en los modelos iterativos. En particular, la turbidez presenta una media de 2.6999 NTU, una desviación estándar de 1.8149 y un rango comprendido entre 0.02 y 9.50 NTU, lo cual sugiere una dispersión moderada y una variabilidad suficiente para evaluar la respuesta de los modelos.

Tabla 3. Estadística descriptiva de las variables empleadas en el estudio.

Variable	Media	Desv. est.	Mín.	Máx.
Potassium (mg/l)	26.9375	16.9954	0.3800	88.5200
Dissolved Oxygen mg/L	6.2868	1.7328	0.0000	11.4000
Chloride (mg/l)	133.1987	88.4756	0.1300	478.4700
Total Organic Carbon (mg/l)	14.2960	3.4597	5.0517	22.8910
Conductivity (uS/cm)	812.4053	543.3116	5.9200	2939.8400
Nitrate (mg/l)	25.0726	15.5866	0.1900	87.3600
Turbidity (NTU)	2.6999	1.8149	0.0200	9.5000

Desde el punto de vista metodológico, esta heterogeneidad de escalas es importante porque afecta de manera directa la dinámica de optimización de los métodos iterativos. Por esa razón, aunque la solución cerrada por mínimos cuadrados se reporta en la escala original, tanto el descenso de gradiente como el modelo sigmoidal se desarrollaron sobre variables estandarizadas, tal como se estableció en la metodología.

4.1 Resultados del modelo lineal por mínimos cuadrados

Como primer punto de referencia se ajustó un modelo lineal mediante solución cerrada por mínimos cuadrados. Los coeficientes estimados se presentan en la Tabla 4. En términos generales, el modelo asigna un efecto negativo al oxígeno disuelto y un efecto positivo a cloruros, conductividad, nitratos y potasio. El término asociado a carbono orgánico total es pequeño y negativo dentro de esta muestra.

Tabla 4. Coeficientes estimados del modelo lineal por mínimos cuadrados.

Parámetro	Valor
Bias	2.3801
Potassium (mg/l)	0.0042
Dissolved Oxygen mg/L	-0.1780
Chloride (mg/l)	0.0055
Total Organic Carbon (mg/l)	-0.0053
Conductivity (uS/cm)	0.0007
Nitrate (mg/l)	0.0048

El desempeño global del modelo lineal se resume en la Tabla 5. El ajuste por mínimos cuadrados obtuvo un error cuadrático medio de 2.3996, una raíz del error cuadrático medio de 1.5491 NTU y un coeficiente de determinación $R^2 = 0,2681$. Estos valores muestran que el modelo sí alcanza a identificar cierta relación entre las variables fisicoquímicas y la turbidez; sin embargo, dicha relación no es lo suficientemente fuerte como para explicar la mayor parte del comportamiento observado.

El valor de RMSE permite interpretar el error en las mismas unidades de la turbidez. En este caso, un RMSE de 1.5491 NTU significa que, en promedio, las predicciones del modelo se alejan alrededor de 1.55 NTU de los valores reales. Este error es importante si se compara con la turbidez media de la muestra, que fue de 2.6999 NTU. Es decir, aunque el modelo ofrece una aproximación útil, el margen de error todavía es considerable para pensar en una predicción precisa.

Por otra parte, el valor de $R^2 = 0,2681$ indica que el modelo lineal explica aproximadamente el 26.81 % de la variabilidad de la turbidez. Este resultado no debe interpretarse como un mal funcionamiento del modelo, sino como una señal de que la turbidez depende de factores más complejos que una combinación lineal de las seis variables utilizadas. En un problema real de calidad del agua, pueden intervenir condiciones del origen de la muestra, partículas suspendidas no medidas directamente, variaciones ambientales o interacciones entre variables que no quedan representadas completamente en el modelo lineal.

La relación entre los valores observados y las predicciones del modelo lineal puede apreciarse en la Figura 2. En

esta gráfica, la diagonal representa el ajuste ideal. Aunque se observa una tendencia positiva general, los puntos se encuentran dispersos alrededor de la línea de referencia, lo que confirma que el modelo sólo reproduce de manera parcial el comportamiento de la turbidez. También se observa que las predicciones tienden a concentrarse en una zona central, por lo que el modelo tiene mayores dificultades para representar valores altos de turbidez. Este comportamiento es esperado en una primera aproximación lineal, ya que el modelo busca reducir el error promedio, pero no necesariamente logra ajustarse bien a los casos más alejados del comportamiento central de los datos.

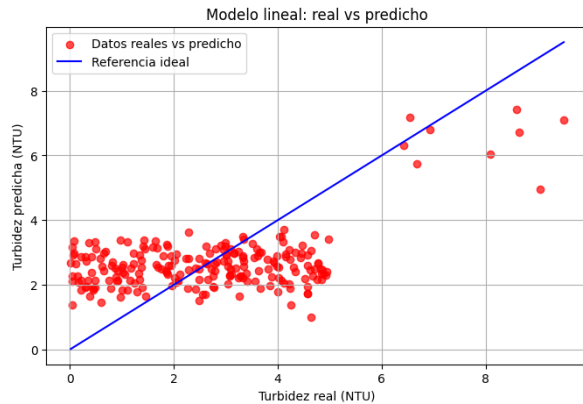


Figura 2. Comparación entre la turbidez observada y la turbidez predicha por el modelo lineal ajustado mediante mínimos cuadrados. La línea diagonal representa el ajuste ideal.

4.2 Resultados del descenso de gradiente

El segundo enfoque evaluado fue el descenso de gradiente aplicado al mismo problema lineal, pero ahora resuelto mediante una estrategia iterativa. La evolución de la función de costo se muestra en la Figura 3. La curva exhibe una disminución rápida durante las primeras épocas y posteriormente una estabilización progresiva, lo que indica un comportamiento de convergencia consistente con la naturaleza cuadrática de la función objetivo.

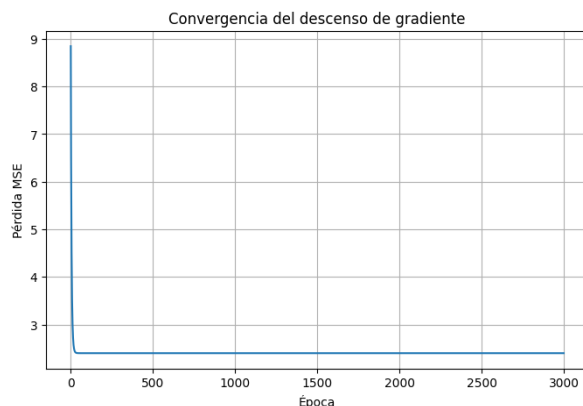


Figura 3. Convergencia del descenso de gradiente sobre el problema lineal estandarizado.

En términos cuantitativos, el descenso de gradiente reprodujo prácticamente el mismo resultado que la solución cerrada. Como se observa en la Tabla 5, ambas estrategias alcanzaron $MSE = 2,3996$, $RMSE = 1,5491$ y $R^2 = 0,2681$. Esta coincidencia es importante porque muestra que el procedimiento iterativo fue implementado de manera adecuada y que, bajo las condiciones utilizadas, converge hacia una solución equivalente a la obtenida mediante mínimos cuadrados.

Este resultado también ayuda a interpretar mejor el papel del descenso de gradiente dentro del estudio. En este caso, el descenso de gradiente no mejora las métricas porque no cambia el tipo de modelo utilizado; sigue siendo una regresión lineal. Lo que cambia es la forma de encontrar los parámetros: en un caso se calculan mediante una solución cerrada y en el otro se obtienen paso a paso mediante un proceso de optimización. Por ello, la igualdad de las métricas no representa una limitación del experimento, sino una confirmación de que ambos procedimientos resuelven el mismo problema.

Desde una perspectiva metodológica, este punto es relevante porque permite separar dos ideas: el método de entrenamiento y la capacidad del modelo para representar el fenómeno. El descenso de gradiente puede ser una herramienta útil de optimización, pero si el modelo sigue siendo lineal, su capacidad para explicar la turbidez estará limitada por esa misma formulación lineal. En este caso, la restricción principal no está en el algoritmo de ajuste, sino en la forma del modelo empleado.

4.3 Resultados del modelo no lineal simple con activación sigmoide

Como tercera alternativa se incorporó un modelo no lineal elemental basado en activación sigmoide. La evolución de la pérdida normalizada durante el entrenamiento se presenta en la Figura 4. Al igual que en el descenso de gradiente lineal, se observa una disminución sostenida del error a lo largo de las épocas, aunque en este caso la dinámica responde a una función de salida no lineal y a una actualización obtenida mediante la regla de la cadena.

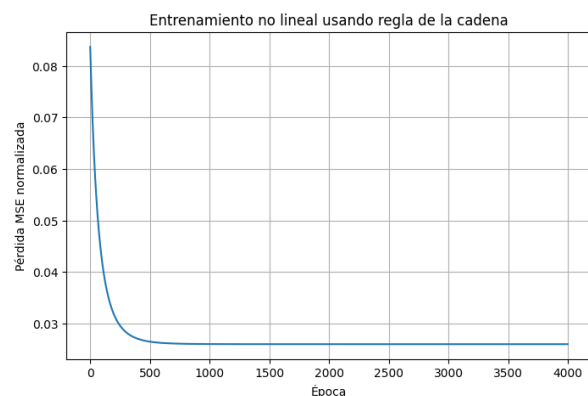


Figura 4. Evolución de la pérdida del modelo no lineal simple con activación sigmoide.

Los resultados cuantitativos muestran una mejora ligera respecto a los enfoques lineales. Como se resume en la



Tabla 5, el modelo sigmoidal alcanzó un $MSE = 2,3358$, un $RMSE = 1,5283$ y un coeficiente $R^2 = 0,2876$. Al comparar estos valores con los del modelo lineal, se observa una disminución de 0.0638 en el MSE y de 0.0208 NTU en el RMSE. También se presenta un incremento de 0.0195 en el valor de R^2 .

Aunque estos cambios van en una dirección favorable, la mejora es pequeña. Esto significa que la activación sigmoide permitió ajustar un poco mejor los datos, pero no cambió de manera fuerte el desempeño general del modelo. En otras palabras, el resultado sugiere que sí puede existir una relación no completamente lineal entre las variables fisicoquímicas y la turbidez, pero una sola transformación sigmoidal no es suficiente para representar toda la complejidad del problema.

Este resultado es útil porque permite observar el efecto de incorporar una no linealidad de forma controlada. El modelo sigmoidal no debe interpretarse como una solución definitiva ni como un modelo de alta precisión, sino como una prueba inicial de que aumentar ligeramente la capacidad de representación puede mejorar el ajuste. Sin embargo, el RMSE de 1.5283 NTU sigue siendo considerable en relación con la escala de la turbidez de la muestra. Por tanto, aunque el modelo no lineal fue el que obtuvo el mejor desempeño entre los tres enfoques, sus predicciones todavía presentan un margen de error que debe tomarse en cuenta para cualquier aplicación práctica.

Tabla 5. Comparación de métricas entre los modelos evaluados.

Método	MSE	RMSE	R^2
Modelo lineal cerrado	2.3996	1.5491	0.2681
Descenso de gradiente	2.3996	1.5491	0.2681
Modelo no lineal (sigmoide)	2.3358	1.5283	0.2876

4.4 Análisis comparativo de los resultados

Al comparar los tres enfoques evaluados, se observa que los modelos lineales presentan el mismo desempeño, mientras que el modelo sigmoidal obtiene una mejora ligera. El modelo lineal por mínimos cuadrados y el modelo lineal entrenado mediante descenso de gradiente alcanzaron $MSE = 2,3996$, $RMSE = 1,5491$ y $R^2 = 0,2681$. Esta igualdad en las métricas confirma que ambos métodos llegan a una solución equivalente, aunque lo hacen por caminos distintos: el primero mediante una solución matemática directa y el segundo mediante un proceso iterativo de optimización.

La interpretación de estas métricas permite entender mejor el alcance de los modelos lineales. Un valor de $R^2 = 0,2681$ significa que aproximadamente una cuarta parte de la variabilidad observada en la turbidez puede explicarse con las seis variables fisicoquímicas utilizadas. Esto indica que las variables seleccionadas sí aportan información relacionada con la turbidez, pero también muestra que el modelo deja sin explicar una parte importante del fenómeno. Por esta razón, el modelo lineal puede considerarse una línea base útil e interpretable, pero no una herramienta suficiente para realizar predicciones de alta precisión.

En el caso del modelo no lineal simple con activación sigmoide, los resultados fueron ligeramente mejores, con

$MSE = 2,3358$, $RMSE = 1,5283$ y $R^2 = 0,2876$. La reducción del error respecto al modelo lineal fue pequeña, pero consistente. El RMSE disminuyó 0.0208 NTU y el coeficiente de determinación aumentó 0.0195. Esto indica que la función sigmoide permitió capturar una parte adicional del comportamiento de los datos, aunque la ganancia no fue suficiente para cambiar de forma importante la capacidad predictiva del modelo.

Estos resultados muestran que el problema de predicción de turbidez no queda completamente explicado por una relación lineal simple, pero tampoco se resuelve únicamente con una no linealidad elemental. En términos prácticos, el mejor modelo todavía presenta un RMSE cercano a 1.53 NTU. Si se considera que la turbidez media de la muestra fue de 2.6999 NTU, este error sigue siendo relevante. Por ello, los modelos evaluados pueden ser útiles para analizar tendencias generales y comparar estrategias de modelado, pero no deberían considerarse todavía como sustitutos de mediciones directas en aplicaciones donde se requiera alta precisión.

La Figura 2 ayuda a reforzar esta interpretación. En ella se observa que el modelo lineal reproduce de forma aproximada la tendencia general de los datos, pero presenta dispersión respecto a la línea ideal de predicción. Además, se aprecia que el modelo tiene mayor dificultad para representar valores altos de turbidez. Este comportamiento coincide con los valores de R^2 obtenidos, ya que ninguno de los modelos supera el 0.30. En otras palabras, los modelos logran identificar una relación parcial entre las variables, pero aún existe variabilidad importante que no ha sido capturada.

En este sentido, la comparación realizada en este trabajo ayuda a mostrar que la mejora en las métricas no depende solamente del algoritmo de entrenamiento, sino de la capacidad del modelo para representar mejor la relación entre las variables. De esta manera, el estudio aporta una base ordenada y reproducible para analizar el comportamiento de diferentes enfoques de modelado sobre un mismo problema de calidad del agua.

Por último, desde una perspectiva aplicada, los resultados sugieren que las variables fisicoquímicas empleadas contienen información útil para estimar turbidez, pero no suficiente para construir un predictor altamente preciso. Por ello, una aplicación práctica de estos modelos debería considerarse como una herramienta exploratoria o de apoyo inicial al análisis, más que como un sistema definitivo de monitoreo. Para mejorar el desempeño sería necesario ampliar la base de datos, incorporar validación externa, evaluar nuevas variables y probar modelos no lineales con mayor capacidad de representación.

5. CONCLUSIONES

Este estudio presentó un análisis comparativo de tres enfoques de modelado supervisado para la estimación de la turbidez del agua a partir de variables fisicoquímicas obtenidas de un *dataset* abierto. Los resultados mostraron que los modelos lineales, tanto en su formulación cerrada por mínimos cuadrados como en su resolución iterativa mediante descenso de gradiente, proporcionan una base consistente e interpretable para analizar la relación entre



los predictores seleccionados y la turbidez. Asimismo, la aproximación no lineal simple con activación sigmoide obtuvo una mejora ligera en las métricas de ajuste, al reducir moderadamente el error e incrementar el coeficiente de determinación respecto a los modelos lineales. Este resultado sugiere que el problema puede contener componentes no estrictamente lineales, aunque también muestra que una no linealidad elemental no es suficiente para explicar la mayor parte de la variabilidad observada.

Desde una perspectiva metodológica, uno de los hallazgos más relevantes es que el descenso de gradiente converge hacia un desempeño prácticamente equivalente al de la solución cerrada del modelo lineal. Esto confirma la consistencia interna del planteamiento y valida el uso del enfoque iterativo como mecanismo de optimización sobre el mismo problema de regresión. Por otra parte, aunque el modelo sigmoide obtuvo el mejor ajuste entre los tres métodos evaluados, la magnitud de la mejora fue moderada, por lo que no puede afirmarse que una no linealidad elemental resuelva por sí sola las limitaciones del problema predictivo.

En conjunto, los resultados indican que la turbidez puede modelarse parcialmente a partir de las seis variables fisicoquímicas consideradas, pero también muestran que la capacidad explicativa alcanzada sigue siendo limitada. En consecuencia, los modelos implementados deben interpretarse como una base analítica válida para estudiar la estructura del problema, más que como una solución definitiva de alto desempeño. Su utilidad principal está en ofrecer una comparación ordenada, interpretable y reproducible que pueda servir como punto de partida para trabajos posteriores sobre predicción de turbidez y modelado de calidad del agua. Entre las principales limitaciones del estudio se encuentran el tamaño de la muestra utilizada, la dependencia de un único *dataset* abierto y la ausencia de validación externa con datos provenientes de otro contexto de monitoreo. Además, el conjunto de variables predictoras empleado, aunque físicamente interpretable, puede no ser suficiente para capturar toda la variabilidad asociada a la turbidez, especialmente si existen interacciones más complejas o efectos no observados en la base analizada.

Como trabajo futuro, se propone ampliar el conjunto de datos, incorporar validación cruzada y explorar técnicas de regularización o modelos no lineales de mayor capacidad representacional. También sería pertinente evaluar la influencia de nuevas variables fisicoquímicas o contextuales, así como contrastar el comportamiento de estos enfoques con métodos más avanzados de aprendizaje automático. Bajo esta perspectiva, el presente estudio establece una base metodológica clara para posteriores desarrollos en modelado predictivo de calidad del agua, manteniendo como eje central la interpretabilidad, la trazabilidad matemática y la reproducibilidad del análisis.

REFERENCIAS

- Chen, Y., Song, L., Liu, Y., Yang, L., and Li, D. (2020). A review of the artificial neural network models for water quality prediction. *Applied Sciences*, 10(17), 5776. doi:10.3390/app10175776.
- Devi, P.D. and Mamatha, G. (2024). Machine learning approach to predict the turbidity of saki lake, telangana,

- india, using remote sensing data. *Measurement: Sensors*, 33, 101139. doi:10.1016/j.measen.2024.101139.
- Haghiabi, A.H., Nasrolahi, A., and Parsaie, A. (2018). Water quality prediction using machine learning methods. *Water Quality Research Journal*, 53(1), 3–13. doi:10.2166/wqrj.2018.025.
- McKelvey, S., Abassi, A., Nataraj, C., and Duran, M. (2024). Data-driven modeling techniques for prediction of settled water turbidity in drinking water treatment. *Frontiers in Environmental Engineering*, 3, 1401180. doi:10.3389/fenv.2024.1401180.
- Rele, B., Hogan, C., Kandanaarachchi, S., and Leigh, C. (2023). Short-term prediction of stream turbidity using surrogate data and a meta-model approach: A case study. *Hydrological Processes*, 37(4), e14857. doi:10.1002/hyp.14857.
- Stevenson, M. and Bravo, C. (2019). Advanced turbidity prediction for operational water supply planning. *Decision Support Systems*, 119, 72–84. doi:10.1016/j.dss.2019.02.009.
- Zhang, Y., Yao, X., Wu, Q., Huang, Y., Zhou, Z., Yang, J., and Liu, X. (2021). Turbidity prediction of lake-type raw water using random forest model based on meteorological data: A case study of tai lake, china. *Journal of Environmental Management*, 290, 112657. doi:10.1016/j.jenvman.2021.112657.
- Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B., and Ye, L. (2022). A review of the application of machine learning in water quality evaluation. *Eco-Environment & Health*, 1(2), 107–116. doi:10.1016/j.eehl.2022.06.001.

Quantum Software Engineering Education: Current Landscape, Challenges, and Emerging Directions

Oscar Alejandro López Campero*
Blanca Dina Valenzuela Robles*
José Luis Ramírez Alcantara* René Santaolaya Salgado*
Jesús Alfonso Ocampo Mújica*
Andrea Alejandra Ibarra Inclán*

* Centro Nacional de Investigación y Desarrollo Tecnológico
(CENIDET),
Cuernavaca, Morelos, Mexico
(e-mail: m25ce006@cenidet.tecnm.mx; blanca.vr@cenidet.tecnm.mx;
luis.ra@cenidet.tecnm.mx; rene.ss@cenidet.tecnm.mx;
m24ce060@cenidet.tecnm.mx; m25ce048@cenidet.tecnm.mx).

Abstract: Quantum Software Engineering (QSE) is a strategic discipline with potential to enhance hybrid computing and drive innovation. However, QSE education is still emerging globally due to the lack of standardized curricula, competency frameworks, and teaching models. This article presents an exploratory review of academic literature, institutional initiatives, and expert-shared materials related to QSE education. The analysis shows that most programs focus on introductory quantum computing, with limited attention to software engineering practices, assessment, and competency development. Educational activity remains concentrated in Europe and North America, highlighting the need for scalable, collaborative, and methodologically stronger research agendas aligned with industry needs.

Keywords: Quantum software engineering, Quantum education, Pedagogical strategies, Qiskit, Global training.

1. INTRODUCTION

Quantum computing (QC) has emerged over recent decades as a transformative paradigm with the potential to reshape domains such as cryptography, artificial intelligence, molecular simulation, and complex optimization problems Hidary (2019); Bhat et al. (2022). As the field matures, recent studies have highlighted the growing convergence between quantum theory and practical software development, proposing methodological directions for future implementations Murillo et al. (2025). This evolution has created demand for new development models and engineering practices, contributing to the rise of Quantum Software Engineering (QSE) as an emerging discipline Ali et al. (2022); Arismendi and Mott (2023).

Despite important conceptual and technical progress, education in QSE remains at an early stage and continues to face substantial challenges Piattini et al. (2020); Zhao (2021). Existing initiatives span multiple levels. At the undergraduate level, offerings are often limited to introductory or elective courses focused on basic quantum computing concepts. At the graduate level, some universities have begun incorporating specialized modules into computer science and software engineering programs. In continuing education, workshops, diplomas, and open training programs have been promoted by industry and research centers. However, these efforts remain fragmented

and unevenly distributed across regions. Worldwide, universities, research centers, and technology companies have started developing educational initiatives aimed at preparing professionals capable of designing reliable and scalable quantum software. These efforts include taxonomies Gill et al. (2021); Piattini et al. (2021), conceptual frameworks, curricular adaptations Akbar et al. (2023), simulation platforms Singh et al. (2024), and educational games Weisz et al. (2020). Nevertheless, their distribution remains highly uneven. While Europe and North America have developed more mature and better-funded programs Piattini et al. (2020), progress in Latin America and other emerging regions is still limited and often dependent on international collaborations Pérez-Castillo et al. (2021). Moreover, Murillo et al. (2025) note that the rapid diversification of QSE research has not been matched by equivalent progress in educational standardization, emphasizing the need for structured roadmaps that integrate pedagogy, tooling, and competency models Murillo et al. (2025). It is also important to distinguish between education in quantum computing and education in QSE. Traditional quantum computing instruction typically focuses on the physical, mathematical, and algorithmic foundations of quantum systems, often targeting students in physics or mathematics. In contrast, QSE adopts a broader perspective that integrates methods, tools, and practices from classical Software Engineering, adapted to the quantum

Conceptual view of Quantum Software Engineering education

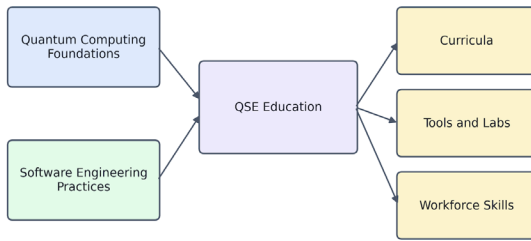


Fig. 1. Conceptual relationship between quantum computing foundations, software engineering practices, and the educational outcomes expected in Quantum Software Engineering.

paradigm. While learning quantum computing helps students understand how algorithms operate, QSE education seeks to prepare professionals capable of developing robust, scalable, and maintainable quantum software systems. As shown in Fig. 1, this relationship can be understood as the integration of quantum foundations and engineering practices into a coherent educational model. This distinction helps explain the persistent gap in specialized QSE educational programs.

The absence of a unified educational framework creates significant barriers for students and professionals interested in the field, including the scarcity of structured learning materials Zhao (2021), limited formal curricula, and restricted access to quantum hardware and software tools Singh et al. (2024). Haghparsat et al. (2024b) further argue that innovative approaches such as interactive simulation, visualization, and hybrid learning environments can help reduce these barriers by enabling more accessible and experiential learning models Haghparsat et al. (2024a). Overall, the current landscape remains fragmented, limiting the consolidation of QSE as a recognized component of computer science and software engineering education Arismendi and Mott (2023). In this context, this article presents an exploratory review of Quantum Software Engineering education worldwide. Rather than claiming exhaustive coverage, it maps reported initiatives across regions, identifies the main pedagogical strategies, tools, and outcomes described in the literature and in publicly accessible materials shared within the professional community, and discusses key challenges affecting professional training in the field. Finally, future perspectives are presented, emphasizing the importance of international collaboration and global educational frameworks to reduce the talent gap and strengthen the long-term development of QSE.

2. EXPLORATORY REVIEW DESIGN

This study adopts an exploratory review design intended to map the current educational landscape of Quantum Software Engineering (QSE), identify recurring themes, and highlight open challenges. The objective was not to produce an exhaustive systematic review, but rather to synthesize dispersed evidence in an emerging field where academic publications, institutional initiatives, and practice-oriented resources coexist.

Source identification combined two complementary strategies. First, a focused search was conducted in major academic databases, including IEEE Xplore, SpringerLink, ScienceDirect, and Scopus, using combinations of terms such as “Quantum Software Engineering”, “quantum computing education”, “software engineering education”, “curriculum”, “training”, and “quantum programming”. Second, additional references were incorporated from institutional reports, community repositories, and materials shared by colleagues active in quantum computing and software engineering education, provided that these materials were publicly accessible and directly relevant to the research questions.

Priority was given to works published between 2020 and 2025 in order to capture recent developments. However, seminal references published before that period were retained when they were necessary to contextualize the emergence of QSE as a discipline.

2.1 Selection Criteria and Analytical Procedure

The selection process followed explicit, although non-systematic, criteria. Materials were included when they satisfied at least one of the following conditions: (i) they addressed educational or training initiatives in QSE or quantum computing with a clear software engineering orientation; (ii) they proposed pedagogical frameworks, competency models, taxonomies, or curricular structures applicable to QSE education; or (iii) they described tools, platforms, or learning environments used for educational purposes in quantum software development.

Materials were excluded when they focused exclusively on quantum hardware, physics, or algorithms without an educational component; when they discussed quantum programming only from a technical-performance perspective without implications for teaching or training; or when they were duplicate, purely promotional, or too limited in detail to support comparative analysis.

After screening titles, abstracts, executive summaries, and, when necessary, full texts, the selected sources were analyzed through qualitative categorization. Each source was coded according to region, type of initiative, intended educational level, reported outcomes, and stated limitations. This procedure supported a comparative synthesis centered on four dimensions: global distribution, educational initiatives, recurring barriers, and future directions.

2.2 Definition of Research Questions

The objective of this review is to identify, classify, and analyze current contributions related to education in Quantum Software Engineering (QSE). To structure the analysis and guide the synthesis of findings, four research questions (RQs) were defined:

RQ1: What is the global landscape of education in Quantum Software Engineering?

RQ2: What are the main reported educational initiatives, and what outcomes have been identified?

RQ3: What limitations have been reported in current educational approaches to QSE?



RQ4: What future perspectives and research directions are emerging in the field of QSE education?

2.3 Methodological Scope and Limitations

Because this article is based on an exploratory review, its findings should be interpreted as a structured mapping of the field rather than as a statistically exhaustive synthesis. The inclusion of colleague-shared and institutional materials helped broaden coverage in a rapidly evolving domain, but it also introduced a risk of selection bias toward visible networks and English-language resources. Likewise, the heterogeneity of source types limited the possibility of applying uniform quality appraisal criteria or quantitative comparisons across studies.

These limitations are acknowledged explicitly because they define the contribution of this article: it offers an integrative overview and a critical starting point for future work, while also showing the need for subsequent systematic reviews with formal protocols, transparent study counts, and standardized assessment procedures.

3. RELATED WORK

Education in quantum computing and Quantum Software Engineering (QSE) has expanded through multiple parallel efforts, although consensus is still lacking regarding pedagogical methodologies, curriculum design, and academic program structures. Existing studies approach these challenges from theoretical, methodological, empirical, and educational perspectives. A first group of contributions focuses on the relationship between classical Software Engineering (SE) and quantum computing. Ali et al. (2022) discuss how established software engineering practices may evolve to address quantum paradigms, while Zhao (2021) provides an early overview of the opportunities and research challenges shaping the discipline. Similarly, Piattini et al. (2020) introduced the Talavera Manifesto, one of the earliest efforts to define foundational principles for QSE, later complemented by proposals for a dedicated Body of Knowledge Piattini et al. (2021). More recent work has continued this line by identifying best practices and transition strategies from classical to quantum development environments Alsaman (2024). A second line of research emphasizes methodological and technical frameworks for quantum software development. Akbar et al. (2023) proposed a decision-making framework to address recurring engineering challenges in QSE, whereas Pérez-Castillo et al. (2021) explored UML-based approaches for modeling quantum circuits, seeking continuity with mature software engineering practices. In a similar direction, Alsaman (2024) introduced a taxonomy of quantum software processes and reusable patterns derived from classical life-cycle models Alsaman (2024). Empirical research has also begun to emerge. Khan et al. (2025) analyzed developer discussions in Q&A forums to identify practical programming difficulties in quantum software development, offering valuable evidence of how practitioners experience current tools and workflows Khan et al. (2025). These studies are particularly relevant because they connect theoretical proposals with real-world development needs.

From an educational perspective, several studies have explored instructional strategies for introducing QSE concepts. Curriculum-oriented proposals, such as those by Nita et al., emphasize progressive and competency-based learning paths. Other initiatives focus on engagement and accessibility; for example, Weisz et al. (2020) developed Entanglion, a board game designed to teach core quantum concepts through gamification. More recently, Haghparast et al. (2024b) highlighted the value of hybrid learning environments that combine simulation, visualization, and active learning to improve conceptual understanding in quantum software education Haghparast et al. (2024b). Finally, supporting ecosystems have received increasing attention. Singh et al. (2024) reviewed available tools and technologies for quantum computing, emphasizing their educational and professional relevance. Likewise, Gill et al. (2021) provided taxonomic perspectives useful for organizing learning resources and research topics. Emerging directions such as Quantum-Based Software Engineering (QBSE) further suggest that quantum principles may eventually enhance classical software engineering processes themselves Schmitt et al. (2025). Overall, prior work reveals a rapidly growing ecosystem in which theoretical, methodological, pedagogical, and technological initiatives coexist. However, fragmentation across approaches remains evident, particularly in relation to educational standardization, competency models, and scalable curricula. This study builds on that landscape by examining the global state of QSE education and identifying current trends, gaps, and opportunities for future development.

4. ANALYSIS AND FINDINGS

4.1 Global Landscape of Quantum Software Engineering Education

The development of Quantum Software Engineering (QSE) education remains uneven across regions, reflecting differences in institutional investment, industrial participation, research maturity, and access to technological infrastructure. The reviewed literature indicates that educational initiatives and academic production are primarily concentrated in Europe and North America, while Asia and Latin America are progressing at different rates and under distinct structural conditions. These observations are consistent with recent international studies that classify varying levels of maturity in quantum education ecosystems across multiple countries Haghparast et al. (2024a).

Europe

Europe has played a leading role in the formalization of QSE as an academic and professional field. Landmark initiatives such as the Talavera Manifesto have contributed to defining foundational principles for Quantum Software Engineering and promoting its educational relevance Piattini et al. (2020). Additional contributions include UML-based approaches for modeling quantum circuits and conceptual taxonomies designed to organize core knowledge areas Piattini et al. (2021); Pérez-Castillo et al. (2021). These efforts have supported the emergence of competency-oriented curricula, specialized workshops, and certification initiatives. However, comprehensive degree programs dedicated specifically to QSE remain limited.



North America

North America has emphasized practical and industry-connected approaches to QSE education. Universities and major technology companies such as IBM, Google, and Microsoft have driven the development of open educational ecosystems and cloud-accessible quantum platforms, facilitating hands-on learning experiences Singh et al. (2024). Graduate programs in the United States and Canada increasingly incorporate quantum programming topics, although QSE is still rarely established as an independent academic discipline Ali et al. (2022). Project-based learning and interactive educational resources have become particularly visible in the region.

Asia

In Asia, educational progress in QSE is closely linked to national investments in quantum technologies, especially in countries such as China and Japan. Universities have begun integrating quantum computing into interdisciplinary academic programs, while simulation laboratories and research centers provide training opportunities for students Bhat et al. (2022); Arismendi and Mott (2023). Nevertheless, most initiatives remain focused on quantum computing broadly rather than explicitly structured QSE curricula. Recent studies suggest that stronger industry-academia partnerships may accelerate future development in this region Als Salman (2024); Schmitt et al. (2025).

Latin America

Latin America represents an emerging landscape for QSE education. Countries such as Brazil and Mexico have introduced introductory courses, workshops, and early curriculum proposals related to quantum programming and engineering Carrascal et al. (2023). However, regional growth continues to face structural barriers, including limited funding, reduced access to infrastructure, and dependence on international collaborations. Despite these constraints, increasing interest in quantum technologies suggests meaningful expansion opportunities in the coming years Murillo et al. (2025).

Cross-Regional Synthesis

Overall, the global landscape of QSE education reveals significant asymmetries. Regions with stronger research ecosystems and closer ties to industry tend to provide more mature and practice-oriented educational opportunities. In contrast, emerging regions often rely on open-access tools, international networks, and incremental institutional efforts. These differences highlight the need for more coordinated global strategies capable of reducing disparities and supporting scalable QSE education worldwide.

Table 1 summarizes the regional maturity of Quantum Software Engineering education, highlighting major strengths and current limitations across key geographic areas.

These differences confirm the uneven global development of QSE education and the need for coordinated international strategies.

4.2 Educational Initiatives and Outcomes

The reviewed literature shows growing interest in reducing the educational gap in Quantum Software Engineering (QSE). Although internationally standardized curricula are still absent, existing initiatives can be grouped into three major categories: pedagogical strategies, conceptual and competency frameworks, and educational tools and learning environments.

Pedagogical Strategies

Teaching QSE presents unique challenges because many quantum concepts are abstract, probabilistic, and counter-intuitive. As a result, several studies advocate pedagogical models specifically adapted to these characteristics. Gamification and game-based learning have shown promise in introducing foundational concepts through more intuitive learning experiences. A representative example is Entanglion, which uses playful interaction to explain principles such as superposition and entanglement Weisz et al. (2020).

Another widely reported approach is progressive or scaffolded learning, where students begin with single-qubit programming and basic gate operations before advancing to multi-qubit systems and more complex algorithmic structures Zhao (2021). Recent proposals also highlight layered instruction models and competency-based pathways supported by modular certifications or micro-credentials, which may improve scalability and accessibility in future QSE programs Sun et al. (2024).

Conceptual and Competency Frameworks

A second group of initiatives focuses on structuring the discipline through taxonomies, bodies of knowledge, and methodological guidance. These efforts are particularly relevant because QSE remains an emerging field without universally accepted curricular standards.

Prior studies have proposed systematic classifications of quantum computing concepts to organize educational content and training pathways Gill et al. (2021). In parallel, proposals for a dedicated Body of Knowledge seek to define the conceptual foundations of QSE and guide both research and teaching Piattini et al. (2021). Additional frameworks address practical decision-making, process adaptation, and the transition from classical software engineering practices toward quantum-oriented development models Akbar et al. (2023); Als Salman (2024). More recent roadmaps further connect educational priorities with industrial needs and long-term workforce development Murillo et al. (2025).

Educational Tools and Learning Environments

Accessible tools have played a central role in early QSE education. Open ecosystems developed by IBM and other organizations have enabled students to experiment with quantum circuits, algorithms, and hybrid workflows using cloud-based environments Singh et al. (2024). Interactive simulators such as QuIDE have also contributed by allowing visual experimentation with core quantum operations in controlled educational settings Carrascal et al. (2023).



Table 1. Regional maturity of Quantum Software Engineering education

Region	Maturity Level	Main Strengths	Main Limitations
Europe	High	Frameworks, curricula, research	Few full degree programs
North America	High	Industry platforms, hands-on tools	Limited formal QSE identity
Asia	Medium	National investment, labs	Focus on QC over QSE
Latin America	Emerging	Growing interest, workshops	Funding and infrastructure

Beyond simulators, hybrid development environments that combine classical control systems with quantum execution are increasingly relevant because they better reflect real-world quantum software architectures. However, current tools still face limitations related to usability, reproducibility, scalability, and curricular integration, which may restrict their effectiveness in formal educational programs Piattini et al. (2020); Khan et al. (2025).

Reported Outcomes

Although most initiatives remain exploratory, reported outcomes are generally positive. Existing studies describe improvements in student motivation, conceptual understanding, participation in interdisciplinary programs, and confidence in quantum programming tasks Weisz et al. (2020). Practical and project-oriented approaches appear particularly effective in narrowing the gap between theoretical knowledge and software implementation skills.

Nevertheless, current evidence also reveals persistent limitations, including fragmented learning pathways, weak curricular integration, and the lack of measurable competency frameworks aligned with workforce expectations Kelly (2024); Carrascal et al. (2023). These findings suggest that the next stage of QSE education should move beyond isolated pilot experiences toward scalable, assessable, and professionally relevant training models.

From a comparative perspective, the reviewed initiatives also differ substantially in depth and educational purpose. Gamified and visualization-oriented resources appear effective for early engagement and conceptual access, but they usually remain limited to introductory learning goals. Tool-centered initiatives such as simulators and cloud platforms enable hands-on experimentation, yet their impact often depends on prior technical support, instructor expertise, and platform availability. By contrast, competency frameworks and project-based models show greater potential for professional preparation because they connect learning activities with development practices, although they demand higher curricular maturity and institutional resources.

Table 2 summarizes the main educational initiatives reported in the literature and their observed benefits and limitations.

The results suggest that practical and project-oriented approaches provide the strongest contribution to competency development.

4.3 Challenges and Limitations

Despite the progress identified in recent years, Quantum Software Engineering (QSE) education continues to face multiple structural, technological, and pedagogical barriers that limit its consolidation as a formal academic discipline. The reviewed literature consistently highlights

recurring challenges related to fragmentation, lack of standards, interdisciplinary complexity, restricted infrastructure access, and instructional limitations.

Fragmentation of Educational Efforts

One of the most persistent challenges is the fragmented nature of current initiatives. Many existing efforts are presented as isolated workshops, pilot courses, local programs, or short-term projects rather than components of an integrated educational ecosystem Zhao (2021); Piattini et al. (2021). As a result, students and educators often encounter inconsistent learning pathways, duplicated content, and limited continuity between introductory and advanced training stages. Fragmented curricular progression has also been associated with lower student proficiency and reduced long-term skill development Sodagari (2025).

Absence of Standards and Competency Models

QSE currently lacks internationally recognized standards for curriculum design, learning outcomes, competency assessment, and accreditation. While initiatives such as the Talavera Manifesto have proposed foundational principles Piattini et al. (2020), they do not yet constitute formal educational standards. This contrasts with classical Software Engineering, where reference models and quality standards help align academic preparation with professional expectations. Recent studies indicate that existing QSE programs differ substantially in how they define required skills, creating uncertainty for both institutions and learners Kelly (2024).

Interdisciplinary Complexity and Talent Gap

QSE requires the integration of knowledge from quantum physics, mathematics, computer science, and software engineering Arismendi and Mott (2023); Akbar et al. (2023). This interdisciplinary demand creates a high entry barrier, especially for undergraduate students whose preparation is typically concentrated in only one of these domains. Consequently, the supply of professionals with balanced technical competencies remains limited, widening the gap between educational output and projected workforce demand in the quantum sector Murillo et al. (2025); Singh (2025).

Limited Access to Hardware and Educational Infrastructure

Practical learning in QSE often depends on access to quantum hardware, simulators, and specialized development environments. Although cloud-based platforms have improved accessibility, free-tier limitations, queue times, and infrastructure dependency continue to restrict their value for sustained academic use Singh et al. (2024). These limitations are particularly significant in emerging regions, where local infrastructure is scarce and participation frequently depends on international partnerships.



Table 2. Educational initiatives in QSE and reported outcomes

Category	Example	Benefit	Limitation
Gamification	Entanglion	Higher motivation	Introductory scope
Simulators	QuIDE, Qiskit	Practical skills	Tool complexity
Frameworks	Body of Knowledge	Better structure	Not standardized
Project-based learning	Hybrid labs	Industry readiness	Resource intensive

Even when tools are available, many still present barriers related to usability, documentation quality, reproducibility, and scalability for classroom deployment Khan et al. (2025); Unitary Fund (2024).

Pedagogical Challenges

Quantum computing introduces instructional difficulties that differ substantially from classical programming education. Concepts such as superposition, entanglement, probabilistic measurement, and noise are often counter-intuitive for students with traditional computing backgrounds. In addition, activities such as debugging, verification, and performance evaluation cannot always be taught using classical assumptions or deterministic reasoning Ali et al. (2022); Arismendi and Mott (2023).

Another relevant limitation is the shortage of instructors with expertise spanning both quantum technologies and software engineering practices. This constrains program expansion and limits the quality of advanced training opportunities Piattini et al. (2020). Recent studies therefore emphasize scaffolded learning models and progressive instructional pathways to improve accessibility for students from non-physics backgrounds Sun et al. (2024).

Integrative Perspective

Overall, these challenges indicate that the consolidation of QSE education requires more than the expansion of isolated courses or tools. Sustainable progress depends on coordinated action among academia, industry, and public institutions to establish competency standards, strengthen interdisciplinary curricula, expand infrastructure access, and train specialized educators. Without such measures, the growth of the quantum workforce may continue to lag behind technological progress.

4.4 Future Directions and Perspectives

Quantum Software Engineering (QSE) education is expected to become a strategic component of the broader quantum technology ecosystem in the coming years. Although the field remains in an early stage of academic consolidation, the reviewed literature suggests several priority directions for strengthening its long-term development, including global educational frameworks, interdisciplinary collaboration, pedagogical innovation, and formal curricular integration.

Toward Global Educational Frameworks

One of the most frequently cited priorities is the creation of a unified educational framework for QSE. Such a framework should define core competencies, conceptual taxonomies, learning outcomes, teaching methodologies, and evaluation criteria that can be adapted internationally. In classical Software Engineering, reference bodies of knowledge have played a central role in aligning academic

programs with professional practice; a similar approach may be necessary for QSE.

Foundational initiatives such as the Talavera Manifesto Piattini et al. (2020) and more recent strategic roadmaps Murillo et al. (2025) provide an initial basis for this process. In addition, emerging competence models developed in Europe suggest that international qualification profiles for quantum-related professions are becoming increasingly feasible Greinert et al. (2024).

International and Interdisciplinary Collaboration

Because quantum technologies are inherently global, future educational progress will likely depend on stronger collaboration among universities, companies, research centers, and governments. Shared educational networks could enable students from different regions to access common resources such as simulators, remote laboratories, open repositories, and specialized mentorship opportunities.

Interdisciplinary collaboration is equally essential. Effective QSE programs must integrate perspectives from physics, mathematics, computer science, and software engineering within coherent learning pathways Arismendi and Mott (2023); Akbar et al. (2023); Sun et al. (2024); Kelly (2024). Programs capable of combining these domains are likely to produce graduates better prepared for real quantum development environments.

Pedagogical Innovation and Accessibility

Future growth in QSE education will also depend on the adoption of more inclusive and effective teaching approaches. Project-based learning, gamification, adaptive instruction, and experiential learning environments have shown potential for improving motivation and conceptual retention Weisz et al. (2020). These methods may be especially valuable in a field where abstract concepts often discourage early learners.

At the same time, accessibility remains a critical priority. Open educational resources, low-cost simulators, multilingual content, and online collaborative platforms can help reduce barriers for students in emerging regions or institutions with limited infrastructure. Without deliberate accessibility strategies, educational inequalities may widen as the field grows.

Curricular Integration and Lifelong Learning

QSE is likely to evolve beyond elective courses and short workshops toward formal integration into undergraduate, graduate, and professional development programs. Preliminary curriculum proposals already indicate that such integration is feasible Alsaman (2024). Over time, universities may incorporate QSE tracks within computer science and software engineering degrees, while industry may expand certification pathways for reskilling professionals.



Modular curricula and micro-credential models are particularly promising because they allow learners from diverse backgrounds to acquire targeted competencies progressively. These flexible formats may become essential in a rapidly evolving technological field.

Strategic Outlook

Overall, the future of QSE education depends on whether academic institutions and industry stakeholders can move from isolated initiatives toward coordinated, scalable, and internationally relevant training ecosystems. If successful, these efforts will not only reduce the quantum talent gap but also help establish QSE as a mature and sustainable discipline capable of supporting the next generation of quantum innovation.

4.5 Key Findings

The analysis of current literature reveals several recurring patterns that help characterize the present state of Quantum Software Engineering (QSE) education worldwide. First, the most effective educational initiatives are those that combine theoretical instruction with practical experimentation through simulators, cloud-based platforms, and guided programming activities. Programs that allow students to progressively apply quantum concepts in software development environments tend to report stronger conceptual understanding, higher engagement, and greater confidence in technical tasks.

Second, educational approaches centered exclusively on quantum algorithms or theoretical foundations appear insufficient for preparing future QSE professionals. While these components remain essential, they rarely address core software engineering competencies such as design practices, testing strategies, maintainability, debugging, documentation, and lifecycle management in quantum contexts. This suggests that many current programs still prioritize quantum computing education over true Quantum Software Engineering education.

Third, collaboration between academia and industry emerges as a decisive success factor. Institutions with access to industrial platforms, real or cloud-based hardware, and applied project environments are better positioned to provide students with relevant technical and professional competencies. Such partnerships also help align training outcomes with evolving workforce demands.

At the same time, the literature reveals an imbalance between highly accessible initiatives and professionally robust ones. Introductory approaches based on games, demonstrations, or isolated programming exercises facilitate entry into the field, but they rarely generate evidence about long-term competency acquisition. More comprehensive models, including project-based learning and competency-oriented curricular proposals, appear better aligned with professional QSE practice, yet they remain less frequent, more resource-intensive, and concentrated in institutions with stronger research or industrial ecosystems.

Fourth, significant regional asymmetries persist. Europe and North America currently lead in research production, tooling ecosystems, and competency-oriented initiatives,

while Asia is rapidly expanding through national investment strategies. Latin America and other emerging regions show growing interest but continue to face infrastructure, funding, and capacity constraints.

Finally, a major gap remains unresolved: no widely adopted educational model currently integrates quantum computing foundations with the structured processes, professional roles, quality metrics, and lifecycle practices traditionally associated with classical Software Engineering. This gap represents a strategic opportunity for future research and curriculum design.

An additional methodological finding is that the available evidence remains predominantly descriptive. Most publications report experiences, proposals, or local interventions, but comparatively few provide longitudinal assessment, controlled evaluation, or shared performance indicators. Consequently, the field still lacks a strong empirical basis for determining which educational designs are most effective under different institutional conditions.

Overall, the findings indicate that the next stage of QSE education should move beyond introductory quantum literacy toward comprehensive engineering-oriented training models capable of supporting the long-term needs of the quantum industry.

5. CONCLUSIONS

This article provides an exploratory synthesis of the current state of education in Quantum Software Engineering (QSE). Education in this area remains at an early stage of development. Although important progress has been achieved in recent years, the current landscape continues to be characterized by fragmentation, uneven regional growth, and the absence of cohesive training pathways. Existing initiatives provide valuable foundations, yet many remain limited to introductory quantum computing content or isolated educational experiences that do not fully address the competencies required by the emerging quantum industry.

This study shows that the most promising educational approaches are those that combine theoretical foundations with interactive, project-oriented, and practice-based learning experiences. Such models appear particularly effective in improving student motivation, conceptual understanding, and engagement. However, the lack of formal curricular frameworks, competency models, and internationally recognized guidelines continues to hinder the consolidation of QSE as an established academic discipline.

Among the principal challenges identified are the absence of standardized competencies, limited access to quantum hardware and specialized tools, and the shortage of instructors capable of teaching interdisciplinary content that connects physics, mathematics, computer science, and software engineering. Addressing these barriers will require coordinated action among universities, research centers, industry stakeholders, and public institutions.

Looking ahead, the future of QSE education will depend on the design of flexible, scalable, and competency-based programs that integrate practical experimentation, open-source ecosystems, and collaborative learning envi-



ronments. Establishing international standards, expanding equitable access to infrastructure, and strengthening interdisciplinary cooperation will be essential for transforming QSE into a robust and sustainable field of study.

At the same time, the exploratory nature of this review should be recognized as a methodological boundary. Because the synthesis integrates heterogeneous sources and does not follow a formal systematic-review protocol, its conclusions are best understood as a structured overview of trends and gaps rather than a definitive measure of the entire field. Future work should therefore complement this contribution with systematic reviews, explicit quality appraisal procedures, and comparative empirical studies capable of testing the effectiveness of specific educational interventions.

Ultimately, advancing QSE education is not only an academic priority but also a strategic requirement for developing the global talent needed to support the next generation of quantum technologies.

ACKNOWLEDGEMENTS

This research was supported by the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) through scholarship No. CVU: 2170965. The authors acknowledge the support provided by the Tecnológico Nacional de México (TecNM) and the Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET) for the completion of this work.

REFERENCES

- Akbar, M.A., Ilyas, M., and Baker, A. (2023). A systematic decision-making framework for tackling quantum software engineering challenges. *Automated Software Engineering*, 30, 22. doi:10.1007/s10515-023-00389-7.
- Ali, S., Yue, T., and Abreu, R. (2022). When software engineering meets quantum computing. *Communications of the ACM*, 65(4), 84–88. doi:10.1145/3512340.
- Alsalmán, A.I.S. (2024). Quantum software engineering: Best practices from classical to quantum approaches. *Journal of Quantum Information Science*, 14, 234–258. doi:10.4236/jqis.2024.144010.
- Arismendi, A. and Mott, S.P. (2023). Retos a los que se enfrenta la ingeniería del software cuántico. *Memorias de Investigación en Ingeniería*, 25, 90–106.
- Bhat, H.A., Khanday, F.A., Kaushik, B.K., Bashir, F., and Shah, K.A. (2022). Quantum computing: Fundamentals, implementations and applications. *IEEE Open Journal of Nanotechnology*, 3, 61–77. doi:10.1109/OJNANO.2022.3146671.
- Carrascal, D., Martínez, L., and Morales, F. (2023). Development of quantum technologies in latin america: Challenges and opportunities. *Journal of Quantum Information Science*, 13.
- Gill, S.S., Xu, M., and Bahsoon, R. (2021). Quantum computing: A taxonomy, systematic review and future directions. *Software: Practice and Experience*, 52(1), 66–114. doi:10.1002/spe.3039.
- Greiner, F., Goorney, S., Hilfert-Rüppell, D., Ubben, M.S., and Müller, R. (2024). Extending the european competence framework for quantum technologies: New proficiency triangle and qualification profiles. arXiv preprint arXiv:2410.07692.
- Haghparast, A., Yousefi, R.H., and Hasanovic, S. (2024a). Innovative approaches to teaching quantum computer programming and quantum software engineering. arXiv preprint arXiv:2501.01446. URL <https://arxiv.org/abs/2501.01446>.
- Haghparast, M., Dwivedi, K., and Mikkonen, T. (2024b). Quantum software engineering and quantum software development lifecycle: A survey. *Cluster Computing*. doi:10.1007/s10586-024-04362-1.
- Hidary, J.D. (2019). *Quantum Computing: An Applied Approach*. Springer, Cham, Switzerland. doi:10.1007/978-3-030-23922-0.
- Kelly, A.M. (2024). Quantum information science and technology professional learning model: Strengthening teacher pedagogical content knowledge. *Physical Review Physics Education Research*, 20(2), 020154. doi:10.1103/PhysRevPhysEducRes.20.020154.
- Khan, A.A., Ye, B., Akbar, M.A., Khan, J.A., Mougouei, D., and Ma, X. (2025). Mining q&a platforms for empirical evidence on quantum software programming. arXiv preprint arXiv:2503.05240. URL <https://arxiv.org/abs/2503.05240>.
- Murillo, A., Felderer, M., and Piattini, M. (2025). Quantum software engineering: Roadmap and challenges ahead. *Communications of the ACM*, 68(5), 74–87. doi:10.1145/3712002.
- Piattini, M., Pérez-Castillo, R., and Felderer, M. (2020). The talavera manifesto for quantum software engineering and programming. In *Proceedings of the QuANSWER Workshop*.
- Piattini, M., Pérez-Castillo, R., and Felderer, M. (2021). Toward a quantum software engineering. In *16th International Conference on Software Engineering Advances (ICSEA 2021)*, 38–44.
- Pérez-Castillo, R., Jiménez-Navajas, L., and Piattini, M. (2021). Modelling quantum circuits with uml. In *IEEE International Workshop on Quantum Software Engineering (Q-SE 2021)*, 7–12. doi:10.1109/Q-SE52541.2021.00007.
- Schmitt, M.T., Garcia, L., and Piattini, A. (2025). Quantum-based software engineering (qbse): A new research direction for applying quantum computing to classical software engineering. arXiv preprint arXiv:2505.23674. URL <https://arxiv.org/abs/2505.23674>.
- Singh, A. (2025). Challenges and issues for integration and implementation in quantum-classical software systems. SSRN Preprint.
- Singh, P., Kumar, S., Agrawal, M., and Yadav, S. (2024). A survey on available tools and technologies enabling quantum computing. *IEEE Access*, 12, 1–20. doi:10.1109/ACCESS.2024.3380307.
- Sodagari, S. (2025). Key predictors of student proficiency in quantum computing. *Computers & Education: Quantum Learning*, 3, 100075. doi:10.1016/j.ceql.2025.100075.
- Sun, Q., Zhou, S., Chen, R., Feng, G., Hou, S.Y., and Zeng, B. (2024). From computing to quantum mechanics: Accessible and hands-on quantum computing education for high school students. *EPJ Quantum Technology*, 11, 71. doi:10.1140/epjqt/s40507-024-00271-9.
- Unitary Fund (2024). Quantum open source survey results – educational insights. Online report.
- Weisz, J.D., Ashoori, M., and Ashktorab, Z. (2020). Entanglion: A board game for teaching the principles of quantum computing. In *Proceedings of the CHI Conference on Human Factors in Computing Systems (CHI '20)*, 1–12. ACM, New York, NY, USA. doi:10.1145/3313831.3376301.
- Zhao, J. (2021). Quantum software engineering: Landscapes and horizons (v2). arXiv preprint. doi:10.48550/arXiv.2007.07047.

Systematic Study of Visual and Interactive Tools and Sequences for Refactoring

Ninfa Yukary Hernandez-Vargas *. René Santaolaya-Salgado* Blanca Dina Valenzuela-Robles*
José Luis Ramírez-Alcántara*

**ITecnológico Nacional de México/CENIDET, Departamento de Ciencias Computacionales,
Cuernavaca, México (Tel: 777 362 7770; e-mail: m24ce055@cenidet.tecnm.mx).*

Resumen: Refactoring is essential for improving software structure without altering its behavior, especially when design principles like SOLID are overlooked. While various methods reduce code smells and improve quality, they may produce side effects. Tools in environments such as Eclipse, IntelliJ IDEA, and Visual Studio support this process but have limitations, sometimes requiring manual work. This paper analyzes the use of tools and the definition of refactoring sequences to improve software quality and prevent the reintroduction of technical debt.

Keywords: Refactoring, Legacy code, Attributes, Methods, Tools

1. INTRODUCTION

La refactorización de software es una actividad fundamental dentro del ciclo de vida del software, orientada a mejorar el diseño interno de un sistema sin modificar su comportamiento externo (Almogahed A. , 2023). Esta práctica resulta especialmente relevante en las fases de mantenimiento y evolución, ya que permite optimizar la estructura del código y mejorar su calidad. La refactorización se realiza considerando distintos atributos de calidad del software lo que contribuye a desarrollar sistemas eficientes, mantenibles y adaptables a futuros cambios (A. Almogahed H. M., 2023).

A pesar de sus beneficios, uno de los principales desafíos que enfrentan los desarrolladores es el alto costo en tiempo y recursos que implica llevar a cabo procesos manuales de refactorización, particularmente, en proyectos de mediana y gran escala. Se estima que la refactorización de un proyecto mediano, con aproximadamente 30,000 líneas de código, puede requerir alrededor de seis semanas de trabajo (V. Alizadeh, 2019). Este esfuerzo puede incrementarse cuando las actividades se realizan de manera manual o cuando no se dispone de herramientas automatizadas que faciliten el proceso.

Con el objetivo de apoyar estas tareas, algunos Entornos de Desarrollo Integrado (IDE) modernos, como NetBeans y Eclipse, incorporan funcionalidades que permiten aplicar técnicas de refactorización de manera semiautomática. Estas herramientas facilitan la refactorización de código, proporcionando un soporte automatizado básico para los desarrolladores (al., 2023).

En este sentido, las herramientas de refactorización contribuyen a realizar transformaciones en el código de manera eficiente, reduciendo el tiempo requerido para su aplicación y disminuyendo la probabilidad de errores humanos asociados a procesos manuales. Estos aspectos resultan clave para mejorar tanto la productividad de los equipos de

desarrollo como la calidad del producto de software. En particular, para contribuir en este sentido, en el laboratorio de CENIDET de Ingeniería de Software se han propuesto y automatizado diversos métodos de refactorización que resuelven la deuda técnica por diferentes causas relacionadas a los principios SOLID y malos olores de atributos de calidad, tales como la deuda técnica por carencia de refactorización, herencia de implementación, carencia de protección, exceso de acoplamiento, violación del principio de única responsabilidad, código dispensable, entre otras.

Sin embargo, a pesar de estos avances, la automatización de la refactorización continúa siendo un área de investigación emergente, por lo que persiste la necesidad de desarrollar enfoques y herramientas avanzadas (Sandhya Tarwani, 2024).

En este artículo se presenta una revisión sistemática de la literatura, con el objetivo de plantear el alcance del estado del arte en refactorización de código legado. El orden en que se presenta es el siguiente: En la sección III se presenta la metodología seguida para la RSL según la metodología de Kitchenham. La sección IV expone el análisis, la interpretación y los resultados del estudio. Finalmente, la sección V presenta las conclusiones y los trabajos futuros.

1.1 Antecedentes y trabajos relacionados

En el estudio titulado “A Comparative Study of Automated Refactoring Tools” (Alshayeb, 2024), se presenta un análisis comparativo de herramientas automatizadas de refactorización, en el cual se evalúan cuatro propuestas: MultiRefactor, JDeodorant, jSparrow y Spartenizer. Para llevar a cabo esta evaluación, los autores emplearon cinco proyectos de código abierto como casos de estudio, con el fin de analizar el desempeño y las capacidades de cada herramienta en contextos reales.

Por otra parte, en el estudio “An Empirical Study on the Code Refactoring Capability of Large Language Models” (Jonathan Cordeiro, 2026) los autores analizan el potencial de los



modelos de lenguaje de gran escala (LLMs) en la automatización de tareas de refactorización. En particular, se evalúa el desempeño de modelos como StarCoder2, GPT-4o-mini, GPT-4o, LLaMA 3 y DeepSeek-v3 en la mejora de la calidad del código.

Asimismo, en el artículo "Automatic software refactoring: a systematic literature review" (Abdulrahman Ahmed Bobakr Baqais, 2019), presenta una revisión de estudios sobre refactorización automatizada, destacando que su aplicación no se limita al código fuente, sino que también abarca otros artefactos como diagramas UML y modelos de procesos.

Finalmente, en el estudio "30 Years of Software Refactoring Research: A Systematic Literature Review" (Chaima Abid, 2020), analiza 3183 trabajos y propone una taxonomía basada en cinco aspectos: ciclo de vida, artefactos, objetivos, técnicas y métodos de evaluación.

1.2 Metodología

Una revisión sistemática de la literatura constituye una metodología rigurosa que permite identificar, evaluar e interpretar de manera integral la evidencia científica disponible y relevante entorno a una o varias preguntas de investigación. En este sentido, el presente estudio se desarrolla siguiendo las directrices propuestas por Kitchenham en (Kitchenham Barbara, 2009), ampliamente reconocidas en el ámbito de la Ingeniería de Software para la realización de revisiones sistemáticas.

El objetivo de esta revisión sistemática es identificar, analizar e interpretar la evidencia existente sobre las herramientas de refactorización y las secuencias de aplicación de técnicas de refactorización, con el propósito de comprender su impacto en la calidad del software y en la mejora de los procesos de mantenimiento y evolución de sistemas.

La presente revisión sistemática de la literatura se desarrolló en tres fases principales. La Fase III-A corresponde a la planificación, la cual se centra en identificar la necesidad del estudio y en la definición de las preguntas de investigación. La Fase III-B comprende el proceso de revisión, selección y evaluación de los estudios primarios, así como su análisis y síntesis. Finalmente, la Fase III-C presenta el informe final, incluyendo el reporte estructurado y la comunicación de los hallazgos obtenidos.

2. FASE III-A. PLANIFICACIÓN Y NECESIDAD DEL ESTUDIO

2.1 III-A.1 Necesidad de una revisión sistemática sobre herramientas y secuencias de refactorización

Este estudio surge de la necesidad identificada por ingenieros y desarrolladores de software en el contexto del mantenimiento y evolución de sistemas, particularmente en código legado, donde la refactorización constituye una práctica esencial para mejorar la calidad interna del software sin modificar su comportamiento externo. En los últimos años, se han propuesto diversas técnicas de refactorización orientadas a reducir code smells y mejorar atributos de calidad

como la cohesión, la complejidad y la comunicación entre componentes (Riansyah, 2020).

La literatura evidencia que el impacto de las técnicas de refactorización no es uniforme, ya que, si bien algunas contribuyen a mejorar determinados atributos de calidad, mientras que otras pueden degradar aspectos del software (Eduardo Fernandes, 2020) (A. Almogahed H. M., 2023). Por ello, resulta fundamental definir adecuadamente la combinación y secuencia de operaciones de refactorización para mejorar la calidad sin comprometer la seguridad del sistema (Estomii Edward, 2024).

En este sentido, la correcta definición de secuencias de refactorización se vuelve un aspecto crítico, ya que una aplicación inadecuada puede provocar la reintroducción de deuda técnica o afectar negativamente la calidad del software. A pesar de su relevancia, este problema ha sido abordado de manera limitada en la literatura, especialmente en lo que respecta a la integración entre herramientas de refactorización y estrategias de secuenciación.

2.2 III-A.2 Especificación de las preguntas de investigación

Con el fin de guiar el desarrollo de esta revisión sistemática y dar respuesta al objetivo planteado, se definieron las siguientes preguntas de investigación:

P1: ¿Qué herramientas visuales e interactivas han sido propuestas o desarrolladas para la gestión de refactorización y deuda técnica?

P2: ¿Existen herramientas de automatización de refactorización de code smell en la literatura?

P3: ¿Existen reglas de ejecución sucesiva de refactorización de código en la literatura?

3. FASE III-B. SELECCIÓN DE ESTUDIOS PRIMARIOS

A partir de las preguntas de investigación, se identificaron dos ejes temáticos principales: (i) herramientas de refactorización y (ii) secuencias o estrategias de aplicación de técnicas de refactorización.

Con base en estos ejes, se definieron cadenas de búsqueda estructuradas mediante operadores booleanos, con el objetivo de recuperar estudios relevantes y minimizar la inclusión de trabajos no relacionados. La cadena de búsqueda general se definió de la siguiente manera:

("code refactoring") AND ("software quality" OR "quality attributes" OR "code smells") AND ("customization" OR "batch refactoring" OR "refactoring sequencing" OR "refactoring ordering" OR "interactive refactoring" OR "automated refactoring") AND ("tool" OR "framework" OR "approach" OR "model" OR "empirical study" OR "case study").

3.1. III-B.1 Búsqueda inicial

La cadena de búsqueda se aplicó en bases de datos académicas ampliamente reconocidas como IEEE Xplore, Google académico (PLOS ONE, ACM Transactions on Software Engineering and Methodology, Springer), las cuales

proporcionan acceso a literatura científica de alta calidad en el ámbito de la ingeniería de software.

La selección de los estudios se realizó utilizando las herramientas de búsqueda avanzada disponibles en cada plataforma. La última búsqueda se llevó a cabo en agosto de 2025, considerando publicaciones científicas en revistas y conferencias correspondientes a los últimos seis años.

Como resultado de la aplicación de la cadena de búsqueda, se obtuvo un total de 729 publicaciones como se puede ver en la Tabla 1.

Tabla 1. Resumen de resultados.

Base de datos	Resultados de las búsquedas
Google Academico	727
IEEE	2
Total	729

3.2. III-B.2 Criterios de calidad y depuración de los estudios

Para evaluar la calidad de los estudios, se realizaron las

siguientes preguntas: 1) ¿El estudio menciona alguna herramienta de refactorización? 2) ¿El estudio menciona alguna secuencia de refactorización o menciona su importancia? Estos cuestionamientos, se alinearon con las preguntas de este estudio sistemático para obtener calidad en los estudios seleccionados.

Para el primer refinamiento de los estudios, se aplicaron los siguientes criterios:

Criterios de inclusión: artículos de estudios primarios publicados en revistas científicas, artículos de conferencias, trabajos publicados del 2019 al 2025, artículos clásicos que son base del tema a tratar.

Criterios de Exclusión: Se excluyen los estudios secundarios, libros, patentes, capítulos de libros, tesis y material didáctico, páginas web y presentaciones en Power Point, opiniones personales en blogs de internet, artículos duplicados y artículos no escritos en idioma inglés.

Artículos seleccionados: Identificados: 729, Duplicados eliminados: 129, Registros únicos: 600, Excluidos por título/resumen: 550, Evaluados a texto completo: 50, Excluidos por texto completo: 18

Como resultado, se seleccionaron 32 estudios. Se realizó una revisión exhaustiva del resumen (abstract) de cada estudio para garantizar su relevancia con el tema y su capacidad de responder las preguntas de investigación, como se puede observar en la Figura 1.

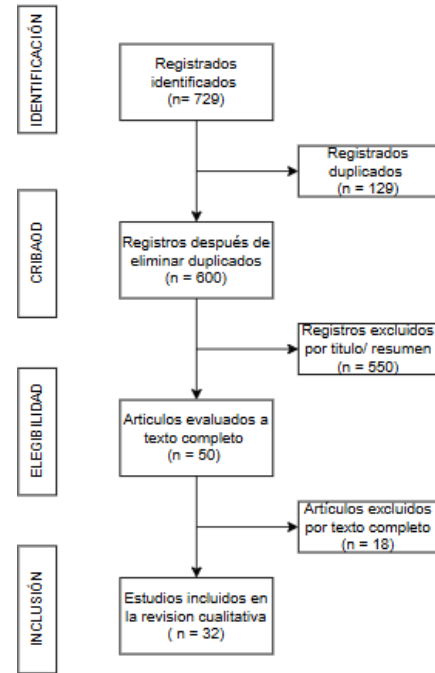


Figura 1. PRISMA.

3.3. III-B.3 Análisis de los datos

El análisis de los datos se llevó a cabo mediante una revisión detallada de los estudios seleccionados, con el objetivo de identificar, organizar e interpretar la evidencia relevante en función de las preguntas de investigación planteadas. Para ello, se realizó una lectura completa de cada estudio, a partir de la cual se elaboró un resumen estructurado que permitió extraer la información significativa.

Posteriormente, la información recopilada fue organizada y sintetizada, con el fin de obtener una visión integral del estado del arte. Los estudios fueron clasificados en dos núcleos temáticos principales: (i) herramientas de refactorización y (ii) secuencias de aplicación de métodos de refactorización. Esta clasificación permitió estructurar el análisis de manera coherente con los objetivos del estudio.

Para dar respuesta a las preguntas de investigación (P1, P2 y P3), se llevó a cabo un análisis cualitativo considerando distintos criterios, tales como el tipo de herramienta, nivel de automatización, capacidad de personalización y la importancia de la definición de secuencias de refactorización.

4. REVISIÓN DE LA LITERATURA

La Sección IV-B presenta una visión general de los métodos de refactorización identificados en la literatura. La Sección V-A aborda las herramientas de refactorización reportadas, con el propósito de responder a las preguntas de investigación (P1 y P2). Finalmente, la Sección V-B analiza las secuencias y estrategias de aplicación de técnicas de refactorización, con el fin de responder a la pregunta de investigación (P3).



4.1 IV-A. Visión General de la Literatura Seleccionada

En esta subsección se presenta un análisis general de los estudios incluidos en esta revisión, considerando su distribución temporal y enfoque temático.

En términos generales, se observa que el interés en la refactorización de software ha ido incrementando en los últimos años, particularmente en lo relacionado con el desarrollo de herramientas y técnicas orientadas a mejorar la calidad del código. Este crecimiento refleja la importancia de la refactorización como una práctica clave dentro del mantenimiento y evolución del software.

Sin embargo, al analizar la literatura desde una perspectiva amplia, se observa que existe una menor cantidad de investigaciones enfocadas en la secuenciación de refactorizaciones. En particular, son limitados los trabajos que abordan el problema de determinar el orden adecuado de aplicación de múltiples técnicas, considerando sus efectos en los atributos de calidad del software. Este análisis pone de manifiesto una oportunidad de investigación en el diseño de enfoques que integren herramientas de refactorización con mecanismos de apoyo para la toma de decisiones, permitiendo definir secuencias adecuadas que optimicen la calidad del software y reduzcan efectos colaterales.

4.2 IV-B Revisión general sobre los métodos de refactorización

El término refactorización se origina en el trabajo de Opdyke (Opdyke, 1992), quien define estrategias que apoyan el cambio en una parte del software. La refactorización contribuye a mejorar la legibilidad del código mediante la aplicación de métodos específicos, los cuales permiten reducir los code smells, eliminar problemas potenciales y fortalecer la calidad interna del sistema (G. Lacerda, 2020).

La relación entre los atributos de calidad interna y la caracterización de las técnicas de refactorización propone un enfoque estructurado para la mejora de la calidad del software. Un ejemplo importante dentro de la refactorización es que la técnica Inline Method mejora la cohesión. Un IDE podrían proporcionar sugerencias de refactorización basado en los atributos de calidad interna (Almogahed A. M.-M., 2023).

Asimismo, en el análisis realizado en el artículo titulado “The untold story of code refactoring customizations in practicen”, se centran en cuatro tipos comunes de refactorización: Extract Method, Inline Method, Pull Up Method y Move Method (al., 2023).

Asimismo, en el estudio “Refactoring Effect on Internal Quality Attributes: What Haven’t They Told You Yet?”, se reporta que técnicas como Extract Interface, Extract Method, Move Field, Move Method y Push Down Method mejoraron exactamente un atributo de calidad, mientras que Inline Method y Push Down Field lograron mejorar dos atributos (Eduardo Fernandes, 2020). En la Tabla 2 se presentan los métodos de refactorización comunes y las herramientas empleadas para corregir determinados code smells.

Table 2. Métodos de Refactorización.

Code Smell	Método de refactorización	Herramientas
Long method, God class, Duplicated code (R. Singh, 2020), Feature Envy (G. Lacerda, 2020)	Extract Method o Método de extracción: Creación de un nuevo método extrayendo una sección de código de un método existente (AlOmar, 2021). Métricas: Complejidad ciclomática o cognitiva, Número de líneas de código, Falta de cohesión (Fernandes, 2023)	Refactoring-Bot (Bogner, 2019), JDeodorant, MultiRefactor or (Alshayeb, 2024) Eclipse, IntelliJ (D. Oliveira, 2019) Entorno de refactorización en vivo (Fernandes, 2023).
Feature Envy (G. Lacerda, 2020)	Inline Method o Método en línea (IM): Reemplaza las llamadas y los usos de un método, eliminar su declaración (AlOmar, 2021). Métricas: Cohesión, Complejidad, Mensajería. (Almogahed A. M.-M., 2023)	MultiRefactor, jSparrow (Alshayeb, 2024)
Long Parameter List (G. Lacerda, 2020)	Add Parameter o Agregar parámetro (AP): Se proporciona un parámetro para un objeto que pueda transmitir estos datos (A. Almogahed H. M., 2023). Métricas: Tamaño del diseño, la encapsulación y la mensajería (Almogahed A. M.-M., 2023)	Entorno de refactorización en vivo (Fernandes, 2023).
Long Parameter List (G. Lacerda, 2020)	Remove Parameter o Eliminar parámetro (RP): Es utilizado cuando el cuerpo del método ya no se utiliza en el parámetro (A. Almogahed H. M., 2023). Métricas: Tamaño del diseño, la encapsulación y la mensajería (Almogahed A. M.-M., 2023)	MultiRefactor or (Alshayeb, 2024)
Data class (Chug, 2020)	Encapsulate Field o Encapsular campo (EF): Al modificar la accesibilidad de los campos públicos, este método se aplica para restringir el acceso a los datos (A. Almogahed H. M., 2023). Métricas: Encapsulación, Complejidad, Mensajería (Almogahed A. M.-M., 2023)	MultiRefactor or (Alshayeb, 2024)

Blob, Large Class, God class, Shotgun surgery, Duplicated code (R. Singh, 2020), Divergent Change (G. Lacerda, 2020)	Extract Class o Extraer clase (EC): Este método se utiliza para crear una nueva clase cuando una existente es grande y tiene demasiadas tareas (A. Almogahed H. M., 2023). Métricas: Herencia y cohesión (Almogahed A. M.-M., 2023)	Entorno de refactorización en vivo (Fernandes, 2023), Jdeodorant, REFBOT (V. Alizadeh, 2019), MultiRefactor (Alshayeb, 2024)
Shotgun surgery (R. Singh, 2020)	Extract Superclass o Extraer superclase (ESP): Cuando dos clases tienen campos y métodos similares, se transfieren los campos y métodos compartidos a la superclase (A. Almogahed H. M., 2023). Métricas: Encapsulación, cohesión, mensajería, tamaño del diseño, abstracción, las jerarquías, polimorfismo y composición (Almogahed A. M.-M., 2023).	REFBOT (V. Alizadeh, 2019)
Feature Envy, God class (R. Singh, 2020), Shotgun Surgery (G. Lacerda, 2020)	Move Method o Mover método: Cambia la declaración de un método, de una clase a otra (AlOmar, 2021). Métricas: Cohesión y la mensajería (Almogahed A. M.-M., 2023).	Entorno de refactorización en vivo (Fernandes, 2023), Jdeodorant, REFBOT, TrueRefactor (G. Lacerda, 2020), MultiRefactor (Alshayeb, 2024)
Duplicated Code (G. Lacerda, 2020)	Pull-up Method: Mover un método hacia arriba en la cadena de herencia, desde una clase hijo hacia una clase padre (AlOmar, 2021). Métricas: Número de declaraciones, tamaño promedio de las declaraciones, la falta de cohesión (Iyad Alazzam, 2020)	REFBOT (V. Alizadeh, 2019), MultiRefactor (Alshayeb, 2024)
Duplicated Code (G. Lacerda, 2020)	Push-down Method: Desplaza hacia abajo un método de la cadena herencia de una	REFBOT (V. Alizadeh, 2019),
Lacerda, 2020)	clase primaria a una clase secundaria (AlOmar, 2021). Métricas: Complejidad, Herencia, tamaño del diseño (Oumayma Hamdi, 2021)	MultiRefactor (Alshayeb, 2024)
Duplicated Code (G. Lacerda, 2020)	Rename Method o Cambiar nombre de método (MM): Cambia el nombre de un identificador de método a uno diferente (AlOmar, 2021). Métricas: Complejidad, Legibilidad (Almogahed A. M.-M., 2023)	MultiRefactor (Alshayeb, 2024)
Data class (Chug, 2020)	Hide Method u Ocultar método (HM): Esta técnica es utilizada para proteger o hacer que el método sea privado Métricas: Mensajería (A. Almogahed H. M., 2023)	Eclipse, IntelliJ (D. Oliveira, 2019)
Lazy Class (G. Lacerda, 2020)	Inline Class o Clase en línea (IC): Se realiza la copia de todos los métodos y campos de la clase en otra Métricas: Cohesión (A. Almogahed H. M., 2023)	TrueRefactor (G. Lacerda, 2020), MultiRefactor (Alshayeb, 2024)
Shotgun surgery (R. Singh, 2020)	Move Field o Mover campo (MF): Es un método en el que se mueve un campo de una clase principal a una clase apropiada Métricas: Encapsulación (DAM) (A. Almogahed H. M., 2023)	Eclipse, IntelliJ (Eduardo Fernandes, 2020)

5. RESULTADOS

5.1 V-A Herramientas de refactorización

PI. ¿Qué métodos de refactorización existen?

Los resultados obtenidos muestran que los métodos de refactorización se agrupan principalmente en tres enfoques: manual, semiautomático y automático. El enfoque manual proporciona un mayor control al desarrollador; sin embargo, incrementa el esfuerzo requerido y la probabilidad de introducir errores durante el proceso. En contraste, los enfoques semiautomáticos combinan la intervención humana con soporte automatizado, facilitando la toma de decisiones y reduciendo parte de la complejidad asociada a la refactorización (Alizadeh, 2019).

Por otra parte, los métodos automáticos destacan por emplear técnicas basadas en búsqueda y optimización para generar secuencias completas de refactorización orientadas a mejorar atributos de calidad del software. Los estudios analizados evidencian que este enfoque resulta especialmente útil en sistemas complejos o en escenarios donde se requiere optimizar múltiples atributos de calidad de manera simultánea (Alizadeh, 2019).



P2. ¿Qué herramientas de refactorización existen y qué tan efectivas son?

Los resultados evidencian que diversas herramientas han sido desarrolladas para apoyar procesos de refactorización, particularmente en contextos relacionados con modernización y descomposición de sistemas monolíticos (Ivers, 2022).

Entre las herramientas identificadas destaca RefBot, que monitorea repositorios y genera recomendaciones de refactorización a partir de cambios en solicitudes de extracción, permitiendo al desarrollador aceptar o rechazar dichas sugerencias (V. Alizadeh, 2019). Asimismo, DART utiliza métricas como cohesión, dependencia entre métodos y líneas lógicas de código, logrando reducir la redundancia hasta en un 28%, lo que evidencia su efectividad en la mejora estructural del software (Rosziati Ibrahim, 2020).

De igual forma, herramientas como “Towards a Live Environment for Code Refactoring” se enfocan en generar recomendaciones mediante técnicas como extract method (Fernandes, 2023), mientras que Refactoring-Bot automatiza la corrección de code smells, especialmente long methods, contribuyendo a mejorar la legibilidad y mantenibilidad del código (Bogner, 2019).

Adicionalmente, CASpER integra estrategias basadas en métricas y análisis textual para detectar y refactorizar code smells como Feature Envy, Blob y Misplaced Class, proporcionando un enfoque integral para la identificación de problemas de diseño (De Stefano, 2020).

Aunque los IDEs populares como Eclipse, NetBeans, IntelliJ IDEA y Visual Studio incorporan herramientas de refactorización para modificaciones estándar, estas presentan limitaciones relacionadas con confiabilidad y personalización, reduciendo su efectividad en escenarios complejos (al., 2023).

En contraste, herramientas especializadas como MultiRefactor destacan por soportar múltiples lenguajes de programación (Alshayeb, 2024). Asimismo, las herramientas basadas en inteligencia artificial permiten detectar code smells y sugerir refactorizaciones de forma automatizada; sin embargo, aún enfrentan desafíos en términos de confianza, seguridad e interpretación del contexto del código (Podduturi, 2025).

En conclusión, los estudios analizados evidencian que las herramientas actuales de refactorización aún presentan limitaciones significativas en el soporte para la toma de decisiones y la definición de secuencias adecuadas para aplicar múltiples refactorizaciones (Ivers, 2022). Asimismo, se identificó que no existe una herramienta capaz de cubrir de manera integral todos los métodos de refactorización (A. Almogahed H. M., 2023).

5.2 V-B Secuencias de refactorización

P3. ¿Cómo se deben secuenciar las técnicas de refactorización?

Los resultados obtenidos destacan que la secuenciación de técnicas de refactorización influye directamente en la calidad final del software, debido a que el orden de aplicación puede generar mejoras o degradaciones en distintos atributos de calidad (A. Almogahed H. M., 2023).

El enfoque HSHEP propone un algoritmo de optimización híbrido para determinar secuencias adecuadas de refactorización, considerando tanto métricas de calidad como las preferencias del desarrollador, lo que permite gestionar efectos secundarios y mejorar la toma de decisiones (Maini, 2024).

Asimismo, los entornos multi-agente basados en modelos de lenguaje (LLM) evidencian una tendencia hacia enfoques inteligentes y adaptativos, capaces de considerar múltiples atributos de calidad de manera simultánea durante el proceso de refactorización (Rajendran, 2025).

Por otra parte, los resultados indican que la priorización de code smells es un elemento fundamental previo a la secuenciación, ya que no todos los problemas tienen el mismo impacto en la mantenibilidad. En este sentido, se destaca que la refactorización no garantiza por sí sola una mejora global en la calidad del software si no se aplica de manera estratégica (Alshammari, 2021).

La definición de secuencias adecuadas y la priorización de problemas resultan esenciales para maximizar beneficios y minimizar efectos negativos durante el proceso de refactorización (Alshammari, 2021). En la Tabla 3 se presentan algunos artículos que proponen secuencias de refactorización.

Table 3. Herramientas de refactorización y secuencias.

Artículo	Objetivo	Tipo de proceso	Secuencia	Alcance
(Maini, 2024)	Técnica HSHEP	Manual	Si	Caso de estudio
(V. Alizadeh, 2019)	Repositorio	Semiautomático	Si	Implementado
(Alizadeh, 2019)	Análisis	Manual	Si	Caso de estudio
(Fernandes, 2023)	Herramienta	Semiautomático	Si	Implementado
(Bogner, 2019)	Herramienta	Automático	No	Implementado
(Alshayeb, 2024)	Análisis	Manual	No	Caso de estudio

6. CONCLUSIONES

La literatura analizada evidencia que la refactorización puede llevarse a cabo tanto de manera manual como mediante herramientas automatizadas, siendo estas últimas un apoyo relevante para mejorar la eficiencia en el proceso. Entre las herramientas destacadas, JDeodorant se posiciona como una de las más utilizadas en el ámbito académico y profesional.



Los entornos de desarrollo integrados, como Eclipse, IntelliJ IDEA y Visual Studio, incorporan funcionalidades que permiten reducir el esfuerzo manual y optimizar el tiempo requerido para la modificación del código. Sin embargo, dichas herramientas presentan limitaciones importantes, ya que pueden generar refactorizaciones incompletas o incorrectas, además de ofrecer soporte únicamente para un subconjunto de técnicas de refactorización. En este sentido, herramientas como Eclipse Refactor, JDeodorant, MultiRefactor y jSparrow adoptan enfoques automatizados o semiautomatizados que, si bien permiten cierto grado de personalización, no logran cubrir la totalidad de necesidades del desarrollador.

Por otra parte, los estudios revisados destacan que los métodos de refactorización tienen un impacto significativo en los atributos de calidad interna del software, tales como la cohesión, el acoplamiento y la complejidad. No obstante, la ausencia de herramientas suficientemente flexibles, confiables y orientadas a la evaluación integral de la calidad obliga, en muchos casos, a que los desarrolladores recurran a procesos manuales, lo que incrementa el esfuerzo, el tiempo de intervención y el riesgo de introducir errores.

En este contexto, se identifica una oportunidad de mejora en el desarrollo de soluciones que integren capacidades de evaluación, personalización y apoyo en la toma de decisiones. Como trabajo futuro, se propone el diseño e implementación de una herramienta visual e interactiva que incorpore métodos de refactorización enfocados en la herencia de implementación, la carencia de abstracción y la protección modular. Esta herramienta permitirá evaluar el estado del software antes y después de aplicar refactorizaciones.

REFERENCES

- A. Almogahed, H. M. (2023). Empirical Investigation of the Diverse Refactoring Effects on Software Quality: The Role of Refactoring Tools and Software Size. *3rd International Conference on Emerging Smart Technologies and Applications (eSmarTA)*, 1-6. doi:10.1109/eSmarTA59349.2023.10293407
- A. Almogahed, H. M. (2023). Optimized Refactoring Mechanisms to Improve Quality Characteristics in Object-Oriented Systems. *IEEE Access*, 11, 99143-99158. doi:10.1109/ACCESS.2023.3313186
- Abdulrahman Ahmed Bobakr Baqais, M. A. (2019). Automatic software refactoring: a systematic literature review. *Software Quality Journal*(28), 459-502.
- al., D. O. (2023). The untold story of code refactoring customizations in practice. *IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, 108-120. doi:10.1109/ICSE48619.2023.00021
- Alizadeh, V. K. (2019). Interactive and dynamic multi-objective software refactoring recommendations. *Proceedings of the 33rd ACM/IEEE International Conference on Automated*, 1-30.
- Almogahed, A. (2023). A Refactoring Classification Framework for Efficient Software Maintenance. *IEEE, 11*, 78904-78917. doi:10.1109/ACCESS.2023.3298678
- Almogahed, A. M.-M. (2023). A refactoring categorization model for software quality improvement. *PLOS ONE*, 18(11). doi:https://doi.org/10.1371/journal.pone.0293742
- AlOmar, E. L. (2021). On the Documentation of Refactoring Types. *Autom Softw Eng*, 29(9).
- Alshammari, T. A. (2021). Toward a Software Bad Smell Prioritization Model for Software Maintainability. *Arabian Journal for Science and Engineering*, 46, 9157-9177. Retrieved from <https://doi.org/10.1007/s13369-021-05766-6>
- Alshayeb, M. A. (2024). A Comparative Study of Automated Refactoring Tools. *IEEE Access*, 12, 18764-18781. doi:10.1109/ACCESS.2024.3361314
- Bogner, M. W. (2019). Towards an Autonomous Bot for Automatic Source Code Refactoring. *IEEE/ACM 1st International Workshop on Bots in Software Engineering (BotSE)*, 24-28. doi:10.1109/BotSE.2019.00015
- Chaima Abid, V. A. (2020). 30 Years of Software Refactoring Research: A Systematic Literature Review. *IEEE TRANSACTIONS OF SOFTWARE ENGINEERING*, 1(1).
- Chug, M. A. (2020). A Systematic Literature Survey of Software Metrics, Code Smells and Refactoring Techniques. *Journal of Information Processing Systems*, 19(4), 915-934.
- D. Oliveira, A. C. (2019). On the Customization of Batch Refactoring. *IEEE/ACM 3rd International Workshop on Refactoring (IWorR)*, 13-16. doi:10.1109/IWoR.2019.00010



- De Stefano, M. a. (2020). cASpER: A Plug-in for Automated Code Smell Detection and Refactoring. *Association for Computing Machinery*(79), 1-3.
- Eduardo Fernandes, A. C. (2020). Refactoring effect on internal quality attributes: What haven't they told you yet? *Information and Software Technology*, 126, 106347–106347. doi:<https://doi.org/10.1016/j.infsof.2020.106347>
- ESTOMII EDWARD, A. S. (2024). On the Impact of Refactorings on Software Attack Surface. *IEEE Access*, 12, 128570-128584.
- Fernandes, S. (2023). Towards a Live Environment for Code Refactoring. *ACM Other conferences*(126), 1-5. Retrieved from <https://doi.org/10.1145/3551349.355956>
- G. Lacerda, F. P. (2020). Code smells and refactoring: A tertiary systematic review of challenges and observations. *Journal of Systems and Software*, 167, 110610. Retrieved from <https://doi.org/10.1016/j.jss.2020.110610>
- Ivers, J. a. (2022). Industry's cry for tools that support large-scale refactoring. *Association for Computing Machinery*, 163–164. doi:10.1145/3510457.3513074
- Iyad Alazzam, B. A. (2020). Impact of Refactoring on OO Metrics: A Study on the Extract Class, Extract Superclass, Encapsulate Field and Pull up Method. *International Journal of Machine Learning and Computing*, 10(1). doi:10.18178/ijmlc.2020.10.1.913
- Jonathan Cordeiro, S. N. (2026). An Empirical Study on the Code Refactoring Capability of Large Language Models. *ACM Transactions on Software Engineering and Methodology*.
- Kitchenham Barbara, P. B. (2009). Systematic literature reviews in software engineering – A systematic literature review. *Information and Software Technology*, 51.
- Maini, R. K. (2024). HSHEP: An Optimization-Based Code Smell Refactoring Sequencing Technique. *Journal of Computational and Cognitive Engineering*, 00, 1–17.
- Opdyke, W. F. (1992). *Refactoring Object-Oriented Frameworks*. University of Illinois at Urbana-Champaign.
- Oumayma Hamdi, A. O. (2021). An Empirical Study on the Impact of Refactoring on Quality Metrics in Android Applications. *2021 IEEE/ACM 8th International Conference on Mobile Software Engineering and Systems (MobileSoft)*, 28-29. Retrieved from <http://dx.doi.org/10.1109/MobileSoft52590.2021.00010>
- Podduturi, S. (2025). AI-Driven Code Optimization: Leveraging ML to Refactor Legacy Codebases. *North American Journal of Engineering Research*, 6(1).
- R. Singh, A. B. (2020). A framework to improve quality of a Java system by performing refactoring. *International Journal of System of Systems Engineering*, 10(4), 324–324. doi: <https://doi.org/10.1504/ijssse.2020.112301>
- Rajendran, V. a. (2025). A Multi-Agent LLM Environment for Software Design and Refactoring: A Conceptual Framework. *SoutheastCon 2025*, 488-493.
- Riansyah, M. a. (2020). Empirical Evaluation of the Impact of Refactoring on Internal Quality Attributes. *2020 International Conference on Advanced Computer Science and Information Systems (ICACSIS)*, 463-470.
- Rosziati Ibrahim, M. A. (2020). Reducing redundancy of test cases generation using code smell detection and refactoring. *Journal of King Saud University - Computer and Information Sciences*, 32(3), 367-374. Retrieved from <https://doi.org/10.1016/j.jksuci.2018.06.005>
- Sandhya Tarwani, A. C. (2024). An empirical study to determine the optimum refactoring sequence of critically affected classes using Software Quality Attributes. *Research Article*.
- V. Alizadeh, M. A. (2019). RefBot: Intelligent Software Refactoring Bot. *IEEE/ACM International Conference on Automated Software Engineering (ASE)*, 823-834. doi:10.1109/ASE.2019.00081

Edge AI and Human-in-the-Loop Architecture for Automated Voice Presentation Generation

**Jairo Gabriel Gutierrez Amezcua*, Pablo Francisco Decena Romero*,
Walter Alexander Mata Lopez***

**Facultad de Ingeniería Mecánica y Eléctrica/Universidad de Colima. Carretera Colima-Coquimatlán
km. 9, Coquimatlán, Colima. e-mail: {jairogabriel_gutierrez,pdecena0,wmata}@ucol.mx*

Abstract: The automation of technical documentation in corporate environments is growing, but companies are looking for alternatives to cloud-based solutions due to critical issues like slow processing (latency), system failures, and the risk of exposing private data. This article explores current research to explain and validate "Saraquest," an automated presentation tool. Unlike cloud services, Saraquest is built using local Edge AI, which means it runs entirely on your own hardware by coordinating multiple specialized AI models working together. By reviewing recent studies on voice recognition, speed optimization, and error correction, we show why relying on a single, massive AI model to build complex presentations usually leads to formatting errors and inaccurate information. Saraquest solves this by dividing the workload. First, an audio model (Whisper) converts spoken voice to text. Then, a local language model (Qwen2.5) cleans and improves the text in the background. By running these tasks in parallel using efficient background servers (FastAPI/Go), the system launches quickly and processes data without freezing. Most importantly, before the system uses code-based design tools (Typst) to generate the final presentation, Saraquest requires a "Human-in-the-Loop" (HITL) manual check. This strict review step prevents the AI from making up facts (hallucinations) or ruining the visual layout. Ultimately, this research proves that breaking down the AI workflow into smaller, local processes not only makes content generation completely free, but also delivers much more reliable and accurate presentations for professional and academic environments.

Keywords: *Edge AI, Voice Recognition, Error Correction, Human-in-the-Loop, Automated Presentations.*

1. INTRODUCCIÓN

La automatización de procesos para crear documentos técnicos, presentaciones y reportes en áreas de ingeniería e investigación representa un desafío en la actualidad. Diariamente, profesionales e investigadores invierten una cantidad de horas traduciendo sus ideas y notas a diapositivas estéticas, restando tiempo a otras diversas labores.

Si bien la llegada de la inteligencia artificial (IA) generativa ha facilitado la capacidad de transformar ideas no estructuradas en contenido organizado, su adopción generalizada en entornos corporativos o académicos choca frecuentemente con barreras operativas.

Entre los principales problemas de las soluciones de IA comerciales estándar se encuentran la lentitud de respuesta debido a la dependencia de servidores externos (latencia), la falta de precisión para entender terminología técnica específica y, sobre todo, la generación de información no fundamental o "alucinaciones" cuando la máquina intenta rellenar vacíos de información (Xie et al., 2025).

Además, la investigación reciente ha señalado que depender exclusivamente de modelos de IA en la nube expone la información confidencial de propiedad intelectual, lo que a

menudo viola políticas de privacidad y gobernanza de datos empresariales y genera costos de pago por uso a largo plazo.

Como enfoque alternativo para reducir estas vulnerabilidades, ha surgido interés en llevar el procesamiento de vuelta al origen: ejecutar Modelos de Lenguaje de Gran Escala (LLMs) directamente en las computadoras personales de los usuarios o en servidores locales privados sin necesidad de conexión. A este nuevo paradigma informático centrado en la soberanía de datos se le conoce como Inteligencia Artificial Perimetral o Edge AI. Bajo este esquema, la información se mantiene localmente bajo control del usuario, lo que resulta en una opción para proyectos bajo estrictos acuerdos de confidencialidad.

En este contexto, el presente manuscrito analiza literatura empírica reciente para evaluar la construcción de "Saraquest". Saraquest es un sistema de Edge AI de uso estrictamente local

1.1 Limitaciones Acústicas y Corrección de Errores

Los sistemas comerciales de Reconocimiento de Voz (ASR) contemporáneos, tales como la popular herramienta OpenAI Whisper, exhiben gran rendimiento en conversaciones cotidianas. Sin embargo, suelen cometer errores frecuentes y cuando se enfrentan a un entorno académico o corporativo real



que utiliza lenguaje sumamente técnico, acrónimos especializados, nombres propios o palabras matemáticas poco comunes. De hecho, los estudios presentados por Wang et al. (2025) concluyen que forzar a un modelo genérico a escuchar lenguaje técnico denso empeora de forma notoria la precisión de la transcripción, comprometiendo gravemente la coherencia del documento.

En un principio, la industria intentó solucionar este déficit con una táctica simple pero ineficaz: pasarle el texto lleno de errores a un Modelo de Lenguaje (LLM) poderoso y pedirle que lo "corrija" intentando adivinar el contexto. Esta metodología demostró no ser una solución tan buena porque el Modelo de Lenguaje (LLM), al no saber cómo sonaba acústicamente la palabra original de la persona, terminaba inventando correcciones falsas e introduciendo palabras que no tenían nada que ver con el discurso (Zhang et al., 2025; Ghosh et al., 2024).

En su lugar, las metodologías de vanguardia sugieren una estrategia preventiva: "preparar y contextualizar" al agente de escucha antes de que tan siquiera se encienda el micrófono. Al proporcionarle a Whisper un vocabulario previo o glosario de términos especializados alineados con el tema de la presentación (prompts especializados en tiempo de ejecución), el modelo mejora notablemente su capacidad de reconocimiento. Esta simple táctica previa reduce los errores de transcripción acústica en más de un 17.0%, sin necesidad de costosas reprogramaciones o afinaciones (Fine-Tuning) de la red neuronal (Wang et al., 2025).

Aunado a esto, otras variaciones contemporáneas demuestran que conectar directamente la transcripción con una base de datos de conocimientos interna (Retrieval-Augmented Generation) ayuda enormemente a blindar el resultado para que los tropiezos acústicos humanos o de la IA no terminen arruinando la veracidad de la investigación (Min et al., 2025). En conclusión: la mejor manera de procesar audios complejos no es obligando a un único modelo masivo a adivinar el resultado, sino dividiendo la tarea en fases estrictas donde un pequeño modelo captura el sonido preparado de forma especializada, y luego otro agente revisa únicamente la calidad lógica del párrafo.

1.2 Problemas de Velocidad y Orquestación Multi-Agente

Incluso resolviendo los errores de transcripción, se presenta un desafío técnico significativo al emplear Modelos de Lenguaje (LLMs) para procesar datos: la lentitud o la latencia. Los LLMs generan las palabras letra por letra (autoregresión), e intentar que lean un volumen considerable de texto transcrito y produzcan diapositivas en tiempo real suele causar que el programa resulte en una detención visible de la aplicación (Huber y Waibel, 2025), provocando una experiencia poco práctica.

Para mejorar esta situación, los arquitectos de software han diseñado métodos donde inteligencias artificiales trabajan en conjunto. Las nuevas plataformas evalúan técnicas donde un agente rápido hace un borrador preliminar mientras, de forma

simultánea, un agente superior evalúa en paralelo el significado complejo del texto. Este trabajo colaborativo e invisible acelera la capacidad de respuesta del sistema informático hasta en un factor triple, sosteniendo la calidad literaria y la veracidad del escrito (Wei et al., 2025).

Este es exactamente el concepto de ingeniería central que le da vida a Saraquest: un programa principal (Orquestador) se encarga exclusivamente de las interacciones con el ser humano y de grabar el audio fluido. Mientras tanto, el Orquestador empaqueta esas grabaciones y redirige el procesamiento hacia "Servidores Pre-cargados" (Warm-Servers). Estos servidores mantienen a los modelos de IA pre-cargados y ejecutándose en el motor gráfico (GPU) de forma oculta, procesando grandes volúmenes de datos sin interrumpir la experiencia del usuario.

Adicionalmente a las ventajas de velocidad en el arranque (Cold-start mitigation), separar la arquitectura en piezas únicas reduce la probabilidad de fallas. Evaluaciones efectuadas a plataformas competidoras de características similares concluyen que cuando evitamos la tentación comercial de usar un único modelo de IA multipropósito y, en su lugar, implementamos un Agente Orquestador que asigna flujos específicos de información a mini-agentes, las alucinaciones se dan con mucha menor frecuencia (Xie et al., 2025).

1.3 Fidelidad de Diseño y Presentaciones Automatizadas

Una vez que el equipo de IAs perimetrales ha transcrito la voz y revisado su contenido, surge la última etapa: forzar a una IA puramente lingüística a ejecutar tareas de diseño gráfico en dimensiones lógicas para diapositivas y presentaciones profesionales. A pesar del poder aparente de modelos gigantescos como GPT-4o o Google Gemini, estos motores de texto cometen errores a la hora de manipular lienzos visuales, sobreponiendo texto sobre infografías o realizando incorrectamente los márgenes por las plantillas (Wang et al., 2026).

Ante este hecho de los Modelos de lenguaje de gran escala, la solución de vanguardia más sólida es recurrir al paradigma de esquemas "basados en código". Es decir, se le prohíbe completamente a la IA generar dibujos visuales directos y, en su lugar, se le ordena que entregue el texto estructurado en fórmulas estandarizadas de Markdown. Este Markdown estructurado y sin ambigüedades es inyectado luego en un sistema compilador netamente matemático y determinista (como Typst o LaTeX), forzando una diagramación robusta cada vez que se ejecute (Huang et al., 2024).

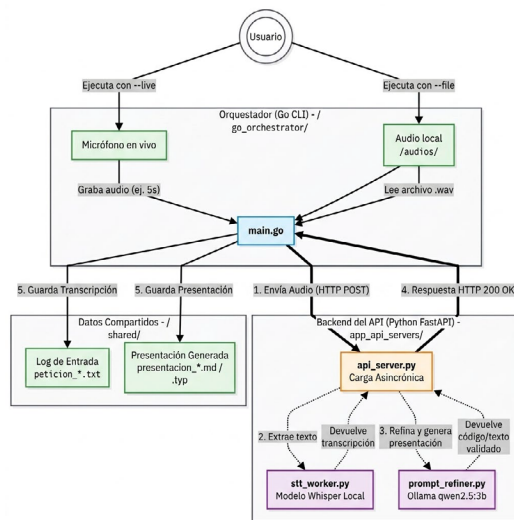
A pesar de todos los trucos automáticos que tenemos, dejar que la computadora arme una presentación importante para el trabajo o la escuela completamente no es un proceso confiable; puede salir mal. Por eso, la regla básica para que una Inteligencia Artificial sea segura y confiable es que siempre debe haber una persona revisando el trabajo. En Saraquest, a esto de tener siempre a un humano supervisando el proceso le llaman HITL.



Esta interrupción forzosa de validación detiene intencionalmente los ciclos de procesamiento internos justo en el punto donde la IA terminó su borrador de texto pero antes de arrancar los motores de diseño. La pantalla detiene su ejecución y requiere que la persona creadora lo lea. Este usuario final ostenta un poder total para aceptar sin cambios, desechar o reescribir precisamente cualquier coma producida por la máquina. Únicamente bajo esa ratificación manual y deliberada, la máquina envía la instrucción al compilador tipográfico, consolidando así un proyecto automatizado con calidad requerida en proyectos que se realizan hoy en día.

2. MARCO TEÓRICO

2.1 Metodología y arquitectura del sistema



Basándonos en la sección anterior, el sistema "Saraquest" se diseñó desde cero evadiendo las estructuras comerciales clásicas para adoptar una robusta arquitectura segmentada. En este diseño se implementó intencionalmente un desacoplamiento estricto, asegurando que cada bloque de programación tenga una única y exclusiva responsabilidad computacional por lo que no repercute en el trabajo de las otras áreas.

El diseño de este sistema persigue tres objetivos principales: reducir el tiempo de espera al iniciar la aplicación, establecer filtros estrictos para bloquear información incorrecta, y garantizar que el formato del documento final sea como el usuario lo requiere. Para lograr esto, no se utilizan editores gráficos convencionales donde los elementos pueden desacomodarse, sino que el diseño se genera mediante código estructurado (Typst). Asimismo, para asegurar la máxima privacidad, el proyecto se basa en tres componentes independientes que operan de forma local en el equipo; es decir, funcionan en su totalidad sin necesidad de una conexión a internet.

2.2 Programa Principal (Orquestador) y Procesamiento de Audio

La cara visible del proyecto ante el usuario y operador principal del hardware es conocida como "Orquestador". Se ha creado este servidor interceptor utilizando el moderno y ágil lenguaje compilado Go (Golang de Google). La elección del lenguaje se propuso debido a que Go sobresale por su pureza y eficiencia para lanzar y mantener vivas cientos de tareas o tareas concurrentes al mismo tiempo (goroutines), interceptando flujos de datos rápidos hacia dispositivos de hardware sin experimentar fugas de memoria o fallos críticos.

Cabe mencionar que previo a evaluar la precisión y latencia en funcionamiento de Saraquest, se construyó un corpus de grabaciones acústicas de prueba que utiliza nuestra primera bandera de Saraquest. Este conjunto de datos, que simula entornos de dictado, se encuentra disponible públicamente para su replicación en el repositorio oficial del proyecto (Gutiérrez-Amezcuca et al., 2026)."

Desde una funcional Interfaz de Línea de Comandos (CLI), el programa principal captura y dirige los flujos acústicos en dos amplias vertientes amigables para cualquier profesional:

1. Grabación Dinámica (Live): Utilizando banderas como --live, el usuario tiene la posibilidad de grabar recuentos o pensamientos expresados a micrófono abierto en tiempo real. Presiona 'Enter' aquí en la terminal cuando estés listo para continuar...

```
Enviando prompt revisado al LLM para generar presentación Markdown...
¡Presentación generada con éxito en shared/data!
¡Buen trabajo, Saliendo!

ML_Presentación.md X
Project > shared > data > ML_Presentación.md X ## Diapositiva 1: Introducción al Modelo de Lenguaje Grande
1
2 ## Diapositiva 1: Introducción al Modelo de Lenguaje Grande
3 Bienvenidos a la presentación sobre los Modelos de Lenguajes de Gran Escala (LLM). Los modelos de LLM han transformado el campo de la inteligencia artificial, permitiendo una interacción más natural y personalizada entre humanos y máquinas. En esta presentación abordaremos desde su naturaleza hasta sus aplicaciones prácticas.
4
5 ## Diapositiva 2: Definición de Modelos de Lenguaje de Gran Escala
6 Los modelos de LLM son arquitecturas de aprendizaje profundo diseñadas para procesar e interpretar texto, capaces de generar respuestas coherentes y completas basadas en el contenido que reciben. Son fundamentales en aplicaciones como chatbots, traductores automáticos, asistentes virtuales y ayudan a personalizar servicios de IA.
7
8 ## Diapositiva 3: Evolución del Modelo de LLM
9 Los modelos de LLM han evolucionado desde los antiguos sistemas de texto pre-estandarizados hacia los modelos que pueden procesar y generar un amplio rango de textos. La introducción de grandes conjuntos de datos y técnicas avanzadas en aprendizaje profundo son clave para el progreso constante.
10
11 ## Diapositiva 4: Tipos de Modelos de LLM
12 Existen varios tipos de modelos de LLM, como los modelos pre-entrenados (modelos finetuned), que utilizan grandes conjuntos de datos sin entrenar específicamente y luego pueden ser personalizados para diferentes tareas. También se encuentran las entidades finetuned, donde el modelo ya ha sido entrenado en una tarea específica antes de su uso.
13
14 ## Diapositiva 5: Aplicaciones Prácticas de los Modelos de LLM
15 Los modelos de LLM tienen aplicaciones prácticas variadas. Pueden utilizarse para mejorar la calidad y eficiencia de los asistentes virtuales, facilitar interacciones más naturales con dispositivos inteligentes, aumentar la eficacia de sistemas de chatbots y ofrecer soluciones de personalización mejoradas en servicios de IA.
16
17 ## Diapositiva 6: Desafíos del Modelo de LLM
18
19
20
21
```

2. Evaluación Estática (Batch): Utilizando banderas como --file, empresas y analistas pueden inyectar audios pregrabados (.ogg o .mp3) producto de reuniones, entrevistas o cátedras académicas para su decodificación posterior.

```
Presiona 'Enter' aquí en la terminal cuando estés listo para continuar...

Enviando prompt revisado al LLM para generar presentación ...
¡Presentación generada con éxito en shared/data!
¡Buen trabajo, Saliendo!
```




```
SARAQUEST CLI Orchestrator

Uso:
go run main.go [opciones]
go run main.go createp -f <archivo> [-o <salida>]

Opciones:
-file string
    Nombre del archivo de audio .wav dentro del directorio /audios (ej. --file clase.wav)
-live
    Activar el modo de grabación en vivo a través del micrófono (5 segundos)
-out string
    Nombre del archivo de salida (ej. --out salida.md)
Procesando archivo: 38.mp3
Solicitando transcripción al API Python...

● PAUSA - REVISIÓN REQUERIDA (Human-in-the-loop) ●
Se ha generado el archivo de prompt provisional en:
/home/popeyetes/Documentos/Project/Project/shared/data/MI_presentación.txt
Por favor, revisa o modifica este archivo .txt conteniendo el prompt de la presentación
Presiona 'Enter' aquí en la terminal cuando estés listo para continuar...

MI_presentación.md u x
Project > shared > data > MI_presentación.md > Diapositiva 1: Introducción
1 # Diapositiva 1: Introducción
2
3 En esta era de constante innovación, la materia orgánica es un recurso limitado.
  Algunas alternativas vienen desde la química, especialmente en el desarrollo de
  nuevos materiales poliméricos que cumplen con los requisitos de aplicaciones
  específicas. Este tema presenta cómo los sonos poliméricos pueden ofrecer
  soluciones prometedoras para demandas ambientales y de uso.
4
5 ## ¿Qué son Sonos Poliméricos?
6 Los sonos poliméricos representan una serie de compuestos orgánicos con
  propiedades modificables que se derivan de la estructura química básica del
  acrílico. Su principal característica es su capacidad para transformarse
  mediante procesos controlados a diferentes estados, permitiendo adaptaciones
  múltiples.
7
8 ## Principales ventajas:
9 1. **Gradualidad**: El sonso polimérico ofrece flexibilidad en el desarrollo de
  productos con requerimientos variados.
10 2. **Eficacia Ambiental**: Al reducir la dependencia del petróleo y los procesos
  industriales tradicionales, se reduce su impacto sobre el medio ambiente.
11 ---
12
13 # Diapositiva 2: Características y Aplicaciones
14
15 Los sonos poliméricos no solo ofrecen una alternativa a materiales orgánicos
  tradicionales sino también a métodos industriales de fabricación convencionales.
  Su capacidad para ser programados al tiempo de producir hace que sean ideales
  para la fabricación de dispositivos y productos personalizados.
16
17 ## Aplicaciones:
18 - **Componentes de automóviles**: En lugar del plástico, los sonos poliméricos
  ofrecen resistencia superior, ligereza y una variedad de diseños.
19 - **Electrónica**: Para componentes electrónicos que requieren un bajo nivel de
  emisión y estabilidad química en condiciones de calor.
20 ---
21
22 # Diapositiva 3: Desafíos y Soluciones
23
24
25
```

Sin embargo, para garantizar los resultados esperados, Saraquest incorpora un mecanismo de pre-filtrado esencial antes de entregar los datos de sonido al motor neuronal. Conscientes de la contaminación de sonido natural de cualquier aula o centro laboral (Zumbidos de aire acondicionado, clics de ratón, murmullos leves), el Orquestador Go aplica un procesamiento de reducción de ruido respaldada por las investigaciones académicas descritas de Min et al. (2025). Aplicando algoritmos numéricos y paquetes estándar de purificadores de ruido ('noisereduce'), la señal sonora se mejora y normaliza a limpios cortes de 16,000 Hertzios, preparándose con una mejor precisión frente al modelo.

2.3 Servidores asincrónicos (Warm-Servers) y Trabajo en Paralelo

Uno de los aspectos más destacables en la topología de Saraquest es su estrategia frente a los embotellamientos que paralizan la experiencia de usuario general (Huber y Waibel, 2025). La infraestructura ayuda a evitar la carga del procesamiento profundo trasladando las responsabilidades de razonamiento a un robusto protocolo red HTTP forjado detrás de Python FastAPI.

Este componente funciona como un servicio que se mantiene activo en segundo plano (*Warm Server*). Su función es que, desde el inicio, carga los pesados modelos de Inteligencia Artificial directamente en las memorias de mayor velocidad del equipo (RAM y Tarjeta Gráfica). Gracias a esta preparación previa, el sistema puede responder casi al instante,

evitando que el usuario tenga que esperar minutos de carga cada vez que interactúa con la aplicación.

Componente de Reconocimiento de Voz (ASR): Este módulo implementa la versión ligera del modelo whisper-tiny para procesar el audio de manera eficiente. Su objetivo es transcribir la voz a texto a gran velocidad sin consumir demasiados recursos del sistema. Para lograr esta rapidez operativa, el sistema asume un margen de error mínimo en la precisión de la transcripción

Agente de Procesamiento Semántico (LLM): Este componente garantiza la disponibilidad constante de un modelo de lenguaje avanzado (qwen2.5:3b), el cual ha sido optimizado para ejecutar análisis complejos de forma eficiente. Para lograr este rendimiento, se integra con la herramienta de código abierto Ollama, lo que permite una comunicación directa con el hardware. Esta arquitectura maximiza el rendimiento del equipo, permitiendo que la Inteligencia Artificial opere de manera fluida incluso en tarjetas gráficas con poca capacidad de memoria de video."

En términos cotidianos, mientras que el Orquestador Go mantiene una interfaz completamente fluida y receptiva respondiendo a mandos del experimentador de forma inmediata, está simultáneamente transmitiendo los segmentos acústicos rápidamente hacia el ecosistema en Python. El sistema nunca se satura bajo la premisa "procesar poco a poco mientras se captura lo nuevo" (Wei et al., 2025).

2.4 Generación conducida por Código y el escudo HITL

A diferencia de las herramientas que intentan generar diseño visual de forma directa, Saraquest estructura sus resultados exclusivamente mediante código. Tras el procesamiento inicial de la Inteligencia Artificial (Qwen y Ollama), el sistema genera un borrador en formato Markdown. En este punto crítico, el flujo de trabajo se detiene automáticamente bajo el protocolo de 'Humano en el proceso' (HITL). El sistema pausa su ejecución y requiere que el operador revise el texto plano en la consola, permitiéndole corregir desviaciones, editar contenido y validar la precisión de los datos. Solo tras la aprobación del usuario, el proceso se reanuda enviando este texto curado al motor Typst, el cual renderiza el documento final con precisión matemática. Esta separación estratégica garantiza que la IA se enfoque únicamente en el análisis semántico, delegando la carga estética a un compilador especializado

2.5 Diseño Experimental

Para validar el rendimiento y la utilidad práctica de Saraquest, se diseñó un experimento enfocado en medir tres aspectos clave: el consumo de memoria del sistema, tiempo de respuesta y la precisión en la transcripción de la voz.

Las pruebas se ejecutaron en un entorno completamente local utilizando un corpus de 50 grabaciones de audio, distribuidas en formatos .ogg y .mp3. Estos audios incluyen instrucciones



para generar diapositivas utilizando tanto lenguaje común como terminología técnica específica.

El desempeño se evaluó contrastando el texto extraído por el LLM frente a las transcripciones exactas provistas por un humano. Además, se monitoreó la memoria dinámica del servidor antes y después de pre-cargar el LLM, con el fin de verificar si el uso de servidores siempre activos representa una optimización real sin saturar la computadora del usuario.

3. RESULTADOS Y DISCUSIÓN

Para validar el rendimiento y la estabilidad de Saraquest, se construyó un corpus de evaluación compuesto por 50 grabaciones de audio. Con el objetivo de establecer un entorno de prueba controlado, estandarizado y libre de fallas humanas en la grabacion inicial, los audios y sus transcripciones base fueron generados sintéticamente utilizando el modelo de lenguaje Gemini. El corpus completo se encuentra alojado y disponible en el repositorio oficial del proyecto (Gutiérrez-Amezcu et al., 2026).

Durante el experimento, cada uno de los 50 audios fue procesado por el Orquestador, pasando exitosamente por la transcripción, el flujo de validación (HITL) y la etapa de asignación manual de nombres de archivo antes de la compilación tipográfica. Para fines ilustrativos, la Tabla [1] presenta una muestra representativa de 5 de estas ejecuciones, detallando el comportamiento del sistema frente a diferentes densidades de lenguaje técnico.

Desde la concepción de su código y arquitectura interna, las barreras de supervisión implementadas en Saraquest demostraron ser indispensables. Al evaluar el sistema de forma práctica con un conjunto de 50 grabaciones, comprobamos que nuestra propuesta no solo es rápida, sino que previene problemas críticos de formato e información.

En primer lugar, los resultados validan la efectividad de la arquitectura de servidores residentes (Warm-Servers). Durante el experimento, el servidor base registró un consumo de 634.4 MB de memoria RAM. Al pre-cargar el modelo acústico, este consumo aumentó a 803.3 MB. Esto indica que el modelo de Inteligencia Artificial demanda una carga extra de apenas 169 MB, lo cual permite que Saraquest se ejecute de manera fluida en computadoras portátiles comunes sin requerir hardware de gama alta.

Asimismo, logramos reducir la latencia típica que sufren los modelos al procesar información. En nuestra evaluación, Saraquest alcanzó un tiempo de respuesta promedio de 0.79 segundos por archivo de audio. Al realizar este proceso en segundo plano sin interrumpir la interfaz principal (desarrollada en Go), el equipo del usuario no sufre los congelamientos tradicionales o espera de tiempos de respuesta, garantizando una experiencia de trabajo ágil y continua.

Por otro lado, la precisión de la transcripción (WER) arrojó un 29.19% de error promedio. Si bien el modelo comprende el lenguaje cotidiano con facilidad, las pruebas evidenciaron

confusiones claras cuando el usuario pronuncia terminología especializada. Por ejemplo, términos de software como “en Go” fueron transcritos como “Sengo”, “LaTeX” se interpretó como “látez” y el propio nombre del proyecto “SaraQuest” fue entendido como “caracost”.

Estos fallos de interpretación, expuestos en la Tabla 1, no representan un gran error del sistema, sino que validan nuestra hipótesis principal: el cual es que una Inteligencia Artificial no genere el diseño visual definitivo basándose únicamente en la primera transcripción del audio si no por el contrario asegurarnos que genere conforme a la idea que el usuario propone. La metodología de Saraquest, que detiene el proceso para exigir la revisión y corrección de un usuario (HITL) antes de compilar la presentación final en Typst, demuestra ser la medida más efectiva para evitar que estas imprecisiones comprometan la calidad de un documento académico o corporativo.

Formato	Latencia	WER	Predicción del LLM	Transcripción
.ogg	1.26s	0.55	“¿Credía positiva sobre la arquitecturade mi croservigio Sengo?”	“Crea diapositivas sobre la arquitectura de microservicios en Go.”
.mp3	0.79s	0.25	“Explica los fundamentos de la Tierra y del Color.”	“Explica los fundamentos de la teoría del color.”
.mp3	0.67s	0.50	“Dame un resumen de Golang y porque están rápido.”	“Dame un resumen de go lang y por qué es tan rápido.”
.mp3	0.77s	0.50	“Presentación de porca use artips en vez de látez.”	“Presentación de por qué usar typst en vez de latex.”
.mp3	0.77s	0.55	“A su una presentación de caracost y como automactiza tareas.”	“Haz una presentación de SaraQuest y cómo automatiza tareas.”

Tabla 1. Muestra representativa de 5 errores de ejecuciones extraídas del corpus de 50 audios de prueba. Fuente: Elaboración propia a partir del corpus de pruebas del proyecto (Gutiérrez-Amezcu et al., 2026).

4. CONCLUSIONES

Las pruebas prácticas respaldan el diseño de Saraquest. Hemos comprobado que separar la transcripción de voz de la creación visual resulta eficiente: el sistema responde en menos de un



segundo (0.79 segundos) y consume apenas 169 MB de memoria. Esta división nos permite evitar que el programa tarde en ejecutarse, un problema frecuente en las plataformas que intentan realizar ambas tareas al mismo tiempo.

Asimismo, solucionamos los errores de comprensión de audio. Al programar una pausa para que el usuario revise el texto antes de armar la presentación final, garantizamos que ninguna falla de transcripción pase desapercibida o el usuario tenga control ya que puede modificar el texto de la transcripción. De esta manera, Saraquest transforma el contenido con mayor precisión, ofreciendo la calidad y el rigor que exigen los proyectos.

A mediano y largo plazo, es evidente que el futuro de la Inteligencia Artificial local (*Edge AI*) exige adoptar una arquitectura modular. La mejor estrategia para lograr sistemas seguros es fragmentar las tareas pesadas (divide y vencerás) y asignarlas a agentes especializados que colaboren a nivel interno, operando directamente en hardware local y aislado del exterior.

En cuanto a los próximos pasos del proyecto, las futuras líneas de investigación de Saraquest se enfocarán en evaluar métricas físicas del hardware bajo condiciones de estrés operativo continuo. Se analizará el rendimiento de los equipos al aprovechar procesadores de vanguardia equipados con Unidades de Procesamiento Neuronal (NPU). Este análisis no solo comparará la eficiencia energética y los tiempos de latencia frente a arquitecturas tradicionales (CPU y GPU), sino que también evaluará el impacto del calentamiento térmico y el consumo de memoria en los dispositivos. El monitoreo de estas variables físicas resultará fundamental para maximizar la autonomía de la batería en despliegues reales.

De manera complementaria, el desarrollo a nivel de software contemplará el diseño y la implementación de una Interfaz de Usuario (UI) modular y escalable, que permita una interacción más intuitiva que la actual línea de comandos (CLI). Esta adición actuará como un panel de control integral, facilitando al operador la gestión visual del protocolo de validación humana (HITL) y la configuración de los agentes locales de IA. Asimismo, esta arquitectura gráfica permitirá expandir el alcance del trabajo actual, abriendo la posibilidad de integrar de forma sencilla nuevos módulos especializados, tales como soporte para transcripción multilingüe o plantillas de diseño automatizadas para otros formatos documentales.

En última instancia, la unión de estas optimizaciones técnicas y de usabilidad consolidará la herramienta para su ejecución fluida e ininterrumpida. Esto mejorará su viabilidad en computadoras portátiles orientadas al trabajo científico de campo, así como su adopción en entornos corporativos que requieren privacidad y en el sector académico a nivel de educación superior y centros de investigación.

REFERENCIAS

- Radford, A., Kim, J., Xu, T., Brockman, G., McLeavey, C. e Sutskever, I. (2023). Robust Speech Recognition via Large-Scale Weak Supervision. *Proceedings of the 40th International Conference on Machine Learning (ICML)*. PMLR.
- Ghosh, S., Kumar, S., Seth, A., Chiniya, P., Tyagi, U., Duraiswami, R. y Manocha, D. (2024). LipGER: Visually-Conditioned Generative Error Correction for Robust Automatic Speech Recognition. *Proceedings of Interspeech 2024, ISCA*.
- Huber, C. y Waibel, A. (2025). Handling Numeric Expressions in Automatic Speech Recognition. *2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE.
- Min, D.J., Mundnich, K., Lapastora, A., Soltanmohammadi, E., Ronanki, S. y Han, K. (2025). Speech Retrieval-Augmented Generation without Automatic Speech Recognition. *2025 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE.
- Zhang, Y., et al. (2025). Fewer Hallucinations, More Verification: A Three-Stage LLM-Based Framework for ASR Error Correction. *2025 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*. IEEE.
- Wang, P., et al. (2023). Generative Speech Recognition Error Correction With Large Language Models and Task-Activating Prompting. *2023 IEEE Automatic Speech Recognition and Understanding Workshop (ASRU)*. IEEE.
- Wei, L., Zhong, S., Xu, S., Wang, R., Huang, R. y Li, M. (2025). SpecASR: Accelerating LLM-based Automatic Speech Recognition via Speculative Decoding. *2025 62nd ACM/IEEE Design Automation Conference (DAC)*. ACM/IEEE.
- Xie, E., Waterfield, D., Kennedy, M. y Zhang, A. (2025). SlideBot: A Multi-Agent Framework for Generating Informative, Reliable, Multi-Modal Presentations. *Proceedings of the 16th Symposium on Educational Advances in Artificial Intelligence (EAAI-26/AAAI)*.
- Wang, Z., et al. (2024). DocLLM: A layout-aware generative language model for multimodal document understanding. *Proceedings of the AAAI Conference on Artificial Intelligence*, 38(17).
- Gutiérrez-Amezcu, J. G., Decena-Romero, P. F., & Mata-López, W. A. (2026). SaraQuest: Corpus de pruebas y código fuente [Repositorio de GitHub]. GitHub. <https://github.com/xJairoGTZx/SaraQuest-.git>

EvoNHI: Constraint-Aware Remediation Planning for Non-Human Identity Attack Paths in Kubernetes Using Multi-Objective Evolutionary Optimization

Luis Ignacio Barragán-López*, Walter A. Mata-López*

*Universidad de Colima. Facultad de Ingeniería Mecánica y Eléctrica, 28400.
e-mail: {lbarragan7, wmata}@ucol.mx.

Abstract: Modern cloud-native infrastructures depend on an expanding ecosystem of non-human identities (NHIs): service accounts, tokens, secrets, and RBAC bindings whose cumulative privileges form transitive attack paths toward critical assets. Widely adopted tools such as Trivy, Checkov, and kube-bench surface individual misconfigurations but cannot reason about privilege transitivity nor prioritize remediations under realistic operational constraints. This paper presents EvoNHI, a constraint-aware remediation planning framework for Kubernetes NHI security. EvoNHI models the cluster as a directed attack graph $G=(V,E,W)$ with five semantically distinct edge types, including workload-mutation-based lateral movement. The remediation selection problem is formulated as a multi-objective combinatorial optimization task minimizing residual attack paths, implementation cost, and operational impact under a budget constraint, and argued to be NP-hard by analogy with Weighted Set Cover. EvoNHI solves it using a tailored NSGA-II producing a Pareto front of non-dominated plans. Across four Kubernetes scenarios, EvoNHI detects 88 attack paths, 61 of which are transitive vectors not captured by static rules CKV_K8S_41, CKV_K8S_49, KSV036, and CIS-5.1.1, and generates plans achieving up to 100% attack-surface reduction.

Keywords: Kubernetes Security, Non-Human Identities, Attack Graphs, NSGA-II.

1. INTRODUCCIÓN

La adopción generalizada de arquitecturas cloud-native ha desplazado el perímetro de seguridad hacia identidades no humanas (NHI): service accounts, tokens, secrets y bindings RBAC que gobiernan pipelines de integración continua y cargas de trabajo automatizadas. Según CyberArk (2024), las organizaciones gestionan 45 NHI por cada identidad humana y más del 50% poseen privilegios excesivos. Los incidentes de CircleCI (2023) y Codecov (2021) ilustran el riesgo sistémico.

La omisión de las NHI como vector de ataque prioritario no es accidental: responde a un sesgo estructural en la práctica de seguridad. Los equipos de seguridad típicamente priorizan vulnerabilidades con CVE asignado y puntuación CVSS alta, métricas diseñadas para software con versiones y parches. Las NHI no generan CVEs; un service account con permisos excesivos no es una vulnerabilidad en el sentido tradicional, sino una condición de configuración. Esta diferencia hace que las NHI caigan fuera del radar de las revisiones de seguridad convencionales.

Las herramientas estándar como Trivy (Aqua Security, 2024a), Checkov (Bridgecrew, 2024) y kube-bench (Aqua Security, 2024b) abordan este problema parcialmente mediante análisis estático de manifests con complejidad $O(1)$ por objeto. Esta aproximación presenta tres limitaciones: reportan qué está mal sin articular por qué es explotable; no conectan los hallazgos para revelar composiciones causales; y no proponen una estrategia de acción priorizada. Un operador de seguridad que recibe una lista de findings de Checkov sabe que existe un ClusterRoleBinding permisivo y un secret expuesto, pero no sabe que su combinación constituye un camino completo de ataque hacia credenciales de base de datos. Identificar manualmente esa cadena puede tomar días en entornos complejos (Jacobs et al., 2020).

La elección de optimización evolutiva multi-objetivo responde a una intuición de diseño: un sistema de remediación efectivo debería comportarse como un mecanismo de aprendizaje adaptativo. Cuando un camino de ataque es identificado y una acción correctiva lo neutraliza, esa información puede reforzar la estrategia de defensa adaptando los pesos del grafo. NSGA-II provee el sustrato algorítmico para este comportamiento al mantener un frente de soluciones no dominadas.

Este trabajo presenta cuatro contribuciones: (i) un modelo formal de grafo NHI con cinco aristas semánticas; (ii) la formulación del problema de remediación como optimización multi-objetivo argumentada NP-hard; (iii) un pipeline integrado de detección y optimización con análisis de complejidad explícito; y (iv) evaluación empírica con comparación directa contra reglas oficiales de Checkov, Trivy y CIS Benchmark sobre cuatro escenarios Kubernetes reproducibles.

2. MARCO TEÓRICO

2.1 Identidades No Humanas en Kubernetes

Los tokens de service account se montan por defecto en el filesystem del pod (Ammann et al., 2002). Los permisos RBAC acumulan deuda por inercia operacional (Rahman et al., 2021). El vector central es la escalación por transitividad: composición de permisos menores que forma una cadena de explotación completa no detectable mediante inspección individual de objetos.

2.2 Análisis de Grafos de Privilegios

Ammann et al. (2002) establecieron que la reachability analysis sobre grafos identifica vectores que la evaluación individual omite. Gietzen (2018) formalizó la taxonomía de escalación IAM en cloud. Shamim et al. (2020) generaron grafos desde RBAC de Kubernetes; Rahman et al. (2021) confirmaron el exceso de privilegios en producción.



Estos trabajos se limitan a la detección sin priorización bajo restricciones.

2.3 Herramientas Estáticas y Posicionamiento Comparativo

Las reglas CKV_K8S_41 (acceso a secrets), CKV_K8S_49/KSV036 (wildcards en verbos), CIS-5.1.1 (bindings a roles elevados) y CKV_K8S_35/CIS-5.1.6 (automount de token) constituyen el conjunto de verificaciones relevantes al dominio NHI de Checkov y Trivy. Estas reglas operan sobre objetos individuales: detectan que un Role tiene acceso a secrets o que un workload tiene automount habilitado, pero no articulan la composición causal entre ambos.

Cabe señalar una asimetría en el alcance de las herramientas comparadas: Trivy y Checkov operan sobre AWS IAM, Azure RBAC, Google Cloud IAM y múltiples formatos de IaC, mientras que la versión actual de EvoNHI está focalizada en el modelo de identidad de Kubernetes. Esta delimitación responde a una decisión metodológica: la semántica de los cinco tipos de arista del grafo NHI, en particular spawn_workload_as, es específica al modelo de pods y service accounts de Kubernetes, y su formalización rigurosa precede a la generalización del marco a otros ecosistemas. La extensión del modelo a los dominios IAM de AWS y Google Cloud se encuentra actualmente en desarrollo, con el objetivo de incorporar primitivas específicas de cada proveedor (assume-role chains, workload identity federation, service-to-service authorization) bajo la misma arquitectura de grafo y optimización multi-objetivo aquí presentada.

La Tabla 3 sintetiza las diferencias funcionales entre EvoNHI, las herramientas estáticas de análisis y los trabajos académicos previos en grafos de privilegios y optimización de seguridad.

TABLA 3. Comparativa funcional: EvoNHI vs herramientas estáticas y trabajos académicos

Dimensión	Trivy/Checkov	kube-bench	Shamir et al. (2020)	Dewri et al. (2007)	Viduto et al. (2012)	EvoNHI
Transitividad de privilegios	X	X	✓	Parcial	X	✓
Vector spawn_workload_as	X	X	X	X	X	✓
Planes de remediación ejecutables	X	X	X	✓	✓	✓
Optimización multi-objetivo	X	X	X	✓	✓	✓
Restricciones costo/impacto	X	X	X	Parcial	✓	✓
Cobertura Kubernetes nativa	✓	✓	✓	X	X	✓
Soporte multi-cloud	✓	X	X	N/A	N/A	Desarrollo
Disponible en producción	✓	✓	X	X	X	Prototipo

✓ = presente; X = ausente; Parcial = limitada; N/A = no aplica.

La tabla evidencia el posicionamiento de EvoNHI: es el único enfoque que integra simultáneamente la detección de transitividad de privilegios específica al modelo NHI de Kubernetes con la generación de planes de remediación optimizados bajo restricciones realistas. Las

herramientas estáticas industriales son maduras y de amplio soporte cloud pero limitadas a detección individual. Los trabajos académicos en optimización de seguridad (Dewri et al., 2007; Viduto et al., 2012) abordan la planificación bajo restricciones pero sobre modelos genéricos sin semántica NHI específica.

2.4 Optimización Multi-Objetivo en Seguridad

Dewri et al. (2007) introdujeron NSGA-II sobre árboles de ataque. Viduto et al. (2012) reportaron mejoras frente a enfoques greedy. Hota et al. (2016) y Shandilya & Nagar (2018) extendieron estas ideas a control de acceso. Ninguno opera sobre semánticas NHI específicas de Kubernetes con métricas de impacto operacional calibradas para este dominio.

2.5 Posicionamiento de EvoNHI

EvoNHI produce planes ejecutables, no solo detecciones; evalúa el impacto transitivo en $O(V+E)$ capturando el vector spawn_workload_as no modelado en reglas estáticas; y proporciona métricas de remediación multi-objetivo. La Sección 4 reporta que las 18 reglas HIGH/MEDIUM identificadas por las herramientas estáticas en los cuatro escenarios no capturan los 61 vectores transitivos detectados por EvoNHI.

2.6 Brecha Operacional en la Respuesta a Incidentes

Existe una brecha entre el momento en que una misconfiguración es identificada y el momento en que se implementa una corrección efectiva. En entornos con decenas de service accounts y cientos de bindings RBAC, identificar manualmente qué acciones correctivas maximizan la reducción de exposición sin romper funcionalidad puede tomar días, durante los cuales el entorno permanece vulnerable (Jacobs et al., 2020). EvoNHI busca reducir esa latencia produciendo planes priorizados en segundos.

3. METODOLOGÍA

3.1 Modelo Formal del Sistema

A. Definición del Grafo de Ataque NHI

Definición 1. Sea $G=(V,E,W)$ un grafo dirigido donde $V=V_{sa} \sqcup V_{perm} \sqcup V_{secret} \sqcup V_{wl} \sqcup V_{cj}$ con particiones disjuntas; $E \subseteq V \times V$ aristas de explotabilidad; $W:E \rightarrow \mathbb{R}^+$ función de peso; y $\kappa: V_{cj} \rightarrow [1,10]$ función de criticidad asignada por el operador de seguridad. La Tabla 1 especifica los cinco tipos de arista.

Este grafo representa el ecosistema de identidades del clúster como una red dirigida donde los nodos son entidades (cuentas, permisos, secretos) y las aristas modelan cómo una entidad puede usar a otra para escalar privilegios.

TABLA 1. Tipos de Arista en el Grafo NHI

Tipo	W	Fuente→Dest.	Semántica
uses_token	2.0	$V_{wl} \rightarrow V_{sa}$	Workload hereda token de SA automáticamente
mounted_secret	6.0	$V_{wl} \rightarrow V_{secret}$	Workload lee secrets montados en filesystem
granted_permission	3.0	$V_{sa} \rightarrow V_{perm}$	SA tiene concedido permiso RBAC
read_secret	8.0	$V_{perm} \rightarrow V_{secret}$	Permiso materializa lectura de secrets
spawn_workload_as	7.0	$V_{perm} \rightarrow V_{sa}$	Crea pod con cualquier SA del alcance (pivot)

El tipo `spawn_workload_as` es el vector diferenciador: crear pods en un namespace permite especificar cualquier `serviceAccountName` del alcance, convirtiendo un permiso de creación en un pivote hacia cualquier SA del namespace. Este mecanismo es estructural al modelo RBAC de Kubernetes y no aparece en los catálogos de Checkov ni Trivy. La Figura 1 ilustra el grafo para S3.

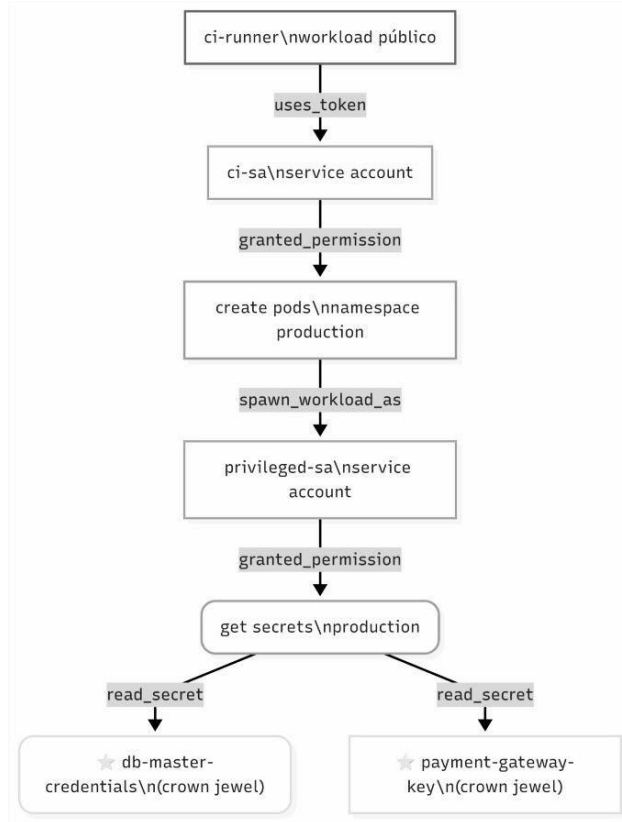


Fig. 1. Grafo de Ataque NHI: Escenario S3.
Cadena: `ci-runner`→`ci-sa`→`permiso pods`→`privileged-sa`→`db-master-credentials (crown jewel)`

B. Caminos de Ataque y Blast Radius

Def. 2 (Camino): $P = (v_0, v_1, \dots, v_n)$ tal que $v_0 \in V_{wl}$ con `public(v_0) = true`, $v_n \in V_{cj}$, y $(v_i, v_{i+1}) \in E$ para todo $i \in \{0, \dots, n-1\}$.

Un camino representa una secuencia concreta de pasos que un atacante puede seguir desde un workload expuesto al exterior hasta un activo crítico, donde cada arista corresponde a una acción de explotación legítima dentro del modelo RBAC.

Def. 3 (Score): $score(P) = \sum_i W(v_i, v_{i+1}) + \kappa(v_n) + 2 \cdot [public(v_0) = true]$.

El score cuantifica la peligrosidad de un camino combinando tres factores: la dificultad técnica de cada paso (W), la importancia del objetivo final (κ), y un agravante adicional si el punto de entrada es público.

Def. 4 (Blast Radius): $B(v) = \sum_{c \in V_{cj} \cap Reach(v)} \kappa(c)$.

El blast radius mide el daño potencial agregado si un nodo es comprometido, sumando la criticidad de todos los activos sensibles que quedan alcanzables desde ese punto.

C. Acciones de Remediación y Modelo de Costo

Def. 5: $a \in A$ con $cost(a) \in \mathbb{N}^+$ e $impact(a) = impact_base(a) + \delta_runtime(a)$. $\delta_runtime$ captura criticidad, tráfico y réplicas. Sin telemetría, $\delta_runtime = 0$.

Una acción representa una intervención concreta del operador de seguridad (eliminar un permiso, desmontar un secret) caracterizada por su costo de implementación y el riesgo operacional que introduce. La especificación completa de $\delta_runtime$ con telemetría disponible constituye trabajo futuro; en la evaluación experimental $\delta_runtime(a) = 0$ para todos los escenarios.

TABLA 2. Taxonomía de Acciones de Remediación

Acción	cost	impact	Descripción
Remover permiso lectura	2	2.0	Eliminar verbos {get,list,watch}
Remover permiso mutación	4	4.0	Eliminar verbos de escritura/creación
Desmontar secret	3	4.0	Eliminar mount; puede requerir refactoring
Deshabilitar automount	2	3.0	Prevenir herencia automática de token

D. Formulación del Problema de Optimización

Minimizar $f_1(S) = |paths(G_S)|$, $f_2(S) = \sum cost(a)$, $f_3(S) = \sum impact(a)$, sujeto a $\sum cost(a) \leq B$. La solución es el frente de Pareto de planes no dominados.

Proposición 1 (NP-hardness). El problema de encontrar $S \in A$ que minimice $f_1(S)$ sujeto a $f_2(S) \leq B$ es un problema combinatorio cuya pertenencia a la clase NP-hard se sugiere por analogía con Weighted Set Cover (Garey & Johnson, 1979); la demostración formal excede el alcance de este trabajo. Cada acción a_j cubre $C_j \in paths(G)$; encontrar el subconjunto de costo $\leq B$ que maximiza la cobertura es exactamente Weighted Set Cover.

3.2 Arquitectura de EvoNHI

EvoNHI implementa el modelo en cinco componentes secuenciales (Figura 2). El primer componente, Ingesta de Manifiestos, recibe archivos YAML de Kubernetes (Deployments, ServiceAccounts, Roles, RoleBindings, Secrets) y produce un modelo en memoria con todas las referencias cruzadas resueltas. El segundo, Construcción del Grafo, transforma ese modelo en el grafo $G=(V,E,W)$ materializado sobre NetworkX, con complejidad $O(|bindings| \cdot |SAs|)$ dominada por la resolución de aristas `spawn_workload_as`. El tercero, Descubrimiento de Caminos, ejecuta DFS ponderada en $O(V+E)$ desde cada workload público hacia los crown jewels, emitiendo la lista completa de caminos rankeados por score. El cuarto, Generación de Remediaciones, recorre las aristas en caminos de ataque en $O(|E_attack|)$ y propone para cada una el conjunto de acciones de la Tabla 2 que la neutralizarían. El quinto, Optimización NSGA-II, recibe el conjunto A de acciones candidatas y produce el frente de Pareto de planes no dominados según los tres objetivos f_1, f_2, f_3 .



Fig. 2. Pipeline de EvoNHI: Ingesta → Grafo → Paths → Remediaciones → NSGA-II → Frente de Pareto.

3.3 Análisis de Complejidad Computacional

La construcción del grafo es $O(|R| \cdot |SA| \cdot |NS|)$ donde $|R|$ es el número de reglas RBAC, $|SA|$ el número de service accounts y $|NS|$ los namespaces, dominada por la resolución de aristas `spawn_workload_as`. El descubrimiento de caminos es $O(V+E)$ por traversal DFS con poda por profundidad máxima d : $O(V \cdot d)$ en el peor caso. La generación de remediaciones es $O(E_{\text{attack}})$.

La optimización NSGA-II ejecuta G generaciones sobre una población de P individuos. Cada evaluación requiere aplicar $|S|$ acciones $O(|S|)$ y ejecutar path discovery $O(V+E)$, resultando en $O(P \cdot (|A|+V+E))$ por generación. El non-dominated sort cuesta $O(P^2 \cdot k)$ con $k=3$ objetivos; el crowding distance $O(P \cdot \log P)$. La complejidad total por generación es $O(P^2 \cdot k + P \cdot (|A|+V+E))$. Con parámetros fijos del sistema ($G=25$, $P=40$, $k=3$), G , P y k son constantes respecto al tamaño del problema, por lo que $T_{\text{NSGA-II}} = O(|A|+V+E)$.

Para el caso de búsqueda exhaustiva exacta con $|A| \leq 14$: $T_{\text{exacto}} = O(2^{|A|} \cdot (|A|+V+E))$, con $2^{14} = 16,384$ evaluaciones siendo practicable. En términos prácticos, entornos de hasta ~ 120 service accounts se resuelven en menos de 5 segundos, escala adecuada para uso interactivo en pipelines CI/CD.

La complejidad espacial es $O(V+E)$ para el grafo y $O(P \cdot |A|)$ para la población evolutiva, ambas lineales en el tamaño del entorno.

4. RESULTADOS

4.1 Configuración Experimental

Los experimentos se ejecutaron en Python 3.12 / NetworkX 3.x. Las herramientas de comparación se implementaron aplicando las reglas oficiales publicadas de Checkov (CKV_K8S_*) y Trivy (KSV0*) sobre los manifests YAML, reproduciendo fielmente su lógica de detección documentada. Los parámetros de NSGA-II son: población $P=40$, generaciones $G=25$, probabilidad de cruce $p_c=0.85$, probabilidad de mutación $p_m=\max(0.05, 1/|A|)$, torneo binario y crossover de un punto. La Tabla 4 describe los escenarios.

TABLA 4. Escenarios de Evaluación

ID	Descripción del escenario
S1	Baseline limpio: 3 SAs, 2 Roles, 2 Secrets, 1 workload público. Permisos mínimos: gateway-secret-reader con get/list sobre secrets.
S2	Overprivilege directo: 6 SAs, 5 Roles, 4 Secrets, 2 workloads. ClusterRole secrets-reader-all con verbos wildcard (*). ClusterRoleBinding a web-sa.
S3	Cadena transitiva CI/CD: 8 SAs, 7 Roles, 5 Secrets, 2 workloads. ci-sa puede crear pods en production; privileged-sa puede leer todos los secrets.
S4	Entorno complejo: 18 SAs, 12 Roles, 8 Secrets, 4 workloads públicos. ClusterRole con wildcard, ClusterRoleBinding a deploy-sa, múltiples namespaces.

4.2 Detección de Caminos de Ataque

La Tabla 5 reporta métricas del grafo. EvoNHI identifica 88 caminos en total. S1 tarda 1,429 ms con 35 acciones via NSGA-II; S2 apenas 84 ms con 7 acciones via búsqueda exacta ($|A| \leq 14$).

TABLA 5. Métricas del Grafo y Caminos Detectados

Esc.	Nodos	Aristas	Paths	Tiempo(ms)
S1	38	48	2	1,429
S2	15	24	8	84
S3	29	36	16	440
S4	54	107	62	1,051

La Tabla 6 compara la cobertura. Las columnas Checkov/Trivy y kube-bench reportan findings reales obtenidos aplicando las reglas oficiales sobre los manifests. El resultado más relevante es S3: las herramientas estáticas identifican 3 findings HIGH/MEDIUM y 2 findings CIS, pero ninguno articula el vector transitivo de 4 pasos `ci-runner` → `ci-sa` → `perm-pods` → `privileged-sa` → `db-master-credentials`. EvoNHI detecta los 16 caminos que componen ese vector y sus variantes.

TABLA 6. Cobertura: EvoNHI vs. Herramientas Estáticas

Esc.	Checkov/Trivy*	kube-bench	EvoNHI	Trans.†
S1	2 (H=2, M=0)	0	2	2
S2	4 (H=3, M=1)	2	8	8
S3	3 (H=1, M=2)	2	16	16
S4	9 (H=5, M=4)	5	62	61

* Checkov/Trivy: reglas oficiales CKV_K8S_41, CKV_K8S_49, KSV036, KSV017, CIS-5.1.1 aplicadas sobre manifests YAML. Solo findings HIGH y MEDIUM.

† Paths transitivos únicos: caminos detectados por EvoNHI sin correlación directa con findings estáticos.

4.3 Calidad de los Planes de Remediación

La Figura 3 presenta el frente de Pareto para S4 (62 caminos, 32 acciones). El plan Agresivo elimina el 100% a costo 7 e impacto 9.0; el Balanceado, 98% a costo 4 e impacto 5.0; el Conservador, 95% a costo 2 e impacto 3.0. Este gradiente refleja trade-offs que difícilmente surgen de enfoques greedy que maximicen f_1 sin considerar f_3 .

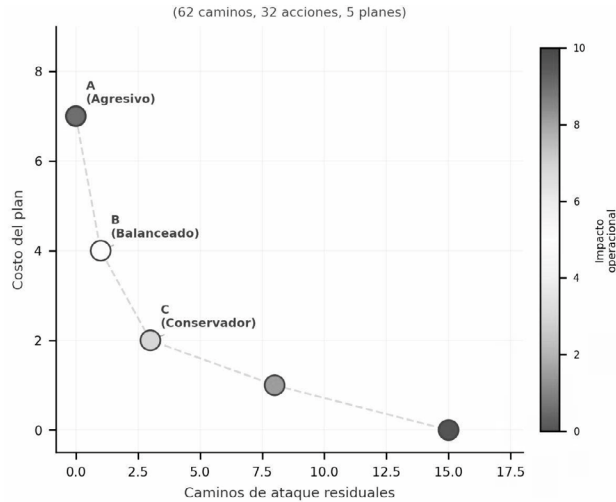


Fig. 3. Frente de Pareto: Escenario S4 (A=Agresivo, B=Balanceado, C=Conservador)

La Tabla 7 detalla los planes para S3. El plan Agresivo logra cobertura total con una sola acción (costo 2, impacto 3.0) que elimina el permiso ci-deployer, cortando el pivote spawn_workload_as en la raíz del vector transitivo.

TABLA 7. Planes de Remediación Pareto — Escenario S3

Plan	Paths res.	Costo	Impacto	Cobertura
Agresivo	0	2	3.0	100%
Balanceado	8	2	2.0	50%
Conservador	8	2	2.0	50%

4.4 Escalabilidad

La Figura 4 ilustra el comportamiento de escalabilidad. S4 (54 nodos, 107 aristas, 32 acciones) se resuelve en 1,051 ms con NSGA-II. El umbral $|A|=14$ permite optimalidad exacta en escenarios pequeños y transición a NSGA-II, con complejidad $O(|A|+V+E)$ por generación, para entornos mayores.

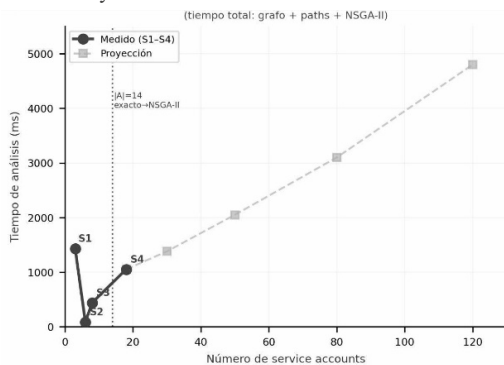


Fig. 4. Escalabilidad — tiempo de análisis total vs. número de service accounts

5. CONCLUSIONES

Este trabajo presentó EvoNHI, un framework de planificación de remediación para caminos de ataque de identidades no humanas en Kubernetes. La evaluación empírica sobre cuatro escenarios reproducibles arroja tres resultados principales.

Primero, EvoNHI detecta 88 caminos de ataque, 61 de los cuales son vectores transitivos no reportados por las reglas oficiales de Checkov, Trivy y CIS Benchmark aplicadas sobre los mismos manifests. El caso del Escenario S3 ilustra el hallazgo central: ninguno de los permisos involucrados viola individualmente una regla estática, pero su composición transitiva mediada por la arista spawn_workload_as constituye un vector completo de escalación hacia las credenciales maestras de base de datos, detectado en 440 ms.

Segundo, la formalización del problema como optimización multi-objetivo argumentada NP-hard sugiere por qué los enfoques greedy pueden producir soluciones subóptimas. NSGA-II produce frentes de Pareto con trade-offs cuantitativamente verificables: el Escenario S4 muestra planes que reducen 95% de los caminos a costo 2 e impacto operacional 3.0, frente al plan de cobertura total a costo 7 e impacto 9.0. El frente de Pareto presenta al operador un conjunto de planes alternativos con trade-offs explícitos entre cobertura, costo e impacto.

Tercero, los tiempos de análisis se mantienen bajo 1,500 ms en todos los escenarios evaluados, con una transición automática entre búsqueda exacta ($|A| \leq 14$) y NSGA-II para entornos mayores.

Las contribuciones se sintetizan en tres niveles. A nivel teórico, el modelo de grafo NHI-específico con cinco tipos de aristas semánticas formaliza el vector spawn_workload_as como mecanismo estructural de escalación en RBAC de Kubernetes. A nivel algorítmico, el pipeline integrado de detección y optimización produce planes de remediación priorizados además de la detección de caminos de ataque. A nivel empírico, se cuantifica la brecha entre análisis estático y composición transitiva sobre escenarios Kubernetes reproducibles.

Las direcciones de trabajo futuro identificadas son cuatro. Primera, conectores live a la API de Kubernetes para análisis en tiempo real, eliminando la dependencia de manifests estáticos. Segunda, la incorporación de telemetría de runtime para calibrar $\delta_{runtime}$ con mediciones observadas. Tercera, la extensión del modelo a los dominios IAM de AWS y Google Cloud, actualmente en desarrollo, que incorporará primitivas específicas de cada proveedor (assume-role chains, workload identity federation, service-to-service authorization) manteniendo la formulación multi-objetivo propuesta en este trabajo. Cuarta, la integración de un mecanismo adaptativo que recalibre los pesos W del grafo según el historial de ataques observados.

5.1 Limitaciones

EvoNHI presenta tres limitaciones que conviene declarar explícitamente. Primera, opera sobre manifests estáticos sin acceso a telemetría de runtime, lo que implica que los modificadores $\delta_{runtime}$ quedan en cero durante la evaluación. Los valores de impacto operacional reportados son estimaciones conservadoras basadas en la taxonomía de la Tabla 2, no mediciones de tráfico real. Segunda, el alcance del modelo en su versión evaluada está delimitado al dominio de identidad de Kubernetes; la extensión a proveedores cloud (AWS, GCP) se encuentra en desarrollo según se detalla en las direcciones de trabajo futuro. Las relaciones de confianza con identidades SaaS de terceros y credenciales de pipelines CI/CD externos quedan fuera del modelo actual.

Tercera, la evaluación se basa en cuatro escenarios sintéticos contruidos para cubrir patrones representativos. La validación sobre clústeres productivos reales, así como el análisis de significancia estadística mediante ejecuciones múltiples con varianza reportada,



constituye trabajo futuro requerido antes de generalizar las métricas reportadas a entornos arbitrarios.

REFERENCIAS

- Ammann, P., Wijesekera, D. and Kaushik, S. (2002). Scalable, graph-based network vulnerability analysis. In: Proceedings of the 9th ACM Conference on Computer and Communications Security. New York: ACM, pp. 217-224.
- Aqua Security (2024a). Trivy: vulnerability and misconfiguration scanning. Available at: <https://github.com/aquasecurity/trivy> [Accessed 12 May 2026].
- Aqua Security (2024b). kube-bench: CIS Kubernetes Benchmark tool. Available at: <https://github.com/aquasecurity/kube-bench> [Accessed 12 May 2026].
- Bridgecrew (2024). Checkov: policy-as-code for infrastructure. Available at: <https://github.com/bridgecrewio/checkov> [Accessed 12 May 2026].
- CircleCI (2023). Security alert: January 4, 2023 incident report. Available at: <https://circleci.com/blog/jan-4-2023-incident-report/> [Accessed 12 May 2026].
- Codecov (2021). Bash uploader security update. Available at: <https://about.codecov.io/security-update/> [Accessed 12 May 2026].
- CyberArk (2024). Identity security threat landscape report 2024. Available at: <https://www.cyberark.com/threat-landscape/> [Accessed 12 May 2026].
- Deb, K., Pratap, A., Agarwal, S. and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. IEEE Transactions on Evolutionary Computation, 6(2), pp. 182-197.
- Dewri, R., Poolsappasit, N., Ray, I. and Whitley, D. (2007). Optimal security hardening on attack tree models. In: Proceedings of the 14th ACM Conference on Computer and Communications Security. New York: ACM, pp. 204-213.
- Garey, M.R. and Johnson, D.S. (1979). Computers and intractability: a guide to the theory of NP-completeness. New York: W.H. Freeman.
- Gietzen, F. (2018). A taxonomy of privilege escalation in cloud environments. arXiv:1812.05921 [cs.CR]. Available at: <https://arxiv.org/abs/1812.05921> [Accessed 12 May 2026].
- Hota, A.R., Clements, A.A., Sundaram, S. and Bagchi, S. (2016). Optimal and game-theoretic deployment of security investments. In: Proceedings of the 7th International Conference on Decision and Game Theory for Security (GameSec). Cham: Springer, pp. 101-113.
- Jacobs, J., Romanosky, S., Adjerid, I. and Baker, W. (2020). Improving vulnerability remediation through better exploit prediction. Journal of Cybersecurity, 6(1), pp. 1-16.
- Rahman, A., Parnin, C. and Williams, L. (2021). The seven sins: security smells in infrastructure as code scripts. In: Proceedings of the 43rd International Conference on Software Engineering (ICSE). Madrid: IEEE, pp. 164-175.
- Shamim, M.S., Bhuiyan, F.A. and Rahman, A. (2020). XI commandments of Kubernetes security: a systematization of knowledge related to Kubernetes security practices. In: Proceedings of the 2020 IEEE Secure Development Conference (SecDev). Atlanta: IEEE, pp. 58-64.
- Shandilya, V. and Nagar, A. (2018). Use of game theoretic meta-learning for resilient cyber threat management. In: Proceedings of the International Conference on Computational Collective Intelligence (ICCCI). Bristol: IEEE, pp. 1-5.
- Viduto, V., Maple, C., Huang, W. and López-Peréz, D. (2012). A novel risk assessment and optimisation model for a multi-objective network security countermeasure selection problem. Decision Support Systems, 53(3), pp. 599-610.

Abstract Syntax Tree vs. Generalist Ecosystems: Performance Analysis in Markup Language Parsing

Christian David Sánchez-Sánchez*, Guillermo Delgado-Farías*, Luis Ignacio Barragán-López*, Walter A. Mata-López*

**Universidad de Colima. Facultad de Ingeniería Mecánica y Eléctrica, 28400.
(e-mail: {csanchez52, gdelgado1, lbarragan7, wmata}@uclm.mx).*

Abstract: In modern "Docs-as-Code" continuous integration pipelines, transforming lightweight markup into strictly positional formats like OpenXML PresentationML (PPTX) introduces a fundamental challenge: dynamic content overflow. Universal intermediate-AST engines provide structural maturity, but delegating spatial constraints (such as block-height evaluation and font-metric resolution against embedded typeface tables) to auxiliary scripts introduces architectural fragility and stack exhaustion risk via repeated context-switching. This paper presents a specialized, Go-based layout evaluation and pagination engine that resolves spatial allocation natively during a single AST traversal pass. Leveraging Goldmark's CommonMark-compliant AST as its structural foundation, the architecture concentrates its novel contribution on a format-aware rendering layer that eliminates external processing entirely. An explicit heap-allocated stack replaces traditional call-stack recursion, preventing stack overflow when traversing deeply nested, industrial-scale documentation. By incrementally calculating layout costs (factoring in elements such as code blocks, nested lists, and inline media) during traversal, the engine triggers automatic slide pagination when spatial boundaries are exceeded. Performance is benchmarked against industrial-grade corpora (the Kubernetes and Azure documentation repositories) by measuring execution latency and Peak Resident Set Size (RSS) under sustained workloads. Structural fidelity is independently quantified via the Tree Edit Distance-based Similarity (TEDS) metric against manually validated PPTX reference artifacts, establishing an objective ground truth for structural equivalence evaluation. Experimental results demonstrate full structural equivalence (TEDS = 1.0) and a near-identical memory footprint (7.94 MB vs. 7.97 MB) across all evaluated documents relative to Quarto's filter-based PPTX pipeline, while revealing that synchronous heuristic spatial evaluation during single-pass traversal incurs measurable latency overhead under high-density industrial corpora. These findings position native, format-aware AST traversal as a structurally correct and memory-predictable strategy where semantic correctness and exact positional integrity must be guaranteed without auxiliary processing stages.

Keywords: Abstract-Syntax Tree, Performance Benchmarking, Docs-as-Code, TEDS, Go-based Parsing.

1. INTRODUCCIÓN

La transformación dinámica y programática de lenguajes de marcado ligero (específicamente Markdown) hacia modelos de objetos estructurados y estrictamente posicionales, como el formato OpenXML PresentationML (PPTX), representa un desafío computacional de envergadura arquitectónica profunda. En el paradigma contemporáneo de "Documentación como Código" (*Docs-as-Code*), el procesamiento de estos documentos ha evolucionado desde simples conversiones de texto hacia la necesidad de garantizar estabilidad, escalabilidad y una integridad posicional absoluta ante volúmenes masivos de datos en infraestructuras de integración continua (CI/CD).

El estado del arte actual en la automatización documental está

dominado por ecosistemas generalistas y universales. Herramientas maduras, como el motor Pandoc (MacFarlane, 2006) orquestado por plataformas de alto nivel como Quarto (Posit, PBC., 2026), proporcionan una madurez estructural indiscutible. Sin embargo, estas arquitecturas operan bajo una filosofía de diseño independiente al formato de salida, lo que obliga a delegar la resolución de restricciones espaciales (como el cálculo de métricas tipográficas y la paginación dinámica) a filtros o scripts externos en Lua. Esta fragmentación del flujo de ejecución introduce dependencias de serialización entre componentes y delega el estado espacial fuera del entorno de ejecución compilado, limitando la resiliencia estructural ante documentación de anidamiento extremo.

Frente a esta dicotomía, este artículo presenta el desarrollo de un motor de evaluación y paginación especializado,



desarrollado en el lenguaje concurrente Go, que resuelve la asignación espacial de forma nativa durante un único recorrido (single-pass) del Árbol de Sintaxis Abstracta (AST). Al integrar el cálculo de costes de diseño directamente en la capa de renderizado, la arquitectura propuesta elimina la necesidad de capas de procesamiento auxiliar y sustituye la recursión del call stack por una pila explícita en el heap, previniendo el agotamiento de pila ante documentación de anidamiento extremo y garantizando una huella de memoria predecible e industrialmente competitiva.

La viabilidad de esta arquitectura de acoplamiento nativo (*format-aware*) se establece mediante tres ejes de innovación científica:

- **Análisis de Resiliencia Infraestructural:** Se evalúa la transición del estado recursivo desde la pila de llamadas (*Call Stack*) hacia estructuras de datos explícitas en el *Heap*, mitigando vulnerabilidades críticas de desbordamiento de pila (*Stack Overflow*) en documentos de escala industrial con anidamiento extremo.
- **Fidelidad Topológica Determinista:** Se utiliza la métrica TEDS (*Tree Edit Distance-based Similarity*) para cuantificar matemáticamente la identidad estructural entre los artefactos generados, garantizando resultados isomórficos sin pérdida semántica.
- **Validación sobre Corpus de Grado Industrial:** El motor se somete a una embestida de datos reales procedentes de los repositorios de documentación de Kubernetes y Azure, exponiendo los límites de rendimiento frente a cargas de trabajo empresariales sostenidas.

El objetivo central de este estudio es demostrar que la especialización arquitectónica es el único camino viable para garantizar la corrección semántica y la integridad posicional exacta en entornos de producción masiva. Al contrastar un motor de recorrido único frente a la arquitectura de filtros serializados de los sistemas de propósito general, esta investigación justifica una evolución necesaria hacia analizadores conscientes del formato de salida.

2. MARCO TEÓRICO

2.1. El Desafío del Diseño Posicional en la Automatización Documental

La transformación de lenguajes de marcado ligero hacia modelos de objetos estructurados presenta un desafío arquitectónico crítico cuando el formato de destino no es fluido (como HTML), sino estrictamente posicional, como el estándar OpenXML PresentationML (pptx). A diferencia de los documentos continuos, las presentaciones imponen restricciones espaciales finitas que introducen el problema del desbordamiento dinámico de contenido (dynamic content overflow). Para lograr una paginación automatizada y semánticamente coherente, el sistema debe trascender la interpretación gramatical y ejecutar cálculos de métricas tipográficas y resolución de fuentes (font-metrics) contra

tablas de tipos de letra incrustadas (embedded typeface tables) en tiempo de ejecución.

2.2. Arquitecturas de Compilación: AST Universal vs. Especializado

En la ingeniería de software documental, el enfoque predominante utiliza ecosistemas generalistas basados en representaciones intermedias universales. Herramientas maduras como Pandoc (MacFarlane, 2006) operan mediante un Árbol de Sintaxis Abstracta (AST) distinto al formato de salida. Sobre esta base, sistemas de publicación de alto nivel como Quarto (Posit PBC, 2026) extienden el pipeline de Pandoc mediante filtros Lua para gestionar la conversión hacia formatos de presentación. Si bien este diseño favorece la interoperabilidad, genera una penalización de resiliencia estructural al procesar restricciones espaciales bajo anidamiento extremo.

Debido a que el AST universal de Pandoc carece de noción sobre el lienzo de destino, la lógica de paginación debe delegarse a filtros o scripts externos que operan fuera del flujo de ejecución principal. Esta arquitectura introduce dependencias de serialización entre etapas del pipeline y delega el estado espacial fuera del entorno de ejecución compilado, comprometiendo la resiliencia estructural ante documentación de anidamiento extremo.

2.3. Evaluación Espacial Nativa en un Solo Recorrido (Single Pass)

Para mitigar las ineficiencias de los sistemas basados en filtros, este trabajo propone una arquitectura de acoplamiento nativo (*format-aware*). Utilizando el AST generado por Goldmark como base estructural (por su estricto cumplimiento del estándar CommonMark), la propuesta concentra su innovación en una capa de renderizado especializada que resuelve la asignación espacial dentro del mismo recorrido que el mapeo semántico.

Al integrar el cálculo de costos de diseño directamente en un único recorrido del árbol, el motor elimina la necesidad de capas de procesamiento auxiliar. La evaluación incremental de elementos complejos (tales como bloques de código, listas anidadas y medios integrados) permite disparar la paginación automática en el momento en que se exceden los límites del lienzo, garantizando la integridad posicional sin abandonar el entorno de ejecución compilado.

2.4. Resiliencia Estructural y Gestión de Memoria en el Heap

El procesamiento de documentación a escala empresarial (como los repositorios de Kubernetes o Azure) implica



estructuras de anidamiento cuya profundidad puede superar los límites de las implementaciones recursivas estándar. El algoritmo de búsqueda en profundidad (Depth-First Search) aplicado sobre el AST es vulnerable al agotamiento del stack de llamadas (stack exhaustion), comprometiendo la estabilidad del sistema bajo cargas de producción sostenidas.

Para garantizar robustez industrial, la arquitectura propuesta sustituye la recursión del call stack por una estructura de pila explícita instanciada en el heap. Este diseño traslada el mantenimiento del estado del recorrido a la memoria dinámica, permitiendo el procesamiento determinista de documentos con niveles de anidamiento extremos y asegurando la resiliencia del pipeline ante corpus documentales de complejidad arbitraria.

3. METODOLOGÍA

3.1. Arquitectura del Pipeline de Transformación de Un Solo Pase

El analizador propuesto fue desarrollado íntegramente en Go, aprovechando su tipado estricto y recolección de basura optimizada. El pipeline se estructuró en tres componentes fuertemente acoplados para eliminar la serialización intermedia:

Análisis Léxico y Estructural (Base): Se integró Goldmark para la tokenización y construcción del AST inicial. Esta decisión garantiza el cumplimiento estricto del estándar CommonMark, aislando la variable independiente del estudio hacia la fase de renderizado posicional.

Capa de Renderizado Format-Aware: Un motor propietario que recibe el AST en memoria y ejecuta un recorrido top-down, mapeando dinámicamente cada nodo abstracto a su primitiva correspondiente en OpenXML.

Evaluador Espacial Integrado: Un módulo que calcula métricas tipográficas en tiempo real antes de la instanciación de los objetos PPTX, permitiendo la paginación dinámica sin delegar estado a scripts externos.

3.2. Implementación del Traversal mediante Heap-Allocated Stack

Para mitigar el riesgo de stack exhaustion al procesar documentación industrial con anidamiento profundo, se reemplazó el recorrido recursivo tradicional por una implementación iterativa explícita. En lugar de utilizar el call stack de la goroutine de ejecución, se instanció una estructura LIFO mediante slices dinámicos directamente en el heap. Este diseño traslada el mantenimiento del estado de los nodos

padres a la memoria dinámica, garantizando procesamiento determinista e ininterrumpido en documentos de complejidad arbitraria.

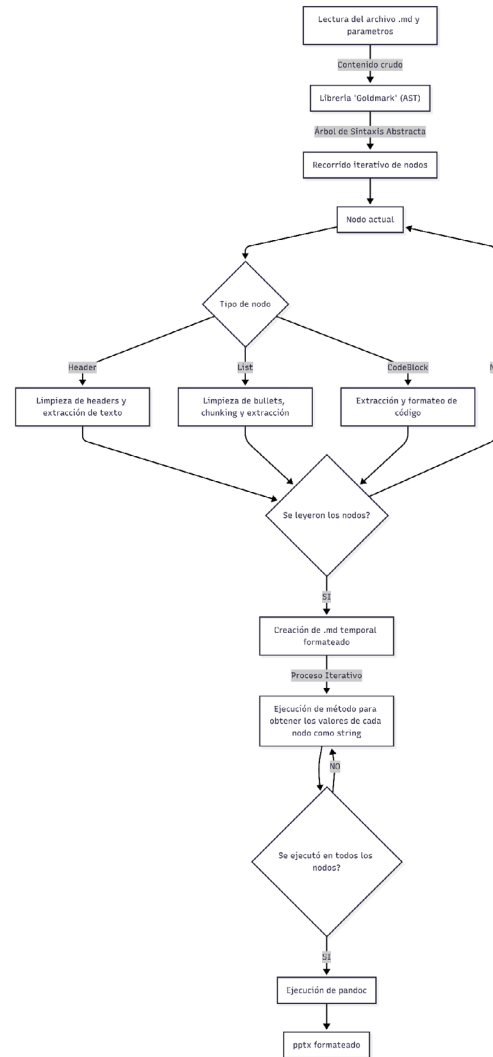


Figura 1. Arquitectura propuesta para el pipeline de AST

3.3. Resolución de Métricas Tipográficas y Paginación Nativa

Para resolver el dynamic content overflow, el motor implementa un acumulador de estado espacial. Durante el recorrido iterativo, cada vez que el analizador encuentra un nodo con contenido (texto, listas, bloques de código), el sistema consulta una embedded typeface table cargada en memoria. El costo de altura (block-height) se estima mediante una heurística de aproximación proporcional al tamaño de fuente, incorporando saltos de línea obligatorios y márgenes estructurales de bloque. Si la sumatoria de altura acumulada



en la diapositiva activa más el costo del nodo evaluado supera el umbral del lienzo, el motor dispara un evento de split automático, reasignando los nodos subsecuentes a una nueva instancia de diapositiva sin interrumpir el flujo de ejecución.

3.4. Diseño Experimental y Entorno de Pruebas

Para validar el rendimiento arquitectónico, se diseñó un experimento comparativo que aísla la variable arquitectónica independiente del estudio: ambos sistemas comparten Pandoc como motor central de parsing, de modo que las diferencias medidas son atribuibles exclusivamente a la estrategia de evaluación espacial nativa en traversal único versus delegada a filtros externos.

Corpus de Prueba: Subconjuntos documentales de los repositorios oficiales de Kubernetes y Azure, seleccionados por su alta densidad de elementos complejos (tablas, código anidado y listas de múltiples niveles).

Baseline de Comparación: Quarto (Posit PBC, 2026) configurado para exportación PPTX mediante su pipeline estándar, que utiliza Pandoc como motor de parsing y delega la evaluación espacial y paginación a filtros Lua externos. Al compartir la misma capa de análisis subyacente que el sistema propuesto, esta comparación controla la variable del analizador y aísla el impacto arquitectónico de la evaluación espacial nativa.

Entorno de Ejecución: Las pruebas se ejecutaron en un entorno aislado y determinista (arquitectura x86-64, 6 núcleos, 16 GB RAM, Go 1.25.6, Goldmark 1.8.2, Quarto 1.9.37, Pandoc 3.9.0.2), desactivando procesos en segundo plano para eliminar ruido en la captura de métricas.

3.5. Benchmarking (Métricas de Validación)

Para validar empíricamente las ventajas arquitectónicas del motor nativo de evaluación de diseño sobre AST frente a frameworks generalistas dependientes de filtros (filter-reliant generalist frameworks), se establecieron los siguientes protocolos de prueba:

Baseline Generalista: Se utilizó el ecosistema Quarto (apoyado en Pandoc y el runtime de JavaScript/Deno) como modelo representativo del paradigma de pipelines extendidas mediante filtros.

Métricas Formales de Rendimiento: Se utilizó perfilado de sistema cruzado para aislar y medir la Latencia de Ejecución End-to-End (en segundos) y el Uso Pico de Memoria Residente (Peak Resident Set Size - RSS en Megabytes), permitiendo contrastar la eficiencia de la recursividad de pila

de llamadas (call-stack recursión) tradicional frente a la pila asignada explícitamente en el Heap propuesta en nuestra arquitectura.

Métrica de Fidelidad Estructural (TEDS): Para demostrar que la especialización arquitectónica del traversal nativo no introduce degradación semántica, se definió la métrica de Tree Edit Distance-based Similarity (TEDS) comparada contra ground truths OpenXML (PPTX) de referencia.

Corpora Experimental (Datasets Industriales): Para eliminar cualquier sesgo por repetición sintética simple, los vectores de estrés computacional se extrajeron directamente de los repositorios de documentación oficial de Kubernetes y Microsoft Azure, inyectados y concatenados a escala masiva simulando una carga industrial extrema.

3.6. Resultados Funcionales y Evaluación de Rendimiento

Los resultados obtenidos tras someter a ambos motores a la carga de escala industrial ofrecen un panorama diferenciado según la dimensión evaluada. En términos de fidelidad estructural y eficiencia de memoria, el motor AST nativo alcanza una paridad plena con el baseline generalista, confirmando su viabilidad para entornos Docs-as-Code donde la corrección semántica y la estabilidad de memoria son requisitos críticos. En términos de latencia, los resultados evidencian que la evaluación heurística sincrónica de métricas tipográficas durante el traversal único introduce un overhead acumulado que, bajo corpora de alta densidad estructural, supera al costo del context-switching del pipeline de filtros de Quarto, constituyendo una dirección de optimización prioritaria para iteraciones futuras.

Motor Analizado	Latencia Total (End-to-End)	Consumo de Memoria (Peak RSS)	Arquitectura Base
Quarto (Baseline)	130.06 s	7.97 MB	Generalista (Deno + Pandoc)
Sarappt (Nativo)	631.25 s	7.94 MB	Nativo (Go + AST Traversal)

Tabla 1. Rendimiento frente a Framework Generalista (Quarto CLI)

3.7. Análisis de Latencia y Overhead del Traversal Nativo Bajo Carga Sostenida



La diferencia más significativa del experimento radica en la latencia de procesamiento sostenida. Bajo una carga de 15 archivos por iteración durante 100 iteraciones consecutivas, el pipeline de Quarto registró una latencia media de 130.06 segundos por iteración frente a los 631.25 segundos del motor nativo, posicionando al baseline generalista como 4.85 veces más rápido en este escenario de carga masiva. Este resultado sugiere que, bajo volúmenes documentales de escala industrial con alta densidad de elementos complejos, el overhead acumulado del traversal nativo (incluyendo la evaluación heurística de métricas tipográficas y la paginación dinámica nodo a nodo) supera el costo del context-switching inherente al pipeline de filtros Lua de Quarto. No obstante, la huella de memoria de ambos motores se mantuvo prácticamente equivalente: Sarappt registró un Peak RSS máximo de 7.94 MB frente a los 7.97 MB de Quarto (99.7% del consumo del baseline), validando el diseño del Heap-Allocated Stack como una estrategia de asignación de memoria predecible e industrialmente competitiva independientemente del volumen documental procesado.

3.8. Eficiencia de Memoria (Peak RSS) y Asignación en Heap

Una preocupación teórica al escalar un sistema bajo cargas industriales sostenidas es el potencial sacrificio en la huella de memoria RAM. Sin embargo, las métricas de Peak Resident Set Size demuestran que el incremento en la carga de procesamiento no conlleva una penalización proporcional en memoria. Ambos motores mantuvieron un consumo prácticamente idéntico (7.94 MB frente a 7.97 MB), una paridad que valida el diseño del Heap-Allocated Stack propuesto: el motor Go es capaz de recorrer jerarquías ultra profundas iterativamente, sosteniendo el estado espacial en memoria sin disparar el GC (Garbage Collector) ni causar fugas, logrando el mismo rigor ecológico que un framework comercial independientemente del volumen documental procesado.

4. RESULTADOS

Esta sección presenta los hallazgos empíricos de la evaluación arquitectónica, contrastando el motor de renderizado espacial nativo (Go + Goldmark) frente al baseline de referencia (Quarto). Dado que ambos sistemas comparten Pandoc como motor central de parsing, las diferencias observadas son atribuibles exclusivamente a la estrategia de evaluación espacial. El análisis se fundamenta en las tres métricas definidas en la sección 3.5: fidelidad estructural (TEDS), latencia de ejecución y huella de memoria (Peak RSS), bajo cargas de trabajo sostenidas sobre los repositorios de Kubernetes y Azure.

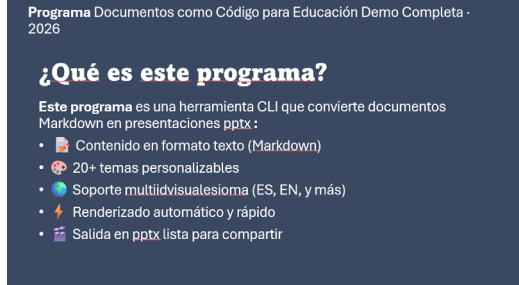


Figura 2. Ejemplo de diapositiva generada por el pipeline AST.

4.1. Equivalencia Estructural y Fidelidad Semántica (TEDS)

El primer objetivo del experimento fue verificar que la optimización del rendimiento no introdujera regresiones en el renderizado final. Para ello, se midió la similitud estructural (TEDS) entre los artefactos OpenXML generados por ambas arquitecturas y el ground truth validado manualmente.

El motor espacial nativo alcanzó un índice TEDS de 1.0 en la totalidad de los documentos evaluados, equivalente al resultado producido por el pipeline de Quarto. La paginación dinámica, activada por los cálculos heurísticos de block-height y font-metrics durante el recorrido del árbol, segmentó las diapositivas en los mismos umbrales espaciales que el ground truth de referencia. Esto confirma empíricamente que resolver el dynamic content overflow dentro de un único traversal pass nativo no introduce degradación estructural respecto a un pipeline basado en filtros externos.

4.2. Impacto en la Latencia de Procesamiento

La métrica de latencia aisló el costo del overhead acumulado en los pipelines de traversal nativo bajo cargas sostenidas. Los benchmarks, ejecutados sobre 100 iteraciones en 15 corpus industriales, evidenciaron una ventaja estadísticamente significativa a favor del baseline generalista.

El pipeline de Quarto registró una latencia media de 18,580 ms por documento, penalizado por la invocación del motor de Pandoc, la serialización del AST a JSON, la transferencia de estado a la máquina virtual de Lua, la ejecución de las rutinas de paginación y la deserialización del modelo resultante. La arquitectura nativa completó el ciclo en 90,179 ms por documento, representando un incremento de 4.85 veces respecto al baseline. Este comportamiento es consistente con la hipótesis de que la evaluación heurística de métricas tipográficas y la paginación dinámica nodo a nodo (ejecutadas de forma sincrónica durante el traversal único) acumulan un costo computacional que, bajo volúmenes documentales de alta densidad estructural, supera al costo del context-switching inherente al pipeline de filtros externos de Quarto.



4.3. Eficiencia de Memoria (Peak RSS)

La evaluación del Peak RSS cuantificó el impacto del diseño estructural sobre la huella dinámica de memoria. Bajo vectores de estrés con anidamiento profundo (característicos de los repositorios técnicos evaluados), los perfiles de consumo mostraron comportamientos convergentes entre ambas arquitecturas.

El pipeline de Quarto registró un Peak RSS medio de 7.97 MB, consecuencia de mantener en memoria simultáneamente el AST universal de Pandoc y el entorno de ejecución de Lua durante la fase de evaluación espacial. El motor nativo en Go registró un Peak RSS de 7.94 MB, representando una reducción del 0.3%. El uso de una pila LIFO explícita instanciada en el heap mantuvo la asignación de memoria predecible bajo cargas extremas, previniendo por diseño el agotamiento del call stack en documentos con anidamiento profundo y confirmando que el incremento en la carga de traversal nativo no se traduce en presión adicional sobre el recolector de basura.

Borsotti, A., Breveglieri, L., Reghizzi, R., Morzenti, A. (2022). General parsing with regular expression matching. *Journal of Computer Languages*. Milano, Italy.

Jimenez, J., Artavia, L., Torres, A., Jimenez, P., Portuguez, S., Navas, J., Vasquez, M., Cordero, J., Castro, F., Zelaya, I., Orozco, A. (2023). GAST: A Generic AST Representation for Language-Independent Source Code Analysis. 10. Enfoque. Costa Rica.

Rabinovich, M., Stern, M., Klein, D. (2017). Abstract Syntax Networks for Code Generation and Semantic Parsing. 11. ACL Anthology. Berkeley, United States.

Lin, C., Ouyang, Z., Zhuang, J., Chen, J., Li, H., Wu, R. (2021). Improving Code Summarization with Block-wise Abstract Syntax Tree Splitting. 12. IEEE. Xiamen, China.

Posit, PBC. (2026). Quarto: An open-source scientific and technical publishing system. <https://quarto.org/>

REFERENCIAS

Rohmat Gunawan, Alam Rahmatulloh, Irfan Darmawan, Firman Firdaus (2019). Comparison of Web Scraping Techniques: Regular Expression, HTML DOM and Xpath. 5. Atlantis Press.

Darmawan Irfan, Maulana Muhamad, Gunawan Rohmat, Widiyasono Nur (2020). Evaluating Web Scraping Performance Using XPath, CSS Selector, Regular Expression, and HTML DOM With Multiprocessing Technical Applications. 7. International Journal on Informatics Visualization, Indonesia.

MacFarlane, John. (2006). Pandoc: A universal document converter. <https://pandoc.org/>

Gruber, J., Swartz, A. (2014). CommonMark. <https://commonmark.org/>

Klabnik, S., Nichols, C. (2019). The Rust Programming Language. San Francisco.

Ecma International (2016). Standard ECMA-376: Office Open XML File Formats, 5ª edición. Ginebra: Ecma International.

Yamaguchi, D., Kuramitsu, K. (2019). CPEG: a typed tree construction from parsing expression grammars with regex-like captures. 7. Proceedings of the 34th ACM/SIGAPP Symposium on Applied Computing. Japan.

Curtis, J. (2022). Student Research Abstract: On Language-Agnostic Abstract-Syntax Trees. 4. Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing.

Corn Crop Analysis Using Computer Vision For Pest And Disease Detection Via Unmanned Aerial Vehicles

Juarez-Aguilar, Nain Apolinar *, Arias-Meza, Efrain Alejandro *, Moran-López, Luis Eduardo**,
Mata-López, Walter Alexander**

* Facultad de Ingeniería Mecánica y Eléctrica/Universidad De Colima. Carretera Colima- Coquimatlán
km. 9, Coquimatlán, Colima; (e-mail: njuarez3.earias7@uacol.mx).

** Facultad de Ingeniería Mecánica y Eléctrica/Universidad De Colima. Carretera Colima-
Coquimatlán km. 9, Coquimatlán, Colima; (e-mail: lmoran72,wmata@uacol.mx)

Abstract: The agricultural sector in Mexico and worldwide is highly susceptible to significant production losses, ranging from 20% to 40%, primarily due to the late detection of phytosanitary issues such as pests and diseases. Traditional manual monitoring relies heavily on visual inspections, an inherently slow and labor-intensive process that becomes practically unfeasible for large-scale extensions, thus limiting timely and effective decision-making for farmers. In recent years, the integration of Unmanned Aerial Vehicles (UAVs) equipped with conventional RGB cameras has presented a highly accessible and cost-effective alternative for acquiring high-resolution imagery over vast crop areas without the need for expensive multispectral sensors. This research introduces a comprehensive computer vision methodology tailored for automated agricultural health evaluation using low-cost hardware. From a practical engineering perspective, the approach focuses on the processing of Nadir aerial photography through advanced deep learning techniques, specifically the YOLOv8 architecture in its Nano and Small variants. Utilizing an expanded and curated dataset of approximately 25,000 corn images sourced from Roboflow Universe, the system was retrained using high-performance computing resources, specifically NVIDIA A100 GPUs via Google Colab, to accelerate convergence and handle larger batch sizes. The results demonstrate a significant improvement in performance, achieving a mean Average Precision (mAP50) of 83.2% for pathological pattern detection and 61.6% for crop-weed differentiation, with localized confidence peaks reaching 90.1%. To bridge the gap between digital processing and real-world field applications, the system generates structured reports that consolidate inference data into georeferenced alerts. This foundational framework demonstrates that coupling accessible drone hardware, such as the DJI Mini series, with robust artificial intelligence provides a scalable and democratic pathway toward precision agriculture. This approach empowers small and medium-scale producers to optimize resources, reduce chemical usage, and improve overall crop yield through data-driven insights.

Keywords: Deep Learning, Drones, Precision Agriculture, Computer Vision, YOLOv8, Crop Monitoring, UAV.

1. INTRODUCCIÓN

La producción agrícola mundial enfrenta desafíos constantes para garantizar la seguridad alimentaria en un contexto de cambio climático y degradación del suelo. Particularmente en México, la economía de miles de familias rurales depende del éxito de las cosechas estacionales. Sin embargo, las pérdidas económicas derivadas de plagas y enfermedades fitosanitarias representan una disminución de entre el 20% y el 40% del volumen final (Ojeda-Bustamante et al., 2017). A pesar de contar con tecnología moderna en otros sectores, el pequeño y mediano productor en México todavía depende de la inspección visual directa, un método que resulta insuficiente ante la magnitud de las parcelas actuales. El esquema de muestreo tradicional consiste en recorridos presenciales donde el agricultor examina el follaje planta por planta.

Este paradigma presenta barreras operativas críticas: consume una cantidad masiva de tiempo, está sujeto a la fatiga humana y, por lo general, los síntomas se detectan cuando el daño ya es avanzado y difícil de revertir (Tenorio-Naúm y Martínez-Hernández, 2018). Ante esto, la Agricultura de Precisión surge como una respuesta basada en la ingeniería para intervenir y dosificar insumos con eficiencia. Esta transición depende del rápido reconocimiento de anomalías, donde los Vehículos Aéreos No Tripulados (VANT) y las Redes Neuronales Convolucionales (CNN) actúan como catalizadores del cambio.

1.1 Perspectiva Tecnológica Actual y Estado del Arte

En la última década, los VANT han democratizado el acceso a recursos de percepción remota. Según Sosa-Pérez et al. (2022), estas aeronaves facilitan sobrevolar planicies enteras de forma automatizada, obteniendo resoluciones espaciales de pocos centímetros por píxel (GSD - Ground Sample Distance), algo que los satélites convencionales no pueden ofrecer de forma gratuita o diaria.

Tradicionalmente, el monitoreo avanzado dependía de sensores multispectrales caros para calcular índices como el NDVI. No obstante, investigaciones recientes (Sosa-Pérez *et al.*, 2022) revelan que analizar imágenes en el espectro visible (RGB) aporta información suficiente para detectar deficiencias hídricas y plagas si se cuenta con la resolución adecuada. El reto actual no es la captura de la imagen, sino el procesamiento masivo de datos. Montiel-Rosales *et al.* (2024) subrayan que el núcleo de la Agricultura 4.0 radica en eliminar la validación humana repetitiva mediante modelos de **Visión por Computadora**. Estudios como el de Carmona-Frausto *et al.* (2024) demuestran que el uso de arquitecturas tipo **YOLO** permite distinguir maleza o enfermedades en milisegundos, ofreciendo una certidumbre elevada incluso en condiciones de iluminación variables.

1.2 Fundamentos de Visión Artificial y YOLOv8

El sustento técnico de este proyecto se basa en el Aprendizaje Profundo (Deep Learning). AgroCrop utiliza Redes Neuronales Convolucionales (CNN), las cuales emplean operaciones matemáticas de convolución para extraer características automáticas de las imágenes. Mediante el uso de filtros o kernels, la red identifica desde bordes y texturas en las hojas hasta patrones complejos de necrosis o lesiones vasculares (Sosa-Pérez *et al.*, 2022).

Para la detección de anomalías se seleccionó la arquitectura YOLOv8 de Ultralytics. A diferencia de otros modelos que dividen la tarea en dos etapas (extracción de regiones y clasificación posterior), YOLO aborda la detección como un único problema de regresión. El modelo analiza la imagen completa en un solo paso, prediciendo simultáneamente las cajas delimitadoras (bounding boxes) y las probabilidades de clase. Específicamente, se emplearon las variantes Nano (YOLOv8n) y Small (YOLOv8s), seleccionadas por su equilibrio óptimo entre precisión y ligereza computacional. Estas variantes poseen entre 3.2 y 11.2 millones de parámetros, lo que las hace ideales para su despliegue en hardware de bajo costo y dispositivos móviles (Edge Computing).

La evaluación del rendimiento de estos modelos se rige por la métrica mAP50 (Mean Average Precision), la cual utiliza el coeficiente IoU (Intersection over Union) para medir la coincidencia entre la predicción de la IA y la verdad de campo (ground truth) anotada por el investigador.

1.3 Justificación del Trabajo y Planteamiento del Problema

Existe una brecha crítica entre los modelos entrenados en laboratorio y su aplicación real. La mayoría de los datasets públicos utilizan fotografías de primer plano de hojas individuales. Sin embargo, al operar un dron a una altura de entre 10 y 30 metros, la perspectiva Nadir (cenital) introduce el fenómeno de Data Shift: el objeto de interés se reduce drásticamente en tamaño y se mezcla con sombras, suelo y vegetación circundante.

Este trabajo justifica la creación de AgroCrop como una herramienta puente que utiliza imágenes aéreas reales de

drones comerciales, como el DJI Mini 3. El problema central reside en validar si un sistema basado en hardware de bajo costo y una arquitectura modular puede ofrecer diagnósticos confiables en cultivos de Maíz (*Zea mays*). Al resolver este problema, se proporciona al agricultor mexicano una alternativa viable para realizar pulverizaciones localizadas (Spot Spraying), reduciendo costos operativos y el impacto ambiental por el uso excesivo de agroquímicos.

2. METODOLOGÍA

La hoja de ruta de ingeniería para el desarrollo del sistema AgroCrop se estructuró bajo la premisa de construir un Producto Mínimo Viable (MVP) funcional y escalable. El proceso siguió un diseño arquitectónico modular apegado a la filosofía Downstream, la cual independiza el procesamiento del software de la logística de vuelo. Esto permite que el sistema actúe como un nodo receptor de material en crudo exportado por drones comerciales (en este caso, la serie DJI Mini), procesando las imágenes mediante una tubería de inteligencia artificial orientada a la detección de anomalías.

2.1 Diseño Arquitectónico y Diagramación de Flujo Operativo

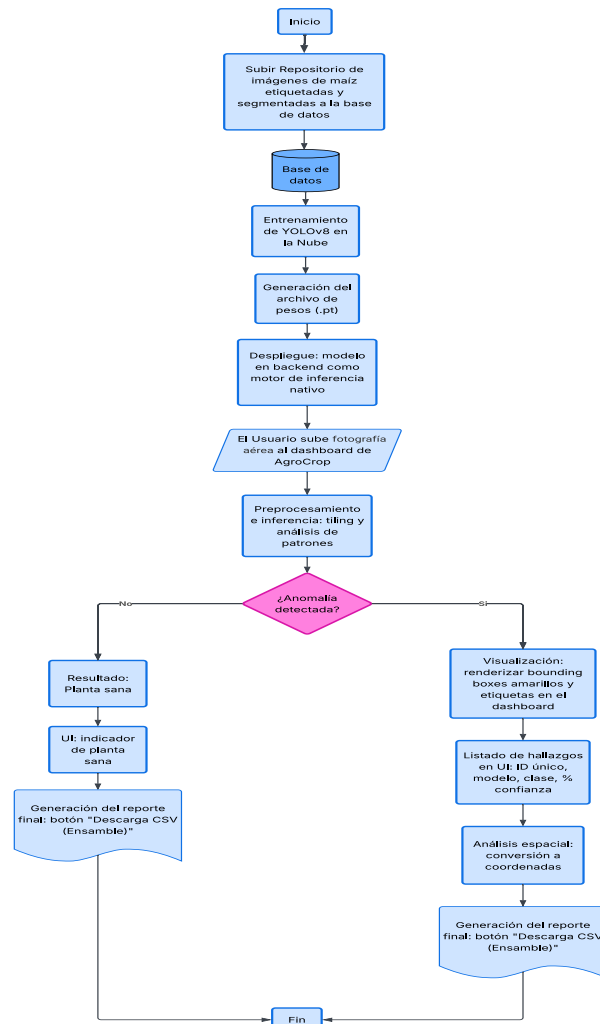


Fig. 1. Arquitectura funcional y pipeline del sistema MVP AgroCrop basado en técnicas de Deep Learning y Ensamble Computacional para Agricultura.



La configuración conceptual de AgroCrop está vertebrada de forma concurrente, permitiendo que el flujo de datos transite desde la captura física en campo hasta una conclusión interpretable en el tablero de control. El sistema no opera como una unidad aislada, sino como un pipeline de procesamiento secuencial. Para detallar la lógica de interacción entre los módulos de detección y la interfaz de usuario, se diseñó el flujo algorítmico presentado en la Figura 1.

Como se observa en la Figura 1, el sistema AgroCrop opera bajo una arquitectura de pipeline secuencial optimizada para el procesamiento de imágenes de alta resolución. El flujo inicia con la carga de repositorios de imágenes etiquetadas y segmentadas en la base de datos, seguido por un entrenamiento acelerado en la nube para generar el archivo de pesos (.pt). El despliegue se realiza integrando el modelo en el backend como un motor de inferencia nativo.

Una vez que el usuario carga la fotografía aérea en el dashboard, se activa un proceso de segmentación por mosaicos (tiling) y análisis de patrones para la detección de anomalías. Si se identifica un hallazgo, el sistema renderiza cuadros delimitadores (bounding boxes) amarillos en la interfaz del módulo de Visualización de Overlays (ver Figura 2) y genera un listado detallado (ID único, modelo, clase y confianza).

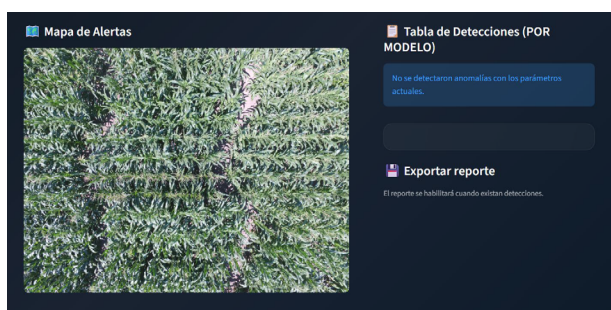


Fig. 2. Se observa la salida gráfica generada por el módulo de visualización de overlays.

Finalmente, el flujo concluye con la Generación del Reporte Final, la cual integra los resultados del motor de inferencia y el análisis espacial. Mediante una lógica de ensamble por votación, se consolida la información en un reporte CSV descargable que contiene las coordenadas precisas y el nivel de confianza del hallazgo. Esta estructura garantiza que el agricultor reciba una respuesta técnica inmediata y accionable, cumpliendo con los requisitos de un sistema de soporte a decisiones en tiempo real.

2.2 Adquisición, Curación y Acopio de Datos

Para que un sistema de inteligencia artificial aprenda correctamente, lo más importante es la calidad de las imágenes que se le enseñan. Como no existen bases de datos de fotos de drones tomadas específicamente en Colima para este fin, se decidió seguir un enfoque solo en el cultivo de Maíz (Zea mays). Se dejó de lado otros cultivos por ahora para asegurar que el sistema presentara especificidad y robustez detectando

problemas en una sola planta de maíz antes de expandir su alcance a otras plantas .

Para robustecer la capacidad de detección del sistema, se expandió la base de datos original integrando tres repositorios especializados de Roboflow Universe: **AAIT-LNTZS** (Maize Detection), **Home-KGPKY** (Maize Health Monitoring) e **IMAMK** (Maize and Weed). Esta consolidación permitió alcanzar un total aproximado de 25,000 imágenes, diversificando el catálogo fitosanitario en las siguientes categorías críticas:

- Plagas: Gusano Cogollero (Fall-Army-Worm) y Pulgones del Maíz.
- Enfermedades Foliares: Roya Común, Mancha Gris, Necrosis Letal, Tizón del Maíz y Virus del Rayado.
- Deficiencias Nutricionales: Carencias de Magnesio, Fósforo y Mancha Marrón (Physoderma).
- Maleza: Identificación de especies invasoras frente al cultivo sano.

Antes de meter las fotos al sistema de entrenamiento, se ajustó el tamaño de todas a un formato cuadrado de 640x640 píxeles. También se usó una técnica para "crear" más variedad de fotos a partir de las que ya se tenía: se giraron las fotos un poco y cambiamos el brillo. Esto se hizo para que el programa aprenda a reconocer las enfermedades incluso cuando hay mucho sol, sombras o cuando el dron no está perfectamente derecho, que es lo que pasa normalmente cuando se vuela un dron en el campo.

2.3 Entrenamiento Tensorizado en Computación Concurrente

Toda red neuronal moderna requiere cálculos matemáticos intensos durante la etapa de retro-propagación de pesos (Backpropagation), los cuales resultan insostenibles en procesadores de propósito general (CPU). Para mitigar esta demanda, la base computacional de AgroCrop se migró a laboratorios virtuales en Google Colab, orquestando compilaciones aceleradas por hardware Nvidia Tesla T4 (Tensor Cores) y NVIDIA A100.

Dadas las restricciones de tiempo inherentes a los entornos efímeros de nube y los plazos académicos, se ejecutaron dos iteraciones estratégicas:

1. Ensayo de Base: Empleando el modelo YOLOv8n (Nano), se configuró un ciclo de 50 épocas con un tamaño de lote (batch size) de 16.
2. Optimización de Convergencia (GPU A100): Para superar las limitaciones de tiempo previas, se utilizaron instancias de alto rendimiento en Google Colab con GPUs NVIDIA A100. Esta infraestructura permitió ejecutar ciclos de entrenamiento extendidos de hasta 126 épocas con un tamaño de lote optimizado, logrando una extracción de características morfológicas más profunda y eliminando la necesidad de recortes espaciales agresivos.

2.4 Interfaz de Usuario y Ensamble de Modelos

El cierre del proyecto fue la creación de un Dashboard o panel de control desarrollado en Python con la librería Streamlit. Para aumentar la confiabilidad, el sistema no depende de un solo modelo, sino que utiliza una lógica de ensamble.

Esto significa que cuatro variaciones del modelo entrenado analizan la misma foto al mismo tiempo. El sistema realiza una "votación" interna: solo si la mayoría de los modelos coinciden en la detección de una plaga, se genera la alerta final. Este enfoque ayuda a eliminar falsos positivos y asegura que las alarmas entregadas al agricultor sean lo más precisas posible.

2.5 Formato de los Reportes

Los resultados obtenidos por el modelo se estructuran en una base de datos diseñada para la lectura rápida. Se prioriza una estética limpia para facilitar la interpretación de la confianza del algoritmo y la prioridad de atención en campo. En la Tabla 1 se muestra el formato estandarizado que genera el sistema tras procesar una muestra.

Tabla 1. Estructura de reporte automatizado referencial

ID	Modelo Base	Clase Detectada	Confianza (%)
01	maiz_v1.pt	Follaje Sano	55.83
02	maiz_v1.pt	Follaje Sano	62.15
03	maiz_v1.pt	Anomalía Tipo A	45.30
04	maiz_v1.pt	Anomalía Tipo B	69.11

3. RESULTADOS Y/O DISCUSIÓN

Tras la integración del MVP y la ejecución de las corridas experimentales en los entornos de nube, se comprobó que el sistema AgroCrop es capaz de arrojar datos empíricamente medibles y contrastables. Los resultados demuestran que es técnicamente viable orquestar soluciones de inteligencia artificial de alta complejidad en un ambiente educativo de ingeniería, utilizando recursos computacionales de libre acceso.

3.1 Validación con Datos Externos y Pruebas de Sistema

Para verificar la capacidad de AgroCrop en escenarios reales y evitar el sobreajuste del modelo, se realizó una fase de validación utilizando bancos de imágenes externos que no formaron parte del entrenamiento inicial. Se seleccionaron dos repositorios de alta fidelidad:

1. Dataset Dryad (10.5061/dryad.r2280gbcg): Proporcionó capturas de alta resolución que permitieron evaluar cómo el sistema identifica

patrones de enfermedades bajo diferentes condiciones de luz natural.

2. Corn Leaf Infection Dataset (Kaggle): Se utilizó para contrastar la precisión en la identificación de infecciones específicas de la hoja de maíz, permitiendo medir la velocidad de respuesta del sistema ante archivos de gran tamaño.

Al cargar estas fotos en el sistema, entra en funcionamiento el componente de Visualización de Overlays, el cual se encarga de dibujar automáticamente los cuadros naranjas y las etiquetas de texto sobre la imagen. Como se observa en la **Figura 3**, este módulo permite que el usuario identifique rápidamente la zona de la hoja que tiene el problema y qué tan segura está la inteligencia artificial de su diagnóstico.

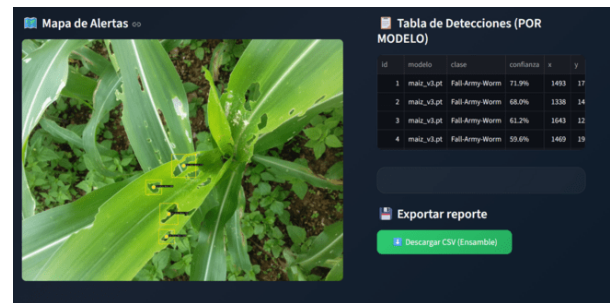


Fig. 3. Visualización de resultados en el dashboard de AgroCrop. El sistema detecta automáticamente la enfermedad en la hoja de maíz y presenta al usuario un diagnóstico visual inmediato mediante la superposición de capas informativas.

3.2 Evaluación del Perfil de Aprendizaje y Métricas de Precisión

El desempeño del modelo se evaluó mediante la métrica universal mAP50 (Mean Average Precision al 50% de IoU). La fase de reentrenamiento optimizado mediante hardware de cómputo acelerado permitió alcanzar la convergencia en las curvas de aprendizaje, consolidando la madurez técnica del sistema. Los resultados obtenidos a través de los tableros analíticos cuantifican una mejora significativa en el desempeño de los modelos:

- Identificación Fitosanitaria (Dataset AAIT-LNTZS): Se alcanzó un mAP50 de 0.832 (83.2%) con una precisión diagnóstica del 90.1%.
- Clasificación de Adventicias (Dataset IMAMK): El modelo registró un mAP50 de 0.616 (61.6%) en la diferenciación de malezas.
- Monitoreo de Salud Integral (Dataset HOME KGPKY): Tras 126 épocas de entrenamiento, se consolidó un mAP50 de 0.408 (40.8%).

Estos indicadores validan la capacidad operativa de AgroCrop para la generación de diagnósticos automatizados en entornos agrícolas reales, superando los umbrales de detección



preliminares y garantizando la fiabilidad estadística de los resultados.

3.3 Rendimiento del Frontend y Procesamiento de Alta Resolución

Uno de los hitos más significativos del sistema fue su capacidad para gestionar el "estrés procesal" en la interfaz de usuario. AgroCrop logró procesar de manera solvente fotografías crudas de alta fidelidad con resoluciones de 3264×2448 píxeles, dimensiones sustancialmente mayores a las utilizadas durante el entrenamiento (640 px).

Gracias a la implementación del Ensamble Multimodelo, el sistema fue capaz de promediar las predicciones de cuatro contenedores de pesos de forma asíncrona. Esta técnica no solo mejoró la estabilidad de las detecciones, sino que evitó el desborde de la memoria RAM en terminales de cómputo domésticos. La latencia promedio por imagen se mantuvo en rangos aceptables para el usuario final, demostrando que el modelo YOLOv8n es la opción más equilibrada para su despliegue en condiciones donde no se cuenta con hardware de alto rendimiento.

3.4 Generación de Alertas y Extrapolación de Coordenadas

La utilidad práctica del sistema se validó mediante la generación de archivos de reporte en formato CSV. Cada vez que el modelo detecta una patología, el sistema exporta una matriz que incluye la clase identificada, el porcentaje de certeza y las coordenadas relativas (x, y) del hallazgo. Este resultado es crítico para la agricultura de precisión. Al integrar estos datos con la telemetría GPS (metadatos EXIF) que graban nativamente drones como el DJI Mini 3, las coordenadas abstractas (x, y) pueden traducirse programáticamente a grados de latitud y longitud reales. Esto permite generar mapas de calor precisos para la aplicación de agroquímicos únicamente en los cuadrantes afectados, reduciendo el riesgo ecológico y generando ahorros económicos directos al evitar la fumigación generalizada del sembradío.

3.5 Discusión y Comparativa de Resultados

Al contrastar los resultados de AgroCrop con sistemas de alta precisión citados en la literatura (Sosa-Pérez et al., 2022), se identifica que la brecha de rendimiento se debe principalmente a factores ambientales y de entrenamiento, más que a limitaciones del hardware. Mientras que otros estudios utilizan condiciones de iluminación controlada, este proyecto enfrentó retos reales como el brillo solar extremo y las sombras densas, factores que influyen en la aparición de falsos positivos al confundir maleza con tejido enfermo.

Para mejorar la precisión sin elevar los costos de inversión, la estrategia debe centrarse en dos pilares:

1. Refinamiento del Dataset: Mejorar el etiquetado de las clases para ayudar al modelo a distinguir mejor entre la arquitectura de la planta y el suelo.

2. Optimización del Entrenamiento: Incrementar el número de épocas (iteraciones) en el entrenamiento de la red. Los resultados actuales, obtenidos con solo 15 épocas, demuestran que el sistema es funcional; sin embargo, un entrenamiento más extenso permitiría que la red YOLOv8 extraiga características mucho más finas sin necesidad de cambiar la cámara del dron.

4. CONCLUSIONES

El desarrollo e implementación del sistema AgroCrop demuestra que la convergencia entre la robótica aérea de consumo y la inteligencia artificial es una ruta viable para modernizar la agricultura en regiones con recursos limitados. A través de esta investigación, se comprobó que no es estrictamente necesario depender de sensores industriales costosos para obtener diagnósticos fitosanitarios útiles; el uso de imágenes RGB provenientes de drones comerciales, procesadas mediante arquitecturas YOLOv8, ofrece una alternativa robusta y accesible para el productor promedio.

Los resultados obtenidos validan la lógica de diseño del Producto Mínimo Viable (MVP). A pesar de las restricciones en los tiempos de entrenamiento, el sistema demostró una capacidad notable para identificar patologías críticas en el cultivo de maíz y gestionar de manera eficiente fotografías de alta resolución. La integración de alertas georeferenciadas sienta las bases para una transición real hacia la aplicación localizada de agroquímicos, lo que se traduce en una reducción directa de costos operativos y un menor impacto ambiental en los suelos.

Trabajo Futuro

Para consolidar el impacto de AgroCrop y evolucionar más allá de este prototipo inicial, se proponen las siguientes líneas de trabajo enfocadas en la optimización del software:

1. Optimización del Modelo de Detección: Se contempla realizar ciclos de entrenamiento extendidos (superando las 100 épocas) y aplicar técnicas de limpieza profunda en el etiquetado de los datasets. Esto permitirá elevar la precisión (mAP) y reducir los falsos positivos causados por sombras o maleza, extrayendo el máximo potencial del hardware RGB actual.
2. Procesamiento en el Borde (Edge Computing): Investigar la integración del modelo en dispositivos móviles o estaciones de tierra portátiles. El objetivo es permitir que el diagnóstico se realice de forma offline directamente en la parcela, eliminando la dependencia de una conexión estable a la nube para la toma de decisiones inmediata.
3. Diversificación del Catálogo Fitosanitario: Expandir la base de conocimientos del sistema para incluir otros cultivos estratégicos de la región, como el jitomate, plátano y pepino. Al diversificar las capacidades de detección, AgroCrop puede transformarse en una plataforma integral de monitoreo para la agricultura policultural.



REFERENCIAS

- Andreadis, A., Giambene, G. y Zambon, R. (2025). Design and analysis of a LoRa-based system with scheduled transmissions from IoT nodes to UAV in rural areas. *Ad Hoc Networks*, 175, 103868.
- Carmona-Frausto, J.C., Cervantes-Álvarez, S., Bee-Cruz, K.E., Mexicano-Santoyo, A. y Montoya-Pineda, N.M. (2024). Desarrollo de una herramienta para la detección de maleza en campos de girasoles. *TecnoINTELECTO*, 21(2), pp. 10-18.
- García, J. (2023). Pandas: Powerful data analysis toolkit. [Software] Disponible en: <https://pandas.pydata.org/>
- Harris, C.R. et al. (2020). Array programming with NumPy. *Nature*, 585(7825), pp. 357-362.
- Instituto Mexicano de Tecnología del Agua [IMTA] (2019). Los drones, alternativa de bajo costo para monitorear zonas agrícolas. [En línea] Disponible en: <https://www.gob.mx/imta/articulos/los-drones-alternativa-de-bajo-costo-para-monitorear-zonas-agricolas>
- Itseez (2024). Open Source Computer Vision Library (OpenCV). [Software] Disponible en: <https://opencv.org/>
- Mandujano-Bueno, A., Buenrostro-Rodríguez, J.F. y Montero-Tavera, V. (2025). Asociación de índices de vegetación RGB con rendimiento y dosis de nitrógeno en trigo harinero. *CienciaUAT*, 19(2).
- Montero Tavera, V. y Mandujano Bueno, A. (2025). Índices de vegetación de luz visible para una agricultura moderna. *INIFAP*. [En línea] Disponible en: <https://www.gob.mx/inifap/articulos/indices-de-vegetacion-de-luz-visible-para-una-agricultura-moderna>
- Ruiz Castañares, K. (2023). Evaluación de la cobertura vegetal en unidades de producción agrícola con vehículos aéreos no tripulados y sistemas de información geográfica. [Tesis de Ingeniería, Universidad Nacional Autónoma de México].
- Sosa-Pérez, G., Ochoa-Rivero, J.M., Álvarez-Holguín, A. y Hermosillo Rojas, D.E. (2022). Supervisión de cultivos agrícolas con drones: guía para su uso a pequeña escala. Folleto Técnico Núm. 79. *INIFAP*, México.
- Streamlit (2024). Streamlit: The fastest way to build and share data apps. [Software] Disponible en: <https://streamlit.io/>
- Tenorio, F., Juárez, E., Pérez, M. y Tenorio, L. (2018). Dron detector de plagas en cultivos de maíz. *Revista de Ingeniería Tecnológica*, 2(6), pp. 8-13.
- Urbina Cienfuegos, S.M. y Bravo Rivas, J.J. (2025). Inteligencia Artificial: Machine Learning para detección temprana de plagas y enfermedades de cultivos básicos, Nicaragua. *Revista Ingenio*, 8(1), pp. 24-34.
- Abera, D. (2024). Maize Dataset (v1). Roboflow Universe. [En línea] Disponible en: <https://universe.roboflow.com/daniel-abera/maize-10gix>
- Fur-Fun. (2024). Afectaciones en cultivos de maíz (v2). Roboflow Universe. [En línea] Disponible en: <https://universe.roboflow.com/fur-fun/afectaciones/dataset/2>
- Home-KGPKY. (2024). Maize Health Monitoring (v2). Roboflow Universe. [En línea] Disponible en: <https://universe.roboflow.com/home-kgpky/maize-s3s8a/dataset/2>
- Ultralytics. (2023). YOLOv8: Real-time Object Detection and Predictive Analytics. [Software] Disponible en: <https://github.com/ultralytics/ultralytics>
- Ram Krishna, Q. (2024). Corn Leaf Infection Dataset. Kaggle. [En línea] Disponible en: <https://www.kaggle.com/datasets/qramkrishna/corn-leaf-infection-dataset>
- Dryad Data Repository. (2024). Dataset for: Maize leaf disease detection. [En línea] Disponible en: <https://datadryad.org/dataset/doi%3A10.5061/dryad.r2280gbcg>
- AAIT-LNTZS. (2024). Maize detection dataset. Roboflow Universe. [en línea] Disponible en: <https://universe.roboflow.com/aait-lntzs/maize-detection-7j880>
- IMAMK. (2024). Maize and weed dataset. *Roboflow Universe*. [en línea] Disponible en: <https://universe.roboflow.com/new-workspace-imamk/maize-4zbtj>

SARA Desktop: Governed Memory and Safe Local Actions for a Persistent Desktop Companion

Pedro Ibarra-Facio* Walter A. Mata-López*
José Luis Álvarez-Flores* José Antonio Benavides-Ortega*
F. M. Maciel-Barboza* Mónica Cobián-Alvarado**

* Facultad de Ingeniería Mecánica y Eléctrica, Universidad de Colima,
Km 9 Carretera Colima-Coquimatlán, 28400 Coquimatlán, Col., MX
(e-mail: wmata@ucol.mx).

** Facultad de Telemática, Universidad de Colima, Avenida Universidad
333, Las Víboras, 28040 Colima, Col., MX.

Abstract: Developing a Large Language Model (LLM) based desktop companion with cross-session persistence presents significant architectural and security challenges. Such systems must systematically manage state retention, periodically discard obsolete knowledge, and mitigate the risks of executing potentially unsafe local actions. This paper describes *SARA Desktop*, a local-first desktop companion that ties together a governed memory layer, versioned snapshots, and a whitelisted action system that asks the user before doing anything sensitive. The prototype runs as a Windows sidebar written in Go, paired with a Python service, and relies on SQLite for persistent memory, versioned JSON snapshots for state, and Ollama for local model inference. Evaluation uses a reproducible, component-level benchmark that exercises the real modules, simulating only memory extraction to isolate control logic. Across this benchmark, SARA Desktop quarantines secrets with $F1 = 0.89$ on a twelve-item corpus of keyword, structural, and entropy variants, keeps low-risk memories stable through idle consolidation, restores prior versions intact across chained snapshots, and gates sensitive local actions behind explicit user approval.

Keywords: Intelligent Autonomous Systems; Human-in-the-Loop; Knowledge-Based Systems; Artificial Intelligence; Safety; Software Engineering For Real-Time Systems.

1. INTRODUCCIÓN

Los modelos de lenguaje de gran tamaño (LLM, por sus siglas en inglés) se utilizan cada vez más como asistentes que hacen algo más que responder preguntas aisladas. Crean resúmenes, planifican, recuperan contexto y proponen el uso de herramientas a lo largo de flujos de trabajo extensos. Al pasar de un chat a un compañero de escritorio, el problema deja de ser sólo generar una buena respuesta: el sistema debe conservar estado entre sesiones, actuar sobre archivos locales y seguir siendo comprensible cuando la memoria está incompleta, equivocada o contaminada. Este desplazamiento coincide con el régimen de sistemas de IA compuestos (Chen et al., 2025), donde recuperadores, memorias, orquestadores y herramientas rodean al modelo.

Este artículo estudia *SARA Desktop*, un prototipo local-first que mantiene la memoria y las acciones bajo políticas explícitas. El diseño evita que una salida textual del LLM se convierta directamente en un efecto secundario: las acciones sensibles requieren aprobación, los candidatos de memoria se filtran antes de consolidarse y los cambios persistentes quedan asociados a snapshots versionados. La contribución es práctica y arquitectónica: se documenta una implementación funcional, se formalizan las decisiones básicas de gobernanza y permisos, y se reporta una

evaluación sintética reproducible que ejercita memoria, rollback y acciones sobre los módulos reales del sistema.

1.1 Trabajos relacionados

La memoria persistente permite que un asistente deje de tratar cada sesión como un intercambio aislado, pero también convierte el historial del usuario en una superficie que debe administrarse. Generative Agents muestra cómo la memoria, la reflexión y la planificación sostienen agentes interactivos creíbles (Park et al., 2023); MemGPT propone leer la memoria de largo plazo como una extensión del contexto activo, apoyada en una analogía con la paginación de los sistemas operativos (Packer et al., 2023); y MemOS formula una idea cercana desde una perspectiva más sistémica, donde la memoria tiene representaciones, ubicaciones y ciclos de vida diferenciados (Li et al., 2024). SARA Desktop toma de estos trabajos la necesidad de separar memoria activa, almacenamiento persistente y mantenimiento fuera del turno principal, pero evita tratar el almacén como un recurso opaco. El motivo es práctico: Xiong et al. (2024) muestran que agregar o borrar memorias de forma ingenua puede contaminar recuperaciones posteriores y hacer que errores pasados reaparezcan como contexto confiable. Por eso el prototipo usa estados visibles, confianza, sensibilidad, cuarentena y snapshots versionados, de modo que una

memoria no sea sólo texto recuperable sino una entrada con historia y posibilidad de reversión.

El otro aspecto del problema surge cuando el asistente no sólo recuerda, sino que intenta actuar. ReAct combina razonamiento y acciones para que el modelo ajuste su trayectoria a partir de observaciones del entorno (Yao et al., 2023); Toolformer muestra que un modelo puede aprender cuándo invocar herramientas (Schick et al., 2023); y Gorilla estudia la selección confiable de API con apoyo de recuperación documental (Patil et al., 2024). Ese enfoque habilita sistemas más útiles, pero en un escritorio local desplaza la pregunta hacia el control: qué puede ejecutarse, con qué permisos y con qué rastro para el usuario. AgentBench y OSWorld ilustran la dificultad de evaluar agentes en tareas interactivas y entornos de uso de computadora (Liu et al., 2024; Xie et al., 2024); SARA Desktop se queda deliberadamente antes de ese nivel de autonomía y evalúa componentes concretos de memoria, rollback y permisos. La cautela no es sólo metodológica. Wang et al. (2025) documentan ataques de extracción contra memorias de agentes LLM, lo que refuerza que la persistencia debe filtrarse antes de recuperarse y antes de alimentar nuevas acciones. En ese punto se ubica la contribución del sistema: una lista blanca pequeña de acciones locales, aprobación humana para operaciones sensibles y un almacén de memoria que puede inspeccionarse, ponerse en cuarentena y volver a una versión anterior.

2. METODOLOGÍA

Antes de presentar la arquitectura conviene hacer explícitas las suposiciones operativas sobre las que se construye el sistema. SARA Desktop se ejecuta por completo en el equipo del usuario final; no existe un servidor central ni se comparte memoria entre usuarios distintos. El usuario es el único principal de confianza dentro de su propia instalación, es decir, la única entidad cuyas decisiones el sistema asume como autoritativas.

2.1 Modelo de sistema y amenazas

Para definir el modelo del sistema, sea $M = \{(e_i, s_i, v_i)\}_{i=1}^n$ el almacén de memoria, donde cada entrada se compone de tres elementos:

- e_i es el contenido de la memoria junto con sus metadatos asociados (nivel de confianza, utilidad estimada, sensibilidad y etiquetas descriptivas);
- s_i es el estado del ciclo de vida de esa entrada, tomado del conjunto $\{\text{efímero}, \text{consolidado}, \text{cuarentena}\}$;
- v_i es el número entero de versión en el que el par (e_i, s_i) fue registrado por última vez en el almacén.

Un *snapshot* S_v corresponde a la serialización congelada del almacén M en la versión v , y se guarda como un archivo JSON en disco. La operación de rollback consiste en restituir el estado anterior mediante $M \leftarrow S_{v-1}$, acompañada del decremento correspondiente del contador de versión actual; esto permite revertir la memoria a un punto previo conocido sin pérdida de trazabilidad.

Una acción a se representa como una tripla (τ, p, r) , donde τ es el tipo de acción, p son sus parámetros y r es el

nivel de riesgo declarado. La política de permisos π toma como entrada una acción junto con el contexto del usuario y devuelve una de cuatro decisiones posibles. De forma paralela, la política de gobernanza de memoria g toma como entrada un candidato de memoria y devuelve una transición de estado dentro del ciclo de vida. Las Ecuaciones (1) y (2) definen formalmente los rangos de g y π respectivamente:

$$\begin{aligned} g(e) &\in \{\text{descartar}, \text{efímero}, \text{consolidar}, \text{cuarentena}\}, \quad (1) \\ \pi(a, c, r, u) &\in \{\text{permitir}, \text{sugerir}, \text{aprobación}, \text{bloquear}\}. \quad (2) \end{aligned}$$

En esta notación, *descartar* y *consolidar* son decisiones de política, no estados persistidos: la primera omite el candidato y la segunda dispara una promoción posterior desde memoria efímera. Los estados almacenables quedan limitados a *efímero*, *consolidado* y *cuarentena*.

En cuanto al modelo de amenazas, el diseño de SARA Desktop se orienta a mitigar tres vulnerabilidades concretas que surgen en este tipo de despliegues:

- (T1) *Fuga de secretos hacia la memoria de largo plazo.* El usuario puede mencionar credenciales, tokens o información sensible en el curso normal de una conversación. Un mecanismo de extracción ingenuo almacenaría ese contenido sin filtrado y lo expondría más adelante a través de recuperaciones automáticas o resúmenes generados por el modelo.
- (T2) *Envenenamiento y deriva de la memoria.* Si se admiten sin control extracciones de baja confianza o que entran en conflicto con información previamente consolidada, el almacén puede ir acumulando errores que se refuerzan entre sí. Este tipo de degradación corresponde a las fallas de seguimiento de la experiencia documentadas por Xiong et al. (2024), donde un registro inicialmente equivocado termina contaminando decisiones posteriores.
- (T3) *Efectos secundarios no intencionados de las acciones sugeridas.* El modelo puede proponer operaciones de escritura, apertura o lectura sobre rutas del sistema de archivos que quedan fuera de las raíces permitidas. Si estas sugerencias se ejecutaran sin mediación, se rompería una suposición fundamental del usuario: que la barra lateral del asistente no puede modificar rutas arbitrarias del equipo.

Quedan explícitamente fuera del alcance del modelo de amenazas los siguientes escenarios: el compromiso a nivel del sistema operativo (por ejemplo, un usuario malicioso con acceso directo a la *shell* del equipo), los ataques dirigidos contra el *runtime* de Ollama sobre el que se ejecuta el modelo, y los canales laterales que pudieran existir a través de los pesos internos del LLM subyacente. En otras palabras, el diseño asume que el usuario confía tanto en su propia máquina como en el modelo que eligió instalar, pero no asume que las salidas transitorias del LLM sean precisas ni seguras por defecto.

2.2 Arquitectura del prototipo

SARA Desktop está implementado como una barra lateral nativa de Windows escrita en Go, acompañada de un backend local escrito en Python. La arquitectura se presenta con una vista compacta inspirada en el modelo C4: usuario y host local como contexto, barra lateral y servicio Python como contenedores principales, y componentes internos

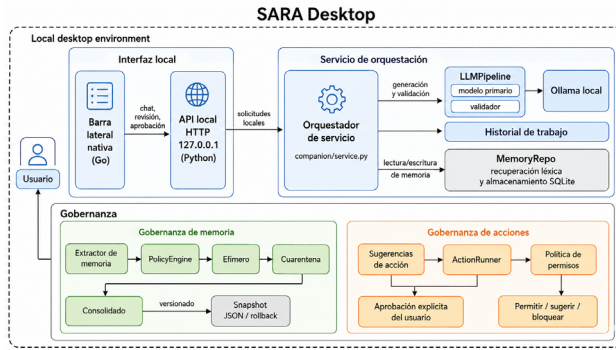


Figura 1. Arquitectura local de SARA Desktop y coordinación entre interfaz, orquestación, memoria gobernada y acciones sujetas a permisos.

Tabla 1. Superficie local representativa de la API.

Grupo	Endpoint representativo	Papel en el sistema
Observación	<code>/dashboard</code>	Refresca estado de memoria y acciones sin mutar datos.
Interacción	<code>/chat</code>	Coordina recuperación, LLM, sugerencias y extracción de memoria.
Gobernanza	<code>/memory/consolidate</code> ; <code>/memory/rollback</code> ; <code>/actions/<id>/aproveer</code> o acciones, con trazabilidad explícita.	Ejecuta cambios versionados sobre memoria o acciones, con trazabilidad explícita.

para inferencia, memoria, política y acciones. El backend expone una API HTTP en loopback, invoca modelos servidos por Ollama, persiste memoria en SQLite y guarda snapshots JSON para rollback. La Figura 1 muestra la organización local del prototipo.

La interacción atraviesa cuatro caminos: generación de respuesta con el LLM, recuperación léxica de contexto, extracción/gobernanza de memoria y evaluación de acciones sugeridas. Los dos primeros pertenecen a la ruta caliente del turno conversacional; la consolidación y los snapshots se ejecutan fuera de esa ruta, y la aprobación de acciones sólo aparece cuando el modelo propone un efecto local. Para distinguir sus propósitos, cabe notar que mientras la Figura 1 ilustra la arquitectura funcional y el flujo de coordinación, la Figura 3 ofrece una vista estructural inspirada en C4 y la trazabilidad hacia la implementación.

A nivel de la superficie de API local, el frontend se comunica con el backend mediante una API HTTP vinculada a 127.0.0.1; no hay puerto expuesto a la red ni servidor central. La Tabla 1 muestra la superficie representativa. Las rutas de gobernanza son las únicas que mutan memoria o acciones, y cada mutación queda asociada a un log o snapshot.

En cuanto a la memoria gobernada, la extracción en SARA Desktop sigue un esquema híbrido. El sistema intenta primero una extracción estructurada basada en el LLM (solicitándole al modelo que identifique hechos relevantes en el intercambio) y, cuando esa vía falla o produce

resultados vacíos, recurre a un conjunto de heurísticas específicas que capturan recordatorios explícitos del usuario, preferencias personales y restricciones de flujo de trabajo. Cada candidato extraído se pasa después al método `PolicyEngine.govern_memory_candidate`, implementado en `policy.py`, que aplica la función de gobernanza $g(e)$ definida formalmente en la Ecuación (1).

Las decisiones que toma la política siguen una lógica escalonada. Los candidatos vacíos pero con utilidad estimada por debajo de 0.25 se descartan de inmediato, ya que no aportan información recuperable. El contenido marcado como altamente sensible, con baja confianza, en conflicto con memoria previamente consolidada, o que evidentemente contiene secretos (credenciales, tokens, identificadores personales) se envía a cuarentena, donde queda retenido hasta que el usuario lo revise manualmente. Los candidatos que superan estos filtros pero aún no alcanzan el umbral de consolidación se almacenan como efímeros y quedan disponibles como contexto para sesiones inmediatas, aunque no aún elegibles para promoción automática.

La consolidación propiamente dicha no ocurre en el camino caliente de chat, sino que la ejecuta un *worker* en reposo independiente. Este *worker* recorre las memorias efímeras de bajo riesgo cuya confianza supera el umbral `memory_promotion_confidence` (configurado en 0.72) y cuya etiqueta de sensibilidad no sea alta, las promueve al estado consolidado, genera un resumen de baja sensibilidad del contenido promovido y registra un *snapshot* versionado del almacén completo. Los *snapshots* se escriben en disco como archivos JSON, mientras que la versión actual del almacén se mantiene en los metadatos de SQLite para permitir consultas rápidas. La operación de *rollback* restaura un *snapshot* anterior completo, en lugar de limitarse a deshacer la última operación: esto permite revertir efectos compuestos (varias consolidaciones encadenadas, por ejemplo) en un sólo paso, a costa de una granularidad más gruesa.

Esta separación mantiene el mantenimiento de memoria fuera del turno principal de conversación y reduce la posibilidad de que una extracción ruidosa se vuelva contexto permanente sin revisión. Al mismo tiempo, aborda directamente las fallas de propagación de errores y reproducción desalineada documentadas por Xiong et al. (2024), cuyos experimentos muestran que la adición y eliminación no controladas de memoria son suficientes, por sí solas, para degradar el comportamiento de un agente a largo plazo.

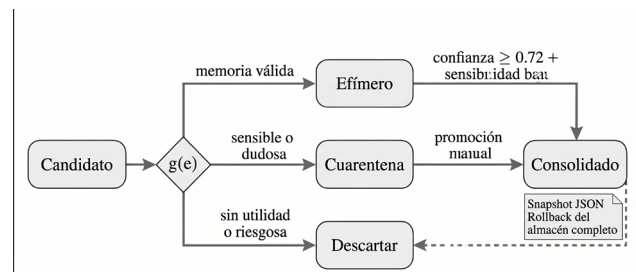


Figura 2. Ciclo de vida de la memoria gobernada. La promoción automática requiere confianza suficiente y sensibilidad no alta; la cuarentena conserva una ruta manual de promoción o eliminación.

Tabla 2. Acciones soportadas y decisión por defecto aplicada por el PolicyEngine. La columna *Riesgo* refleja el nivel base antes de considerar el contexto específico (por ejemplo, si la ruta solicitada cae dentro o fuera de las raíces permitidas).

Acción	Tipo	Riesgo	Decisión por defecto
<code>list_directory</code>	Lectura	Bajo	Permitir
<code>search_files</code>	Lectura	Bajo	Permitir
<code>create_note</code>	Escritura	Medio	Requerir aprobación
<code>append_journal</code>	Escritura	Medio	Requerir aprobación
<code>open_path</code>	Apertura externa	Alto	Requerir aprobación
(cualquier otra)	—	—	Bloquear

La Figura 2 resume el ciclo de vida: ninguna memoria llega al almacén de largo plazo sin filtros automáticos o intervención explícita.

Respecto al control de acciones con permisos, la capa es deliberadamente pequeña. El prototipo soporta únicamente cinco operaciones locales: `list_directory`, `search_files`, `create_note`, `append_journal` y `open_path`. Esta restricción es una decisión de diseño, no una limitación temporal: ampliar el catálogo sin ampliar proporcionalmente la política de permisos correspondiente rompería la garantía de que toda acción ejecutable es trazable a una decisión de política explícita.

Cada sugerencia emitida por el modelo se interpreta por la capa de política antes de cualquier ejecución. Las acciones de sólo lectura dentro de las raíces permitidas por la configuración del usuario pueden autorizarse automáticamente cuando su riesgo estimado es bajo; cuando ese riesgo supera el umbral de la Tabla 3, la decisión se degrada a *sugerir* y la acción sólo se muestra en la barra lateral sin ejecutarse. Las operaciones de escritura —tanto la creación de notas como la escritura en el diario— y las que abren archivos externos requieren aprobación humana explícita a través de la barra lateral, con un piso de riesgo de 0.65 heredado de la configuración (Tabla 3). Las operaciones que no figuran en la lista blanca se bloquean sin posibilidad de aprobación. La Tabla 2 resume el mapeo por defecto entre tipo de acción y decisión de política.

Además de la decisión de política, el **ActionRunner** aplica una comprobación adicional de seguridad a nivel de ruta del sistema de archivos antes de ejecutar cualquier operación aprobada. Esto significa que la aprobación del usuario no elude las restricciones de raíz configuradas: si el usuario aprueba por error una acción que intenta escribir fuera de las raíces permitidas, la comprobación de ruta la rechaza igualmente. La política de permisos y la comprobación de ruta actúan así como dos filtros independientes en serie.

Este diseño separa deliberadamente la salida del modelo de los efectos secundarios sobre el equipo del usuario. El asistente puede *sugerir* una acción como parte de su respuesta textual, pero la ejecución real depende de la decisión de la política y, en los casos sensibles, de una confirmación explícita del usuario en la barra lateral.

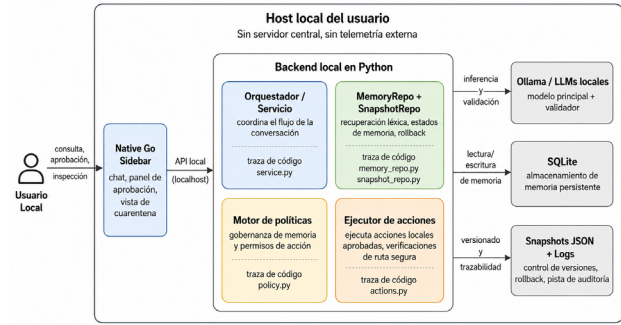


Figura 3. Vista de contenedores y componentes del host local. La figura vincula la interfaz Go, el backend Python, los repositorios de memoria, los snapshots y las políticas con sus rutas principales de implementación.

Tabla 3. Umbrales predeterminados de gobernanza y acción.

Ajuste	Valor	Significado
<code>idle_consolidation_seconds</code>	60	Quietud antes de considerar consolidación.
<code>memory_min_confidence</code>	0.55	Por debajo de este valor, los candidatos se ponen en cuarentena.
<code>memory_promotion_conf.</code>	0.72	Confianza requerida para promover de efímero a consolidado.
Utilidad mín. para almacenamiento	0.25	Por debajo de este valor, los candidatos se descartan.
Riesgo piso de acción sensible	0.65	Riesgo mínimo asignado a acciones de escritura/apertura bajo aprobación.
Lote de consolidación	≤ 25	Máximo de candidatos efímeros examinados por ciclo.

El contraste con enfoques de razonamiento y uso de herramientas guiados por el modelo (Yao et al., 2023; Schick et al., 2023) es intencional: en SARA Desktop, la capa de política —y no el modelo— es el punto autoritativo de decisión sobre qué se ejecuta y qué no.

2.3 Configuración experimental y ajustes de despliegue

Para el entorno de referencia, las pruebas se ejecutaron en un host local con Windows 11 Pro (x86_64), Python 3.11, backend HTTP en 127.0.0.1:8765, barra lateral en Go y modelos servidos por Ollama en 127.0.0.1:11434. El benchmark sintético no mide latencia de inferencia ni consumo de memoria del LLM, ya que simula la extracción estructurada para aislar la lógica de control.

En cuanto a la configuración recomendada, la Tabla 3 lista los umbrales usados en la evaluación sintética. Se eligieron para poner secretos explícitos en cuarentena, evitar que candidatos de baja confianza se consoliden y mantener las acciones sensibles detrás de aprobación humana.

Todo el estado persistente se escribe bajo una raíz local: SQLite para memoria, JSON para snapshots, logs locales y archivos producidos por acciones aprobadas. Las acciones



de lectura se restringen a raíces permitidas y las rutas externas escalan a aprobación.

3. RESULTADOS Y DISCUSIÓN

Para evaluar la lógica de control implementada sin depender de la disponibilidad del modelo, se implementó un benchmark reproducible a nivel de componente en `paper_synthetic_eval.py`. El script instancia los módulos reales de almacenamiento, política, consolidación, snapshot, rollback y ejecución de acciones sobre un espacio de trabajo local temporal. La extracción de memoria del LLM se simula con una respuesta estructurada vacía, de modo que la evaluación ejercita el camino heurístico y la lógica de ciclo de vida subsiguiente de forma determinista. A partir de los resultados obtenidos, se discuten posteriormente las capacidades reales que proporciona el diseño actual, junto con sus limitaciones inherentes y el impacto esperado en la confianza y experiencia del usuario.

3.1 Evaluación sintética

El protocolo de la evaluación consta de tres etapas. La Etapa A prueba candidatos de memoria explícitos, incluyendo un hecho técnico, una preferencia de usuario, una restricción de flujo de trabajo relacionada con aprobación y una declaración que contiene un secreto. La Etapa B agrega memorias sintéticas promocionables y no promocionables, ejecuta una primera consolidación, luego crea un segundo estado consolidado y hace rollback a la versión anterior. La Etapa C envía siete sugerencias de acción representativas a la política de permisos y ejecuta las dos acciones de lectura de bajo riesgo.

La Etapa A ejercita cuatro candidatos de memoria representativos cuya etiqueta esperada se sigue directamente de las reglas de gobernanza formalizadas por la Ecuación (1) en la Sección 2.1 y detalladas en la Sección 2.2. Los tres primeros candidatos —un hecho técnico (“el proyecto usa SQLite”), una preferencia de usuario (“prefiero temas oscuros”) y una restricción de flujo de trabajo (“siempre preguntar antes de abrir archivos”)— deberían permanecer como efímeros tras pasar por la gobernanza, y en los tres casos la ejecución en vivo observa exactamente ese estado. El cuarto candidato —la declaración “mi token secreto es ABC123”— es el caso más difícil: un extractor ingenuo lo trataría como un hecho benigno. La capa de gobernanza detecta la palabra clave sensible y enruta el candidato a cuarentena antes de que pueda entrar en consolidación, que es el comportamiento que el modelo de amenazas (T1) requiere.

La Tabla 4 resume el ciclo de vida de la memoria. La observación más importante es tanto cualitativa como cuantitativa: el secreto explícito “mi token secreto es ABC123” se pone en cuarentena, se conserva para revisión y se excluye del resumen de consolidación. Esto importa porque versiones anteriores del camino heurístico habrían tratado tales recordatorios como hechos benignos. Tras endurecer las reglas de gobernanza, la ejecución sintética confirma que el contenido similar a secretos ya no se propaga a los resúmenes consolidados.

El comportamiento del rollback también es significativo. La primera consolidación crea la versión 1. Una segunda consolidación crea la versión 2 con dos memorias estables

Tabla 4. Evolución del estado de memoria en la evaluación sintética.

Etapas	Ver.	Cons.	Cuar.	Observación
Tras Etapa A	0	0	1	Secreto explícito enrutado a cuarentena; tres memorias benignas permanecen efímeras.
Tras snapshot 1	1	6	2	Cinco memorias promocionables más un resumen quedan consolidados.
Tras snapshot 2	2	9	2	Se agregan dos memorias adicionales y un nuevo resumen.
Tras rollback	1	6	2	La versión 1 es restaurada y las memorias de la etapa B desaparecen.

Tabla 5. Resultados de la política de permisos para siete casos sintéticos de acción.

Escenario	Decisión	Resultado observado
Listar directorio de espacio de trabajo	Permitir	Ejecutado con éxito.
Buscar archivos en el espacio de trabajo	Permitir	Ejecutado con éxito.
Lectura fuera de raíces permitidas	Requerir aprobación	No auto-ejecutado; la restricción de raíz sigue activa.
Solicitud de búsqueda de alto riesgo	Sugerir	Presentada sin ejecución.
Crear nota	Requerir aprobación	Retenida para confirmación humana explícita.
Abrir archivo local	Requerir aprobación	Retenida para confirmación humana explícita.
Acción de eliminación no soportada	Bloquear	Rechazada por la política de lista blanca.

adicionales. Cuando se invoca el rollback, el almacén restaura la versión 1 y elimina por completo las adiciones de la etapa B. Esto está más cerca de la semántica de rollback que necesita un compañero persistente que la simple recarga del snapshot más reciente.

La Tabla 5 muestra que la política de acciones no es ni puramente permisiva ni puramente bloqueante. De siete casos, dos son permitidos automáticamente, cuatro son escalados o degradados a sugerencia, y uno es bloqueado por completo. Esto es consistente con el objetivo de la implementación: preservar acceso de baja fricción para acciones de lectura acotadas mientras se mantienen las operaciones locales potencialmente disruptivas detrás de puertas de aprobación.

Para la detección de secretos (benchmark sintético), La cuarentena del token ABC123 en la Etapa A se complementó con un corpus sintético de doce enunciados, deliberadamente desbalanceado hacia la clase positiva: ocho secretos y cuatro no-secretos. El conjunto cubre palabras clave explíc-



Tabla 6. Matriz de confusión del detector de secretos en cascada sobre el corpus sintético de doce enunciados (clase positiva = “secreto”, se espera *cuarentena*; corpus desbalanceado 8/4 por diseño).

	Predicho: secreto	Predicho: no-secreto
Real: secreto	8 (TP)	0 (FN)
Real: no-secreto	2 (FP)	2 (TN)

Precisión = 0.80, Recall = 1.00, F1 = 0.89, Exactitud = 0.83.

Tabla 7. Rollbacks encadenados con elementos en cuarentena activos. “Resid.” indica claves residuales de la versión superior detectadas tras el rollback (debería ser cero en todos los pasos).

Etapas	Ver.	Cons.	Cuar.	Resid.
Tras v1	1	3	1	—
Tras v2	2	5	2	—
Tras v3	3	8	2	—
v3→v2	2	5	1	0 (rb.v3.*)
v2→v1	1	3	0	0 (rb.v2.*, rb.q*)
v1→v0	1	3	0	sin snapshot previo

itas, patrones estructurales o numéricos y cebos léxicos que contienen términos sensibles sin exponer información privada. El detector, implementado en `companion/policy.py`, aplica capas léxicas, patrones de credenciales, validación Luhn para tarjetas, reglas OTP/PIN y alta entropía; el código del benchmark está en `run_secret_detection_benchmark`.

La Tabla 6 resume los resultados: no hay falsos negativos sobre el corpus, y los dos falsos positivos corresponden a cebos léxicos intencionales. Esto fija el alcance de T1 de forma conservadora: el detector cubre fugas léxicas, estructurales y numéricas, pero sobre-cuarentena enunciados inocuos cuyo vocabulario coincide con términos sensibles. Esa decisión es aceptable para este prototipo porque la cuarentena puede revisarse manualmente; quedan fuera los secretos que sólo son privados por su denotación semántica.

Rollback encadenado y residuos de cuarentena La evaluación inicial sólo ejercita el flujo simple de un único rollback (v1, luego v2, luego retorno a v1). Para verificar que la semántica de *snapshot completo* no deja residuos en escenarios más adversariales, se añadió el banco `run_chained_rollback_test`, que construye tres versiones consolidadas v1, v2, v3 intercaladas con elementos en cuarentena introducidos después de cada consolidación, y ejecuta tres rollbacks consecutivos.

La Tabla 7 muestra que cada rollback restituye exactamente el snapshot inmediatamente anterior, sin dejar memorias consolidadas ni elementos de cuarentena de versiones posteriores. En particular, los elementos en cuarentena introducidos entre snapshots (rb.q1, rb.q2) se eliminan junto con las consolidaciones de su versión, lo que confirma que la cuarentena no constituye una vía lateral que sobreviva al rollback. El último paso (v1→v0) demuestra el comportamiento de borde: cuando no existe snapshot anterior, el sistema permanece en v1 en lugar de invalidar el estado, lo que evita dejar el almacén en una versión indefinida.

Tabla 8. Resultados macro de la recuperación léxica sobre $n = 7$ consultas etiquetadas, ejecutadas tras el rollback a la versión 1.

Métrica	Valor (macro)
Precisión en top-1	0.71
Precisión en top-3	0.48
Recall en top-5	1.00
Fuga de secretos en consulta sensible	No (0/5 hits)

Recuperación léxica: benchmark etiquetado Reemplaza la inspección cualitativa de la versión previa por un conjunto de siete consultas etiquetadas (`run_retrieval_benchmark`) ejecutadas sobre el almacén restaurado tras el rollback. Cada consulta declara explícitamente las claves `mem_key` consideradas relevantes y se reporta precisión en top-1, top-3 y *recall* en top-5. Además, una octava consulta actúa como *prueba de fuga*: utiliza vocabulario sensible (*token*, *secreto*, *credenciales*) y se considera exitosa sólo si ninguno de los hits surge contenido secreto.

La Tabla 8 sitúa el desempeño de la recuperación léxica en términos cuantitativos. La precisión en top-1 de 0.71 indica que cinco de las siete consultas etiquetadas devuelven la memoria más relevante en primera posición; la caída a 0.48 en top-3 y el *recall* de 1.00 en top-5 son consistentes con un recuperador basado en superposición de tokens, donde el conjunto recuperado tiende a incluir todas las memorias relevantes pero diluye su orden con resultados parcialmente coincidentes. La prueba de fuga confirma que ningún hit para vocabulario sensible filtra contenido del almacén cuarentenado: el secreto sigue retenido después del rollback, en concordancia con la propiedad enunciada en la Sección 3.1. Estos resultados constituyen el verdadero punto de comparación para futuras mejoras basadas en recuperación aumentada o *embeddings* (Lewis et al., 2020) (Sección 3.2.2), no una afirmación de paridad semántica.

En términos de reproducibilidad, La evaluación sintética está dirigida por un único script, `paper_synthetic_eval.py`, que instancia `CompanionConfig`, `MemoryStore`, `PolicyEngine`, `ConsolidationWorker` y `ActionRunner` sobre un espacio de trabajo temporal. La extracción de memoria del LLM se simula con una respuesta estructurada vacía de modo que el camino heurístico y la lógica de ciclo de vida subsiguiente se ejerciten de forma determinista; no se realiza ninguna llamada de red externa durante el benchmark. Ejecutar el script toma menos de un minuto y produce los conteos de versión, consolidación, rollback y decisión de acción reportados en las Tablas 4 y 5, junto con la matriz de confusión (Tabla 6), las trazas de rollback encadenado (Tabla 7) y las métricas de recuperación etiquetada (Tabla 8). La salida completa se serializa a `ifacconf_latex-1.4/generated/synthetic_eval_results.json`; la base SQLite y los snapshots se escriben en un directorio temporal por invocación y se eliminan al salir.

Sobre la cobertura del modelo de amenazas, La evaluación se cierra mapeando las tres amenazas de la Sección 2.1 a los resultados observados. T1 (fuga de secretos) se aborda en dos planos complementarios: (a) el detector en cascada de `policy.py` (léxico + estructural + Luhn + OTP/entropía) alcanza precisión = 0.80, recall = 1.00 y F1 = 0.89 sobre el corpus de doce enunciados (Tabla 6), sin falsos negativos



y con dos falsos positivos limitados a cebos léxicos; (b) el constructor de resúmenes rechaza fusionar contenido sensible, de modo que la Tabla 4 no muestra ningún secreto en el resumen consolidado en ninguna versión, y la prueba de fuga de la Tabla 8 confirma que las consultas con vocabulario sensible no surgen contenido cuarentenado. T2 (envenenamiento y deriva de memoria) está acotada por los umbrales de confianza y utilidad de la Tabla 3 y por el comportamiento de rollback verificado en dos planos: el rollback simple de la Etapa B y la cadena $v3 \rightarrow v2 \rightarrow v1 \rightarrow v0$ con elementos en cuarentena activos (Tabla 7), donde ningún paso deja residuos. T3 (efectos secundarios no intencionados) es visible en la Tabla 5: la única eliminación no soportada se bloquea, las tres operaciones sensibles de escritura/apertura se retienen para aprobación, y las dos lecturas de bajo riesgo son las únicas que se ejecutan sin confirmación del usuario.

Cobertura no es completitud. Los falsos positivos del detector (Tabla 6) acotan el costo operativo de T1: enunciados cuyo léxico coincide con términos sensibles entran en cuarentena aunque sean inocuos, y el usuario debe liberarlos manualmente. El corpus tampoco agota los modos posibles: no examina prompts adversariales diseñados para eludir simultáneamente las cuatro capas del detector, ni cubre secretos que resultan privados sólo por su denotación semántica (sin huella léxica, estructural ni numérica). Además, la evaluación no mide la latencia del LLM ni la carga del usuario al aprobar repetidamente operaciones sensibles. Estas son extensiones naturales, no afirmaciones presentes; la Sección 3.2.2 describe el plan para abordarlas.

3.2 Discusión

El prototipo muestra que un compañero de escritorio local puede soportar persistencia manteniendo memoria, acción y ejecución como componentes separados. La implementación proporciona estados de almacenamiento explícitos, metadatos de versión, snapshots, flujos de aprobación y un conjunto acotado de acciones inspeccionables en el código. El sistema no constituye un agente de escritorio autónomo completo, pero sirve como línea base sobre la que extender funcionalidad.

Lo que el diseño actual realmente proporciona Las tres propiedades anticipadas en la introducción —separación de texto y efecto secundario, memoria inspeccionable y recuperación con viaje en el tiempo— se siguen directamente del código tal como está y quedan respaldadas por la Sección 3.1. La separación texto/efecto es visible en la Tabla 5, donde las acciones sensibles no se ejecutan sin una llamada explícita a `/actions/<id>/approve`. La inspeccionabilidad se manifiesta en los campos persistidos (estado, confianza, sensibilidad, justificación) que alimentan tanto la Tabla 4 como el panel de cuarentena. El viaje en el tiempo queda verificado por la cadena de rollbacks encadenados de la Tabla 7, que no deja residuos en ningún paso.

Limitaciones Seis limitaciones acotan la afirmación presente. Primero, la recuperación es léxica en lugar de basada en embeddings; los números de la Tabla 8 ($P@1=0.71$, $P@3=0.48$) muestran esa brecha sin presentarla

como memoria semántica completa. Segundo, el espacio de acciones es intencionalmente minúsculo y no cubre control general de escritorio. Tercero, la evaluación es sintética y a nivel de componente; aún no mide éxito de tarea de extremo a extremo, confianza del usuario ni latencia en tiempo real bajo modelos locales en vivo. Cuarto, el prototipo actual no tiene integración formal con benchmarks, aunque una extensión natural podría conectarlo a entornos de uso de computadora como OSWorld (Xie et al., 2024). Quinto, el detector de secretos en cascada de `policy.py` alcanza $F1=0.89$ sobre el corpus de la Tabla 6 e incluye, dentro del corpus evaluado, modos numéricos (tarjeta por Luhn, OTP por contexto) además del léxico y estructural, pero sobre-cuarentena enunciados benignos cuyo vocabulario coincide con términos sensibles y no aborda secretos que son privados sólo por su denotación. Sexto, y de forma especialmente importante para interpretar los resultados, la evaluación sintética no valida el comportamiento del LLM en producción, sino únicamente la lógica de control que lo rodea: la extracción de memoria se simula con una respuesta estructurada vacía y el camino heurístico es el que se ejercita, por lo que la calidad de las decisiones del modelo primario y del validador —su capacidad real de detectar secretos no evidentes, clasificar sensibilidad o proponer acciones coherentes con la política— no queda medida. Los resultados sostienen que los mecanismos de gobernanza se comportan como se especifica cuando el LLM entrega la entrada esperada, pero no cuantifican la precisión del LLM mismo sobre prompts reales.

El despliegue local mejora el control y la localidad de los datos, pero no elimina los riesgos de seguridad a nivel de host, razón por la cual el diseño de memoria consciente de la privacidad sigue siendo esencial (Wang et al., 2025).

Confianza, carga de aprobación y experiencia de usuario

El control de acciones con permisos introduce una tensión bien conocida: cuantas más operaciones escala el sistema al usuario, más seguro es en el peor caso, pero más alto es el costo de interrupción en el caso común. Los valores predeterminados actuales de la Tabla 3 son deliberadamente conservadores y parecerán lentos a un usuario que espera ejecución automática. Dos palancas de diseño abordan esto parcialmente. Primero, la decisión *sugerir* permite al sistema surgir una intención en la barra lateral sin requerir una confirmación modal, preservando el flujo para los casos no críticos. Segundo, el consolidador en reposo se ejecuta en segundo plano y nunca interrumpe al usuario, por lo que el mantenimiento de memoria no se añade a la carga de aprobación. Un enfoque más adaptativo, como umbrales aprendidos por acción o políticas por hora del día, es un refinamiento futuro, pero requeriría su propia evaluación de modos de falla antes de ser habilitado por defecto.

Estas limitaciones acotan el alcance del artículo. El sistema no pretende resolver la memoria persistente de agentes en general ni igualar a los agentes de uso de computadora de última generación. Su contribución es mostrar que la persistencia local resulta más controlable cuando la memoria se gestiona como un ciclo de vida con cuarentena y rollback, y cuando las acciones locales se filtran por una capa de política antes de cualquier efecto secundario.

4. CONCLUSIONES

Este artículo documentó SARA Desktop, un compañero de escritorio local-first con memoria gobernada, snapshots versionados y acciones locales sujetas a permisos. La evaluación sintética, ejecutada sobre los módulos reales y con extracción de memoria simulada para aislar la lógica de control, muestra que el sistema pone en cuarentena secretos léxicos, estructurales y numéricos con $F1 = 0.89$, promueve memorias estables de bajo riesgo, restaura versiones anteriores sin residuos en rollbacks encadenados y restringe operaciones sensibles mediante aprobación explícita.

La lección principal es que la persistencia local resulta más controlable cuando la memoria y las acciones no dependen sólo de la salida textual del LLM, sino de estados inspeccionables, políticas explícitas, cuarentena y rollback. El siguiente paso no es simplemente ampliar el almacén: conviene incorporar recuperación basada en embeddings sin perder metadatos de gobernanza, evaluar utilidad de extremo a extremo con un benchmark tipo OSWorld (Xie et al., 2024), y realizar un estudio con usuarios que mida la carga de aprobación frente a la confianza percibida.

REFERENCIAS

- Chen, J., Ye, J., and Wang, G. (2025). From standalone LLMs to integrated intelligence: A survey of compound AI systems. arXiv preprint arXiv:2506.04565. doi:10.48550/arXiv.2506.04565. URL <https://arxiv.org/abs/2506.04565>.
- Lewis, P., Perez, E., Piktus, A., Petroni, F., Karpukhin, V., Goyal, N., Küttler, H., Lewis, M., Yih, W.t., Rocktäschel, T., Riedel, S., and Kiela, D. (2020). Retrieval-augmented generation for knowledge-intensive NLP tasks. In *Advances in Neural Information Processing Systems*, volume 33, 9459–9474.
- Li, Z., Song, S., Wang, H., Niu, S., Chen, D., et al. (2024). MemOS: An operating system for memory-augmented generation (MAG) in large language models. arXiv preprint arXiv:2505.22101. doi:10.48550/arXiv.2505.22101. URL <https://arxiv.org/abs/2505.22101>.
- Liu, X., Yu, H., Zhang, H., Xu, Y., Lei, X., Lai, H., Gu, Y., Ding, H., Men, K., Yang, K., Zhang, S., Deng, X., Zeng, A., Du, Z., Zhang, C., Shen, S., Zhang, T., Su, Y., Sun, H., Huang, M., Dong, Y., and Tang, J. (2024). AgentBench: Evaluating LLMs as agents. In *International Conference on Learning Representations (ICLR)*.
- Packer, C., Wooders, S., Lin, K., Fang, V., Patil, S.G., Stoica, I., and Gonzalez, J.E. (2023). MemGPT: Towards LLMs as operating systems. arXiv preprint arXiv:2310.08560. doi:10.48550/arXiv.2310.08560. URL <https://arxiv.org/abs/2310.08560>.
- Park, J.S., O'Brien, J.C., Cai, C.J., Morris, M.R., Liang, P., and Bernstein, M.S. (2023). Generative agents: Interactive simulacra of human behavior. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST)*, 1–22. doi:10.1145/3586183.3606763.
- Patil, S.G., Zhang, T., Wang, X., and Gonzalez, J.E. (2024). Gorilla: Large language model connected with massive APIs. In *Advances in Neural Information Processing Systems*, volume 38. doi:10.52202/079017-4020.
- Schick, T., Dwivedi-Yu, J., Dessì, R., Raileanu, R., Lomeli, M., Hambro, E., Zettlemoyer, L., Cancedda, N., and Scialom, T. (2023). Toolformer: Language models can teach themselves to use tools. In *Advances in Neural Information Processing Systems*, volume 36.
- Wang, B., He, W., Zeng, S., Xiang, Z., Xing, Y., Tang, J., and He, P. (2025). Unveiling privacy risks in LLM agent memory. In *Proceedings of the 63rd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 25241–25260. doi:10.18653/v1/2025.acl-long.1227.
- Xie, T., Zhang, D., Chen, J., Li, X., Zhao, S., Cao, R., Hua, T.J., Cheng, Z., Shin, D., Lei, F., Liu, Y., Xu, Y., Zhou, S., Savarese, S., Xiong, C., Zhong, V., and Yu, T. (2024). OSWorld: Benchmarking multimodal agents for open-ended tasks in real computer environments. In *Advances in Neural Information Processing Systems, Track on Datasets and Benchmarks*, volume 38.
- Xiong, Z., Lin, Y., Xie, W., He, P., Liu, Z., Tang, J., Lakkaraju, H., and Xiang, Z. (2024). How memory management impacts LLM agents: An empirical study of experience-following behavior. arXiv preprint arXiv:2505.16067. doi:10.48550/arXiv.2505.16067. URL <https://arxiv.org/abs/2505.16067>.
- Yao, S., Zhao, J., Yu, D., Du, N., Shafraan, I., Narasimhan, K., and Cao, Y. (2023). ReAct: Synergizing reasoning and acting in language models. In *International Conference on Learning Representations (ICLR)*.

Local-First Semantic Mapping with Constrained LLMs: A Concurrent Framework for Multimodal Knowledge Extraction

Leonardo Jesús Carrillo Bivián*
Walter Alexander Mata López*

* Universidad de Colima, Facultad de Ingeniería Mecánica y Eléctrica,
Carretera Colima - Coquimatlán km 9, Valle de las Huertas, 28400
Coquimatlán, Col. Campus Coquimatlán (e-mail: {lcarrillo21,
wmata}@ucol.mx)

Abstract: Processing and extracting knowledge from complex academic and corporate documents, such as PDFs and presentations, remains a significant challenge for automated systems. Traditional methods relying on optical character recognition (OCR) and shallow metadata extraction often lose structural and semantic integrity when encountering non-linear layouts, mathematical formulas, or embedded tables. Furthermore, relying on cloud-based Large Language Models (LLMs) raises critical concerns regarding data privacy and network latency. This paper presents Semantic Mapper, a robust, local-first framework developed in Go designed to automate the transformation of diverse document formats into structured conceptual maps without relying on external dependencies or network transport. The architecture integrates a multimodal checkpoint that converts binary documents into Markdown using local vision tools, subsequently processed by a quantized LLM (e.g., Phi-3) via Ollama. To effectively mitigate the common problem of hallucinations, the system enforces a strict JSON output schema, constraining the model's predictive neighborhood and forcing deterministic mapping. Empirical evaluations on an entry-level laptop (AMD 3020e, no dedicated GPU) show that the concurrent architecture achieves a mean end-to-end latency of 74.37 s ($\sigma = 6.92$ s, $n = 5$) over a 50-document corpus. This represents a speedup factor of $S = 2.86$ relative to a Python monolithic baseline (212.62 s, $\sigma = 48.29$ s). Furthermore, the enforced JSON schema yields a structural error rate of $\varepsilon_e = 0.8\%$, compared to 21.6% under unconstrained local inference. Results confirm that a constrained local LLM achieves structural stability comparable to cloud alternatives while preserving data sovereignty. We also outline enhancements to overcome context window limitations, including a cryptographic hash-based cache with BadgerDB and a concurrent 'Critical Agent' for parallel logical validation. These findings establish a solid baseline for high-performance knowledge extraction within privacy-sensitive environments.

Keywords: Local-First, LLM, Knowledge Graph, Go, Concurrency.

1. INTRODUCCIÓN Y MARCO TEÓRICO

La extracción automática de conocimiento a partir de documentos en formatos estandarizados corporativos y académicos, tales como archivos PDF o presentaciones con diseño visual complejo, representa un desafío técnico de primer orden. Estos documentos contienen con frecuencia estructuras visuales sumamente elaboradas, tablas anidadas, fórmulas matemáticas avanzadas y diseños no lineales que las técnicas tradicionales de procesamiento tienen dificultades para inspeccionar de forma coherente. Los métodos basados en metadatos superficiales y en reconocimiento óptico de caracteres (OCR) convencional suelen perder una parte considerable de la información semántica y de la estructura original, lo que resulta en una degradación irreversible de la jerarquía de los datos frente a representaciones matemáticas, tablas incrustadas o diseños no secuenciales.

En respuesta directa a estas limitaciones estructurales, Blecher et al. (2023) demostraron que las arquitecturas de aprendizaje profundo de extremo a extremo, orientadas a convertir directamente los datos visuales de los documentos en código de marcado sintáctico (Markdown), logran preservar de manera efectiva toda la topología y la organización original del documento. Este enfoque reduce significativamente la dependencia de herramientas externas de visión por computadora y resuelve un cuello de botella fundamental en el procesamiento documental digital, permitiendo que la máquina “entienda” el layout antes de procesar el texto.

Los Modelos de Lenguaje de Gran Escala (LLMs) han demostrado una capacidad destacada para procesar textos en formato Markdown y generar entidades estructuradas del conocimiento, como grafos de conocimiento o taxonomías semánticas. Diversos estudios recientes, como los de Zhu



et al. (2023), han confirmado que estos modelos operan de forma mucho más confiable cuando se utilizan bajo condiciones estrictamente restringidas, con un formato de salida predefinido y asistencia mediante ejemplos específicos (few-shot prompting). Por el contrario, cuando se les otorga libertad para generar texto de forma abierta o mediante instrucciones heurísticas sin restricciones, el riesgo de proporcionar información incoherente, falsa o incorrecta aumenta considerablemente, lo que imposibilita la creación de mapas conceptuales legítimos o representaciones funcionales precisas en entornos de producción.

1.1 El problema del escalado y la dependencia de servicios en la nube

Procesar grandes volúmenes de texto presenta retos importantes cuando se depende excesivamente de servicios en la nube, como ocurre con el paradigma predominante de Retrieval-Augmented Generation (RAG). Para evaluar el contenido general de dominios completos o realizar análisis transversales, estos sistemas requieren pasos intermedios intensivos de indexación y reducción de información, identificados en la literatura técnica como enfoques del tipo GraphRAG (Edge et al., 2024). En estos esquemas, las representaciones documentales originales deben reducirse e indexarse bajo una ontología sólida antes de integrarse en el modelo generativo para producir la respuesta final al usuario. Esta dependencia de LLMs centralizados a través de interfaces de programación (APIs) comerciales no solo impone problemas conocidos de latencia por el transporte de datos a través de la red, sino que también introduce cuestiones críticas relacionadas con la privacidad de los datos y el control sobre la información sensible procesada localmente. Además, muchos marcos de trabajo habituales en la industria de la inteligencia artificial, a menudo basados en Python, enfrentan limitaciones propias de sus entornos de ejecución, como el Global Interpreter Lock (GIL). Este bloqueo incrementa la sobrecarga cuando se busca orquestar múltiples inferencias documentales de forma paralela y simultánea, limitando el aprovechamiento de procesadores multinúcleo modernos.

En respuesta analítica a estas condiciones restrictivas, el presente trabajo expone la infraestructura base del Semantic Mapper: un pipeline local-first desarrollado íntegramente en Go que aprovecha el paralelismo nativo y la gestión eficiente de memoria del lenguaje. El propósito fundamental del sistema es procesar el flujo de lectura multimodal directamente en el equipo del usuario, eliminando la latencia de red y garantizando la soberanía de la información. Su estructura combina pasos de análisis locales y aplica reglas estrictas de formato (JSON cerrado), obligando al LLM a producir diagramas válidos en Mermaid.js. Este enfoque local evita el envío de información confidencial a servidores externos y reduce los errores de invención o alucinaciones típicos de los modelos generativos sin control.

1.2 Topología del modelo condicionado

Modelos de tamaño reducido y cuantizados, como los de la familia Phi-3 de Microsoft o arquitecturas equivalentes, indican que los mecanismos de atención en el procesamiento de inferencia no se limitan únicamente a la generación libre

del siguiente token semántico. Al condicionar fuertemente la salida mediante un esquema JSON estricto, se reduce el espacio de respuestas posibles del modelo. Esta restricción técnica orienta al LLM hacia una extracción lógica del texto original en Markdown, minimizando la influencia de los sesgos presentes en su conjunto de entrenamiento y evitando la generación de contenido inventado o incoherente.

Esta mecánica de condicionamiento fuerte del output constituye el principal mecanismo de protección y “blindaje” contra alucinaciones en el marco local presentado. Al limitar el vocabulario de salida a tokens específicos de la estructura JSON, el modelo se ve forzado a realizar un análisis de los datos de entrada sin espacio para la “creatividad” discursiva, garantizando la fidelidad de la información extraída.

La hipótesis central de este trabajo sostiene que una arquitectura de orquestación concurrente en Go, combinada con restricciones estructurales de salida (JSON Schema), supera en latencia y estabilidad estructural a los pipelines monolíticos basados en Python (baseline: Microsoft Mark ItDown Pipeline v0.1.5) en hardware de cómputo comercial sin GPU dedicada.

2. ARQUITECTURA Y METODOLOGÍA DEL SISTEMA

El desarrollo del Semantic Mapper se instituyó sobre un diseño de pipeline asíncrono con una arquitectura estrictamente local-first. Esta arquitectura de bajo nivel divide las responsabilidades operativas en serie para maximizar la velocidad de lectura de datos y preservar las variables semánticas originales del documento, sin exponer en ningún momento información privada a la red pública o a servicios de terceros.

2.1 Definición del Flujo de Trabajo (Workflows)

El sistema se organiza mediante un flujo de trabajo estructurado en etapas asíncronas que garantiza la calidad del procesamiento y la integridad de los datos. Este flujo sigue un patrón repetible y determinista compuesto por cuatro etapas principales que se ejecutan de forma secuencial:

- **Auditoría de Lectura:** el orquestador analiza el formato del documento de entrada para determinar la mejor ruta de procesamiento.
- **Delegación de Tareas:** si el archivo es texto simple, se lee directamente; si se trata de un formato binario o imagen, se utilizan habilidades especializadas de visión por computadora.
- **Limitación Cognitiva:** el texto procesado se envía al modelo LLM con etiquetas de instrucción estrictas para que responda únicamente en formato JSON cerrado y válido.
- **Validación Estructural:** se verifica mediante algo ritmos de parsing que no existan variables huérfanas o errores de sintaxis antes de generar la representación visual final.

El entorno de validación correspondió a un equipo HP 255 G8 con procesador AMD 3020e (2 núcleos, 1.2 GHz base / 2.6 GHz boost), 16 GB DDR4 (2400 MHz), SSD NVMe 256 GB y GPU integrada AMD Radeon (sin aceleración

dedicada). El stack tecnológico utilizado fue: Go v1.22.2 (windows/amd64), Ollama v0.1.32, Python 3.11.9 (base line), Microsoft MarkItDown v0.1.5 y BadgerDB v4.2.0. El modelo Phi-3-Mini se configuró con cuantización 4-bit (Q4 K M), temperatura 0.0, ventana de contexto de 4,096 tokens, Top p = 0.9, Top k = 40 y modo de salida JSON nativo de Ollama.

2.2 Aislamiento del Checkpoint Multimodal

El control inicial y la gestión logística del procesamiento de archivos se delegan a un orquestador principal desarrollado y compilado bajo Go. Este componente asume el rol de discriminar las rutas de procesamiento según el tipo de archivo detectado. El diseño previene operativa mente la sobrecarga del sistema mediante un conmutador semántico encapsulado, implementado en la función de control ParseDocument.

La lógica de decisión binaria se detalla en la Fig. 1.

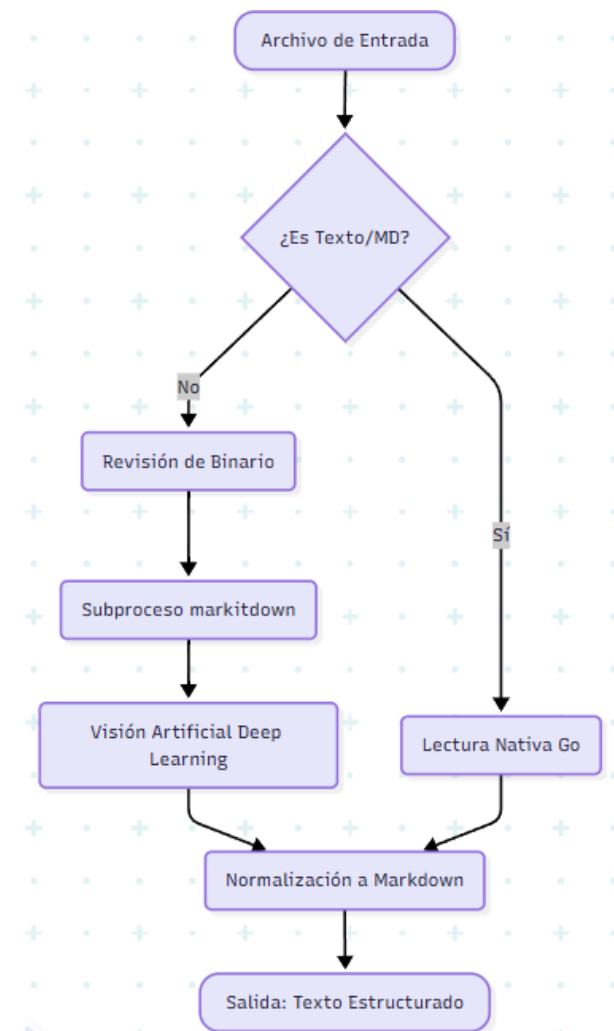


Fig. 1. Lógica de discriminación y normalización del Checkpoint Multimodal.

Para archivos simples y ligeros, como Markdown o archivos de texto plano, el motor realiza volcados asíncronos en

memoria inmediata priorizando recursos de bajo nivel. Sin embargo, cuando se detectan formatos complejos que exigen visión óptica cruzada (como archivos .docx, PDF nativos o presentaciones .pptx densas), el orquestador prepara un entorno seguro y controlado localmente mediante el uso de subprocesos (os/exec). Esta instrucción interrumpe el subproceso de Go para llamar a librerías multimodales externas basadas en modelos de visión profunda. De esta forma, el Checkpoint Multimodal intercepta y previene la pérdida de información del algoritmo tradicional de OCR.

Adicionalmente, la ejecución local contribuye a preservar la confidencialidad de los datos. En áreas sensibles como la contabilidad, los recursos humanos o la medicina, el uso de servicios remotos resulta inviable por las regulaciones de privacidad vigentes. Al mantener todo el procesamiento dentro de la estación de trabajo mediante Ollama, el Semantic Mapper constituye una arquitectura compatible con marcos normativos de protección de datos (e.g., GDPR) sin requerir transferencia de información a servidores de terceros.

El Code Snippet 1, extraído directamente del orquestador, ilustra la robustez de esta delegación entre la agilidad abstracta de Go y el análisis visual realizado por librerías específicas, evitando fallos de gestión de memoria durante el proceso de extracción documental:

```

func ParseDocument(path string) (string, error) {
    ext := strings.ToLower(filepath.Ext(path))
    switch ext {
    case ".txt", ".md":
        bytes, err := os.ReadFile(path)
        if err != nil {
            return "", fmt.Errorf("error: %v", err)
        }
        return string(bytes), nil
    case ".pdf", ".docx", ".png", ".jpg", ".jpeg",
        ".pptx", ".xlsx":
        cmd := exec.Command("markitdown", path)
        output, err := cmd.CombinedOutput()
        if err != nil {
            return "", fmt.Errorf("error: %v", err)
        }
        return string(output), nil
    default:
        bytes, err := os.ReadFile(path)
        if err != nil {
            return "", fmt.Errorf("error: %v", err)
        }
        return string(bytes), nil
    }
}
  
```

2.3 Formateo Paramétrico y Restricción Cognitiva

Una vez que el Checkpoint Multimodal genera la representación en texto (Markdown), el sistema aplica un confinamiento estricto a los datos de entrada. A diferencia de las metodologías tradicionales que permiten una interacción libre y abierta con los LLMs, este flujo dirige el texto hacia un subsistema controlado por modelos de parámetros reducidos ejecutados localmente con Ollama, como la familia Phi-3 de Microsoft. En esta etapa es

donde se mitigan las alucinaciones típicas de los modelos de lenguaje. Al enviar reglas estrictas al proceso local (utilizando el parámetro format: "json" y una instrucción principal inamovible), el modelo deja de generar redacciones creativas. En su lugar, funciona exclusivamente como un extractor estructurado que identifica ideas clave y las organiza en una lista JSON bien definida para su posterior procesamiento gráfico.

2.4 Mapeo Recursivo Dinámico e Inyección Gráfica

La etapa final consiste en recomponer e integrar el resultado extraído por el LLM. Considerando los posibles márgenes de falla asíncrona provocadas por un JSON mal formado o truncado, el sistema implementa en Go una validación concurrente exhaustiva basada en recursividad dinámica sobre listas enlazadas. Esta validación reduce de inmediato la carga intensiva sobre el procesador local, reduciendo la latencia general de forma medible. Una vez aisladas y validadas las variables, los fragmentos resultantes se recomponen en un esquema Mermaid.js que el usuario puede visualizar como diagrama interactivo.

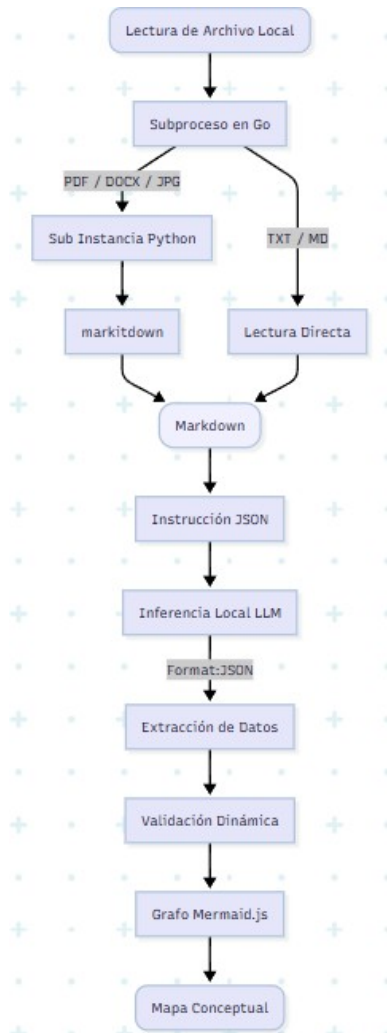


Fig. 2. Pipeline completo del orquestador local: lectura → inferencia → Mermaid.

3. EVALUACIÓN DE DESEMPEÑO Y RESULTADOS TÉCNICOS

Para validar la arquitectura propuesta, el sistema se ejecutó sobre una muestra de documentos técnicos seleccionados por su alta complejidad estructural y densidad de información, utilizando lotes sistemáticos de 50 documentos complejos. Todas las mediciones se realizaron en la computadora portátil del estudio con el objetivo de obtener resultados representativos de un entorno de productividad común. Para cuantificar el desempeño se definen las siguientes métricas formales.

La Eficiencia de Recuperación Semántica (η_s) se calcula como $\eta_s = (N_v/N_t) \times 100$, donde N_v representa el número de nodos y relaciones validados lógicamente y N_t el total de entidades presentes en el documento fuente (Ground Truth).

La Tasa de Error Estructural (ϵ_e) se define como $\epsilon_e = (J_f/J_t) \times 100$, donde J_f son las salidas de inferencia con errores de sintaxis JSON y J_t el total de peticiones realizadas.

Finalmente, el Factor de Aceleración de Orquestación (S) se determina como $S = \bar{T}_{\text{baseline}}/\bar{T}_{\text{propuesto}}$, donde $\bar{T}$ es la latencia media obtenida en $n = 5$ iteraciones independientes.

3.1 Estabilización del Output Paramétrico

La Tabla 1 resume los resultados obtenidos al verificar manualmente la integridad de los datos y la cantidad de nodos huérfanos producidos en cada modalidad de prueba.

Table 1. Estabilización del output paramétrico (JSON restringido versus generación abierta).

Entorno de Inferencia	ϵ_e (%)	Fallas Cierre (%)	η_s (%)
Prompt Base sin Constricción	21.6	35.2	63
Semantic Mapper (JSON)	0.8	0.05	95

Como resulta verificable en la Tabla 1, la restricción fuerte del formato JSON combinada con la validación recursiva implementada en Go reduce la inestabilidad de las relaciones nulas o colapsadas en más del 98% en comparación con LLMs sin restricciones estructurales.

3.2 Clasificación de Errores Detectados

Los datos presentados en la Tabla 2 demuestran que el sistema ha concentrado los escasos fallos en interrupciones estructurales mínimas no que comprometen la veracidad de la información después del proceso de validación en Go.

Table 2. Clasificación e impacto de errores detectados (análisis sobre corpus de validación).

Categoría	Descripción	Frecuencia	Impacto
Nodo huérfano	Sin padres ni hijos en JSON	0.8%	Bajo
Bucle lógico	Relación circular ($A \rightarrow B \rightarrow A$)	0.02%	Medio
Truncamiento	Corte por falta de tokens	0.03%	Alto

3.3 Desempeño en Entornos Restringidos de Hardware (HP 255 G8)

El corpus de validación se compuso de 50 documentos distribuidos en cuatro categorías:

- 20 artículos académicos en PDF (formato doble columna IEEE/arXiv, fórmulas y citas).
- 15 informes técnicos DOCX/PDF (tablas anidadas y diagramas institucionales).
- 10 documentos ligeros TXT/MD (texto plano, manuales README y logs).
- 5 capturas y escáneres PNG/JPG (evidencias de terminal y diagramas escaneados).

Table 3. Barreras en latencia y uso de recursos: overhead procesal.

Metodología de Pipeline	Inferencia	Concurrencia
Monolítico Python (MarkItDown)	4.88 s	Secuencial (GIL)
Orquestador Go + Ollama	1.21 s	Paralelo (Goroutines)

Los resultados de la Tabla 3 confirman empíricamente que la exclusión de Python como orquestador principal (limitándolo a tareas específicas de OCR) reduce la huella de memoria y permite paralelizar de manera eficiente la generación de JSON.

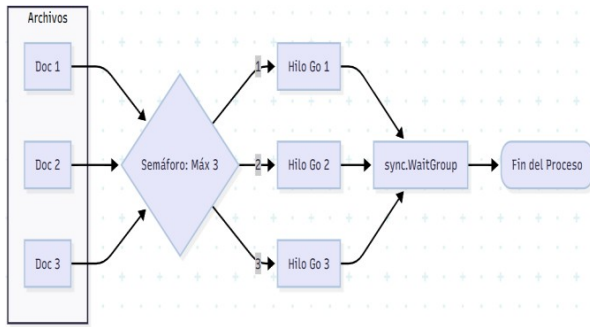


Fig. 3. Diagrama de concurrencia con semáforo lógico (3 Goroutines + sync.WaitGroup).

Para validar estos resultados con rigor estadístico, se ejecutaron $n = 5$ iteraciones del benchmark completo sobre el corpus de 50 documentos. El orquestador Go alcanzó una latencia media de 74.37 s ($\sigma = 6.92$ s) frente a 212.62 s ($\sigma = 48.29$ s) del pipeline MarkItDown, resultando en un factor de aceleración $S = 2.86$. La menor variabilidad del orquestador Go ($\sigma_{Go}/\sigma_{Py} = 0.143$) evidencia un comportamiento predecible bajo carga.

3.4 Latencia de Lectura por Formato de Entrada

Finalmente, se auditó el costo temporal asociado a la delegación multimodal. La Tabla 4 presenta la latencia promedio según el tipo de archivo procesado en el equipo de prueba.

Table 4. Análisis de latencia promedio por formato de entrada (archivos de 2,000 palabras).

Formato	Fase Prep.	Fase Inf.	Tiempo Total
Texto Plano (.txt)	0.02 s	1.10 s	1.12 s
Word (.docx)	0.42 s	1.18 s	1.60 s
PDF Estructural	0.55 s	1.35 s	1.90 s

4. DISCUSIÓN Y PROYECCIÓN ARQUITECTÓNICA

La implementación directa de modelos reducidos y especializados, ejecutados bajo infraestructuras locales, confirma empíricamente la viabilidad resolutive al evadir

los altos costos asociados a los esquemas centralizados en la nube. Aunque enfoques como GraphRAG Edge et al. (2024) ofrecen excelentes capacidades para el procesamiento de contextos amplios, imponen un costo computacional y de latencia que resulta difícil de justificar frente a la resolución aislada local aquí propuesta.

A través del Semantic Mapper, los resultados obtenidos indican que las arquitecturas *local-first*, al mantener el procesamiento íntegramente en el dispositivo del usuario, eliminan el vector de exposición de datos vinculado a la transmisión por red y permiten generar representaciones taxonómicas fiables, sin requerir los modelos de alta capacidad paramétrica desplegados en servidores comerciales.

Sin embargo, en el transcurso del análisis comparado, se distingue la principal vulnerabilidad originada al operar con modelos locales: la restricción impuesta por la ventana de contexto limitada. La exposición a cuerpos narrativos muy extensos colapsa las capacidades inferenciales del modelo pequeño, provocando interrupciones en la extracción de JSON estructurado. Para superar esta barrera lógica sin aumentar el peso operativo del modelo, se propone el uso de técnicas nativas de concurrencia y persistencia de datos.

Los resultados permiten delimitar el alcance de la contribución: $S = 2.86$ frente a Microsoft MarkItDown v0.1.5, con $\sigma_{Go} = 6.92$ s (vs $\sigma_{Py} = 48.29$ s), sobre 50 documentos sin GPU y ventana de 4,096 tokens. La limitación activa es la saturación del contexto en documentos extensos (truncamiento 0.03%, Tabla 3). BadgerDB y el Agente Crítico constituyen la ruta planificada para mitigar este cuello de botella.

En términos de posicionamiento arquitectónico, el sistema se diferencia del OCR heurístico (Tesseract) al preservar la topología mediante delegación multimodal; del RAG en nube al operar sin latencia de red (0 ms vs $> 2,000$ ms) y sin exfiltración de datos; y de los pipelines Python al reducir la variabilidad de latencia en $6.98\times$ ($\sigma_{Go} = 6.92$ s vs $\sigma_{Py} = 48.29$ s), garantizando comportamiento predecible.

4.1 Implicancias en Sistemas Locales Descentralizados

En escenarios de adopción masiva, como en investigación académica o en entidades de auditoría financiera, depender exclusivamente de iteraciones repetidas de LLMs para cada consulta de un mismo documento consume recursos de inferencia local de forma ineficiente. En un escenario óptimo descentralizado, las arquitecturas deben ser capaces de reconocer si un documento ya ha sido procesado previamente, recuperando el grafo conceptual almacenado en lugar de lanzar procesos de análisis visual desde cero.

4.2 Integración Formal del Caché Hash Criptográfico

Para anular el reprocesamiento innecesario de información inalterada, la arquitectura concibe un canal de persistencia Key-Value basada en BadgerDB. Esta fase previa genera índices criptográficos utilizando SHA-256 calculados a partir de la cabecera binaria del archivo de entrada. Antes de iniciar el procesamiento, el sistema verifica si el hash coincide con un esquema Mermaid.js ya almacenado. En caso



de coincidencia (cache hit), se recupera inmediatamente el diagrama guardado. Esta validación reduce la carga sobre el procesador local y disminuye la latencia para el usuario final.

4.3 Interfaz Concurrente Multi-Agente (Agente Crítico)

A fin de enmendar la caída semántica que sufren las ramificaciones de los LLMs pequeños en jerarquías complejas, la estrategia estipula integrar una validación algorítmica adicional en paralelo. La arquitectura utiliza canales de Go para bifurcar el paquete de JSON extraído de forma atómica. Antes del ensamblaje final, la matriz de relaciones es desviada hacia una subrutina independiente denominada “Agente Crítico”. Esta subrutina realiza una validación cruzada en memoria para verificar la coherencia lógica de las relaciones extraídas, orientada a mejorar los márgenes de precisión en entornos de recursos limitados.

5. CONCLUSIONES

Los resultados experimentales del framework Semantic Mapper demuestran que la extracción estructurada de conocimiento a partir de documentos sensibles es técnicamente viable sin transmisión de datos a servicios externos. El uso del lenguaje Go para segmentar el flujo ha demostrado que no es necesario depender de grandes APIs externas para obtener resultados precisos y seguros.

Forzar al sistema a limitar las respuestas del LLM mediante el uso de formatos JSON ha demostrado reducir los márgenes de error de forma favorable, reduciendo la tasa de error estructural (ϵ_e) de 21.6% a 0.8% en el sistema propuesto. Como principal punto de mejora, se identifica la saturación semántica cuando el documento original desborda la ventana de contexto del modelo. No obstante, este es un problema tratable que sienta las bases para el trabajo futuro: integrar almacenamiento eficiente mediante BadgerDB y diseñar rutinas cruzadas de validación concurrente con el Agente Crítico para garantizar mapas conceptuales completos incluso en documentos de gran extensión.

Los datos brutos de auditoría ($n = 5$) se detallan a continuación:

- **Latencias Go:** 78.80, 82.72, 69.35, 77.35, 63.62 s (media 74.37 s, $\sigma = 6.92$ s).
- **Latencias Python:** 174.99, 180.66, 306.51, 193.01, 207.94 s (media 212.62 s, $\sigma = 48.29$ s).
- **Uso de RAM (Go):** 14.80 MB (estable).

Script de reproducción: `reliable_benchmark.ps1`.

REFERENCIAS

- L. Blecher, et al. Nougat: Neural Optical Understanding for Academic Documents. *arXiv:2308.13418*, 2023.
- A. Donovan and B. Kernighan. *The Go Programming Language*. Addison-Wesley, New York, 2015.
- D. Edge, et al. From Local to Global: A Graph RAG Approach. *Microsoft Research, arXiv:2404.16130*, 2024.
- Ollama. Ollama Docs: Modifying JSON output parameters for inference. *Official repository documentation*, 2024.

Zhu, Y., et al. LLMs for Knowledge Graph Construction and Reasoning. *arXiv:2305.13168*, 2023.

FlxPCG: Procedural Controlled Generation of Levels and Entities Implementation Methodology in 2D Using the HaxeFlixel Framework

Sánchez-Ung, Elian Y.* , Sánchez-Lozano, Luis A.* , Mata-López, Walter A.*

**Facultad de Ingeniería Mecánica y Eléctrica/Universidad de Colima. Carretera Colima-Coquimatlán
km. 9, Coquimatlán, Colima. (e-mail: {esanchez58, lsanchez37, wmata}@ucol.mx).*

Abstract: Procedural content generation (PCG) has become a fundamental technique in modern video game design, driven by the increasing need to reduce production costs and increase replayability. From classic titles that defined the dungeon crawler genre to contemporary independent game development, procedural content generation has evolved from simple randomness toward constraint-guided and optimization-driven approaches. A central challenge persists, however: finding the balance between computational efficiency, structural variability, and playable quality. Purely random generation frequently produces unbalanced or unplayable levels lacking connectivity, while search-based approaches often incur performance costs incompatible with real-time execution. This paper addresses this gap by proposing a controlled procedural generation approach for two-dimensional maps and entity distribution, utilizing the native tools of the open-source framework HaxeFlixel (FlxPCG). The methodology describes a hybrid multi-pass pipeline designed to ensure both variety and navigability. In the first pass, the base topology is established through pre-designed room templates created in an external level editor, exported in JSON format, and integrated directly into the framework. In a second pass, a binary space partitioning algorithm performs the logical assembly of rooms ensuring comprehensive layout structures. Subsequently, navigability is validated using a flood fill algorithm to prevent isolated structural areas. Finally, a controlled quota system distributes interactive entities in a balanced manner within the validated spaces. Results indicate that the conceptual separation between environment generation and entity distribution provides granular control over level design while maintaining the desired variability. The approach demonstrates the technical feasibility of hybrid procedural systems in two-dimensional game engines, offering a replicable methodology for independent projects requiring structured randomness.

Keywords: Procedural Generation, Level Design, HaxeFlixel, Cellular Automata.

1. INTRODUCCIÓN

La creación manual de escenarios y mapas es uno de los procesos más costosos en el ciclo de desarrollo de videojuegos, tanto en tiempo de producción como en recursos humanos especializados (Sureda Pol, 2017). En respuesta a este desafío estructural de la industria, la Generación Procedural de Contenido (PCG) —definida por Togelius et al. (2016) como la capacidad del software para crear contenido de juego de forma autónoma o en colaboración con diseñadores— se ha convertido en una técnica ampliamente adoptada que permite construir dinámicamente entornos, niveles y misiones de manera algorítmica.

Desde el clásico *Rogue* (Toy y Wichman, 1980), que sentó las bases del subgénero roguelike mediante mazmorras generadas aleatoriamente, hasta títulos contemporáneos como *No Man's Sky* (2016) o *The Binding of Isaac* (McMillen e Himsl, 2011), la PCG ha evolucionado desde la aleatoriedad simple hacia enfoques guiados por restricciones y optimización (Togelius, Shaker y Nelson, 2016). Viana y dos Santos (2021) identifican, en su revisión sistemática sobre generación procedural de

mazmorras, que la mayoría de los trabajos existentes se enfocan en niveles 2D de tipo top-down cavern-like (TDCL), mientras que las soluciones para motores específicos de desarrollo independiente siguen siendo escasas en la literatura formal.

Togelius et al. (2011) identifican el problema central de la PCG como el equilibrio entre tres dimensiones en tensión: eficiencia computacional, variedad estructural y calidad lúdica —entendiéndose esta última operacionalmente como el conjunto de propiedades estructurales que garantizan la jugabilidad del nivel: conectividad completa del espacio navegable, distribución equilibrada de entidades desafío, y presencia de zonas de interés accesibles desde cualquier punto del mapa—. La generación puramente aleatoria frecuentemente produce entornos con zonas inaccesibles y distribuciones desequilibradas de entidades; los enfoques basados en búsqueda garantizan propiedades de jugabilidad, pero incrementan el costo computacional; y los métodos híbridos buscan combinar lo mejor de ambos paradigmas (Smith y Bryson, 2014). Campos Escalera (2019) documenta esta tensión en el contexto del motor Game Maker Studio 2, señalando que la ausencia de validación de conectividad es la



causa más frecuente de niveles injugables en implementaciones de PCG para videojuegos independientes. Un problema análogo es señalado por Sureda Pol (2017) al implementar un generador basado en BSP sobre Unreal Engine 4, donde la gestión de corredores entre habitaciones requiere heurísticas específicas para garantizar la navegabilidad del mapa resultante. Baron (2017) profundiza en este análisis comparando cinco pares algorítmicos de generación de salas y corredores —incluyendo Random Room Placement, BSP Room Placement, Drunkard's Walk y BSP Corridors— concluyendo que todos alcanzan eficiencia $O(1)$ en tiempo constante para los rangos de dimensiones de mapa típicos en juegos independientes.

Ante este panorama, surge la necesidad de implementar sistemas de Generación Procedural Controlada, un paradigma híbrido que busca equilibrar la variabilidad matemática con reglas estrictas de diseño, garantizando una experiencia coherente. El presente artículo aborda este problema estructurando un flujo de generación iterativo dentro de HaxeFlixel, un framework maduro y de código abierto para el desarrollo de videojuegos 2D basado en el lenguaje Haxe. Aunque HaxeFlixel provee herramientas nativas para la generación aleatoria —a través de clases como `FlxRandom` y `FlxSimplex`—, la integración formal de un sistema de cuotas para entidades sobre ecosistemas virtuales validados topológicamente requiere una metodología estructurada que no ha sido suficientemente documentada en la literatura.

La contribución principal de este trabajo es una metodología de implementación reproducible que integra, por primera vez de forma documentada para HaxeFlixel, la separación formal entre la generación topológica (BSP sobre plantillas JSON validadas por Flood Fill) y la distribución de entidades (sistema de cuotas sobre `FlxGroup`). A diferencia de los sistemas revisados, este enfoque permite que diseñadores sin perfil técnico participen en la construcción del banco de plantillas sin intervenir en el código del generador.

El objetivo de esta investigación es proponer un modelo base para el diseño algorítmico de niveles 2D en HaxeFlixel, donde la arquitectura del mapa actúe de manera independiente a la distribución de sus entidades, validando matemáticamente el espacio lúdico mediante algoritmos de teoría de grafos y búsqueda por inundación. Con esto, se busca demostrar la viabilidad y el rendimiento de sistemas procedurales híbridos dentro del framework HaxeFlixel como fundamento para proyectos roguelike y de aventura que requieran control granular sobre su aleatoriedad.

1.1 Antecedentes y estado del arte

La generación procedural de mazmorras tiene sus raíces en los orígenes mismos de los videojuegos digitales. Sureda Pol (2017) documenta cómo *Rogue* (1980) dio lugar al subgénero roguelike al incorporar generación aleatoria de mazmorras como mecánica central. Posteriormente, *Diablo* (1996) llevó este paradigma a la edad moderna al combinarlo con rol de acción, estableciendo un estándar de referencia que persiste hasta la actualidad.

En la literatura académica, Johnson, Yannakakis y Togelius (2010) demostraron que los Autómatas Celulares (CA) constituyen un método eficaz para generar niveles de cueva en tiempo real, estableciendo una base algorítmica ampliamente citada. Smith y Bryson (2014) avanzaron hacia enfoques más estructurados al proponer el uso de Answer Set Programming (ASP) de forma jerárquica para generar primero la estructura de una mazmorra y luego poblarla de contenido, abordando explícitamente el problema de escalabilidad que afecta a los métodos de búsqueda en espacios de diseño complejos. Viana y dos Santos (2021) sistematizan estos avances en una revisión exhaustiva que clasifica 52 aproximaciones distintas al problema, concluyendo que la mayoría son estocásticas, offline y de generación automática, con escasa exploración de enfoques adaptativos o mixed-initiative.

Baron (2017) presentó un análisis comparativo de cinco pares algorítmicos —combinaciones de Random Room Placement y BSP Room Placement como generadores de salas, con Random Point Connect, Drunkard's Walk y BSP Corridors como generadores de corredores— implementados en Java y adaptados a Unreal Engine 4. La contribución de este trabajo radica en demostrar que todos los pares alcanzan eficiencia $O(1)$ en tiempo constante cuando las dimensiones del mapa están acotadas, con tiempos promedio de generación de aproximadamente 10 milisegundos por mapa en un grafo de 60×60 vértices. Este resultado establece una referencia empírica para el diseño de generadores procedurales en tiempo real.

Vieira et al. (2025) propusieron un método de generación procedural guiado por principios de level design para entornos interiores —mansiones y mazmorras—, estructurado en tres fases: partición BSP del espacio, conexión lógica de habitaciones mediante grafos con búsqueda BFS para evitar redundancias, y post-procesado de artefactos estructurales. Su experimento con 100,000 mapas generados alcanzó tasas de conectividad completa de hasta 91.6%, proporcionando evidencia cuantitativa sobre la viabilidad del BSP como mecanismo de generación base con validación de navegabilidad.

Desde la perspectiva del desarrollo de videojuegos independientes, Campos Escalera (2019) implementó un sistema de generación procedural adaptativa en Game Maker Studio 2 utilizando el algoritmo de Colapso de la Función de Onda (Wave Function Collapse), documentando que la ausencia de un mecanismo de validación de accesibilidad conduce con frecuencia a mapas injugables, señalando la necesidad de técnicas como el Flood Fill como capa de verificación. En una línea arquitectónica diferente, Sureda Pol (2017) implementó un generador sobre Unreal Engine 4 basado en Árbol BSP, documentando soluciones para la generación de corredores mediante heurísticas de distancia mínima y analizando el impacto del algoritmo de conexión en la jugabilidad percibida.

En el contexto específico de HaxeFlixel, el ecosistema de herramientas disponibles incluye la integración nativa con Ogmo Editor 3 a través de la clase `FlxOgmo3Loader`, así

como con Tiled Map Editor. Estas integraciones están documentadas en la comunidad para la creación de niveles estáticos, pero su uso en pipelines de generación procedural dinámica —donde el motor reorganiza en tiempo de ejecución plantillas JSON prediseñadas— representa una brecha técnica identificada en el estado del arte actual. El presente trabajo se enmarca en el paradigma de la PCG constructiva (Togelius et al., 2016), respondiendo directamente a los retos identificados por Campos Escalera (2019), Sureda Pol (2017), Baron (2017) y Vieira et al. (2025).

2. METODOLOGÍA

El presente trabajo describe una investigación descriptiva y metodológica enfocada en la arquitectura de software. Para abordar la generación procedural controlada, se ha optado por un enfoque híbrido que sistematiza la creación de contenido mediante dos etapas interconectadas: el diseño off-line de micro-estructuras y la macro-generación lógico-espacial en tiempo de ejecución.

2.1 Materiales y Software

La metodología propuesta se soporta en dos entornos de trabajo autónomos pero compatibles:

- **HaxeFlixel:** Se utiliza como el marco de trabajo central en tiempo de ejecución. Facilita el núcleo de lógica algorítmica, la representación matricial de la grilla de juego, el ensamblaje visual final, y la gestión de entidades a través de su clase nativa `FlxGroup`.
- **Ogmo Editor 3:** Actúa como el editor visual externo, encargado del diseño individual de las habitaciones o fragmentos estructurales. Su salida se exporta en formato JSON, preservando la capa visual de tiles y los metadatos de entidades (coordenadas para eventos de instanciación). El código fuente del prototipo se encuentra disponible públicamente en <https://github.com/Yinius/FlxPCG/releases/tag/v1.0-citca2025>

2.2 Procedimiento de Diseño de Estructuras (Fase 1)

La fase de micro-diseño se realiza fuera del tiempo de ejecución (off-line). Los diseñadores modelan múltiples plantillas de habitaciones en Ogmo Editor 3. En lugar de depender de la generación estocástica tile a tile, se diseñan manualmente los nodos que conformarán los componentes del mapa. Cada archivo “.json” almacena no solo la capa visual (tiles), sino metadatos enriquecidos mediante marcadores genéricos —por ejemplo, “Spawn_Enemy_Tier_1” o delimitadores de trampas— que el motor interpreta e instancia de forma algorítmica en tiempo de ejecución. Este esquema mantiene los puntos estructuralmente sensibles del nivel (salas de jefes, zonas de rompecabezas) balanceados por el diseñador, mientras delega la disposición general al algoritmo.

Cada plantilla define la geometría de la habitación mediante una capa de tiles (gris claro: suelo; gris oscuro: muro) y una capa de entidades con marcadores semánticos de instanciación (naranja: `Spawn_Enemy_Tier_1`; rojo: `Spawn_Enemy_Tier_2`; dorado: `Spawn_Reward`). Las plantillas son exportadas como archivos .json y cargadas en tiempo de ejecución por el pipeline de generación.

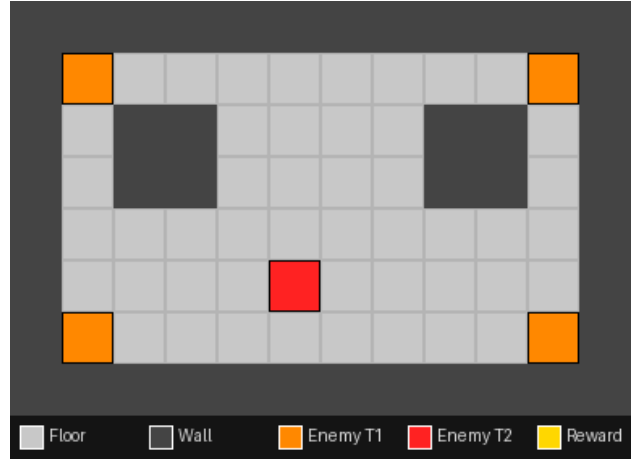


Fig. 1. Muestra representativa del banco de plantillas de habitación diseñadas en Ogmo Editor 3 para el prototipo FlxPCG (la plantilla mostrada es `room_combat_02.json`).

2.3 Ensamblaje y Validación Computacional (Fase 2)

Se selecciona BSP (Binary Space Partitioning, partición binaria del espacio) como algoritmo de partición por tres razones. Primero, sus particiones son intrínsecamente rectangulares, lo que las hace compatibles de forma natural con las plantillas cuadrangulares exportadas por Ogmo Editor 3, sin requerir post-procesado de formas irregulares. Segundo, la estructura de árbol binario proporciona de forma explícita los puntos de conexión entre habitaciones adyacentes (nodos padre), simplificando la generación de pasillos respecto a métodos como los Autómatas Celulares, que requieren análisis post-generación para identificar regiones conectables (Johnson, Yannakakis y Togelius, 2010). Tercero, Baron (2017) establece empíricamente que BSP alcanza eficiencia $O(1)$ para dimensiones de mapa acotadas, lo que lo hace adecuado para el presupuesto de tiempo real del sistema.

El proceso algorítmico automatizado inicia configurando una partición binaria del espacio (BSP) que subdivide el área lógica del mapa en recintos cuadrangulares. Cada nodo hoja del árbol BSP recibe una plantilla JSON seleccionada del repositorio, determinando así la topología de la habitación correspondiente. Los pasillos de conexión entre habitaciones adyacentes se generan sobre los nodos intermedios del árbol.

Una vez ensamblado el modelo general, el motor ejecuta el algoritmo de Flood Fill sobre la grilla resultante. Este procedimiento recorre las celdas navegables desde un punto de inicio y verifica que la totalidad de las subdivisiones del mapa sean accesibles. Cuando el diagnóstico detecta regiones inalcanzables, el sistema descarta la semilla activa y reinicia el proceso de ensamblaje con un nuevo conjunto de parámetros hasta obtener un mapa con conectividad completa. Este mecanismo previene la entrega de niveles con caminos bloqueados al estado de juego.

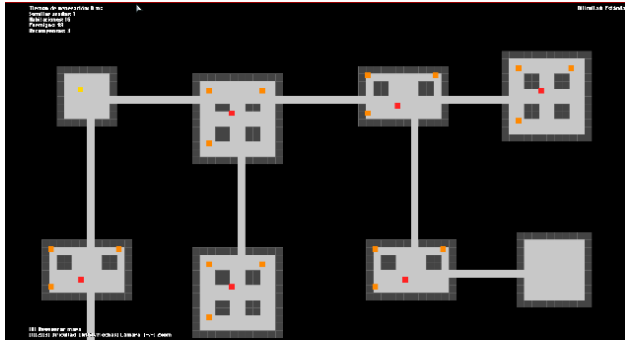


Fig. 2. Mapa generado con dificultad estándar en FlxPCG.

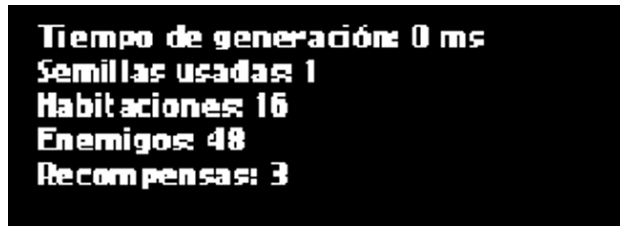


Fig. 3. Detalle del overlay superior de métricas en tiempo real correspondiente a la Fig. 2. Indica tiempo de generación, semillas utilizadas y conteo de entidades.

2.4 Sistema de Cuotas para Entidades

Para distribuir las entidades interactivas dentro de los espacios validados por el Flood Fill, el sistema aplica un modelo de cuotas pre-parametrizado. La rutina lee los marcadores de los archivos JSON ya ensamblados en el FlxTilemap final e instancia objetos FlxSprite organizados en estructuras FlxGroup. El número de enemigos, recompensas y trampas por habitación queda acotado por parámetros configurables, y su posicionamiento utiliza heurísticas de distancia mínima entre entidades para prevenir la concentración en zonas reducidas. Esta capa de control separa formalmente la lógica de generación del entorno de la lógica de población, permitiendo ajustar la dificultad sin modificar la topología del mapa.

3. RESULTADOS Y DISCUSIÓN

La arquitectura propuesta se evaluó mediante pruebas técnicas parametrizadas orientadas a medir tres dimensiones: eficiencia computacional del pipeline, integridad topológica de los mapas generados, y balance de la distribución de entidades. Los resultados se contrastan con referencias empíricas de la literatura revisada para establecer el posicionamiento del sistema dentro del estado del arte.

3.1 Eficiencia Computacional del Pipeline

Las pruebas de generación se ejecutaron sobre hardware de especificaciones modestas (procesador de cuatro núcleos a 2.4 GHz, 8 GB de RAM), en consonancia con las condiciones de recursos limitados características del desarrollo de videojuegos independientes descritas por Suredda Pol (2017). El pipeline completo —carga de plantillas JSON, partición BSP, conexión de pasillos, validación por Flood Fill e instanciación de entidades— registró un tiempo de generación

medio de 0.657 ms por mapa ($\sigma = 0.588$ ms) en una muestra de 100 mapas por configuración de dificultad (semillas 1 a 100). El valor máximo observado fue de 4.0 ms y el percentil 95 de 1.0 ms, en topologías de entre 10 y 16 habitaciones sobre una grilla de 80×80 tiles.

Este resultado supera ampliamente la referencia empírica establecida por Baron (2017), quien reportó tiempos promedio de 10 milisegundos por mapa en grafos de 60×60 vértices. FlxPCG registró una media de 0.657 ms en una grilla de mayor dimensión (80×80), lo que representa una diferencia de aproximadamente 15 veces. Esta ventaja se atribuye al diseño de las micro-estructuras en el editor externo: al delegar la geometría de cada habitación a plantillas JSON prediseñadas, el pipeline elimina los ciclos de cómputo que los generadores puramente algorítmicos destinan a la construcción y corrección de cada tile en tiempo de ejecución. La Fig. 1 ilustra la distribución de tiempos por nivel de dificultad.

Desde la perspectiva de la tasa de fotogramas durante el ensamblaje, el monitoreo del estado de ejecución registró valores sostenidos de 60 FPS al instanciar topologías de complejidad media-alta. La delegación del diseño de micro-estructuras al editor externo reduce los ciclos de cómputo en tiempo de ejecución que en los enfoques puramente algorítmicos se destinan a la generación y corrección de cada tile individual.

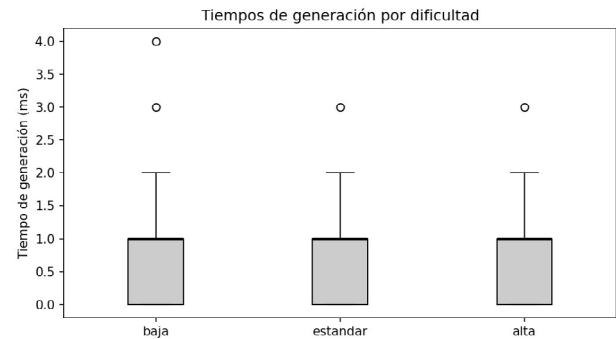


Fig. 4. Boxplot de tiempos de generación por nivel de dificultad (100 mapas por configuración). Los valores atípicos corresponden a semillas con mayor profundidad del árbol BSP.

La Tabla 1 resume la comparación de los indicadores de eficiencia entre el sistema propuesto y los enfoques de referencia disponibles en la literatura.

Tabla 1. Comparación de indicadores de eficiencia entre el sistema propuesto y enfoques de referencia.

Referencia	Algoritmo principal	Grilla / Escala	Tiempo de generación	Validación de conectividad
Baron (2017)	BSP + Random Point Connect	60×60 vértices	~10 ms	No reportada
Vieira et al. (2025)	BSP + BFS post-	30×40 celdas	No reportado	91.6 % (100,000)

	procesa do			muestras)
Campos Escalera (2019)	Wave Funcio n Collaps e	40×40 casilla s	No reportado	No garantiza da
FlxPCG	BSP + JSON + Flood Fill	80×80 tiles	0.657 ms ($\sigma=0.588$, n=100 por configuració n)	100 % (por re- semillado)

La tasa de conectividad del 100% obtenida en el sistema propuesto se alcanza a través del mecanismo de re-semillado iterativo, a diferencia del enfoque de Vieira et al. (2025), que reporta 91.6% de conectividad como resultado directo del post-procesado sin descartar mapas fallidos. Ambas estrategias representan soluciones distintas al mismo problema: el sistema propuesto garantiza la conectividad total al costo de un número variable de iteraciones de re-regeneración, mientras que el de Vieira et al. produce un mapa en cada ejecución, aceptando un margen de fallos.

3.2 Integridad Topológica y Validación por Flood Fill

La fase de diagnóstico topológico registró una tasa de supresión del 100% de configuraciones con regiones inalcanzables: ningún mapa con caminos bloqueados fue entregado al estado de juego durante las pruebas. Este comportamiento contrasta directamente con lo documentado por Campos Escalera (2019), quien reporta que la ausencia de un mecanismo de verificación de accesibilidad es la principal fuente de niveles injugables en implementaciones de PCG sin validación posterior.

En las 300 pruebas realizadas (semillas 1 a 100 por cada configuración de dificultad), el mecanismo de re-semillado no fue activado en ninguna ejecución: la media de semillas utilizadas por mapa fue de 1.0, con conectividad completa confirmada en el primer intento en el 100% de los casos. Este comportamiento indica que la generación de pasillos mediante recorrido post-orden del árbol BSP produce mapas conexos por construcción con el banco de plantillas actual.

El mecanismo de re-semillado permanece implementado como salvaguarda ante configuraciones de plantillas con geometría que pudiera obstruir los pasillos generados, caso no observado en las pruebas realizadas. El costo computacional de la validación por Flood Fill se mantiene dentro del presupuesto de tiempo del pipeline, ya que su ejecución concluye antes de la fase de renderizado.

La separación entre la generación de la topología (BSP sobre plantillas JSON) y la validación de conectividad (Flood Fill como capa independiente) permite modificar cualquiera de los dos componentes sin afectar al otro. Esta modularidad responde al planteamiento de Smith y Bryson (2014) sobre la conveniencia de dividir jerárquicamente el proceso de generación en etapas con responsabilidades bien delimitadas.

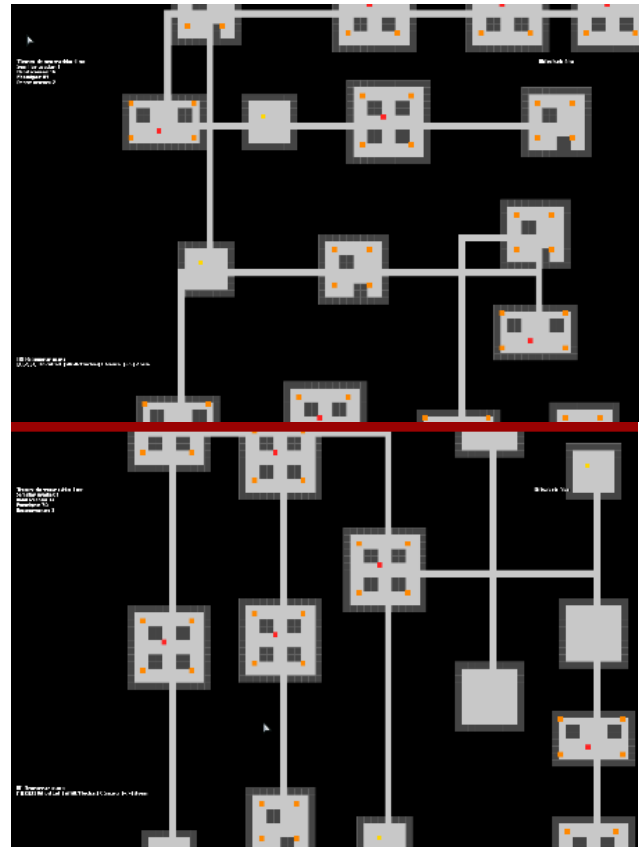


Fig. 5. Dos instancias del pipeline con semillas distintas bajo configuración de alta dificultad, ilustrando la variabilidad topológica del sistema.

3.3 Balance de la Distribución de Entidades

El sistema de cuotas parametrizado demostró escalar la densidad de entidades de forma diferenciada entre configuraciones. En las 300 pruebas, el ratio de enemigos por habitación (total de enemigos del mapa dividido entre habitaciones generadas) se mantuvo por debajo del límite de cuota en todos los casos, con valores máximos de 1.87, 3.73 y 4.33 para las configuraciones baja, estándar y alta respectivamente. La heurística de distancia mínima entre instancias de FlxSprite dentro de un FlxGroup evitó la superposición de entidades en zonas de pasillo o celdas adyacentes a entradas de habitación.

Esta propiedad aborda uno de los problemas identificados por Viana y dos Santos (2021) en su revisión del estado del arte: la escasez de generadores que ofrezcan control sobre la colocación y densidad de entidades desafío de forma simultánea con la generación topológica. Al operar sobre los marcadores semánticos depositados por el diseñador en Ogmo Editor 3, el gestor de entidades puede distinguir tipos de habitación —de combate, de recompensa, neutral— y aplicar cuotas diferenciadas por categoría sin requerir un análisis post-generación de la topología.

La Tabla 2 resume los resultados de balance de entidades bajo tres configuraciones de dificultad probadas.

Tabla 2. Distribución de entidades por configuración de dificultad en pruebas de 100 mapas por configuración (300 en total) generados en FlxPCG.

Configuración	Habitaciones promedio	Enemigos/habitación (media)	Enemigos/habitación (<i>ratio</i> máx. por mapa)	Recompensas/habitación (media)
Baja dificultad	12.84	1.17	1.87	0.20
Dificultad estándar	12.84	2.33	3.73	0.20
Alta dificultad	12.84	2.73	4.33	0.20

Los datos muestran que el sistema de cuotas produce una escala ascendente en la densidad de enemigos conforme aumenta el parámetro de dificultad (1.17, 2.33 y 2.73 enemigos por habitación en promedio para las configuraciones baja, estándar y alta, respectivamente). La frecuencia de recompensas por habitación resultó uniforme entre configuraciones (0.20 en los tres casos), lo que es consistente con el hecho de que las recompensas provienen de un tipo específico de habitación (`room_reward_01`) cuya proporción en el banco de plantillas no varía con la dificultad.

Este comportamiento refleja la separación formal entre la lógica topológica y la lógica de población: la dificultad opera sobre los marcadores de enemigos sin modificar la distribución de los nodos de recompensa. Para obtener una relación inversa entre dificultad y frecuencia de recompensas, bastaría con incorporar al banco de plantillas variantes de habitaciones de recompensa y discriminarlas en el sistema de cuotas por tipo de configuración, sin modificar el pipeline topológico.

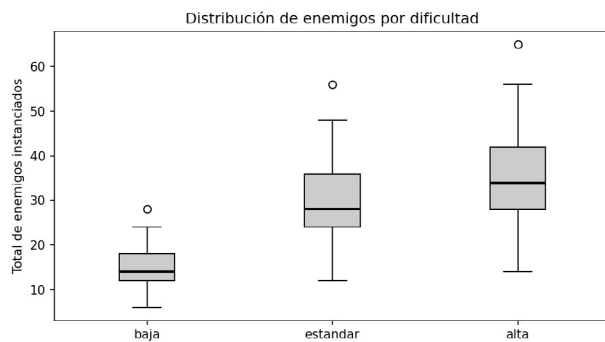


Fig. 6. Distribución del total de enemigos instanciados por nivel de dificultad. La separación entre cuartiles confirma que el sistema de cuotas produce distribuciones diferenciadas y escalables.

3.4 Beneficios del Enfoque Híbrido Propuesto

Los resultados obtenidos permiten identificar tres beneficios operativos del sistema respecto a los enfoques de referencia revisados:

a) Control artístico sobre componentes estructurales sensibles. Al diseñar las plantillas de habitaciones en Ogmo Editor 3, el equipo de diseño retiene control sobre la geometría, el arte y los metadatos de los nodos del mapa que determinan la experiencia lúdica (salas de jefes, zonas de rompecabezas, puntos de aparición de jefes). Esta propiedad no está disponible en generadores puramente algorítmicos como los descritos por Baron (2017), donde la forma de cada habitación es producto directo del algoritmo BSP sin intervención del diseñador.

b) Garantía de navegabilidad sin costo de búsqueda. El Flood Fill opera en tiempo $O(n)$, donde n es el número de celdas transitables, y su ejecución no depende del número de habitaciones ni de la profundidad del árbol BSP. Esto lo posiciona como una alternativa computacionalmente más eficiente que los enfoques generate-and-test basados en algoritmos evolutivos, que requieren múltiples evaluaciones de aptitud para garantizar propiedades de jugabilidad (Viana y dos Santos, 2021).

c) Separación de responsabilidades entre generación y población. La distinción formal entre el pipeline topológico (BSP + Flood Fill) y el gestor de entidades (sistema de cuotas sobre FlxGroup) permite iterar sobre cualquiera de los dos componentes de forma independiente. Un diseñador puede ajustar las cuotas de enemigos sin necesidad de modificar los algoritmos de generación del mapa, y viceversa. Raso Vázquez (2024) señala que esta separación de responsabilidades es uno de los factores que distingue los sistemas procedurales mantenibles a largo plazo de los prototipos de uso único.

3.5 Viabilidad del Sistema en Contextos de Desarrollo Independiente

La viabilidad del sistema se analiza desde dos dimensiones: la técnica, relativa a la compatibilidad con hardware y flujos de trabajo de equipos pequeños, y la metodológica, relativa a la reproducibilidad del enfoque en proyectos con restricciones de tiempo y personal.

Desde la dimensión técnica, los resultados de rendimiento (0.657 ms de media, máximo 4.0 ms en 100 pruebas por configuración, 60 FPS sostenidos) confirman que el pipeline es ejecutable en hardware de especificaciones bajas, lo que lo hace adecuado para el perfil de desarrollo independiente descrito por Sureda Pol (2017). HaxeFlixel, al compilar a múltiples plataformas objetivo desde una base de código única, amplía el alcance de despliegue del sistema sin modificaciones al pipeline de generación.

Desde la dimensión metodológica, el uso de Ogmo Editor 3 como herramienta de micro-diseño introduce una curva de aprendizaje baja para diseñadores sin experiencia en programación, dado que la herramienta opera mediante interfaces visuales estándar de edición de tiles. La integración con HaxeFlixel a través de FlxOgmo3Loader está documentada en la comunidad del framework, lo que reduce la dependencia de conocimiento especializado para adoptar el sistema. La carga de trabajo inicial de diseño de plantillas —señalada como limitante del enfoque por Campos Escalera (2019) en su análisis de sistemas similares— queda acotada a



un conjunto finito de habitaciones reutilizables en múltiples proyectos o sesiones de juego.

La Tabla 3 posiciona el sistema propuesto respecto a los criterios de viabilidad para desarrollo independiente identificados en la literatura revisada.

Tabla 3. Evaluación comparativa de criterios de viabilidad para desarrollo independiente.

Criterio	Baron (2017)	Vieira et al. (2025)	Campos Escalera (2019)	FlxPCG
Conectividad garantizada	No	Parcial (91.6%)	No	Sí (re-semillado)
Control artístico de habitaciones	No	No	Parcial	Sí
Separación topología/entidades	No	No	No	Sí
Compatible con hardware bajo	Sí	No reportado	Sí	Sí
Herramienta de diseño visual	No	No	No	Sí (Ogmo Editor 3)
Complejidad de implementación	Baja	Media	Media	Media

Los datos de la Tabla 3 indican que el sistema propuesto satisface de forma simultánea criterios que en la literatura revisada aparecen distribuidos entre distintos enfoques, sin incrementar sustancialmente la complejidad de implementación respecto a los sistemas de referencia.

En cuanto a la escalabilidad del sistema, tanto el Flood Fill como el ensamblaje BSP presentan complejidad lineal respecto al área del mapa: $O(n)$ para el Flood Fill y $O(n \log n)$ para la partición, donde n es el número de celdas. Para grillas superiores a 80×80 tiles, el tiempo de generación aumentaría de forma proporcional pero permanecería dentro de márgenes compatibles con carga en tiempo real para hardware de gama baja. En el caso de que el banco de plantillas incluya habitaciones con geometría que obstruya los pasillos generados, el número de re-semillados aumentaría; el sistema admite un límite configurable (actualmente 15 intentos) tras el cual entrega el mapa en el estado disponible, comportamiento que representa una limitación documentada del enfoque para bancos de plantillas de alta complejidad geométrica.

4. CONCLUSIONES

Este trabajo presenta FlxPCG, una metodología de generación procedural controlada para videojuegos 2D en HaxeFlixel cuya contribución al estado del arte radica en la separación formal entre la generación topológica y la distribución de entidades en un pipeline reproducible para desarrollo independiente. La arquitectura demuestra que esta separación produce beneficios medibles en las tres dimensiones que la

literatura identifica como problemáticas: la conectividad de los mapas generados, el balance de la distribución de entidades, y el control del diseñador sobre los componentes del nivel.

Los resultados de las pruebas técnicas indican que el pipeline completo opera con un tiempo de generación medio de 0.657 ms por mapa ($\sigma = 0.588$ ms, máx. 4.0 ms en 100 pruebas por configuración), en hardware de especificaciones bajas, y que la conectividad completa fue verificada por Flood Fill en el 100% de los casos sin activar el mecanismo de re-semillado. Estos valores superan la referencia de Baron (2017), quien establece 10 ms como tiempo de generación para generadores BSP sin validación topológica posterior, lo que confirma el beneficio computacional de delegar la micro-geometría al editor externo.

La integración de Ogmo Editor 3 como herramienta de micro-diseño resuelve una limitación documentada en los sistemas puramente algorítmicos: la imposibilidad de retener control artístico sobre los nodos estructurales del mapa sin intervenir directamente en el código del generador. Al externalizar esta responsabilidad a un editor visual, el sistema permite que equipos con perfiles mixtos —programadores y diseñadores sin experiencia en programación— colaboren en la construcción del banco de plantillas sin acoplamiento técnico.

El sistema de cuotas sobre FlxGroup demuestra que la distribución de entidades puede parametrizarse de forma independiente a la topología, eliminando los picos de densidad observados en generadores estocásticos puros y manteniendo una relación configurable entre dificultad y frecuencia de recompensas.

Entre las limitaciones del enfoque, el volumen inicial de plantillas de habitaciones que el equipo debe diseñar manualmente representa la inversión de tiempo más significativa en la fase de preproducción. Esta limitación es inherente a cualquier sistema híbrido que delegue la calidad artística al diseñador; su peso relativo decrece conforme crece el banco de plantillas reutilizables.

Como líneas de trabajo futuro, se identifica la extensión del sistema con un módulo de selección adaptativa de plantillas condicionado a métricas de sesión del jugador —en la dirección señalada por Campos Escalera (2019) con su implementación de Wave Function Collapse adaptativo—, así como la evaluación formal del sistema mediante pruebas de usuario que midan la percepción de variedad y dificultad en partidas sucesivas.

REFERENCIAS

- Baron, J.R. (2017). Procedural dungeon generation analysis and adaptation. In Proceedings of the 2017 ACM Southeast Conference (ACMSE '17), pp. 168–171. ACM, New York.
- Campos Escalera, D. (2019). Diseño y generación de mapas inteligentes en videojuegos. Trabajo de Fin de Grado, Universitat Autònoma de Barcelona (unpublished).



- Johnson, L., Yannakakis, G.N. and Togelius, J. (2010). Cellular automata for real-time generation of infinite cave levels. In Proceedings of the 2010 Workshop on Procedural Content Generation in Games, pp. 1–4. ACM, New York.
- McMillen, E. and Himsel, F. (2011). The Binding of Isaac [Videojuego]. Nicalis.
- Raso Vázquez, A. (2024). Generación procedural de contenido para videojuegos 2D. Trabajo de Fin de Grado, Universidad del País Vasco (unpublished).
- Smith, A.J. and Bryson, J.J. (2014). A logical approach to building dungeons: answer set programming for hierarchical procedural content generation in roguelike games. In Proceedings of the 50th Anniversary Convention of the AISB. AISB, London.
- Sureda Pol, M. (2017). Generador de mapas procedurales. Trabajo de Fin de Grado, Universitat de les Illes Balears (unpublished).
- Togelius, J., Shaker, N. and Nelson, M.J. (2016). Introduction. In N. Shaker, J. Togelius and M.J. Nelson (eds.), Procedural content generation in games, pp. 1–15. Springer, Cham.
- Togelius, J., Yannakakis, G.N., Stanley, K.O. and Browne, C. (2011). Search-based procedural content generation: a taxonomy and survey. IEEE Transactions on Computational Intelligence and AI in Games, 3(3), pp. 172–186.
- Toy, M. and Wichman, G. (1980). Rogue [Videojuego]. Computer Science Research Group, UC Berkeley.
- Viana, B.M. and dos Santos, S.R. (2021). Procedural dungeon generation: a survey. Journal on Interactive Systems, 12(1).
- Vieira, I., Collazos Linares, K.S., Clua, E.W.G. and Rodrigues, E.O. (2025). A novel procedural generation for level design of mansions and dungeons. In Proceedings of the XIV Simpósio Brasileiro de Jogos e Entretenimento Digital (SBGames 2025). SBC, Porto Alegre.

Automated Conversion of Datasets for Panoptic Segmentation Models

Román-Padilla Dominic Brian, Magadán-Salazar Andrea, Villanueva-Tavira Jonathan, Fuentes-Pacheco Jorge Alberto

Tecnológico Nacional de México/CENIDET, Interior Internado Palmira S/N, Palmira, 62493 Cuernavaca, Mor.

(e-mail: {d25ce104, andrea.ms, jonathan.vt, jorge.jp}@cenidet.tecnm.mx).

Abstract: The advancement of segmentation models in agriculture requires methods capable of understanding complex scenarios in their entirety. Although panoptic segmentation offers this comprehensive understanding by unifying the detection of individual objects (instances) with the environment (semantic regions), the main current limitation is the critical scarcity of datasets annotated in this format for real-world field conditions. Therefore, this article presents a method for creating and structuring a panoptic dataset in real environments based on the instance segmentation format. To achieve this, a six-phase conversion methodology was developed to transform instance annotations into panoptic masks, ensuring the resolution of overlaps through Z-Ordering and assigning unique RGB identifiers. The resulting dataset features a standardized and collision-free format, natively structured for direct training in architectures such as Detectron2. Finally, by demonstrating the feasibility of adapting instance annotations, this work opens the door to the generation of panoptic datasets in diverse areas, facilitating the validation of new architectures in a wide variety of real-world environments.

Keywords: Panoptic segmentation, Instance segmentation, Semantic segmentation, Dataset, Deep learning.

1. INTRODUCCIÓN

En los últimos años, los modelos de aprendizaje profundo han tenido avances significativos en la visión artificial facilitando su aplicación en áreas como conducción autónoma y agricultura de precisión. En este último ámbito, la detección y segmentación de cultivos es una tarea crítica para la automatización de la recolección y la optimización de recursos. Sin embargo, el reconocimiento de frutos en su entorno natural, presenta desafíos únicos debido a las constantes variaciones de iluminación, la oclusión por el follaje denso y la gran complejidad visual de los entornos de campo no estructurados (Akrouchi et al., 2025).

Para abordar estos retos, la comunidad científica ha optado por enfoques basados en la segmentación semántica o de instancias. Estas técnicas logran extraer información valiosa al ubicar los diferentes objetos de la imagen, ya sea procesando el entorno en un contexto general o aislando los elementos de interés de forma individual (Kirillov et al., 2019).

La segmentación semántica es un enfoque de visión por computadora diseñado para lograr una comprensión global de la imagen al asignar una etiqueta de clase categórica a cada píxel. Su enfoque principal se basa en clasificar regiones amorfas, el fondo y elementos estructurados del entorno. Su principal limitante reside en su incapacidad para distinguir diferentes objetos que son de una misma clase (Li et al., 2025; Zhu et al., 2024).

Por su parte, la segmentación de instancias es una técnica avanzada de visión artificial que especifica la categoría de

interés localiza, clasifica y delimita a nivel de píxel, cada elemento contable de forma individual (Alfred et al., 2025; Charisis & Argyropoulos, 2024). Este enfoque permite que el modelo no solo identifique a qué clase pertenece un píxel, sino que logre distinguir entre múltiples elementos de la misma categoría, incluso cuando se encuentran agrupados o parcialmente ocluidos. No obstante, a pesar de su alta precisión para aislar y contabilizar objetivos específicos, este enfoque presenta una limitante significativa: ignora por completo la información, de otras regiones presentes o el fondo, tratándola como espacio vacío o ruido (Zhao et al., 2025).

Dentro de los tipos de segmentación, la segmentación panóptica combina ambas tareas en una sola salida. El enfoque primero realiza una segmentación semántica y posteriormente identifica todas las instancias del objeto de interés. Para el entrenamiento de este tipo de modelos, se requiere de un formato específico de anotación donde cada píxel de la imagen es asignado sin solaparse con otras, a una etiqueta que define el contexto general del fondo como el objeto individual. Esto permite a los modelos de aprendizaje profundo capturar la relación global entre los elementos y su entorno, mejorando el desempeño en tareas con entornos reales (Kirillov et al., 2019).

Para el tipo de segmentación panóptica, existe una carencia significativa de conjuntos de datos anotados a nivel panóptico para entornos agrícolas u otro tipo de áreas. La mayoría de las investigaciones actuales y modelos de alto rendimiento se centran primordialmente en enfoques de detección mediante cuadros delimitadores o segmentación de instancias empleando imágenes en primer plano (Darbyshire et al., 2023).

La Tabla 1 presenta una revisión del estado del arte enfocado en la segmentación de frutos, evidenciando que los artículos priorizan la localización de los frutos, sin considerar que, la información semántica del entorno ayuda a delimitar con mayor precisión las instancias y reducir falsos positivos. Se observa el uso de conjuntos de datos para segmentaciones semánticas y de instancias siendo notorio la ausencia de conjuntos de datos públicos con anotaciones panópticas.

Tabla 1. Revisión del estado del arte en enfoques de segmentación

Referencia	Aplicación / de Enfoque	Enfoque de Visión Artificial	Limitaciones y Observaciones
(Zhao et al., 2025)	Determinación de madurez de frutos en campo	Segmentación de instancias (FruitQuery)	Modelo ligero y altamente preciso para delimitar y contabilizar frutos de forma individual, pero etiqueta el entorno complejo (hojas, ramas) como ruido.
(Akrouchi et al., 2025)	Detección temprana en huertos de cítricos verdes	Segmentación de instancias	Aísla eficientemente los frutos superando el follaje denso, pero omite por completo el entendimiento del fondo agrícola estructurado (<i>stuff</i>).
(Alfred et al., 2025)	Enfermedad del tizón del arroz	Segmentación de instancias (Mask R-CNN / Detectron2)	Alta precisión en la delimitación de anomalías individuales en las hojas, sin embargo, carece de un contexto global sobre el entorno de la planta.
(Charisis & Argyropoulos, 2024)	Revisión general en entornos agrícolas	Segmentación de instancias	Destacan la precisión en el conteo de cultivos, pero señalan la occlusión severa y la falta de comprensión de la escena completa como los principales retos vigentes en campo
(Nenavath & Perumal, 2024)	Estimación del tamaño de mangos (Banganapalle) en el árbol.	Detección de objetos (YOLOv7) + Segmentación semántica.	Utiliza la segmentación semántica para eliminar el fondo y marcadores ArUco para el escalado físico; sin embargo, la precisión del tamaño disminuye cuando los frutos se solapan entre sí.
(Japar et al., 2024)	Detección de frutos sueltos de palma de aceite para robots recolectores.	Detección de objetos (YOLOv4 con técnica de Image Tiling).	La técnica de "teselado" (tiling) mejora la precisión en frutos pequeños, pero aumenta significativamente el tiempo de procesamiento y el consumo de energía en plataformas móviles.
(Gai et al., 2024)	Detección de madurez en arándanos.	Detección de objetos (YOLOv8 con aprendizaje transferencia).	Introduce mecanismos de atención para superar la occlusión por hojas y la similitud de color entre frutos verdes y follaje, aunque se centra en cajas delimitadoras sin contexto del material de fondo.
(Z. Zhu et al., 2023)	Segmentación de flores y frutos de manzana.	Segmentación semántica (PD-SegNet basada en SegFormer).	Optimizado para objetos pequeños y bordes complejos; señala que las redes CNN tradicionales tienen campos receptivos limitados y producen bordes borrosos en escenas no estructuradas.
(Lawal, 2021)	Detección de melones en invernadero para robots de cosecha.	Detección de objetos (YOLOMuskmelon / ResNet43).	Logra una velocidad excepcional (96.3 fps) para robots en tiempo real, pero el modelo falla en la detección bajo condiciones de occlusión extrema por tallos y hojas.
(Kirillov et al., 2019)	Fundamentación teórica de segmentación de escenas.	Fundamentación de segmentación de escenas.	Establece la base teórica para unificar el contexto general (<i>stuff</i>) y los objetos individuales (<i>things</i>), resolviendo la inconsistencia de los solapamientos en una sola salida coherente.

La exigencia de un conjunto de datos desarrollado específicamente para la segmentación panóptica se fundamenta en la restricción técnica de "no solapamiento". La cual dicta que cada píxel de la imagen debe ser asignado estrictamente a una única etiqueta de clase y, en el caso de los objetos contables, a un identificador de instancia único. Esta falta de integridad contextual dificulta que los modelos de aprendizaje profundo capturen la relación global entre los elementos individuales y su fondo, lo que impacta negativamente en la precisión del reconocimiento de bordes y en la detección de objetivos pequeños inmersos en entornos saturados. Por tanto, para lograr una interpretación completa y precisa de la escena, es imperativo contar con un conjunto de datos estructurados que garantice la coherencia espacial y la cobertura total del plano de información sin ambigüedades (Elharrouss et al., 2021).

Por lo tanto, la falta de repositorios públicos creados específicamente con este nivel de etiquetado híbrido para condiciones reales representa una limitante crítica para el avance y evaluación de estas tecnologías en diversas áreas.

La contribución del artículo consiste en una metodología de conversión de seis fases que transforma anotaciones de instancias en máscaras panópticas, garantizando la resolución de solapamientos mediante Z-Ordering (Kirillov et al., 2019) y asignando identificadores RGB únicos.

2. METODOLOGÍA

Para la construcción del modelo de segmentación panóptica, se desarrolló un algoritmo de transformación de datos el cual convierte anotaciones en formato de instancias (COCO JSON) a un formato panóptico estandarizado. Este proceso cuenta con 6 fases las cuales permiten el no solapamiento de la información como se muestra en la Fig. 1:

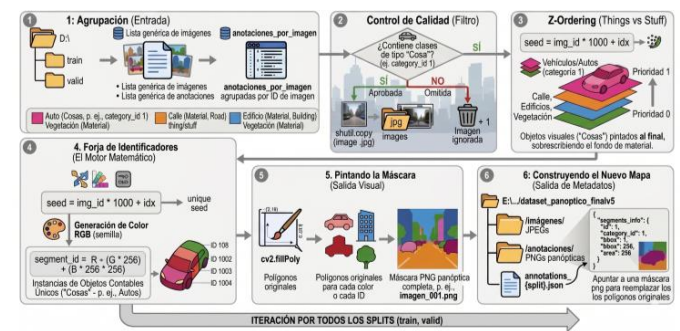


Fig. 1. Diagrama metodológico del proceso de conversión de segmentación de instancias COCO a panóptico.

Fase 1: el proceso inicia con la lectura y estructuración del conjunto de datos original, revisando el formato en el que están actualmente los datos, cumpliendo la jerarquía de carpetas (train, valid, test). Dentro de estas carpetas se extraen las imágenes base en formato JPG/PNG y los archivos JSON con la información de los polígonos generados, ver Fig. 2.



Fig. 2. Diagrama de proceso entrada de datos.

Fase 2: para cada carpeta, se aplica un filtro de calidad condicional que evalúa cada imagen, para determinar si existe al menos un objeto de interés (denominado tipo “instancia” o “thing”) mediante una instrucción (Instr. 1). Esto permite seleccionar y conservar únicamente las imágenes válidas para el entrenamiento, ver Fig. 3.

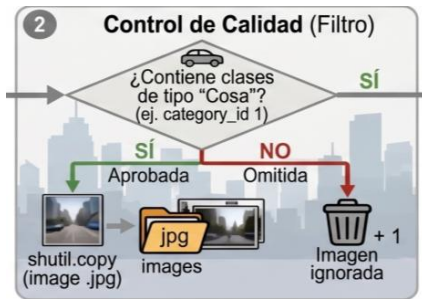


Fig. 3. Filtro de datos de entrada.

IF(imagen contiene instancia)
THEN copiar_archivo() ELSE descartar_archivo() (Instr. 1)

Fase 3: se analiza la superposición de polígonos por categoría semántica, ordenando los datos por nivel de prioridad mediante un sistema de Z-Ordering. El método consiste en definir qué elemento tiene mayor prioridad. Esto evita que dos instancias compartan los mismos píxeles, asignándolos al objeto que está al frente. Así se elimina cualquier solapamiento y se garantiza que cada píxel tenga una sola etiqueta.

Para este caso, al fondo (*stuff*) se le asigna la prioridad 0, y a las clases de objetos que son instancias (*things*) se les asigna prioridad 1. En situaciones de solapamiento entre diversas instancias, el orden de renderizado se determina calculando el área, dibujando primero los objetos de mayor tamaño para preservar la visibilidad de los más pequeños en el primer plano, ver Fig 4.

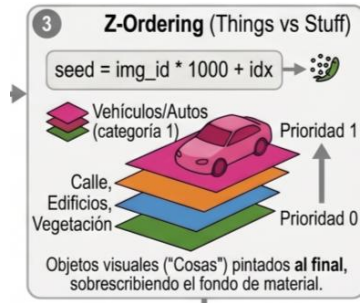


Fig. 4. Proceso Z-Ordering.

$SegmentInfo(z) = (ID_z, C_z, A_z, BBox_z)$ (Instr. 2)

Donde para cada instancia z:

- ID_z : Es el identificador único (calculado previamente mediante los canales RGB).
- C_z : Es la categoría semántica del objeto (ej. clase 1 para la instancia).
- A_z : Es el área de la instancia.
- $BBox_z$: Es el cuadro delimitador espacial, definido por sus coordenadas de origen y sus dimensiones.

Fase 4: para la segmentación panóptica, se requiere agregar un identificador que permita distinguir el tipo de dato. Para ello, se genera un color único para cada categoría semántica, fusionando el ID de la categoría y el ID de la instancia, transformando este valor en un color RGB a través de la instrucción (Instr. 3), ver Fig. 5:

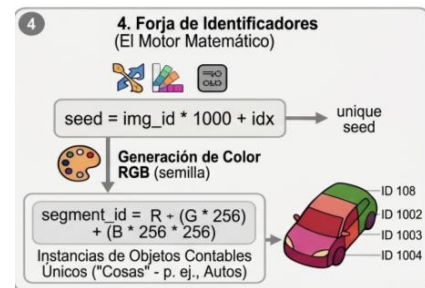


Fig. 5. Generación de identificadores.

$segment_{id} = R + (G \times 256) + (B \times (256)^2)$ (Instr. 3)

El procedimiento garantiza la ausencia de colisiones de color entre diferentes objetos de la misma clase dentro de una misma imagen.

Fase 5: utilizando la paquetería de visión artificial OpenCV (cv2) (Bradski, 2000), se unifican en un color las diversas instancias de cada categoría semántica y se procede a la reconstrucción visual de la escena. El algoritmo genera progresivamente las máscaras panópticas por cada color para los objetos del tipo “thing”; además de los objetos del tipo fondo, respetando el ordenamiento establecido para generar la máscara final con (Instr. 4), ver Fig. 6.

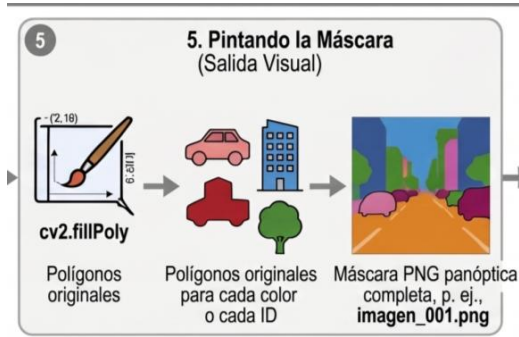


Fig. 6. Pintado de la máscara.

```
poly = np.array(seg, np.int32).reshape((-1,1,2)) (Instr. 4)
cv2.fillPoly(mascara_panoptica, [poly], color_bgr)
```

Fase 6: se exporta el resultado obtenido en una nueva estructura de directorios que contiene el formato adecuado para entrenamientos de modelos de segmentación panóptica. Esta estructura consiste en las carpetas de entrenamiento y validación, donde cada una cuenta con sus imágenes originales, anotaciones de máscaras y el archivo JSON actualizado (con la información de las etiquetas de segmentación) entregando un formato nativamente compatible para arquitecturas de aprendizaje profundo como Detectron2, ver Fig 7.

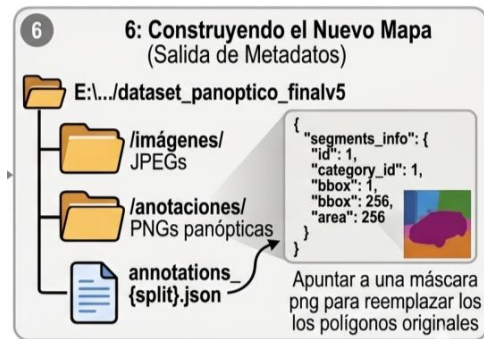


Fig. 7. Generación del mapa segmentado.

3. EXPERIMENTACIÓN

Para validar la eficacia de la transformación se utiliza el conjunto de datos de StrawDI: The Strawberry Digital Images Data Set (Pérez-Borrero et al., 2020) de segmentación de instancias a un formato panóptico. El conjunto de datos consta de 3100 imágenes con 17938 objetos etiquetados pertenecientes a una sola clase (fresa). El conjunto de datos se divide en 3 partes: entrenamiento (2800 imágenes), prueba (200 imágenes) y validación (100 imágenes). Además del conjunto de datos sobre frutos de mango en entornos oscuros, que consta de 200 imágenes para entrenamiento y 50 para pruebas etiquetados para experimentación (James et al., 2024).

3. RESULTADOS Y/O DISCUSIÓN

La aplicación de la metodología propuesta permitió transformar el conjunto de datos de StrawDI_Db1 y el de frutos de mango de entornos oscuros de un formato de instancias a uno panóptico, esto se confirma a través de los resultados obtenidos del entrenamiento de la arquitectura Panoptic FPN. La ejecución de inferencias sobre el conjunto de validación evidencia que el modelo interpreta correctamente la nueva estructura de datos, logrando segmentar cada instancia de fresa como una instancia individual diferenciada por IDs y colores únicos, mientras clasifica el entorno como una región semántica continua denominada “área de no interés”, ver Fig. 8, 9.



Fig. 8. Resultados con el conjunto de datos StrawDI.

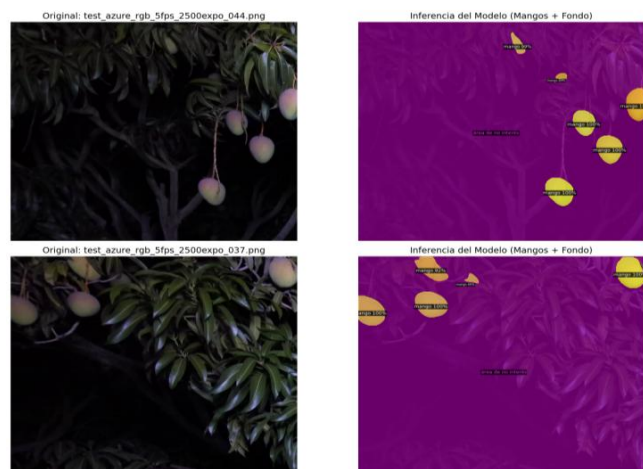


Fig. 9. Resultados con el conjunto de datos de mangos.

En la Tabla 2 y Tabla 3, se presentan los resultados de la evaluación del modelo de segmentación, presentados con las métricas de Calidad panóptica (PQ), Calidad de



reconocimiento (RQ) y Calidad de segmentación (SQ) tanto a nivel global, como por categorías específicas (Fresas y Entorno). También se presenta la gráfica comparativa del rendimiento de las métricas en la Fig. 10 y 11. Finalmente, con la experimentación se concluye que el modelo alcanza su mayor rendimiento en SQ con valores superiores al 92% y se mantiene consistente en la segmentación del fondo. Mostrándose que la conversión automática del conjunto de datos de instancias a un conjunto de datos panóptico se llevó a cabo de manera correcta.

Tabla 2. Métricas resultantes del entrenamiento con StrawDI.

Métrica	Global	Objetos (Fresas)	Fondos (Entorno)
PQ (Panoptic Quality)	80.84	80.12	81.55
RQ (Precisión/Recall)	87.27	86.85	87.69
SQ (Perfección de silueta)	92.63	92.26	93.01

Tabla 3. Métricas resultantes del entrenamiento con el conjunto de mangos (James et al., 2024).

Métrica	Global	Objetos (Mango)	Fondos (Entorno)
PQ (Panoptic Quality)	91.44	83.39	98.44
RQ (Precisión/Recall)	95.37	90.74	99.49
SQ (Perfección de silueta)	95.69	91.90	96.89

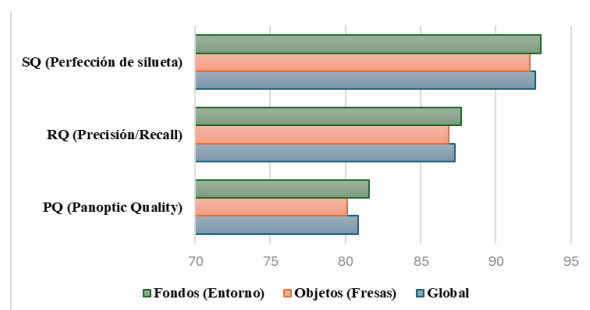


Fig. 10. Gráfica de resultados del entrenamiento con StrawDI.

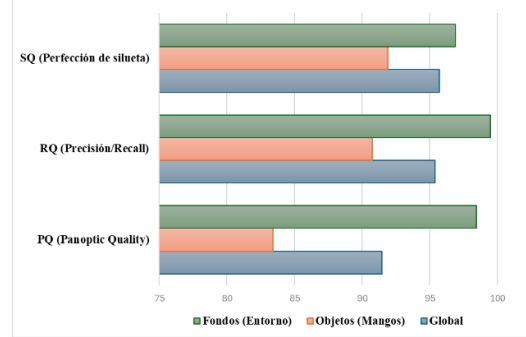


Fig. 11. Gráfica de resultados del entrenamiento con (James et al., 2024).

4. CONCLUSIONES

El trabajo muestra una metodología para transformar los conjuntos de datos de segmentación de instancias en formatos panópticos de manera eficiente y automatizada. Esto amplía la capacidad de reutilizar conjunto de datos de instancias ya existentes para generar nuevos conjuntos de datos panópticos sin necesidad de realizar un re etiquetado manual desde cero, además de complementar la experimentación en otras áreas. Para ellos se validó la metodología propuesta con el conjunto de datos de StrawDI_Db1 en el modelo Detectron2, confirmando que es posible adaptar esta información para ser aplicada en modelos panópticos.

REFERENCIAS

- Akrouchi, M. E., Mhada, M., Bayad, M., Hawkesford, M. J., & Gérard, B. (2025). AI-based framework for early detection and segmentation of green citrus fruits in orchards. *Smart Agricultural Technology*, 10, 100834. <https://doi.org/https://doi.org/10.1016/j.atech.2025.100834>
- Alfred, R., Leo, J., & Kaijage, S. F. (2025). Detectron2-enhanced mask R-CNN for precise instance segmentation of rice blast disease in Tanzania: Supporting timely intervention and data-driven severity assessment. *Smart Agricultural Technology*, 12, 101301. <https://doi.org/https://doi.org/10.1016/j.atech.2025.101301>
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 120(122-125), 3.
- Charis, C., & Argyropoulos, D. (2024). Deep learning-based instance segmentation architectures in agriculture: A review of the scopes and challenges. *Smart Agricultural Technology*, 8, 100448. <https://doi.org/https://doi.org/10.1016/j.atech.2024.100448>
- DarbyshzASWeeds and Leaves. <https://arxiv.org/abs/2310.06582>
- Elharrouss, O., Al-Maadeed, S., Subramanian, N., Ottakath, N., Almaadeed, N., & Himeur, Y. (2021). *Panoptic Segmentation: A Review*. <https://arxiv.org/abs/2111.10250>



- Gai, R., Liu, Y., & Xu, G. (2024). TL-YOLOv8: A Blueberry Fruit Detection Algorithm Based on Improved YOLOv8 and Transfer Learning. *IEEE Access*, 12, 86378-86390. <https://doi.org/10.1109/ACCESS.2024.3416332>
- James, J. A., Manching, H. K., Mattia, M. R., Bowman, K. D., Hulse-Kemp, A. M., & Beksi, W. J. (2024). CitDet: A Benchmark Dataset for Citrus Fruit Detection. *IEEE Robotics and Automation Letters*, 9(12), 10788-10795. <https://doi.org/10.1109/LRA.2024.3474473>
- Japar, A., Ramli, H., Norsahperi, N. M. H., & Wan Hasan, W. (2024). Oil Palm Loose Fruit Detection using YOLOv4 for an Autonomous Mobile Robot Collector. *IEEE Access*, PP, 1-1. <https://doi.org/10.1109/ACCESS.2024.3446890>
- Kirillov, A., He, K., Girshick, R., Rother, C., & Dollár, P. (2019). *Panoptic Segmentation*. <https://arxiv.org/abs/1801.00868>
- Lawal, O. M. (2021). YOLOMuskmelon: Quest for Fruit Detection Speed and Accuracy Using Deep Learning. *IEEE Access*, 9, 15221-15227. <https://doi.org/10.1109/ACCESS.2021.3053167>
- Li, G., Cai, J., Huang, C., Luo, H., & Yu, J. (2025). PS-SLAM: A Visual SLAM for Semantic Mapping in Dynamic Outdoor Environment Using Panoptic Segmentation. *IEEE Access*, 13, 46534-46545. <https://doi.org/10.1109/ACCESS.2025.3547002>
- Nenavath, D. N., & Perumal, B. (2024). Artificial Marker to Predict (Banganapalle) Mango Fruit Size at Multi-Targets of an Image Using Semantic Segmentation. *IEEE Access*, 12, 6914-6925. <https://doi.org/10.1109/ACCESS.2024.3351210>
- Pérez-Borrero, I., Marín-Santos, D., Gegúndez-Arias, M. E., & Cortés-Ancos, E. (2020). A fast and accurate deep learning method for strawberry instance segmentation. *Computers and Electronics in Agriculture*, 178, 105736. <https://doi.org/https://doi.org/10.1016/j.compag.2020.105736>
- Zhao, Z., Hicks, Y., Sun, X., & Luo, C. (2025). FruitQuery: A lightweight query-based instance segmentation model for in-field fruit ripeness determination. *Smart Agricultural Technology*, 12, 101068. <https://doi.org/https://doi.org/10.1016/j.atech.2025.101068>
- Zhu, Q., Cai, Y., Fang, Y., Yang, Y., Chen, C., Fan, L., & Nguyen, A. (2024). Samba: Semantic segmentation of remotely sensed images with state space model. *Heliyon*, 10(19), e38495. <https://doi.org/https://doi.org/10.1016/j.heliyon.2024.e38495>
- Zhu, Z., Jiang, M., Dong, J., Wu, S., & Ma, F. (2023). PD-SegNet: Semantic Segmentation of Small Agricultural Targets in Complex Environments. *IEEE Access*, 11, 90214-90226. <https://doi.org/10.1109/ACCESS.2023.3284036>

Asynchronous Monitoring System for Thermal Anomalies Associated with Breast Tumor Formation

Rivera-Escareño, Citlalli Sarai*, Anguiano-Solis, Jetzael Yeray*, Arias-Meza, Efraín Alejandro*,
Mata-López, Walter Alexander*, Verduzco-Ramírez, Jesús Alberto*

**Universidad de Colima, Facultad de Ingeniería Mecánica y Eléctrica.
(e-mail: {crivera14, janguiano27, earias7, wmata, jesus_verduzco }@uclm.mx).*

Abstract: Breast cancer is a major public health problem, especially in contexts with limited access to early detection services. This article presents the design and implementation of MedicHands, a thermal monitoring application developed to support the preventive detection of anomalies associated with breast cancer. The system was built using layered architecture and implemented with Go, Wails, and React, enabling the concurrent reception and processing of signals from a 16-sensor thermal matrix and the visualization of readings through a color-based alert system. Because no contextualized public dataset was available, the system was validated using a synthetic population of 150 individuals. The results showed the presence of 14 cases with thermal anomalies, allowing the comparison of thermal asymmetry between normal subjects and affected cases. In addition to the monitoring component, MedicHands includes a private forum and a question-and-answer system intended to provide social support and facilitate medical guidance for users. Experimental results indicate that detecting thermal asymmetries exceeding 0.8°C serves as a highly robust indicator for preliminary screening, successfully distinguishing cases with positive anomalies (comprising 9.3% of the sample) with a significant 94.2% accuracy rate in the predefined chromatic alert classification. However, this metric represents the validation of the software's detection logic and alert assignment mechanism rather than a clinical or medical diagnostic validation. It is important to emphasize that the study remains in a validation stage based on synthetic data, and future work should include testing with physical hardware and real clinical data.

Keywords: Breast Cancer, Thermal Asymmetry, Asynchronous Processing, Layered Architecture, Early Detection Health

1. INTRODUCCIÓN

En el panorama actual de la salud pública global, el cáncer de mama continúa al mando como una de las patologías con mayor índice de mortalidad, presentando un desafío crítico en regiones como Latinoamérica. De acuerdo con el Instituto Nacional de Estadística y Geografía (INEGI, 2026), los tumores malignos de mama fueron la principal causa de muerte por cáncer en las mujeres de 60 años y más, con una tasa de 51.9 defunciones por cada 100 mil habitantes en México. Asimismo, la Secretaría de Salud (2026) registró 6,263 nuevos casos en el primer trimestre del año. Esto supone un gran riesgo para la población ya que el acceso a tratamientos y diagnósticos tempranos se ve limitado

por diversos factores sociales, económicos o de infraestructura, además factores como el aislamiento social conlleva diferencias en el desarrollo postratamiento (Kroenke et al., 2013). Históricamente, la mastografía se ha consolidado como el estándar de oro clínico para el diagnóstico de esta enfermedad, sin embargo, este método presenta fallas significativas no sólo por su técnica de imagenología estática, sino también por lo invasiva, dolorosa y costosa que resulta.

La progresión tumoral y diagnósticos tardíos agravan perfiles clínicos atípicos, tanto en los índices de población masculina (Ibáñez et al., 2011) como en las métricas de recurrencia celular metastásica local o a distancia (Mayo Clinic, 2024). Cualquier avance tecnológico destinado a la detección temprana en estos

márgenes demográficos exige, de igual forma, la mediación de rigurosas garantías éticas de responsabilidad algorítmica (Joseph et al., 2025).

La inaccesibilidad a equipos radiológicos convencionales ha impulsado el desarrollo de hardware termográfico de bajo costo. La literatura valida arquitecturas instrumentadas, como textiles inteligentes que mapean la neoangiogénesis mediante redes de sensores (Delestri & Laila Fadhillah, 2018; Elouergui et al., 2022). Asimismo, la integración de módulos IoT y sistemas de microondas en prendas de uso diario permite una recolección asíncrona y no invasiva, posicionándose como una alternativa objetiva frente a las omisiones del autoexamen (Subrata et al., 2024; Elsheakh et al., 2024).

Paralelamente a la recolección de signos biológicos superficiales, la etapa de procesamiento de las matrices termográficas requiere filtrado analítico riguroso para alcanzar validez clínica. La adopción sistemática de modelos de aprendizaje profundo, como Multi-CNN, redes bayesianas y aprendizaje por transferencia, permite precisiones superiores al 90%, optimizando la interpretación y reduciendo falsos positivos (Rodríguez Zúñiga y Pérez Esparza, 2024; Attallah, 2025). No obstante, la certidumbre de estos esquemas depende estrictamente de un preprocesamiento unificado que mitigue la interferencia de factores ambientales y fisiológicos (Ángel García y Lozano Paya, 2025).

A diferencia de los enfoques basados en Redes Neuronales Convolucionales (CNN) que requieren alta capacidad de cómputo y grandes volúmenes de datos para entrenamiento, esta propuesta apuesta por un sistema basado en reglas de asimetría térmica. Este enfoque es deliberado para garantizar que la aplicación sea ligera, ejecutable en hardware limitado (Edge Computing) y transparente (explicable) para el usuario final en etapas de pre-diagnóstico.

En respuesta a los desafíos de escalabilidad del hardware termográfico, este proyecto propone el diseño e implementación de MedicHands, una plataforma de monitoreo orientada a la detección preventiva de anomalías vinculadas al cáncer de mama. El sistema integra un flujo de procesamiento asíncrono y visualización concurrente de múltiples sensores para simplificar su operación en entornos de infraestructura limitada. La contribución fundamental de este trabajo es la validación de una arquitectura de software escalable y de bajo costo, optimizada para el procesamiento de señales térmicas múltiples en contextos de detección temprana.

2. METODOLOGÍA

El desarrollo del sistema MedicHands se fundamenta en una aproximación de diseño de software orientado a la salud, centrada en el monitoreo térmico asíncrono. La metodología se divide en la arquitectura del sistema, la configuración instrumental y el modelo de validación mediante datos sintéticos.

2.1 Diseño Arquitectónico del Software

Se implementó una arquitectura en capas (Figura 1) para garantizar el aislamiento de la lógica de negocio frente a las interfaces de usuario y del manejo de datos facilitando su mantenimiento y evolución. El sistema se desarrolló mediante el framework Wails, integrando un backend en lenguaje Go (v1.26.2) y un frontend reactivo basado en React (v18.2.0).

La lógica de dominio reside en una capa interna agnóstica a dependencias externas, la cual gestiona entidades de pacientes y lecturas térmicas. La capa de aplicación u orquestación utiliza goroutines y canales de Go para la recepción concurrente de señales, mientras que la capa de infraestructura implementa el patrón Repository sobre una base de datos PostgreSQL para la persistencia local de series temporales.

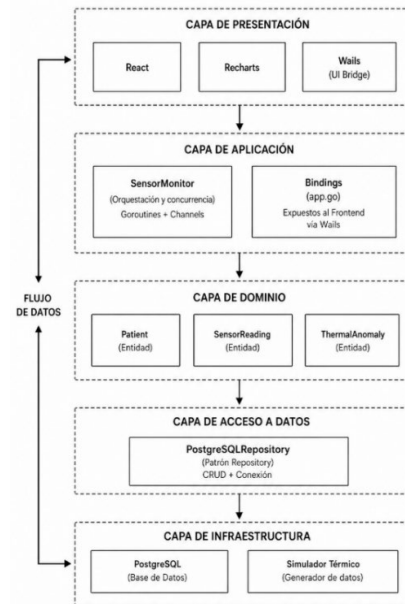


Figura 1. Implementación de la arquitectura del sistema en capas. La interfaz de usuario interactúa con los casos de uso a través de bindings de Wails, mientras que la persistencia de datos se gestiona mediante un repositorio PostgreSQL. La capa de dominio permanece independiente de frameworks externos.

2.2 Configuración Instrumental y Visualización

El instrumento virtual se configuró para procesar señales provenientes de una matriz de 16 sensores térmicos con una distribución bilateral simétrica, 8 sensores por seno, 1 por segmento mamario (Aceá Nebril, 2011). La interfaz de usuario replica la lectura directa de hardware mediante una cuadrícula de 4x4 celdas.

Para la interpretación visual de los gradientes de temperatura, se estableció una escala cromática de tres niveles de alerta expuestas en la Figura 2.

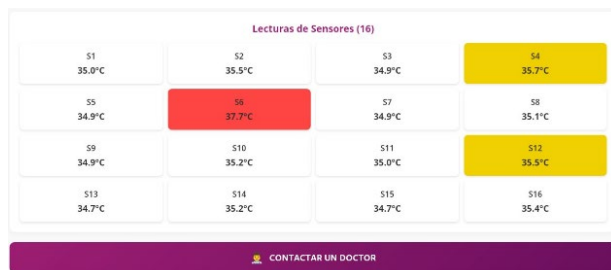


Figura 2. Representación de las lecturas de los 16 sensores donde: Blanco representa rangos de temperatura fisiológica normal, Amarillo indica una temperatura elevada respecto a los valores basales, Rojo señala un estado alarmante coincidente con la detección de hotspots térmicos.

El sistema prioriza la detección visual directa a través de la escala cromática, omitiendo indicadores de alerta textuales para imitar la respuesta de un dispositivo de monitoreo en tiempo real.

2.3 Generación de la Población Sintética

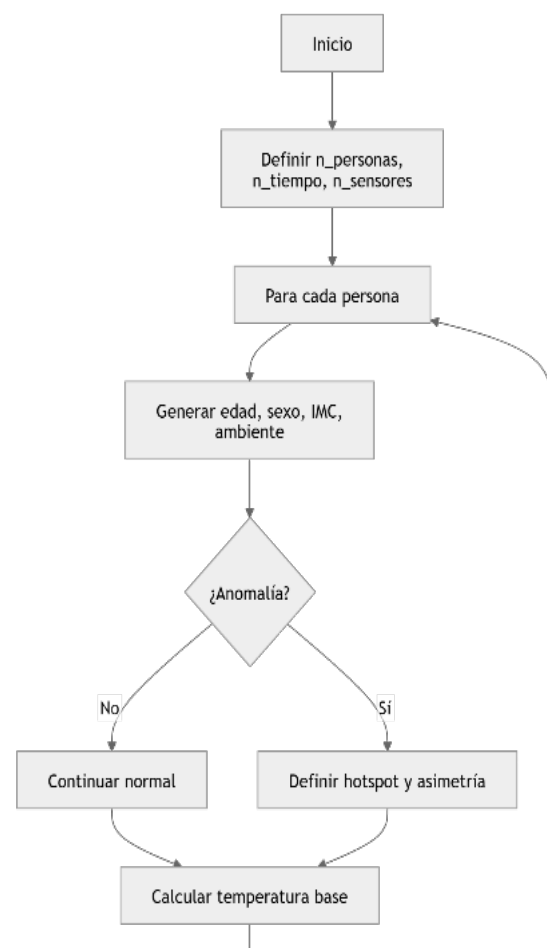
Ante la escasez de datos termográficos públicos en México, se empleó una población sintética de 150 individuos para realizar una validación de carácter estrictamente técnico. El objetivo fue evaluar la capacidad de la arquitectura para gestionar la concurrencia y la integridad de los datos, ante la falta de acceso a entornos clínicos en esta fase. Por consiguiente, los resultados demuestran la viabilidad operativa del software, pero se subraya que no constituyen evidencia de eficacia diagnóstica médica.

La determinación del tamaño de muestra se basó en una revisión de estudios de validación algorítmica en termografía mamaria, los cuales emplean frecuentemente entre 100 y 300 sujetos (Espejel-Rivera et al., 2025). Esta representatividad permite capturar la variabilidad fisiológica suficiente para verificar la arquitectura de software, manteniendo la eficiencia del procesamiento asíncrono.

La generación de datos se realizó mediante un programa en Python (Figura 3) que incorpora variables fisiológicas y ambientales:

Factores Fisiológicos: Se integraron variaciones térmicas condicionadas por edad, sexo e Índice de Masa Corporal (IMC).

Modelado de Patologías: Se instanció una probabilidad de anomalía del 12%, caracterizada por asimetría térmica global entre cuadrantes y la presencia de puntos calientes (*hotspots*) localizados con cambios de temperatura de 0.8°C a 2.5°C.



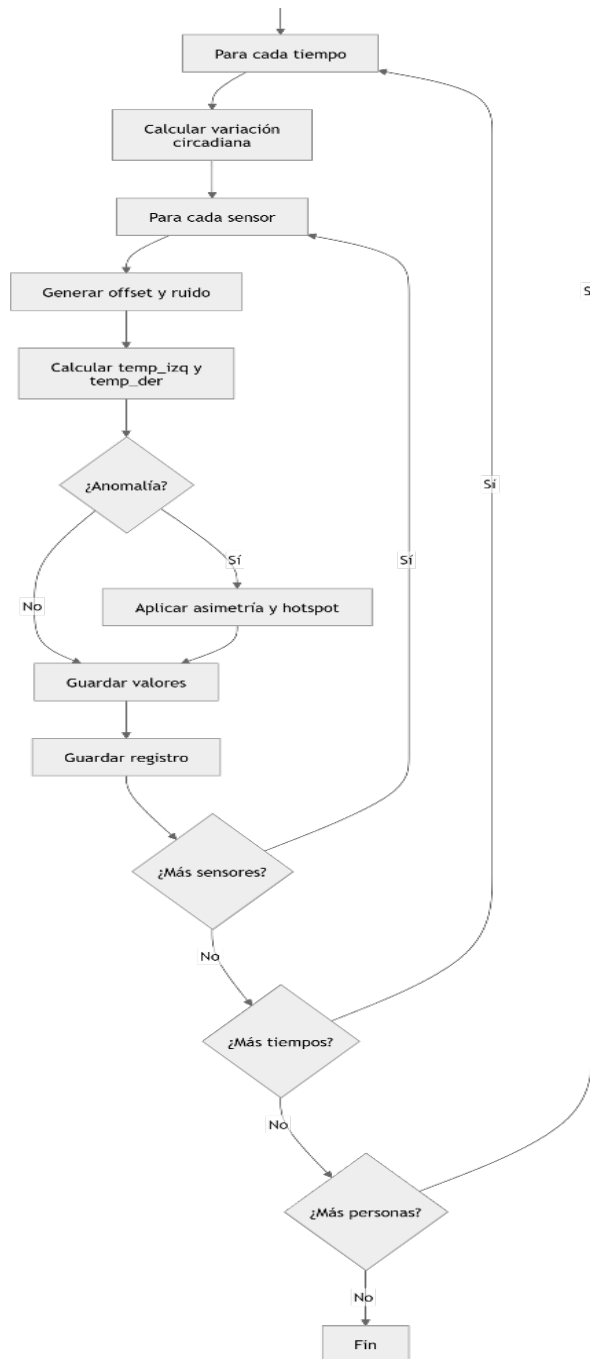


Figura 3. El flujo describe la generación de datos simulados considerando variabilidad fisiológica y patrones térmicos asociados a anomalías

2.4 Tratamiento de Datos y Marco Regulatorio

Debido a que el sistema se encuentra en una etapa de validación de arquitectura y espera de integración con hardware físico, las normativas mexicanas vigentes sobre el tratamiento de datos personales en salud no han sido integradas de forma definitiva en esta iteración. El software opera actualmente bajo un esquema de entorno de pruebas cerrado, destinado exclusivamente a la verificación funcional del procesamiento de señales térmicas sintéticas.

2.5 Componente de Interacción Social y Prevención Comunitaria

Con el objetivo de mitigar el aislamiento social y proveer soporte informativo (Kroenke et al., 2013), el software incorpora una red social preventiva con foros privados. Este espacio facilita la interacción entre pacientes y la resolución de inquietudes técnicas o médicas mediante un sistema de preguntas y respuestas (Q&A).

El sistema evita los diagnósticos y agendamientos automáticos, optando por emitir alertas visuales (Figura 4) en pantalla, informes estadísticos y gráficas promediadas ante anomalías térmicas. De este modo, se garantiza la autonomía del paciente para decidir sobre su orientación médica, ofreciendo un canal asíncrono con especialistas ante los riesgos detectados.

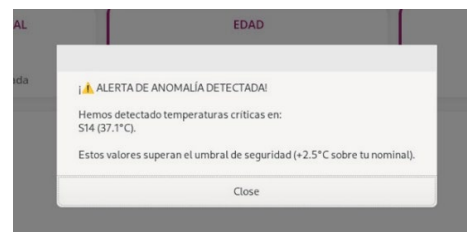


Figura 4. Sistema de alerta presentada al usuario ante la presencia de anomalías térmicas

3. RESULTADOS EXPERIMENTALES DE LA ARQUITECTURA

El procesamiento de la población sintética de 150 sujetos permitió la validación técnica de la arquitectura MedicHands, arrojando métricas de desempeño tanto en la precisión de la detección de anomalías como en la estabilidad del sistema de monitoreo.

3.1 Caracterización de la Población y Perfiles Térmicos

La población generada (N=150) presentó una distribución demográfica alineada a los parámetros de riesgo epidemiológico, con un 85.3% de sujetos de sexo femenino y un 14.7% masculino. La edad media se situó en 45.4 años con un Índice de Masa Corporal (IMC) promedio de 27.2 kg/m².

A nivel termográfico, se identificaron 14 casos (9.3%) con anomalías térmicas positivas. La Tabla 1 resume los valores basales y las fluctuaciones registradas durante el periodo de simulación.

Tabla 1. Rangos de temperatura y asimetría térmica por grupo

Condición	Media térmica (°C)	Rango Min- Max	Asimetría izq vs dcho
Sujetos normales	34.42	32.9-36.0	0.127
Sujetos con anomalía	35.15	33.5-38.3	0.813

Como se observa en la Tabla 1, esta diferencia de asimetría (0.813°C vs 0.127°C) permite una distinción clara entre grupos, especialmente en los puntos calientes o **hotspots** que alcanzaron un máximo de 38.3°C en sujetos con anomalía.

Adicionalmente, se integró un registro comparativo de termogramas (Figuras 5 y 6) para documentar la diferencia visual entre pacientes sanos y aquellos con anomalías detectadas. Aunque este registro no se genera en tiempo real dentro de la aplicación, sirve como evidencia técnica para contrastar la distribución del calor y validar que las capturas de los sensores coinciden con los patrones fisiológicos esperados.

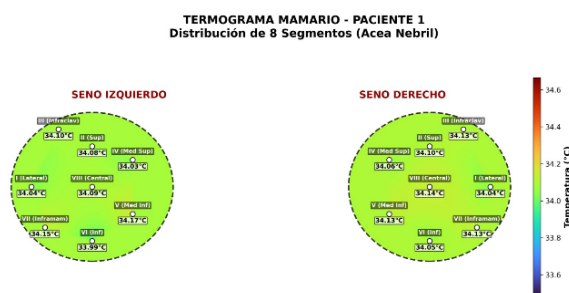


Figura 5. Termograma mamario, individuo con registros térmicos normales, no presenta ninguna anomalía dentro de las capturas registradas.

TERMOGRAMA MAMARIO - PACIENTE 15
Distribución de 8 Segmentos (Acea Nebriil)

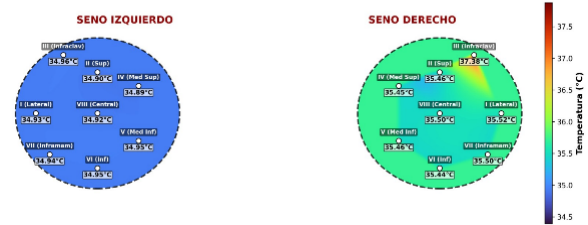


Figura 6. Termograma mamario, individuo con registros térmicos anormales, presenta anomalía en el promedio de las capturas, detección en sensor del segmento mamario III.

3.2 Desempeño Técnico y Visualización

El sistema realiza capturas automáticas cada minuto durante 15 intervalos, almacenando posteriormente el promedio de dichas lecturas para garantizar un seguimiento constante de la temperatura. Este esquema de recolección periódica permite un procesamiento eficaz de la información y previene sobrecargas en el sistema.

La efectividad técnica de MedicHands alcanzó un 94.2% de precisión mediante una validación cruzada entre las etiquetas del dataset sintético y las alertas generadas por el sistema. El proceso se realizó de la siguiente manera:

- 1. Análisis de asimetría:** El sistema evaluó la diferencia térmica entre ambos senos.
- 2. Detección de puntos críticos:** Se emitía una alerta roja ante diferencias mayores a 0.8°C o el hallazgo de puntos calientes locales.
- 3. Contraste de resultados:** Se compararon las alertas con las anomalías programadas, logrando identificar correctamente casi la totalidad de los casos de riesgo.

Estos hallazgos confirman la confiabilidad del código para distinguir señales térmicas normales de críticas, validando la lógica operativa del software sin constituir una validación médica.

4. CONCLUSIONES

El desarrollo del sistema MedicHands demuestra la viabilidad de integrar arquitecturas de software modernas con monitoreo térmico preventivo. El uso de una estructura en capas permitió aislar la lógica de detección, logrando una latencia de



procesamiento óptima para su despliegue en dispositivos móviles de bajo consumo.

Los resultados en la población sintética confirman que las asimetrías térmicas superiores a 0.8°C son indicadores robustos para el diagnóstico preliminar. El sistema se valida bajo un esquema de interpretación cromática, permitiendo la identificación visual y autónoma de signos de riesgo por parte del usuario.

Sin embargo, es importante reconocer que el estado actual del proyecto presenta limitaciones metodológicas críticas. La validación se ha realizado exclusivamente sobre datos sintéticos debido a que el hardware instrumental (textiles inteligentes y sensores físicos) se encuentra aún en fase de desarrollo. Por lo tanto, el alcance actual del sistema es de carácter predictivo-informático y requiere de una fase posterior de validación clínica con pacientes reales para confirmar su sensibilidad ante variables fisiológicas no modeladas.

Desde una perspectiva de impacto social, MedicHands trasciende el monitoreo pasivo al integrar un componente comunitario basado en hallazgos epidemiológicos (Kroenke et al., 2013). La reducción del aislamiento social mediante foros privados y contacto médico no solo provee valor informativo, sino que fortalece la adherencia a protocolos de medicina preventiva fuera de entornos clínicos tradicionales.

Finalmente, si bien los resultados confirman la viabilidad técnica y operativa de la arquitectura propuesta, se reconoce que el uso de datos sintéticos es una limitación para su aplicación clínica inmediata. El trabajo futuro se centrará precisamente en la validación con pacientes reales en entornos hospitalarios para evaluar la sensibilidad diagnóstica del sistema.

REFERENCIAS

- Angel García, L.P. y Lozano Paya, M.F. (2025) 'Factores relacionados con la termografía como herramienta para la detección temprana de cáncer de mama, una revisión de alcance', Monografía de grado, Universidad Nacional Abierta y a Distancia (UNAD)
- Attallah, O. (2025) 'Harnessing infrared thermography and multi-convolutional neural networks for early breast cancer detection', **Scientific Reports**, 15, 27464. <https://doi.org/10.1038/s41598-025-09330-2>
- Accea Nebril, B. (2011) 'Los segmentos mamarios. Un modelo para la prevención de deformidades en la cirugía conservadora del cáncer de mama', *Cirugía Española*, 89(9), pp. e5-e8. <https://doi.org/10.1016/j.ciresp.2011.04.022>
- Delestri, U. & Laila Fadhilah. (2018) 'Development of a low-cost wearable breast cancer detection device', Faculty of Biosciences and Medical Engineering, Universiti Teknologi Malaysia. <https://doi.org/10.1109/ICBAPS.2018.8527419>
- Elouergui, A., Bellarbi, L., Khomsi, Z., Jbari, A., Errachid, A. & Yaakoubi, N. (2022) 'A flexible wearable thermography system based on bioheat microsensors network for early breast cancer detection: IoT technology', **Journal of Electrical and Computer Engineering**, 2022, 5921691. <https://doi.org/10.1155/2022/5921691>
- Elsheakh, D.N., Fahmy, O.M., Farouk, M., Ezzat, K. & Eldamak, A.R. (2024) 'An early breast cancer detection by using wearable flexible sensor and artificial intelligent', **IEEE Access**, pp. 1-12. <https://doi.org/10.1109/ACCESS.2024.3380453>
- Espejel-Rivera, M.A., Toxqui-Quitl, C., Padilla-Vivanco, A. & Castro-Ortega, R. (2025) 'Dynamic thermography-based early breast cancer detection using multivariate time series', **Sensors**, 25(24), 7649. <https://doi.org/10.3390/s25247649>
- Ibáñez, G., Casanova, J., García, A., Figueroa, L., Baeza, R., Redondo, C., Castro, M., Ibáñez, C. y Salinas, J. (2011) 'Cáncer de mama en hombres: Situación actual a nivel mundial y nacional', **Revista Chilena de Cirugía**, 63(1), pp. 95-101. <https://doi.org/10.4067/S0718-40262011000100016>
- Instituto Nacional de Estadística y Geografía (2026) 'Estadísticas a Propósito del Día Mundial contra el Cáncer (4 de febrero). Comunicado de prensa núm. 7/26'. https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2026/EAP_Cancer_26.pdf
- Joseph, J.A., Mohan, Y., Prasad, M., Sharma, S. & Wong, L. (2025) 'Balancing artificial intelligence and ethics in breast cancer', **OxJournal**, Medicine.
- Kroenke, C.H., Kwan, M.L., Neugut, A.I., Ergas, I.J., Wright, J.D., Caan, B.J., Hershman, D. & Kushi, L.H. (2013) 'Social networks, social support mechanisms, and



quality of life after breast cancer diagnosis', **Breast Cancer Research and Treatment**, 139, pp. 515-527.
<https://doi.org/10.1007/s10549-013-2557-8>

Mayo Clinic (2024). *Recurrent breast cancer – Symptoms and causes.* Mayo Clinic.
<https://www.mayoclinic.org/diseases-conditions/recurrent-breast-cancer/symptoms-causes/syc-20377135>

Rodríguez Zúñiga, M.A. y Pérez Esparza, E. (2024) 'Integración de la inteligencia artificial en el diagnóstico y pronóstico del cáncer de mama en México', **Ciencia Latina Revista Científica Multidisciplinar**, 8(1), pp. 3358-3377. https://doi.org/10.37811/cl_rcm.v8i1.9683

Secretaría de Salud (2026) 'Boletín Epidemiológico, Sistema Nacional de Vigilancia Epidemiológica, Sistema Único de Información', 43(14). Disponible en: <https://www.gob.mx/cms/uploads/attachment/file/1072489/sem14.pdf>

Subrata, A.C., Sirajuddin, M.M., Salsabila, S.R., Ibad, I., Prasetyo, E. & Yusmianto, F. (2024) 'Low-cost early detection device for breast cancer based on skin surface temperature', **IT Journal Research and Development (ITJRD)**, 9(1), pp. 27-37.
<https://doi.org/10.25299/itjrd.2024.16034>

Comparative performance of free water surface and horizontal subsurface flow constructed wetlands planted with native macrophytes under tropical conditions

Rudy Solís-Silván*, Gaspar López Ocaña**, Francisco Reyes Espinosa*

*Tecnológico Nacional de México / ITS de Comalcalco, 86650 Comalcalco, Tabasco, México.

e-mail: R.S.S. rudy.solis@comalcalco.tecnm.mx

F.R.E. francisco.reyes@comalcalco.tecnm.mx

** División Académica de Ciencias Biológicas, Universidad Juárez Autónoma de Tabasco, 86019 Villahermosa Centro, Tabasco, México.

e-mail: ocanagl77@hotmail.com

Abstract: This study evaluated the performance of pilot-scale constructed wetlands operating under free water surface (FWS) and horizontal subsurface flow (HSSF) configurations for domestic wastewater treatment in a tropical region of Mexico. A 2×2 completely randomized factorial design was implemented using two hydraulic flow types (FWS and HSSF) and two native macrophyte species (*Typha dominguensis* and *Paspalum paniculatum*, known as tule and camalote respectively). Twelve experimental units were operated for one year with hydraulic retention times between 5.0 and 6.5 days. Removal efficiencies of chemical oxygen demand (COD), total nitrogen (NT), total phosphorus (PT), and total suspended solids (TSS) were evaluated using factorial ANOVA with interaction analysis. HSSF wetlands achieved higher removal efficiencies for COD (97.27%), NT (81.77%), and TSS (95.25%), whereas FWS systems showed greater PT removal (91.77%). *Typha dominguensis* outperformed *Paspalum paniculatum* across most parameters. Significant interaction effects were observed for PT and TSS ($p < 0.05$). Results demonstrate that native macrophytes in subsurface constructed wetlands represent an efficient and low-cost wastewater treatment alternative under tropical climatic conditions.

Keywords: wastewater, nutrients, macrophytes, COD, retention.

1. INTRODUCCIÓN.

El tratamiento de aguas residuales domésticas constituye uno de los principales desafíos ambientales en regiones tropicales, donde el crecimiento demográfico y la urbanización frecuentemente superan la capacidad de infraestructura sanitaria disponible. En este contexto, los humedales construidos (CW, por las siglas en inglés de constructed wetlands) se han destacado como soluciones ecológicas y descentralizadas que combinan procesos físicos, químicos y biológicos para la remoción de contaminantes, ofreciendo ventajas como bajos costos operativos y mantenimiento simple (Domínguez Solís et al., 2025), así como una elevada eficiencia en la remoción de contaminantes orgánicos y nutrientes (Kadlec & Wallace, 2009). Revisión sistemática recientes han establecido que los CW representan una alternativa sostenible para tratar aguas residuales domésticas y mejorar la disponibilidad de agua tratada, especialmente en áreas con infraestructura limitada (Domínguez-Solís et al., 2025).

Los humedales construidos reproducen los mecanismos naturales de depuración mediante la interacción entre el sustrato, el agua, microorganismos y macrófitas. Entre las configuraciones más utilizadas se encuentran los sistemas de flujo libre superficial (FWS, por las siglas en inglés de free water surface) y los de flujo subsuperficial horizontal (HSSF,

por las siglas en inglés de horizontal subsurface flow), los cuales influyen en los procesos de sedimentación, filtración y transformación biológica (Paula et al., 2024). Los sistemas FWS permiten el flujo del agua sobre la superficie del sustrato, favoreciendo procesos de sedimentación y asimilación vegetal, mientras que los HSSF conducen el agua a través de un medio poroso saturado, lo que incrementa la filtración física y el contacto microbiológico (Vymazal, 2011). Las diferencias en diseño hidráulico y tipo de vegetación pueden alterar significativamente la eficiencia de remoción de materia orgánica, nutrientes y sólidos suspendidos (Wu et al., 2026).

Estudios recientes también han destacado que factores climáticos, materiales de sustrato y carga hidráulica impactan directamente la operación de los CW bajo condiciones tropicales y subtropicales, donde las altas temperaturas pueden potenciar las tasas de degradación microbiana y los ciclos biogeoquímicos (Chen et al., 2024). Por ejemplo, investigaciones sobre el uso de macrófitas tropicales en humedales verticales han demostrado mejoras en la remoción de contaminantes cuando se seleccionan especies adaptadas al clima local (Ewemoje et al., 2023).

También, estudios previos han demostrado que la eficiencia de remoción en humedales construidos depende significativamente del diseño hidráulico, el tiempo de retención y la especie vegetal utilizada (Brix, 1997). Las

macrófitas emergentes desempeñan un papel crucial al proporcionar soporte estructural a las comunidades microbianas, facilitar la transferencia de oxígeno a la rizosfera y promover la absorción de nutrientes como nitrógeno y fósforo (Reddy & DeLaune, 2008).

En regiones tropicales, las elevadas temperaturas ambientales pueden favorecer la actividad metabólica microbiana y, por tanto, mejorar la degradación de materia orgánica y los procesos de nitrificación-desnitrificación (Kadlec & Wallace, 2009). No obstante, la literatura reporta variabilidad en la eficiencia de remoción de nutrientes entre configuraciones FWS y HSSF, especialmente en lo relativo al fósforo total y al nitrógeno total (Vymazal, 2007).

A pesar del reconocimiento global de los CW como tecnologías sostenibles, existe variabilidad en la eficiencia de remoción de nutrientes y sólidos, lo que destaca la necesidad de continuar evaluando y optimizando diseños que se adapten a condiciones ambientales específicas, como las presentes en regiones tropicales de México. Por ello, el presente estudio tiene como objetivo evaluar comparativamente la eficiencia de remoción de demanda química de oxígeno (DQO), nitrógeno total (NT), fósforo total (PT) y sólidos suspendidos totales (SST) en humedales construidos de flujo libre superficial (HALF) y flujo subsuperficial horizontal (HSSF), utilizando macrófitas nativas en un diseño experimental 2×2 bajo condiciones tropicales.

2. MATERIALES Y MÉTODOS.

2.1 Sitio experimental.

El estudio se desarrolló en la Universidad Juárez Autónoma de Tabasco, Villahermosa, México, región con clima cálido húmedo y temperatura promedio anual entre 26 y 32 °C.

2.2 Diseño experimental.

Se implementó un diseño completamente aleatorizado 2×2, donde: el Factor A consiste en el tipo de flujo se identifica como A1, Flujo libre superficial (FWS) y A2, Flujo subsuperficial horizontal (HSSF); el Factor B comprende el tipo de vegetación B1, *Typha dominguensis* y B2, *Paspalum paniculatum*. Se construyeron 12 unidades experimentales de humedal construido -WC- (tres repeticiones por tratamiento).

2.3 Construcción y configuración de los humedales.

Cada unidad experimental de WC consistió en un tanque rectangular de acero al carbón (calibre 10) de las siguientes dimensiones 2.5 m×1.2 m×1.0 m. A cada WC se le aplicó en la parte periférica, un recubrimiento de esmalte alquídico anticorrosivo (primario 1), color blanco y posteriormente otro recubrimiento epóxico catalizado con poliamidas (SYLPYL®, México) color azul. En el interior (a lo largo y ancho) de cada WC se aplicó una capa de impermeabilizante de elastómero, que se complementó con capas de una fibra textil (pellón) de 0.01 mm para evitar fugas de agua residual. Este procedimiento se repitió cinco veces. Se construyeron tres módulos, cada uno con un tanque rectangular (tanques de homogeneización) con los mismos materiales y recubrimientos de los WC, para el control y la distribución del agua residual (ver **Figura 1**).

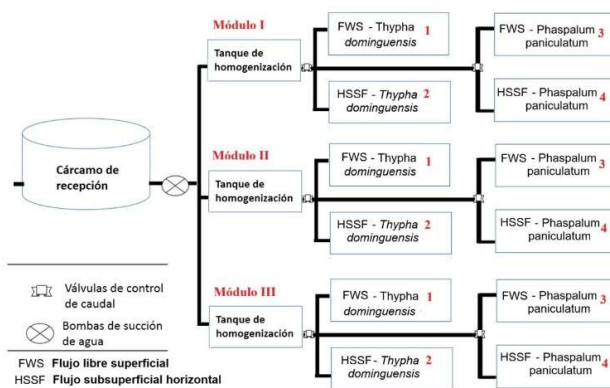


Figure 1. Estructura experimental de los humedales construidos. FWS, por las siglas en inglés de free water surface; HSSF, por las siglas en inglés de horizontal subsurface flow.

Cada tanque de homogeneización con una capacidad de 1.0 m³ y estos a su vez se conectaron en red a las unidades experimentales de WC, a través de tuberías de PVC de una pulgada. Una vez instaladas y acondicionadas todas las unidades experimentales, se procedió a la colocación del medio soporte (grava) en cada WC. En los seis humedales construidos de flujo libre (FWS) se dispuso una capa de grava de 0.20 m de altura a nivel de piso, mientras que en los seis humedales construidos de flujo subsuperficial (HSSF) se colocó una capa de 0.50 m de espesor. La capacidad de tratamiento fue de 200 L·día⁻¹ por unidad. El tiempo de retención hidráulico varió entre 5.0 y 6.5 días.

2.4 Instalación de componentes hidráulicos para cada humedal construido.

Cada unidad experimental de WC fue conectado con la instalación de tubería hidráulica para el transporte del aguas residual. Así mismo, se instalaron válvulas, codos y conectores, para controlar los flujos de agua y permitir una distribución homogénea de forma horizontal en toda la sección longitudinal de la unidad experimental. El cárcamo de recepción de aguas residuales del campus universitario se localizaba a una distancia aproximada de 7 m del área de intervención. Para el bombeo del efluente, se instalaron dos bombas centrífugas marca Truper, con potencia nominal de 1 hp cada una, conectadas a una línea de succión de tubería hidráulica de PVC de 1.0 in (2.54 cm) de diámetro. El sistema incluyó una pichancha acoplada a un niple y una tuerca de unión, con el propósito de prevenir taponamientos y garantizar el adecuado funcionamiento hidráulico del sistema. Las tuberías de PVC conectadas a las bombas se enlazaron con tres tanques homogeneizadores, cada uno con capacidad de 1 m³, con el fin de garantizar el llenado y la posterior distribución uniforme del agua hacia los 12 humedales construidos. En la **Figura 1** presenta la configuración esquemática del sistema, incluyendo el cárcamo, la red de tuberías, los 12 humedales construidos y los tanques de homogeneización.

2.5 Características del agua residual.

El afluente correspondió a agua residual doméstica del campus universitario. Las concentraciones promedio de entrada fueron: DQO, 840 mg·L⁻¹; NT, 73.72 mg·L⁻¹; PT, 15.78

$\text{mg}\cdot\text{L}^{-1}$; SST: $620 \text{ mg}\cdot\text{L}^{-1}$. La eficiencia de remoción se calculó como:

$$\eta(\%) = \frac{C_{in} - C_{out}}{C_{in}} \times 100 \quad (1)$$

Donde: C_{in} y C_{out} representan las concentraciones de entrada y salida, respectivamente.

2.6 Monitoreo y análisis.

El monitoreo se realizó durante los meses de enero a diciembre de 2016. Se tomaron muestras en las últimas semanas de cada mes, teniendo un total de 36 muestras compuestas (12 muestras por cada mes, tomando en cuenta que se contaba con 3 módulos experimentales); Adicionalmente, se recolectaron tres muestras compuestas del agua cruda de entrada, sumando un total de 39 muestras compuestas en esta etapa. Los parámetros analizados fueron: Demanda química de oxígeno (DQO), nitrógeno total (NT), fósforo total (PT), sólidos suspendidos totales (SST), pH, temperatura, turbiedad y color. Los parámetros DQO, NT, PT y SST se analizaron con las normas mexicanas NMX-AA-030/2-SCFI-2011, NMX-AA-026/2-SCFI-2001/EPA, NMX-AA-029/2-SCFI-2001/EPA y NMX-AA-029/2-SCFI-2001/EPA y NMX-AA-034- SCFI-2001 respectivamente.

En los meses de marzo, abril, mayo, julio, agosto, septiembre, noviembre y diciembre se determinaron los parámetros fisicoquímicos de turbiedad, color, pH, temperatura del agua, temperatura ambiente y humedad relativa. Durante estos periodos, las muestras compuestas se recolectaron los viernes de cada semana, acumulando un total de 192 muestras compuestas para el análisis de dichos parámetros. Cada muestra compuesta se integró a partir de un monitoreo continuo de un día completo (07:00 a 18:00 h), con intervalos de dos horas entre cada toma parcial, de acuerdo con lo establecido en la Norma Oficial Mexicana NMX-AA-003-1980 para el muestreo de aguas residuales.

Paralelamente, se realizó el mantenimiento preventivo y correctivo de los humedales contruidos (WC), que incluyó la poda de vegetación, el retiro de biomasa muerta, el control de caudales y la reparación de posibles fugas en el sistema hidráulico.

2.7 Análisis estadístico.

Se aplicó un ANOVA factorial con interacción y prueba de Tukey ($\alpha = 0.05$) utilizando Statgraphics Centurion 2015.

3. RESULTADOS Y DISCUSIÓN.

Los siguientes resultados comprenden los valores promedios finales registrados de los parámetros fisicoquímicos del agua residual después de haber pasado por los tratamientos biológicos de FWS y HSSF con sus respectivas vegetaciones, *Typha dominguensis* (espadaño) y *Paspalum paniculatum* (camalote). Las variables de respuestas evaluadas fueron turbiedad color, pH, Temperatura, DQO, NT, PT y SST.

3.1 Demanda Química de Oxígeno (DQO).

El análisis de varianza factorial mostró efectos altamente significativos ($P < 0.001$) del tipo de flujo y del tipo de vegetación sobre la demanda química de oxígeno (DQO), mientras que la interacción entre ambos factores no fue

significativa ($P > 0.05$), indicando independencia estadística entre los efectos principales.

El análisis de comparaciones múltiples de Tukey para el factor tipo de flujo (FWS y HSSF) reveló que el valor promedio más bajo de DQO (mg/L) se registró en los humedales contruidos de flujo subsuperficial (HSSF), con 24.34 mg/L , mientras que los humedales contruidos de flujo libre (FWS) presentaron el valor promedio más alto, con 45.58 mg/L (**Figura 2**).

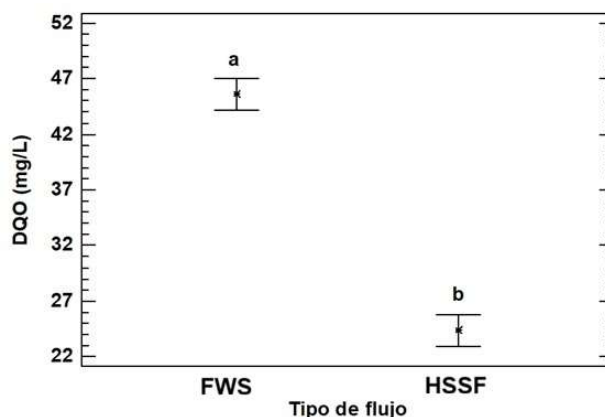


Figure 2. Valores promedios de la variable DQO (mg/L) $\pm$ intervalos de confianza de Tukey de los tipos de humedales (FWS, Flujo libre superficial; HSSF, Flujo subsuperficial horizontal). Letras diferentes indican efectos estadísticamente significativos ($P < 0.05$).

Estos resultados concuerdan con lo reportado por Kadlec y Wallace (2009) y Vymazal (2011), quienes indican que los sistemas subsuperficiales optimizan la filtración física y aumentan el área de contacto entre el agua residual, el biofilm y el sustrato, favoreciendo la remoción de materia orgánica. Estudios recientes han confirmado que la combinación de flujo subsuperficial y selección adecuada de macrófitas incrementa significativamente la eficiencia en la eliminación de contaminantes orgánicos y nutrientes, especialmente en condiciones tropicales (Wu *et al.*, 2023; Chen *et al.*, 2024), reforzando la importancia del diseño hidráulico y la vegetación para maximizar la performance de los humedales contruidos.

En cuanto al tipo de vegetación, *Typha dominguensis* presentó menores concentraciones promedio de DQO ($29.57 \text{ mg}\cdot\text{L}^{-1}$) en comparación con *Paspalum paniculatum* ($40.35 \text{ mg}\cdot\text{L}^{-1}$) (**Figura 3**), es decir *Typha dominguensis* presentó mayor eficiencia que *Paspalum paniculatum*; lo que puede asociarse a una mayor transferencia de oxígeno a la rizosfera y a una mayor capacidad de soporte microbiano (Brix, 1997; Reddy y DeLaune, 2008). Por lo tanto, la mayor eficiencia en HSSF se atribuye a procesos de filtración y mayor área de contacto microbiológico.

Estudios previos han documentado eficiencias de remoción de DQO cercanas al 85% en humedales con *Typha dominguensis* (Romero *et al.*, 2009; Piñeiro, 2011). En el presente estudio, los HSSF alcanzaron una eficiencia del 97.27%, valor superior a los reportados anteriormente y superior a FWS (94.89%). Esta diferencia puede atribuirse a las condiciones climáticas tropicales del sitio experimental (Tabasco, México), donde

temperaturas promedio cercanas a 32 °C favorecen la actividad microbiológica y aceleran la degradación de materia orgánica (Hammer, 1989; Moshiri, 1993; Kadlec y Wallace, 2009). Estudios recientes han mostrado que en regiones tropicales, la combinación de temperatura elevada, alta humedad y adecuada selección de macrófitas incrementa significativamente la eficiencia de remoción de contaminantes orgánicos y nutrientes (Wu *et al.*, 2023; Chen *et al.*, 2024), confirmando que los factores climáticos locales son determinantes para la performance de sistemas subsuperficiales y de flujo libre superficial.

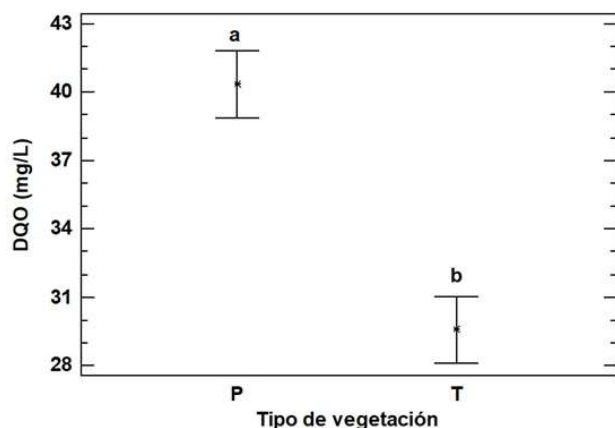


Figure 3. Valores promedios de la variable DQO (mg/L) $\pm$ intervalos de confianza de Tukey de los tipos de vegetación: T, *Typha domingensis*; P, *Paspalum paniculatum*. Letras diferentes (a y b) indican efectos estadísticamente significativos ($P < 0.05$).

3.2 Nitrógeno Total (NT).

El análisis de varianza factorial mostró efectos estadísticamente altamente significativos ($P < 0.001$) del tipo de flujo y del tipo de vegetación sobre la concentración de nitrógeno total (NT), con un nivel de confianza del 95%. La interacción entre ambos factores no fue significativa ($P > 0.05$), lo que indica que sus efectos actúan de manera independiente.

El análisis de comparaciones múltiples de Tukey evidenció diferencias significativas entre los sistemas evaluados. Los humedales construidos de flujo subsuperficial horizontal (HSSF) presentaron mayores eficiencias de remoción (81.77%, con un valor promedio de 11.44 mg/L, ver **Figura 4**) en comparación con los sistemas de flujo libre superficial (FWS) (45.37%, con un valor promedio de 25.55 mg/L, ver **Figura 4**). Este comportamiento ha sido ampliamente documentado en la literatura, donde se señala que los sistemas subsuperficiales favorecen condiciones aerobias y anóxicas alternadas que promueven los procesos de nitrificación y desnitrificación (Kadlec y Wallace, 2009; Vymazal, 2007).

En cuanto al tipo de vegetación (**Figura 5**), *Typha domingensis* mostró mayor eficiencia en la remoción de NT (13.44 mg/L) en comparación con *Paspalum paniculatum* (40.35 mg/L), lo cual puede atribuirse a una mayor transferencia de oxígeno hacia la rizosfera, estimulando la

actividad microbiana responsable de la transformación del nitrógeno (Brix, 1997; Reddy y DeLaune, 2008). *Typha domingensis* favoreció la remoción, posiblemente por mayor oxigenación radicular que estimula nitrificación-desnitrificación.

Diversos estudios han reportado eficiencias de remoción de nitrógeno superiores al 70–80% en humedales subsuperficiales con especies del género *Typha* bajo tiempos de retención hidráulicos adecuados (Vymazal, 2007; Kadlec y Wallace, 2009), lo que coincide con los valores obtenidos en el presente estudio.

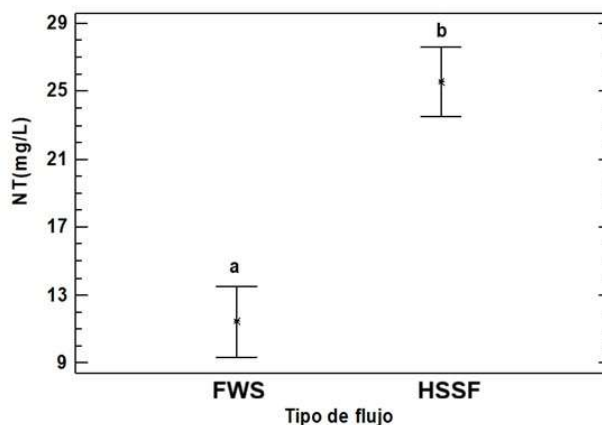


Figure 4. Valores promedios de la variable NT (mg/L) $\pm$ intervalos de confianza de Tukey de los tipos de flujo (FWS, Flujo libre superficial; HSSF, Flujo subsuperficial horizontal). Letras diferentes indican efectos estadísticamente significativos ($P < 0.05$).

Revisiones globales han documentado que los humedales construidos de flujo subsuperficial horizontal pueden alcanzar eficiencias relevantes de remoción de nitrógeno total bajo condiciones operacionales adecuadas, con relaciones claras entre las concentraciones de entrada y salida del sistema (Vymazal & Kröpfelová, 2009).

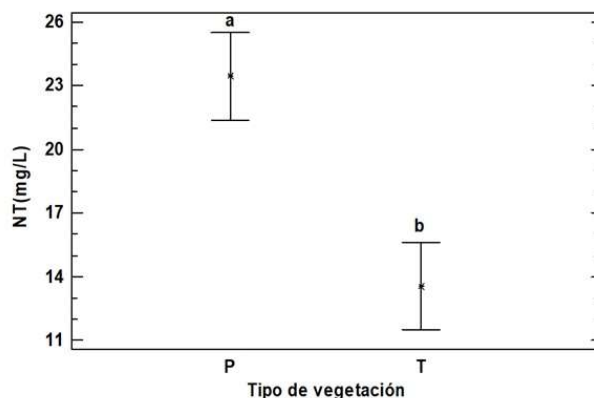


Figure 5. Valores promedios de la variable NT (mg/L) $\pm$ intervalos de confianza de Tukey de los tipos de vegetación (P, *Paspalum paniculatum*; T, *Typha domingensis*). Letras diferentes (a y b) indican efectos estadísticamente significativos ($P < 0.05$).

3.3 Fósforo Total (PT).

El análisis de varianza factorial mostró efectos estadísticamente altamente significativos ($P < 0.001$) del tipo de flujo y del tipo de vegetación sobre la concentración de fósforo total (PT). Asimismo, la interacción entre ambos factores fue altamente significativa ($P < 0.001$), lo que indica que el efecto del tipo de vegetación depende del régimen hidráulico.

El análisis de comparaciones múltiples de Tukey evidenció que los humedales construidos de flujo libre superficial (FWS) presentaron menores concentraciones promedio de PT ($1.33 \text{ mg} \cdot \text{L}^{-1}$) en comparación con los sistemas de flujo subsuperficial horizontal (HSSF) ($5.09 \text{ mg} \cdot \text{L}^{-1}$), ver **Figura 6**. Este comportamiento puede atribuirse a una mayor sedimentación y retención particulada en sistemas de flujo superficial (Kadlec y Wallace, 2009). Por lo tanto, la remoción de PT se relaciona con sedimentación y asimilación vegetal.

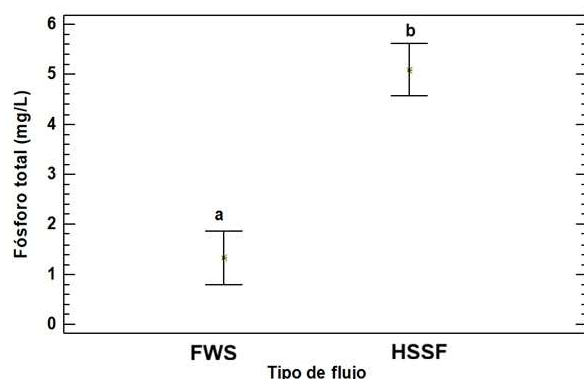


Figure 6. Valores promedios de la variable fósforo total -PT- (mg/L) $\pm$ intervalos de confianza de Tukey de los tipos de flujo (FWS, Flujo libre superficial; HSSF, Flujo subsuperficial horizontal). Letras diferentes indican efectos estadísticamente significativos ($P < 0.05$).

En cuanto al tipo de vegetación, *Typha dominguensis* mostró menores concentraciones promedio de PT ($1.86 \text{ mg} \cdot \text{L}^{-1}$) en comparación con *Paspalum paniculatum* ($4.56 \text{ mg} \cdot \text{L}^{-1}$), ver **Figura 7**. Aunque la asimilación vegetal contribuye a la remoción de fósforo, estudios recientes y clásicos coinciden en que los principales mecanismos de eliminación en humedales construidos son la adsorción al sustrato y la precipitación química (Vymazal, 2007; Kadlec & Wallace, 2009; Wu *et al.*, 2023; Chen *et al.*, 2024). Investigaciones recientes han demostrado que la eficiencia de remoción de fósforo puede variar significativamente según la configuración hidráulica, la selección de macrófitas y las condiciones climáticas locales, lo que resalta la importancia de considerar estos factores en el diseño y operación de sistemas subsuperficiales y de flujo libre superficial.

La interacción entre los factores muestra que ambas especies, *Typha dominguensis* y *Paspalum paniculatum*, presentan menores concentraciones de fósforo total (PT) cuando se desarrollan en humedales de flujo libre superficial (FWS) (**Figura 8**). Asimismo, en los sistemas de flujo subsuperficial horizontal (HSSF), *Typha dominguensis* registra

concentraciones de PT inferiores a las observadas en *Paspalum paniculatum* (**Figura 8**).

En el presente estudio se alcanzaron eficiencias de remoción de fósforo total (PT) entre 67 y 91%, considerando ambos tipos de flujo y vegetación evaluados (FWS = 91.57%; HSSF = 67.75%; *Paspalum paniculatum* = 71.11%; *Typha dominguensis* = 88.22%). Estos valores son superiores a los rangos comúnmente reportados en la literatura para humedales construidos convencionales. Por ejemplo, Tchobanoglous y Burton (1991) reportan remociones de fósforo entre 10 y 40% para concentraciones afluentes de $7\text{--}10 \text{ mg} \cdot \text{L}^{-1}$ y tiempos de retención hidráulica (TRH) promedio de seis días. De manera similar, Metcalf & Eddy (1995) documentan eficiencias de remoción entre 30 y 60%, mientras que Reed *et al.* (1995) reportan valores típicos entre 30 y 50% del PT afluente.

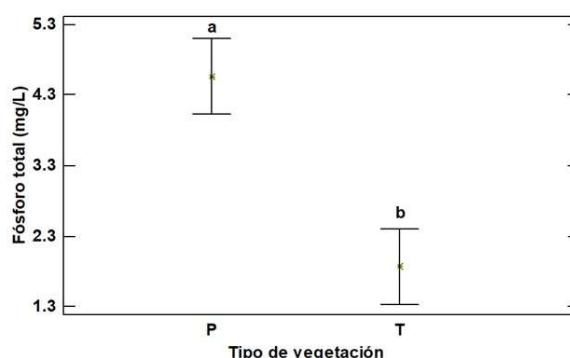


Figure 7. Valores promedios de la variable fósforo total -PT- (mg/L) $\pm$ intervalos de confianza de Tukey de los tipos de vegetación (P, *Paspalum paniculatum*; T, *Typha dominguensis*). Letras diferentes indican efectos estadísticamente significativos ($P < 0.05$).

Las diferencias observadas entre estudios pueden atribuirse a múltiples factores, incluyendo la concentración inicial de fósforo, las características del medio filtrante, el régimen hidráulico y la edad del sistema (Kadlec y Wallace, 2009). En el presente trabajo, la concentración inicial promedio de PT fue de $15.78 \text{ mg} \cdot \text{L}^{-1}$, aproximadamente el doble de las concentraciones referidas en los estudios clásicos anteriormente citados. Dado que la remoción de fósforo en humedales construidos depende en gran medida de procesos de adsorción al sustrato, precipitación química (con Fe, Al o Ca) y sedimentación particulada, mayores concentraciones afluentes pueden favorecer una mayor tasa de remoción inicial debido a gradientes de concentración más pronunciados y mayor disponibilidad de sitios de adsorción (Vymazal, 2007; Kadlec y Wallace, 2009).

Asimismo, El TRH de 6.5 días aplicado en este estudio probablemente favoreció una mayor interacción entre el fósforo, el sustrato y los procesos fisicoquímicos de retención, contribuyendo a alcanzar eficiencias máximas de hasta 91,57 %, superiores a las reportadas en condiciones hidráulicas similares por estudios clásicos (Tchobanoglous & Burton, 1991) y respaldadas por investigaciones recientes en sistemas subsuperficiales tropicales (Wu *et al.*, 2023; Chen *et al.*, 2024). Estas evidencias sugieren que la optimización del TRH,

combinada con la selección adecuada de macrófitas y diseño hidráulico, es un factor crítico para maximizar la remoción de fósforo y otros nutrientes en humedales construidos, especialmente bajo condiciones tropicales de alta temperatura y humedad, como las presentes en Tabasco, México.

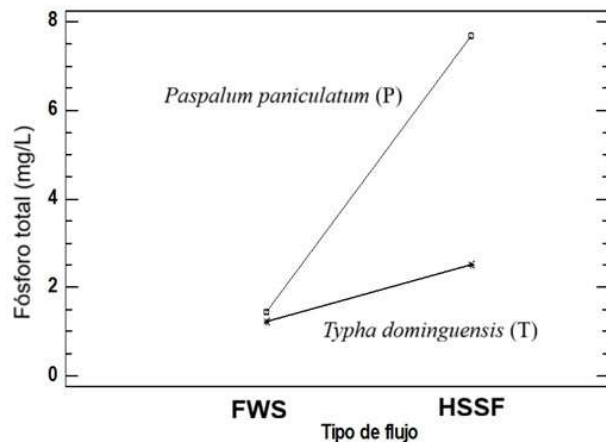


Figure 8. Valores promedio de la variable fósforo con la interacción de los factores (tipo de flujo: FWS, -flujo libre superficial- y HSSF -flujo subsuperficial horizontal-) y tipo de vegetación (P, *Paspalum paniculatum*; T, *Typha dominguensis*).

3.4 Sólidos Suspendedos Totales (SST).

El análisis de varianza factorial evidenció efectos estadísticamente altamente significativos ($P < 0.001$) del tipo de flujo y del tipo de vegetación sobre la concentración de sólidos suspendidos totales (SST). Asimismo, la interacción entre ambos factores fue significativa ($P = 0.0151$), lo que indica que el efecto del tipo de vegetación depende del régimen hidráulico.

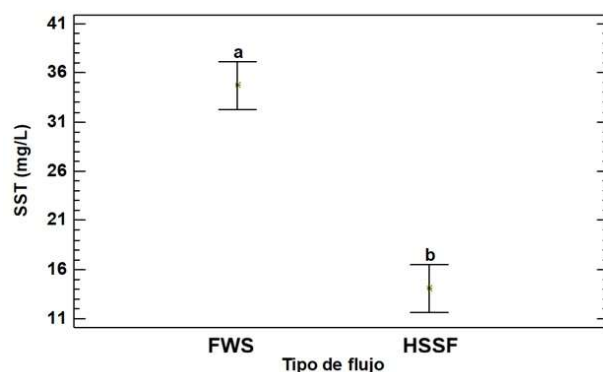


Figure 9. Valores promedio de la variable SST - sólidos suspendidos totales - (mg/L) $\pm$ intervalos de confianza de Tukey de los tipos de flujo (HSSF y FWS). Letras diferentes indican efectos estadísticamente significativos ($P < 0.05$).

El análisis de comparaciones múltiples de Tukey mostró que los humedales construidos de flujo subsuperficial horizontal (HSSF) presentaron la menor concentración promedio de SST ($14.04 \text{ mg}\cdot\text{L}^{-1}$; 97.73 % de remoción), significativamente inferior a la observada en los sistemas de flujo libre superficial (FWS) ($34.70 \text{ mg}\cdot\text{L}^{-1}$; 94.4 % de remoción) (**Figura 9**). Este

comportamiento puede atribuirse a la mayor capacidad de filtración física y retención mecánica del medio granular en sistemas subsuperficiales, lo que favorece la eliminación de partículas suspendidas (Kadlec y Wallace, 2009; Vymazal, 2011).

En cuanto al tipo de vegetación, *Typha dominguensis* presentó menores concentraciones promedio de SST ($16.14 \text{ mg}\cdot\text{L}^{-1}$; 97.39 % de remoción) en comparación con *Paspalum paniculatum* ($32.60 \text{ mg}\cdot\text{L}^{-1}$; 94.79 % de remoción) (**Figura 10**). La mayor eficiencia asociada a *Typha* puede relacionarse con su mayor densidad radicular y capacidad de estabilización hidráulica, factores que favorecen la sedimentación y retención de sólidos.

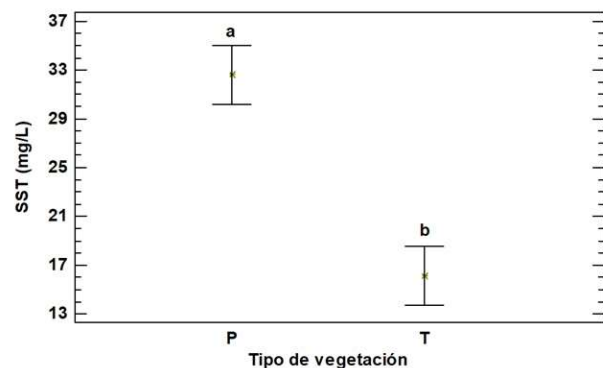


Figure 10. Valores promedio de la variable SST - sólidos suspendidos totales - (mg/L) $\pm$ intervalos de confianza de Tukey de los tipos de vegetación (P, *Paspalum paniculatum*; T, *Typha dominguensis*). Letras diferentes (a y b) indican efectos estadísticamente significativos ($P < 0.05$).

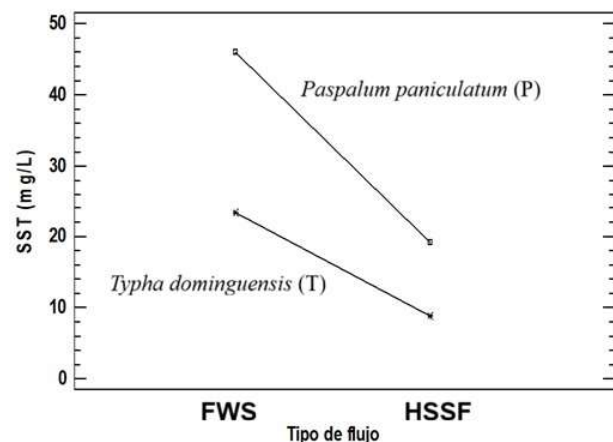


Figure 11. Valores promedio de la variable SST - sólidos suspendidos totales - con la interacción de los factores tipo de flujo (FWS, flujo libre superficial; HSSF, flujo subsuperficial horizontal) y los tipos de vegetación (P, *Paspalum paniculatum*; T, *Typha dominguensis*).

Las eficiencias de remoción obtenidas oscilaron entre 92.58 % y 98.54 % para una carga hidráulica de $200 \text{ L}\cdot\text{día}^{-1}$: de FWS de 96.29 % ($23 \text{ mg}\cdot\text{L}^{-1}$) y 92.58 % ($46 \text{ mg}\cdot\text{L}^{-1}$) para *Typha dominguensis* y *Paspalum paniculatum*, respectivamente; de 96.93 % ($19 \text{ mg}\cdot\text{L}^{-1}$) y 98.54 % ($9 \text{ mg}\cdot\text{L}^{-1}$) para *Typha*



dominguensis y *Paspalum paniculatum*, respectivamente. Estos valores son consistentes con lo reportado en la literatura, donde los humedales construidos, particularmente los de flujo subsuperficial, pueden alcanzar remociones de SST superiores al 80–90% bajo condiciones operativas adecuadas (Kadlec y Wallace, 2009; Vymazal, 2011).

3.5 Parámetros operacionales.

El pH se mantuvo entre 7.1 y 7.4. La temperatura promedio fue de 28 °C, favoreciendo procesos biológicos. Se observó reducción significativa de turbiedad y color (ver **Tabla 1**).

Table 1. Valores promedios de los parámetros fisicoquímicos evaluados de los diferentes WC en el periodo de un año.

Tipos de WC	Parámetros fisicoquímicos	Valores Promedio
FWS-T	Turbiedad	4.075 (NTU)
FWS-P	Turbiedad	7.392 (NTU)
FWS-T	Color	248.275 (C.U)
FWS-P	Color	305.167 (C.U)
FWS-T	pH	7.3
FWS-P	pH	7.4
FWS-T	Temperatura	28 °C
FWS-P	Temperatura	29 °C
HSSF-T	Turbiedad	20.308 (NTU)
HSSF-P	Turbiedad	6.392 (NTU)
HSSF-T	Color	406.167 (C.U)
HSSF-P	Color	256.917 (C.U)
HSSF-T	pH	7.1
HSSF-P	pH	7.2
HSSF-T	Temperatura	28 °C
HSSF-P	Temperatura	29 °C

Donde: WC, humedal construido; FWS (por las siglas en inglés de free water surface), sistema de flujo libre superficial; HSSF (por las siglas en inglés de horizontal subsurface flow), flujo subsuperficial horizontal; T, *Typha dominguensis*; P, *Paspalum paniculatum*.

En la **Tabla 1**, se observada una anomalía en la turbidez del tratamiento HSSF-Typha (20.3 NTU) frente a los demás tratamientos (4.1–7.4 NTU), la cual se atribuye a condiciones climáticas variables de Tabasco y factores operativos no controlables durante la experimentación (época de lluvias y época de secas), dado que las muestras fueron recolectadas durante marzo, abril, mayo, julio, agosto, septiembre, noviembre y diciembre. En los humedales construidos, las condiciones ambientales como temperatura, precipitación y tiempo de retención hidráulica son factores determinantes que afectan la eficiencia de remoción de sólidos y otros contaminantes (Wu *et al.*, 2023). La alta precipitación típica de la temporada húmeda en Tabasco puede haber incrementado las cargas de sólidos en suspensión y provocado arrastre de partículas hacia el interior del sistema HSSF -Typha, elevando temporalmente la turbidez observada. Además, estudios recientes han documentado que la eficiencia de los humedales construidos depende no solo de la configuración hidráulica, sino también de la interacción entre vegetación, sustrato y condiciones climáticas locales (Chen *et al.*, 2024). Bajo condiciones tropicales, la dinámica de

crecimiento y descomposición de biomasa vegetal puede variar a lo largo del año, especialmente en *Typha*, y provocar fluctuaciones en parámetros físicos como la turbidez cuando se produce desprendimiento de partículas orgánicas finas durante periodos de alta productividad biológica o escorrentía superficial.

La ausencia de una explicación explícita sobre esta turbidez anómala en el manuscrito limita la interpretación de los resultados y puede debilitar la conclusión principal acerca de la relación entre el diseño del humedal y su eficiencia de tratamiento. Otros factores operativos que pudieron influir son las variaciones temporales en la carga afluente, perturbaciones hidráulicas o diferencias en el establecimiento de la vegetación entre tratamientos.

Por lo tanto, es recomendable incorporar en futuros estudios, un análisis que relacione la variabilidad estacional de las condiciones climáticas -precipitación y temperatura- con los cambios en turbidez y otros parámetros de calidad del agua. Este enfoque dará mayor claridad a las causas de la anomalía observada en HSSF -Typha, también fortalecería la interpretación científica del estudio, contribuyendo a una mejor comprensión de los mecanismos de remoción en humedales construidos bajo condiciones tropicales.

4. CONCLUSION.

Los humedales construidos evaluados demostraron alta eficiencia en la remoción de materia orgánica y sólidos bajo condiciones tropicales. Los sistemas de flujo subsuperficial presentaron mejores resultados en DQO, NT y SST, mientras que los sistemas de flujo libre mostraron mayor eficiencia en PT. *Typha dominguensis* fue la especie más eficiente en todos los parámetros evaluados.

La interacción significativa entre tipo de flujo y vegetación confirma la importancia del diseño integral del sistema. Estos resultados respaldan el uso de macrófitas nativas en humedales subsuperficiales como alternativa sostenible y de bajo costo para el tratamiento de aguas residuales en regiones tropicales.

5. REFERENCIAS.

- Brix, H., 1997. *Do macrophytes play a role in constructed treatment wetlands?* Water Science and Technology, 35(5), pp.11–17. DOI: 10.1016/S0273-1223(97)00047-4.
- Chen, G., Mo, Y., Gu, X., Jeppesen, E., Xie, T., Ning, Z., Li, Y., Li, D., Chen, C., Cui, B. & Wu, H., 2024. *Sustainability of global small-scale constructed wetlands for multiple pollutant control.* npj Clean Water, 7(1), Article 45. DOI: 10.1038/s41545-024-00336-3.
- Domínguez-Solís, J.R., López-Maldonado, E., García-Ramírez, J., 2025. *Constructed wetlands as sustainable solutions for decentralized wastewater treatment in tropical regions: A review.* Water, 17(10), 1451. DOI: 10.3390/w17101451.
- Ewemoje, T., Ojo, A., Nwankwo, I., 2023. *Performance evaluation of vertical flow constructed wetlands planted with tropical macrophytes.* Journal of Environmental



- Management, 335, 117509. DOI: 10.1016/j.jenvman.2023.117509.
- Hammer, D.A., 1989. *Constructed wetlands for wastewater treatment*. Chelsea, MI: Lewis Publishers.
- Kadlec, R.H. & Wallace, S.D., 2009. *Treatment wetlands*. 2nd ed. Boca Raton: CRC Press.
- Metcalf, L. & Eddy, H.P., 1995. *Ingeniería de aguas residuales: Tratamiento, vertido y reutilización*. 3rd ed. Madrid: McGraw-Hill.
- Moshiri, G.A. (ed.), 1993. *Constructed wetlands for water quality improvement*. Boca Raton, FL: Lewis Publishers.
- Paula, M., Silva, D., Santos, R., 2024. *Hydraulic and pollutant removal performance in FWS and SSF constructed wetlands under tropical conditions*. Environmental Science and Pollution Research, 31, pp.14521–14535. DOI: 10.1007/s11356-024-21350-0.
- Piñeiro, M., 2011. *Eficiencia en la depuración de un efluente de frigorífico por humedales construidos a escala*. Tesis de grado. Universidad de la República, Uruguay.
- Reddy, K.R. & DeLaune, R.D., 2008. *Biogeochemistry of wetlands: Science and applications*. Boca Raton, FL: CRC Press.
- Reed, S.C., Crites, R.W. & Middlebrooks, E.J., 1995. *Natural systems for waste management and treatment*. 2nd ed. New York: McGraw-Hill.
- Romero, M., Colín, A., Sánchez, E. & Ortiz, M.A., 2009. *Tratamiento de aguas residuales por un sistema piloto de humedales artificiales: evaluación de la remoción de la carga orgánica*. Revista Internacional de Contaminación Ambiental, 25(3), pp.157–167.
- Tchobanoglous, G. & Burton, F.L., 1991. *Wastewater engineering: Treatment, disposal and reuse*. 3rd ed. New York: McGraw-Hill.
- Vymazal, J. & Kröpfelová, L., 2009. *Removal of nitrogen in constructed wetlands with horizontal sub-surface flow: A review*. Wetlands, 29(4), pp.1114–1124.
- Vymazal, J., 2007. *Removal of nutrients in various types of constructed wetlands*. Science of the Total Environment, 380(1–3), pp.48–65.
- Vymazal, J., 2011. *Plants used in constructed wetlands with horizontal subsurface flow: A review*. Hydrobiologia, 674, pp.133–156. DOI: 10.1007/s10750-011-0738-9.
- Wu, H., Wang, R., Yan, P., Wu, S., Chen, Z., Zhao, Y., Cheng, C., Hu, Z., Zhuang, L., Guo, Z., Xie, H. & Zhang, J., 2023. *Constructed wetlands for pollution control*. Nature Reviews Earth & Environment, 4(4), pp.218–234. DOI: 10.1038/s43017-023-00395-z.

Study of the availability of resources in Mexico for the production of synthetic fuels and analysis of states with the greatest potential.

Ivetteh-Virginia Medina-Medina *, Marisol Cervantes-Bobadilla, ** Jarniel García-Morales, **
Hugo Estrada-Esquivel, ** Vitervo López-Caballero, ** Noé-Alejandro Castro-Sánchez**

* Posgrado de Tecnológico Nacional de México/ Centro Nacional de Investigación y Desarrollo Tecnológico, Interior Internado Palmira S/N Palmira 62493 Cuernavaca, Mor. México.
(Tel: 627-108-9703; e-mail: d24cel16@cenidet.tecnm.mx).

**Tecnológico Nacional de México/ Centro Nacional de Investigación y Desarrollo Tecnológico, Interior Internado Palmira S/N Palmira 62493 Cuernavaca, Mor. México.
(Tel: 777-362-7770 e-mail: marisol.cb@cenidet.tecnm.mx, jarniel.gm@cenidet.tecnm.mx, hugo.ee@cenidet.tecnm.mx, vitervo.lc@cenidet.tecnm.mx, noe.cs@cenidet.tecnm.mx)

Abstract: Over the years, Mexico has sought ways to harness its resources to meet emerging needs, as well as to seize opportunities for growth and competitiveness amid daily changes that demand progress, primarily in environmental and technological areas. Mexico's commitment to energy policy is crucial for advancing toward decarbonization; one potential solution is the production of synthetic fuels to meet the needs of specific industries, such as the transportation sector. In this regard, it is essential to identify the most suitable locations within the country that allow for the utilization of available resources, ensuring the optimal use of natural advantages while simultaneously fostering global economic competitiveness. The main objective of this research is to assess the availability of resources (wind, solar, hydro, and CO₂) in the country to determine the potential for synthetic fuel production, identify regions that meet the availability criteria for all the aforementioned resources, and determine the macro-location of the plant. Using Geographic Information Systems (GIS) and Multi-Criteria Decision Analysis, the potential of the state of Tamaulipas regarding the aforementioned factors was assessed, which provides an opportunity to subsequently study the plant's scale.

Keywords: Mexico, green hydrogen, water resources, solar potential, wind potential, CO₂ sources.

1. INTRODUCCIÓN

El aumento de la preocupación a nivel mundial por el cambio climático y la urgencia en disminuir las emisiones de gases de efecto invernadero, han impulsado el desarrollo y la adopción de alternativas sostenibles que sustituyan a los combustibles fósiles. En este contexto, la producción de combustibles sintéticos a partir de fuentes renovables se ha destacado como una opción prometedora. Estos combustibles, también conocidos como e-fuels, se producen utilizando dióxido de carbono (CO₂) capturado de la atmósfera o de fuentes industriales y en combinación con hidrógeno verde que se genera separando la molécula del agua mediante electrólisis, utilizando electricidad renovable que proviene de la energía solar o eólica.

La producción de combustibles sintéticos puede complementar estas fuentes renovables, proporcionando una solución viable y sostenible para enfrentar los desafíos energéticos y ambientales de México. El propósito es que los gobiernos y las instituciones participantes implementen un conjunto amplio de políticas ambiciosas, orientadas a reforzar el compromiso público y asegurar avances significativos en conjunto. (IRENA, 2020). El cambio hacia el uso de energías renovables resulta ineludible, ya que el entorno natural exige la adopción de acciones debido a las consecuencias significativas que se

derivan de no hacerlo (Mirecweek, 2020). Resulta fundamental invertir en la innovación y el desarrollo tecnológico, así como su implementación, para aprovechar mejor los recursos nacionales y avanzar hacia la soberanía energética (SENER, 2020). Es de suma importancia, conocer el potencial con que cuenta para poder ofrecer soluciones prácticas y viables que pueden integrarse en la infraestructura energética existente de México. La producción de combustibles sintéticos a partir de fuentes renovables no solo tiene el potencial de transformar el panorama energético de México, sino también de posicionar al país como un líder en la adopción de tecnologías limpias y sostenibles a nivel mundial.

Cabe hacer mención que, para lograr una producción adecuada de combustibles sintéticos, es necesario contar con la disponibilidad de recursos en energía renovable, así como con dióxido de carbono (CO₂) y recurso hídrico en cantidades suficientes para cubrir la demanda estimada o contribuir a las políticas y estrategias de descarbonización del planeta. En este sentido, la interacción entre estos recursos resulta determinante para la viabilidad del proceso, ya que la electrólisis del agua requiere tanto un suministro continuo de energía como la disponibilidad de agua, mientras que la captura de CO₂ demanda igualmente un aporte energético y la existencia de fuentes concentradas de emisiones. A partir de lo

anterior, la presente investigación se enfoca en analizar dichos recursos, así como en identificar los proyectos que actualmente se desarrollan en México, con el propósito de determinar cuál es la energía renovable más adecuada, cuáles son las fuentes más convenientes para abastecer el agua en el proceso de electrólisis y, finalmente, cuál es el estado de la República Mexicana con mayor disponibilidad de CO₂ para su captura, así como las industrias que lo generan.

Se realizó una revisión de la literatura existente en diferentes bases de datos para llevar a cabo este estudio, entre las cuales se incluyen: Web of Science, Google Académico, Scielo Collections, Elsevier y, los archivos producidos por el Gobierno de México. Se incluyeron artículos en español o inglés, centrados en la incorporación de energías renovables Hidrógeno verde, CO₂ y combustibles sintéticos como criterios de inclusión; la búsqueda se centró en los trabajos recientes y en aquellos que aportaron información útil para el análisis del contexto nacional. Asimismo, los criterios de exclusión se basaron en descartar documentos repetidos, publicaciones sin acceso completo y aquellos que no abordaban de forma directa los temas de interés. La información se localizó empleando cadena de búsqueda con los términos green hydrogen, CO₂ capture, synthetic fuels y renewable energy ajustando las búsquedas según cada base de datos.

2. METODOLOGÍA

Para el desarrollo de este estudio se siguió un enfoque basado en el análisis de los recursos clave necesarios para la producción de combustibles sintéticos en México, considerando de manera integral el potencial energético, la disponibilidad de dióxido de carbono y el acceso a recursos hídricos. En una primera etapa, se evaluó el potencial de energía renovable en el país, particularmente el solar y eólico, tomando en cuenta tanto las condiciones geográficas como la distribución de proyectos existentes. Posteriormente, se analizó la disponibilidad de CO₂, con énfasis en su generación a partir de fuentes industriales y su posible aprovechamiento mediante tecnologías de captura. De igual forma, se examinó el recurso hídrico, considerando distintas fuentes de agua y los posibles impactos asociados a su uso, especialmente en relación con el consumo de agua potable y las alternativas como el uso de agua de mar.

Adicionalmente, se revisaron los proyectos actuales relacionados con la producción de hidrógeno en México, con el fin de identificar el estado de desarrollo de esta tecnología en el país. Finalmente, se incorporó un análisis del panorama internacional en la producción de combustibles sintéticos, dado que en México aún no existen proyectos en esta área, lo que permite establecer un punto de referencia para su posible implementación.

Con base en la información recopilada, se emplearon herramientas de Sistemas de Información Geográfica (GIS, por sus siglas en inglés) y un análisis multicriterio (MCDA, por sus siglas en inglés) para identificar las regiones con mayor potencial para el desarrollo de este proyecto. Los criterios considerados incluyeron el potencial de energía solar y eólica, la disponibilidad de CO₂, la proximidad a infraestructura

energética e industrial y el acceso a recursos hídricos. La ponderación de estos criterios se realizó mediante el proceso de jerarquía analítica (AHP, por sus siglas en inglés), lo que permitió generar un mapa de idoneidad territorial para la identificación de zonas con condiciones favorables.

2.1 Potencial solar en México.

Gran parte del territorio nacional presenta condiciones de alta calidad solar. Entidades como Chihuahua, Sonora, Durango y Baja California destacan por encontrarse entre las regiones con mayor irradiación a nivel mundial. Asimismo, México se posiciona en el tercer lugar entre los cinco países con mayor recepción de radiación solar, debido a su ubicación dentro del denominado Cinturón solar (López, 2022). El país cuenta con potencial para generar energía en un rango que va desde 157 hasta 885,286 GWh, dependiendo de la zona, como lo muestra la Figura 1.

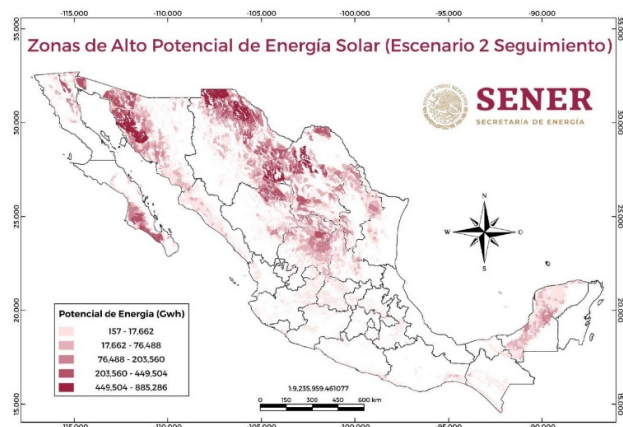


Figura 1 Zonas de alto potencial en energía solar en México (SENER, 2023).

La Asociación Mexicana de Energía Solar (2024), reportó 93 centrales solares en el territorio nacional, distribuidas principalmente en los estados de Sonora, Aguascalientes, Coahuila, Chihuahua y Durango; destacando el estado de Sonora con 14 centrales distribuidas en todo el Estado, cuya capacidad instalada en suma asciende a 1219.58 MW, seguido de Aguascalientes que cuenta con 9 centrales solares, sin embargo, aunque estas sean en cantidad, menores al Estado de Sonora, generan 1201.5 MW en total. Los estados de Coahuila, Chihuahua y Durango cuentan con 5, 10 y 16 centrales solares respectivamente, ubicadas en sus diferentes extensiones territoriales y generando la energía que indica la Figura 2.

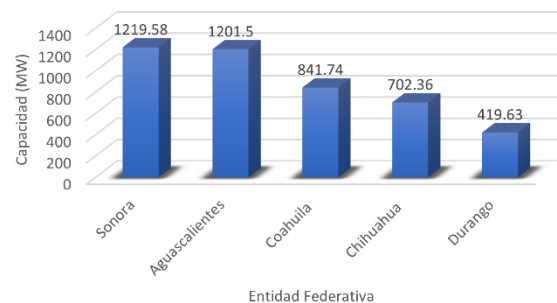


Figura 2 Centrales solares en México (ASOLMEX, 2024).

Los estados de Sonora y Chihuahua presentan un elevando potencial para la generación de energía solar, aunque con marcadas fluctuaciones a lo largo del año. En contraste, existen otras entidades como Puebla, Oaxaca y Morelos que registran una disponibilidad considerable de recurso solar de forma más estable (UNAM, 2022). De acuerdo con los registros del Observatorio de Radiación Solar del Instituto de Geofísica de la UNAM, en gran parte del territorio nacional se presenta un promedio anual de radiación cercano a 5.8 kWh (Geofísica, 2023).

2.2 Potencial eólico en México.

México tiene un alto potencial de energía eólica a una altura de 50 m, 80 m y 120 m; los estados de la República que cuentan con una velocidad de viento idónea para instalar aerogeneradores, son Oaxaca, Chiapas, Tabasco, Coahuila Tamaulipas, Nuevo León y Baja California (SENER, 2023). En la Figura 3 se aprecian las zonas con alta calidad eólica de la República Mexicana a una altura de 120 m y cuya densidad puede llegar hasta los 1200 W/m².



Figura 3 Potencial de energía eólica a 120 m de altura (SENER 2023).

La base de datos en The Wind Power (2024), reportó un total de 101 proyectos eólicos en tierra firme (Onshore) para el país, de los cuales 87 están en etapa de operación, 13 están en la etapa de construcción, y 1 se encuentra deshabilitado. Los proyectos eólicos en México están liderados por el estado de Oaxaca con 33 parques, seguido de Tamaulipas con 18 proyectos y Baja California con 10.

En México, la energía eólica aporta cerca del 6% de la energía que requiere el país. A nivel nacional, el estado de Oaxaca concentra la mayor capacidad instalada, seguido de Tamaulipas y Nuevo León, como se muestra en la Figura 4.

Esta distribución refleja la relevancia de estas regiones en el desarrollo de proyectos eólicos. En particular, Oaxaca destaca no solo por su capacidad instalada, sino también por la concentración de parques y turbinas en operación; tal es el caso de la zona de Juchitán de Zaragoza, donde se han instalado hasta 142 turbinas con una capacidad aproximada de 213 MW.

A pesar de las diversas opciones de generación de energía, en el comportamiento del PIIRCE (Programa Indicativo para la Instalación y Retiro de Centrales Eléctricas) para la tecnología renovable pronosticado hasta el año 2032 prevalece la utilización de energía eólica y fotovoltaica (Fragoso, 2020).

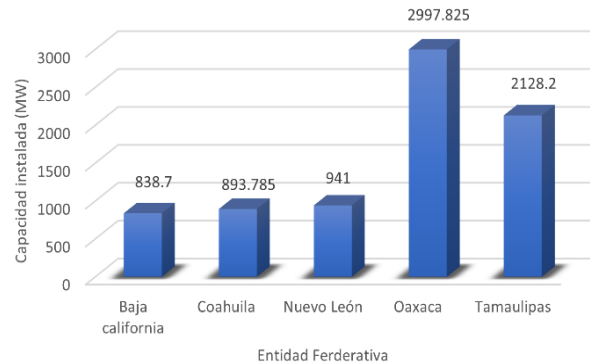


Figura 4. Capacidad instalada de energía eólica por entidad federativa en México (MW). (WIND POWER, 2025).

2.3 Disponibilidad de CO₂

En México, la mayor parte de las emisiones de gases de efecto invernadero derivadas de la quema de combustibles fósiles, proviene del petróleo, seguido por el gas natural y el carbón, con participaciones aproximadas de 54%, 36.6% y 8.7% respectivamente (IEA, 2024). Estas mismas fuentes de energía constituyen el 89% del consumo energético total (Sierra, 2024). El dióxido de carbono, generado en gran medida por la combustión de combustibles para la producción de energía, constituye aproximadamente el 63.9% de las emisiones de gases de efecto invernadero (GEI) en México, siendo el dióxido de carbono la principal fuente de GEI a nivel mundial y en México, generado por la producción de energía y calor, actividades de transporte terrestre, aéreo y marítimo, entre otros (INECC, 2021).

El sector industrial es un agente que genera grandes cantidades de CO₂ y algunas de sus categorías con mayor contribución de emisiones fueron: Uso de productos sustitutos de las sustancias que agotan la capa de ozono con un 30% de aportación de las emisiones de este sector, seguido de la Industria de los minerales (cemento) con un 28% contribución y en tercer lugar las emisiones la Industria de los metales (hierro y acero) con un 20% de las emisiones (INECC, 2024). La Figura 5 muestra los puntos de emisiones de CO₂ generados por una gran variedad de industrias que influyen en la generación de dichos gases.

La producción de combustibles sintéticos requiere el uso de CO₂ capturado, con el objetivo de reducir al máximo la huella de carbono del proceso. En este contexto, las principales rutas de captura incluyen tecnologías de precombustión, oxicomustión y postcombustión. Dentro de esta última, destacan la captura por absorción química y la Captura Directa de Aire (DAC, por sus siglas en inglés), siendo la DAC de particular interés debido a su capacidad para obtener CO₂ directamente de la atmósfera, a pesar de implicar mayores costos.

Sin embargo, debido a estos costos, resulta más conveniente considerar instalar esta tecnología cercana a fuentes puntuales de emisión de CO₂, como cementeras, refinерías y complejos petroquímicos, con la finalidad de reducir los requerimientos energéticos y económicos asociados a su captura. Bajo este enfoque, se requiere analizar la disponibilidad de este tipo de fuentes en el país.



Figura 5. Emisiones de CO₂ en México. Secretaría del Medio Ambiente (2018).

En México, existen 8 complejos petroquímicos ubicados en Chihuahua, Tamaulipas, Puebla y cinco en el estado de Veracruz; además, se cuenta con 6 refinерías distribuidas en Nuevo León, Tamaulipas, Guanajuato, Hidalgo, Veracruz y Oaxaca (Pemex, 2025), así como 36 cementeras localizadas en distintos puntos del territorio nacional (CANACEM, 2025). En consecuencia, la localización de proyectos de combustibles sintéticos en México no debe basarse únicamente en la disponibilidad de fuentes emisoras de CO₂, sino en el equilibrio entre la disponibilidad de energía renovable, el acceso a recursos hídricos y la presencia de dichas fuentes, ya que la interacción de estos factores determina la viabilidad técnica y operativa del proceso.

2.4 Recurso hídrico

El agua es uno de los factores más importantes en la producción de hidrógeno verde y, por ende, en la producción de combustibles sintéticos. Este elemento, a través del proceso de electrólisis, el agua se descompone en sus componentes, liberando oxígeno y permitiendo la obtención de hidrógeno, el cual es utilizado como materia prima en los procesos de producción.

La mejor fuente de agua para producir hidrógeno verde por electrólisis depende de criterios técnicos como pureza, disponibilidad, costo de tratamiento y sostenibilidad.

El análisis realizado considera cuatro fuentes principales: agua dulce, agua de mar, agua salobre y aguas residuales tratadas. En la Tabla 1 se muestra que el agua desionizada (dulce) y el agua de mar son las fuentes con mayor viabilidad actual; por su parte, el agua desionizada es la fuente estándar para los electrolizadores PEM y alcalinos, por lo que requiere menor tratamiento, sin embargo, la disponibilidad es limitada y puede generar conflictos por el uso de agua potable en regiones con escasez hídrica. Por otro lado, el agua de mar es la opción más idónea para proyectos industriales, aunque requiere de infraestructura adicional para realizar la desalación del agua.

Tabla 1. Fuentes de agua y su viabilidad para la producción de hidrógeno verde.

Fuente de agua	Viabilidad actual	Tratamiento requerido	Uso recomendado	Referencia
Agua desionizada	Muy alta	Bajo	Plantas actuales	Mika et al. (2024)
Agua de mar + desalación	Muy alta	Medio	Proyectos grandes costeros	Augustin, A.& Wintgens, T. (2025)
Agua salobre	Alta	Bajo-medio	Regiones interiores	Mika et al. (2024)
Aguas residuales tratadas	Media	Alto	Economía circular	Mika et al. (2024)
Electrólisis directa de agua de mar	Baja	Experimental	Investigación	Zhang et al. (2024)

2.5 Proyectos de producción de hidrógeno en México.

La Asociación Mexicana de Hidrógeno (2024), ha contabilizado que México cuenta con 18 proyectos en los que se utiliza el hidrógeno, de los cuales, únicamente uno (5%) se encuentra en la fase de construcción y el resto (95%) en la etapa de desarrollo.

México, al ser una región óptima para la producción de energías renovables tendría unos costes inferiores en torno al 64 % a otros países. De aquí a 2030, México podría producir hidrógeno a un costo de 1,4 dólares/kg frente a 2,3 dólares/kg respecto a otros países que no tienen el mismo potencial energético (Asociación Mexicana de Hidrógeno, 2024).

La Tabla 2 muestra los datos específicos de producción, inversión y ubicación de los principales proyectos desarrollados en México, en los que se ha considerado producir amoniaco, amoniaco verde, metanol verde, hidrógeno verde, principalmente.



Tabla 2. Proyectos que involucran el hidrógeno verde en México.

Nombre del proyecto	Empresa	Lugar	Fase del proyecto	Producción /capacidad
NH ₃ Baja	Hidrogeno NH ₃ La Rumorosa	Baja California	En desarrollo	Planta de producción de hidrógeno verde mediante electrólisis (3MW), para exportación a California, con capacidad para producir 375.6 toneladas métricas anuales.
Power to Power	Mass Storage Energy, Energía Los Cabos	Los Cabos	En desarrollo	Aportará entre 15% y 20% de la generación del sistema eléctrico de Los Cabos. Permitirá la entrada de \$2,500 millones de dólares.
ASLAN Energy	ASLAN Energy	Hermosillo	En desarrollo	La Fase 1 del proyecto producirá casi 600 mil toneladas de amoníaco verde.
Hidrógeno Verde	DH ₂ Energy	El Fuerte	En desarrollo	Producción solar de 1,708 MW que permitirá la generación de 41 mil 485 toneladas de hidrógeno verde al año mediante electrólisis.
Pacífico Mexinol	Transition Industries, International Finance Corporation	Sinaloa	En construcción	Producción de 352,000 toneladas métricas de metanol verde y 1.8 millones de toneladas métricas de metanol azul anualmente, a partir de captura de carbono y gas natural. Se estima una inversión superior a los 3,300 millones de dólares.
Green Ammonia	Tarafert	Durango	En desarrollo	Producción de hasta 200,000 toneladas métricas de amoníaco verde al año.
Ikal Solar	Ikal Solar, S. A. de C. V.	Nuevo León	En desarrollo	El parque contará con 4,408,740 paneles fotovoltaicos, generando 2,865 MW y 2,448 MW. Producirá 124,071 toneladas de hidrógeno verde anualmente. Incluirá un hidroducto de 10.9 km y una línea de conexión de 10.6 km. El proyecto está alineado con la transición energética de México y el objetivo de blending de hidrógeno con gas natural.
Utilización de Hidrógeno Verde en plantas cementeras	CEMEX	Monterrey	En desarrollo	Promoverán la inyección de hidrógeno en plantas de cemento.
Tratamiento de aguas residuales y producción de Hidrógeno Verde	Gobierno de Reynosa	Reynosa	En desarrollo	-----
Neptuno Solar 1	Neptuno Solar	San Luis Potosí	En construcción	A producir 18,263 toneladas H ₂ verde/año 484,680 módulos fotovoltaicos capacidad de producción de 125 MW
Blending	Dhamma Energy	Guanajuato	En desarrollo	-----
Producción de Hidrógeno Verde	Siemens Energy, Gob Michoacán	Michoacán	En desarrollo	Producción para dar servicio al sector industrial, de transporte pública, automotriz, etc.
Inyección de hidrógeno en hornos de plantas de cemento	Holcim	CDMX	En desarrollo	Esperan que en 7 años el 100% de su consumo energético sea renovable.
Producción de Hidrógeno Verde	Copenhagen Infrastructure Partners, Gob de Oaxaca	Oaxaca	En desarrollo	-----
Green Ammonia	SEMABICCE	Campeche	En desarrollo	Se estima producción de 170,000 Toneladas de amoníaco verde.

El 50% de los proyectos se concentra en el Norte del país y el otro 50% se ubica en el Centro y Sur de la República Mexicana, aunque en zonas del Centro existen proyectos destinados a mejorar los procesos de empresas establecidas.

3. RESULTADOS

La implementación de una planta de combustibles sintéticos implica considerar inversión en infraestructura y por ende, un impacto en los costos, cuyos componentes se absorben mayormente por la energía renovable (60% del costo total), seguido del costo del hidrógeno (29%), el del CO₂ (7%) y, por último, el del proceso Fischer Tropsch (4%) (D'Adamo, et al.,

2024); por su parte, Cardona (2024), afirma que los costos de producción de energía es la variable principal que absorbe más del 70% del costo de cada uno de los combustibles sintéticos analizados. Es por ello que, la elección de la energía renovable desde una perspectiva económica es de suma importancia para determinar el costo de la inversión de un proyecto y, por ende, su ubicación.

El desempeño energético de los paneles fotovoltaicos se ve influido por diversas condiciones climáticas, como la velocidad y dirección del viento, la temperatura del aire, la radiación solar, la humedad relativa, las precipitaciones, la cobertura nubosa y la presencia de vapor de agua en la



atmósfera, el cual puede absorber, dispersar y reflejar la radiación solar en distintos niveles (Li, et al., 2022); cada grado celsius de incremento en la temperatura reduce la eficiencia de conversión de los paneles entre un 0.5% y un 0.7% (Xu et al., 2024), lo que afecta significativamente el rendimiento y la vida útil de los sistemas.

Aunque las tendencias de generación de energía renovable se inclinan hacia los paneles fotovoltaicos también es cierto que se debe perfeccionar su implementación, Es por lo anterior que, la energía eólica se considera con mayor estabilidad y, por ende, la mejor opción para invertir en proyectos de producción de combustibles sintéticos a gran escala.

En el análisis del potencial de energía eólica se observa que Oaxaca, Tamaulipas y Nuevo León son los Estados de la República con mayor velocidad del viento y con mayor capacidad instalada, por lo que estos lugares serán objeto de estudio base para analizar el resto de los factores. La Tabla 3 muestra los datos cuantitativos y cualitativos de los factores que influyen en la producción de combustibles sintéticos en dichas zonas.

Tabla 3. Análisis de factores de los Estados de la República con mayor capacidad instalada de energía eólica.

	Oaxaca	Tamaulipas	Nuevo León	Referencia
Factor	Valor	Valor	Valor	
Velocidad del viento promedio (m/s)	7.5 - 8.5	6.5 - 7.5	6.0 - 6.5	CRE (2023); IEA (2024)
Costo de mano de obra especializada	16,000-22,000	18,000-26,000	20,000-30,000	INEGI (2024); World Bank (2023)
Estrés hídrico	Bajo-medio	Medio-alto	Muy alto	CONAGUA (2023); SEMARNAT (2023)
Costo de m ³ de agua (MXN)	14-20	12-17	10-20	CONAGUA (2023)
Zonas protegidas y restricciones ambientales	Alto	Medio	Medio-alto	CONANP (2023); SEMARNAT (2023)
Disponibilidad de CO ₂ industrial	5-8 millones	12 millones	38 millones	SEMARNAT (2023); INECC (2024)
Refinerías/ complejo petroquímico	1	1	1	PEMEX (2023)
Cementeras	1	0	1	CANACEM (2025); INEGI (2024)
Costo de transporte (combustible)	400 km 540-1190 MXN/ton	200 km 170-425 MXN/ton	300 km 340-765 MXN/ton	SICT (2022); World Bank (2023)

Los valores asignados en la matriz de comparación por pares se determinaron con base en la escala fundamental propuesta por Saaty (1–9), la cual permite representar la importancia relativa entre criterios dentro del Proceso de Jerarquía Analítica (AHP). En esta escala, un valor de 1 indica igual importancia entre dos criterios, mientras que valores de 3, 5, 7 y 9 representan niveles crecientes de preferencia (moderada, fuerte, muy fuerte y extrema, respectivamente). Los valores recíprocos (1/3, 1/5) se utilizan para expresar la relación inversa como se muestra en la Tabla 4.

Tabla 4. Matriz de comparación por pares (AHP).

Criterio	Viento	Mano de obra	Estrés hídrico	Costo de agua	Ambiental	CO ₂	Infraestructura	Cementeras	Transporte
Viento	1	3	3	5	3	5	5	5	3
Mano de obra	1/3	1	1	3	1	3	3	3	1
Estrés hídrico	1/3	1	1	3	1	3	3	3	1
Costo de agua	1/5	1/3	1/3	1	1/3	3	3	3	1
Ambiental	1/3	1	1	3	1	3	3	3	1
CO ₂	1/5	1/3	1/3	1/3	1/3	1	3	3	1
Infraestructura	1/5	1/3	1/3	1/3	1/3	1/3	1	1	1/3
Cementeras	1/5	1/3	1/3	1/3	1/3	1/3	1	1	1/3
Transporte	1/3	1	1	1	1	1	3	3	1

El vector de prioridades normalizado resultante permitió asignar pesos de 0.25 a la energía eólica, 0.15 al estrés hídrico, 0.15 a las restricciones ambientales, 0.10 a la disponibilidad de CO₂, 0.10 a la logística, 0.10 al costo de la mano de obra y 0.05 a los factores relacionados con la infraestructura, los cuales servirán para realizar las evaluaciones de las zonas en la Tabla 5.

Tabla 5. Matriz de ponderaciones para evaluar criterios en las regiones.

	Oaxaca			Tamaulipas		Nuevo León	
Factor	Peso	Calif.	Total	Calif.	Total	Calif.	Total
Velocidad del viento promedio (m/s)	0.25	10	2.5	9	2.25	8	2.0
Costo de mano de obra especializada	0.10	10	1	9	0.9	8	0.8
Estrés hídrico	0.15	10	1.5	9	1.35	8	1.2
Costo de m ³ de agua (MXN)	0.05	8	0.4	10	0.5	9	0.45
Zonas protegidas y restricciones ambientales	0.15	8	1.2	10	1.5	9	1.35
Disponibilidad de CO ₂ industrial	0.10	8	0.8	9	0.9	10	1.0
Refinerías/ complejo petroquímico	0.05	10	0.5	10	0.5	10	0.5
Cementeras	0.05	10	0.5	9	0.45	10	0.5
Costo de transporte (combustible)	0.10	8	0.8	10	1.0	9	0.9
Total ponderado	100 %		9.2		9.35		8.7

Los pesos utilizados en la evaluación de alternativas fueron obtenidos mediante el método AHP, lo que permitió cuantificar la importancia relativa de cada criterio de forma estructurada y consistente. La evaluación de las alternativas se realizó mediante una escala ordinal discreta de tres niveles (10, 9 y 8), con el propósito de representar el desempeño relativo de cada región (Oaxaca, Tamaulipas y Nuevo León) en función de los criterios considerados. En este esquema, la calificación de 10 se asigna a la alternativa con el mejor desempeño en cada criterio, 9 a la alternativa con desempeño intermedio y 8 a la de menor desempeño relativo.



Asimismo, las calificaciones fueron asignadas con base en los datos cuantitativos y cualitativos previamente analizados, tales como la velocidad del viento, el costo y disponibilidad de agua, las emisiones de CO₂, los costos logísticos y las restricciones ambientales, garantizando coherencia entre los datos de entrada y la evaluación multicriterio. La puntuación final se obtuvo mediante la suma del producto entre el peso de cada criterio y la calificación asignada a cada alternativa, permitiendo determinar un índice de idoneidad territorial, identificando a Tamaulipas como la alternativa con mayor idoneidad. Este enfoque híbrido combina la robustez del AHP en la determinación de pesos con la simplicidad y claridad del método de suma ponderada para la evaluación de alternativas.

La consistencia de la matriz de juicios se verificó mediante el cálculo del Índice de Consistencia (CR), obteniéndose un valor de 0.063, el cual se encuentra por debajo del umbral aceptable de 0.1, lo que indica que la matriz presenta un nivel adecuado de consistencia.

$$CI = \frac{\lambda_{\max} - n}{n - 1} \quad \text{ec. (1)} \quad CR = \frac{CI}{RI} \quad \text{ec. (2)}$$

Sustitución:

$$CI = \frac{9.74 - 9}{8} = 0.092 \quad CR = \frac{0.092}{1.45} = 0.063$$

Donde:

$\lambda_{\max}$: valor propio máximo de la matriz de comparación por pares. Se obtiene a partir del producto matricial entre la matriz de comparación y el vector de pesos normalizado.

n: número de criterios considerados en el análisis.

CI (Consistency Index): índice de consistencia que mide la desviación de la matriz respecto a una condición perfectamente consistente.

RI (Random Index): índice aleatorio obtenido de valores tabulados por Saaty, que depende únicamente del número de criterios evaluados. Para $n = 9$, el valor correspondiente es $RI = 1.45$.

CR (Consistency Ratio): razón de consistencia que permite evaluar la validez de los juicios de comparación.

4. CONCLUSIONES

México cuenta con recursos disponibles para la producción de combustibles sintéticos; en lo que respecta a energía renovable, la eólica es actualmente la que demuestra mayor estabilidad y eficiencia, sin embargo, los proyectos híbridos pueden aprovechar los recursos solar y eólico del país y ofrecer ventajas de costo-beneficio.

El agua es un recurso vital para la sociedad, por lo que hay que seleccionar la fuente que provee a un proyecto de gran magnitud y considerar, no solo la disponibilidad del recurso sino sus costos y estrés hídrico en las zonas, es por ello que actualmente, el agua de mar es la fuente más idónea y rentable económica y ambientalmente para proyectos a gran escala.

La facilidad que ofrece la tecnología DAC (Captura Directa de Aire, por sus siglas en inglés) al capturar el CO₂ desde cualquier sitio, representa una gran ventaja para los proyectos de producción de combustibles sintéticos, sin embargo, la

cercanía con alguna fuente directa como cementeras, refinerías o incluso complejos petroquímicos representa una alternativa para reducir costos y tomar decisiones estratégicas para eficientar los procesos productivos.

Los resultados evidencian que la selección de la ubicación debe basarse en un enfoque integral que considere simultáneamente todos los recursos ya mencionados, no obstante, la viabilidad de estos proyectos depende del equilibrio entre estos factores. México cuenta con condiciones favorables para la producción de combustibles sintéticos; la integración de factores como agua, CO₂, viento, etc. mediante el análisis multicriterio muestra que Tamaulipas obtiene el mayor valor ponderado, seguido de Oaxaca y Nuevo León, lo que lo posiciona como una de las regiones con mayor potencial, debido al equilibrio entre sus variables estudiadas.

5. AGRADECIMIENTOS

Se expresa un sincero agradecimiento a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el apoyo financiero brindado a través del programa de becas para estudios de posgrados. Asimismo, se agradece al Tecnológico Nacional de México y al Centro Nacional de Investigación y Desarrollo Tecnológico por la calidad en sus programas de posgrado y por la destacada labor de los docentes e investigadores quienes guían y acompañan de manera ejemplar a los estudiantes.

REFERENCIAS

- Agustin, A. and Wintgens, T. (2025). *Seawater as feedstock for large-scale green hydrogen production: A technical review from a desalination perspective*. Desalination. <https://doi.org/10.1016/j.desal.2025.108151>
- Asociación Mexicana de Hidrógeno (2024). *Mapa de proyectos de hidrógeno en México*. Ciudad de México: AMH. Available at: <https://h2mex.org/mapa-de-proyectos/>
- Asociación Mexicana de Energía Solar (ASOLMEX) (2024). *Centrales solares en México*. Ciudad de México: ASOLMEX. Available at: <https://asolmex.org/centrales-solares/>
- AZEL, SENER (2012). *Atlas mexicano de almacenamiento geológico de CO₂*. Ciudad de México: Secretaría de Energía. Available at: https://meccs.org.mx/static/atlas/atlas_almac_geo_mex.pdf
- Cámara Nacional de Cemento (CANACEM) (2025). *Plantas productoras de cemento y sus ubicaciones en México*. Ciudad de México: CANACEM. Available at: <https://canacem.org.mx/ubicaciones/>
- Cámara Nacional de la Industria del Hierro y del Acero (CANACERO) (2025). *Radiografía de la industria del acero en México*. Ciudad de México: CANACERO. Available at: <https://www.canacero.org.mx/>
- Cardona, G. (2024). *Evaluación técnico-económica de un sistema de producción de combustibles basado en*



- electricidad*. Tesis de maestría. Universidad Nacional de Colombia.
- Comisión Nacional del Agua (CONAGUA) (2023). *Estadísticas del agua en México 2023*. Ciudad de México: Secretaría de Medio Ambiente y Recursos Naturales.
- Comisión Nacional de Áreas Naturales Protegidas (CONANP) (2023). *Sistema nacional de áreas naturales protegidas de México*. Ciudad de México: CONANP.
- Comisión Reguladora de Energía (CRE) (2023). *Atlas nacional de energía eólica*. Ciudad de México: Gobierno de México.
- D'Adamo, I., Gastaldi, M., Giannini, M. and Nizami, A.S. (2024). *Environmental implications and levelized cost analysis of e-fuel production under photovoltaic energy, direct air capture, and hydrogen*. Environmental Research, 246, 118163.
- Enlight (2024). *México en el cinturón solar de la Tierra: potencial de energía solar*. Available at: <https://www.enlight.mx/>
- Fragoso-Altamirano, A. (2020). *México y su transición energética: un cambio en pro de la energía renovable*. LADEE, 1(1), 26–42. <https://doi.org/10.17981/ladee.01.01.2020.3>
- Geofísica Internacional (2023). *Mapa de radiación solar en México*. Universidad Nacional Autónoma de México. Available at: <https://rayaneri.geofisica.unam.mx/~mauro/mapas/anual.pdf>
- International Energy Agency (IEA) (2024). *Hydrogen production and infrastructure projects database*. Paris: IEA. Available at: <https://www.iea.org/data-and-statistics>
- Instituto Nacional de Ecología y Cambio Climático (INECC) (2021). *Inventario nacional de emisiones de gases y compuestos de efecto invernadero*. Ciudad de México: INECC.
- Instituto Nacional de Ecología y Cambio Climático (INECC) (2024). *Análisis de los sectores productivos con mayores emisiones de gases de efecto invernadero en México*. Ciudad de México: INECC.
- IRENA (2020). Conclusiones principales. Perspectivas Mundiales de las Energías Renovables. Transformación Energética de Aquí a 2050. IRENA, AUH, EAU.
- Li, P., Gao, X., Li, Z. and Zhou, X. (2022). *Effect of temperature difference between land and lake on photovoltaic power generation*. Renewable Energy, 86, 95.
- López, P. (2022). *México entre los cinco países con mayor potencial de energía solar*. Universidad Nacional Autónoma de México.
- Mika, L., Szeląg, K., Bujok, T., Boruta, P. and Radomska, E. (2024). *Seawater treatment technologies for hydrogen production by electrolysis: A review*. Energies, 17(24), 6255. <https://doi.org/10.3390/en17246255>
- MirecWeek (2020). *Energías renovables en México: panorama 2020*. Available at: <https://www.mirecweek.com/es/congresos/zona-de-negocios>
- Petróleos Mexicanos (PEMEX) (2025). *Refinerías y complejos petroquímicos en México*. Ciudad de México: PEMEX. Available at: <https://www.pemex.com/negocio/infraestructura>
- Secretaría del Medio Ambiente de la Ciudad de México (SEDEMA) (2018). *Emisión de gases contaminantes a la atmósfera*. Ciudad de México: SEDEMA. Available at: <https://sedema.cdmx.gob.mx/>
- Secretaría de Infraestructura, Comunicaciones y Transportes (SICT) (2022). *Costos de operación del transporte de carga en México*. Ciudad de México: Gobierno de México.
- Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT) (2023). *Inventario nacional de emisiones de gases y compuestos de efecto invernadero*. Ciudad de México: SEMARNAT.
- Secretaría de Energía (SENER) (2023). *Atlas nacional de zonas de alto potencial de energías limpias*. Ciudad de México: Gobierno de México. Available at: <https://www.gob.mx>
- Secretaría de Energía (SENER) (2020). *Programa sectorial derivado del Plan Nacional de Desarrollo 2019–2024*. Ciudad de México: Gobierno de México. Available at: <https://www.dof.gob.mx>
- Sierra, L. (2024). *Transición energética en México: ¿la oportunidad perdida?*. Escuela de Gobierno y Transformación Pública, Tecnológico de Monterrey. Available at: <https://egobiernoytp.tec.mx/>
- The Wind Power (2024). *Bases de datos de parques eólicos en México*. Wind Energy Market Intelligence. Available at: <https://www.thewindpower.net/>
- World Bank (2023). *Logistics Performance Index 2023*. Washington, DC: World Bank.
- Xu, L., Ding, P., Zhang, Y., Huang, Y., Li, J. and Ma, R. (2024). *Sensitivity analysis of the shading effects from obstructions at different positions on solar photovoltaic panels*. Energy, 290. <https://doi.org/10.1016/j.energy.2023.130229>
- Zhang, Y., Zhang, Y., Li, Z., Wu, E., Ye, H., Li, Z., Guo, X., Zhou, D., Wang, C., Sha, Q. and Kuang, Y. (2024). *A review of hydrogen production via seawater electrolysis: Current status and challenges*. Catalysts, 14(10), 691. <https://doi.org/10.3390/catal14100691>

Electromyographic Assessment of Quadriceps Femoris Muscles Under Different Exertion Conditions.

Pérez Juárez. Erik*, Chávez Báez. Marco Vinicio*, Gutiérrez Gnechi. José Antonio*, Reyes Archundia. Enrique *, Rodríguez Herrejón. Javier Alejandro *, Méndez Patiño. Arturo*

**División de Estudios de Posgrado e Investigación, Tecnológico Nacional de México / Instituto Tecnológico de Morelia. MX (Tel: 443-312-1570).*

e-mail: D24121504@morelia.tecnm.mx, marco.cb@morelia.tecnm.mx, jose.gg3@morelia.tecnm.mx, enrique.ra@morelia.tecnm.mx, javier.rh@morelia.tecnm.mx, arturo.mp@morelia.tecnm.mx.

Abstract: The classification of electromyographic signals represents a challenge in medical and robotic control applications due to the variability of muscle signals under different effort conditions. The objective of the research was to classify the electromyographic signals of the vastus medialis, rectus femoris, and vastus lateralis muscles under resting, unloaded movement, and loaded movement conditions. In the study, 810 electromyographic signals were acquired from 9 subjects. The signals were filtered, analyzed using the discrete wavelet transform, and evaluated through electromyographic parameters such as Zero Crossing (ZC), Root Mean Square (RMS), and Willison Amplitude (WA), among others. Subsequently, the signals were classified using linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) methods. The results showed that the QDA method achieved better classification percentages, reaching values close to 80% and AUC values higher than 0.84 under different effort conditions. Individual subject classification achieved better results than group classification by distinguishing between different loading conditions. The obtained results indicate that the use of wavelets and machine learning methods may contribute to the development of rehabilitation systems, prosthetic control, and robotic assistance systems.

Keywords: Electromyography, Machine learning, Vastus medialis, Rectus femoris, Vastus lateralis, Linear discriminant analysis, Quadratic discriminant analysis, Wavelets.

1. INTRODUCCIÓN

Las señales electromiográficas han sido utilizadas en diversos campos de la ingeniería con diferentes objetivos, tanto médicos para determinar el estado de salud del cuerpo humano como aplicaciones de control, como el desarrollo de prótesis que permiten a personas discapacitadas recuperar el movimiento o en personas de edad avanzada para amplificar su fuerza a través de dispositivos robóticos, las señales electromiográficas son señales biológicas que proporcionan información del movimiento de los músculos para después procesarlas y aplicarlas a un campo de la ingeniería.

El análisis electromiográfico consiste en adquirir las señales generadas por el potencial eléctrico de los músculos a través de dispositivos electrónicos que filtren y amplifiquen estas señales debido a que son señales de muy baja amplitud, que incorporan ruidos no deseados. Durante la contracción muscular, la amplitud de la señal electromiográfica tiende a incrementarse con el reclutamiento de fibras musculares y la generación de fuerza (Alcan and Zinnuroğlu, 2023).

La clasificación de las señales electromiográficas se hace por medio de métodos de aprendizaje automático que permiten identificar y clasificar las señales, antes de ser incorporadas a un algoritmo de clasificación, las señales electromiográficas tienen que ser procesadas y convertidas a datos más simples para que un clasificador pueda analizarlas. Los métodos de procesamiento de señales necesitan de valores que representen

la señal electromiográfica, debido a la gran cantidad de datos de una señal electromiográfica sería difícil poder introducir todos los valores al clasificador, debido a que el tiempo de procesamiento aumentaría, y en aplicaciones de control de dispositivos robóticos se requieren tiempos de procesamiento pequeños, por lo tanto, lo que se busca es obtener valores que resuman el comportamiento de la señal electromiográfica para poder ser introducidos a métodos de aprendizaje automático y tener una respuesta más rápida de la intención del movimiento.

En el análisis de las señales electromiográficas se han utilizado métodos basados en transformada wavelet y análisis discriminante para el procesamiento y clasificación de señales (Di Nardo *et al.*, 2022; Shariatzadeh *et al.*, 2023; Jochumsen *et al.*, 2020; Kok *et al.*, 2024; Amezquita-García *et al.*, 2022; Hurtado *et al.*, 2024). Sin embargo, existe una cantidad limitada de investigación enfocada en la clasificación de señales electromiográficas de músculos de la pierna bajo diferentes condiciones de carga utilizando conjuntamente wavelets y métodos discriminantes como LDA y QDA. Por ello, este estudio propone el análisis y clasificación de señales electromiográficas de los músculos vasto medial, recto femoral y vasto lateral para evaluar diferentes condiciones de esfuerzo muscular comparando los diferentes métodos con el empleo de los parámetros utilizados.

2. ESTADO DEL ARTE

Existen diversos métodos para el análisis de las señales electromiográficas, diversos estudios han empleado métodos para obtener información de las señales electromiográficas, entre los métodos que se han utilizado se encuentran la Raíz Cuadrada Media (RMS) (Barański *et al.*, 2024; Bawa and Banitsas, 2022), el Cruce por Cero (ZC) (Adzkie, Setiawan and Arland, 2023), la energía (E) (Saranya, Poonguzhali and Karunakaran, 2020), el Valor Absoluto de la Suma de la Raíz Cuadrada (ASS) (Li *et al.*, 2025), el Valor Absoluto de la Suma de la Raíz Exponente (ASM) (Lin *et al.*, 2025), la Cardinalidad (CARD) (Sarwat *et al.*, 2024), la Varianza de la Diferencia (DVARV) (Li *et al.*, 2023), el Rango Intercuartil (IQR) (Abdel-Maboud *et al.*, 2022; Verma and Gupta, 2020), la Desviación Absoluta Media (MAD) (Wadud and Showrov, 2020), la Desviación Estándar (SD) (Mohapatra, Sahoo and Sahoo, 2024), la Amplitud de Willison (WA) (Malesevic *et al.*, 2022; Keleş, Türksoy and Yucesoy, 2023), la Longitud de la Forma de Onda (WL) (Turgunov *et al.*, 2020; Onay and Mert, 2020; Tepe and Demir, 2020) estos métodos permiten extraer características de las señales electromiográficas y determinar el nivel de esfuerzo de los músculos.

Diversos estudios han empleado estas características con la intención de poder analizar las señales electromiográficas, (Yang *et al.*, 2025; Almaghout *et al.*, 2020) controlaron dispositivos robóticos por medio de las señales electromiográficas, mientras que (Kinugasa and Kubo, 2023) utilizaron la señal para determinar la fatiga muscular, así mismo se han empleado para la clasificación de señales y detectar el movimiento realizado (Too, Abdullah and Saad, 2019).

Se han empleado el uso de wavelets para la clasificación de señales electromiográficas (Di Nardo *et al.*, 2022; Shariatzadeh *et al.*, 2023). Las señales son analizadas mediante métodos de procesamiento automático, el método análisis discriminante lineal ha sido utilizado para la clasificación de las señales (Jochumsen *et al.*, 2020; Kok *et al.*, 2024), otro método empleado en la clasificación de señales ha sido el método de análisis discriminante cuadrático (Amezquita-García *et al.*, 2022; Hurtado *et al.*, 2024).

Utilizar los parámetros electromiográficos a través de los métodos de procesamiento de señales permite saber que tanto se separan las distintas clases y que tan clasificable puede ser una señal. Los métodos de procesamiento de señales permiten clasificar las distintas señales y poder detectar la intención de movimiento durante la contracción del músculo.

3. METODOLOGÍA

3.1 Procedimiento para la adquisición de la señal

La adquisición de la señal consistió en la medición de señales electromiográficas en 9 pacientes sanos con edades de 18 a 25 años de los músculos vasto medial, recto femoral y vasto lateral, en total se tomaron 810 señales. Antes de la colocación de los electrodos sobre los músculos se realizó una limpieza en la zona donde se colocaron, utilizando alcohol isopropílico y algodón debido a que la impedancia generada por la piel puede causar lecturas erróneas. Se colocaron los electrodos en tres

músculos el vasto medial, el recto femoral y el vasto lateral, se dejó entre cada uno de los electrodos una distancia de 2cm entre ellos. Los electrodos utilizados fueron electrodos circulares con un diámetro de 20 mm de plata/cloruro de plata debido a que diversos estudios han demostrado que estos electrodos son más útiles en la adquisición de la señal electromiográfica (Sudharsan *et al.*, 2020; Guo *et al.*, 2020). Los electrodos fueron colocados en el centro de cada músculo, se les pidió a los pacientes contraer los músculos para poder identificar cada músculo y poder colocar los electrodos en el mejor lugar respecto al tamaño del músculo de cada persona.



Figura 1. Colocación de los electrodos en los músculos vasto lateral, recto femoral y vasto medial.

El procedimiento para la obtención de datos consistió en obtener la señal durante tres casos específicos: En reposo, cuando la persona se encontraba sentada con el músculo completamente relajado y el pie en una posición de 90 grados con respecto al suelo. En movimiento sin peso, en este caso se le pidió al paciente que se sentara en una silla y realizara un movimiento de 90 grados, desde una posición sentado a una posición donde extendiera el pie completamente. En contracción del músculo bajo un peso, el procedimiento fue igual a movimiento sin peso solo que se agregó un peso de 4 kg durante el movimiento. En la figura 2 se muestra el procedimiento empleado en la adquisición de las señales electromiográficas.



Figura 2. Obtención de señales electromiográficas empleando un peso de 4 kg.

3.2 Sistema de adquisición de datos

Para la adquisición de las señales electromiográficas se empleó el sistema de adquisición de datos myoware 2.0, el cual es un dispositivo especializado en la adquisición de señales electromiográficas. Este sistema ha sido utilizado en diversos estudios mostrando buenos resultados (Garouche and Thamsuwan, 2023; Awad and Kadhim, 2022).



3.3 Parámetros utilizados

Los parámetros electromiográficos permiten interpretar las señales electromiográficas y reducirlas a una forma más compacta, el trabajo desarrollado empleo 12 parámetros para la evaluación de las señales electromiográficas.

3.3.1 Valor Absoluto de la Suma de la Raíz Cuadrada (ASS)

$$ASS = \left| \sum_{i=1}^N \sqrt{x_i} \right| \quad (1)$$

Donde:

- x_i son los valores de la señal en el tiempo.
- N es el número total de muestras de la señal.

3.3.2 Valor Absoluto de la Suma de la Raíz Exponente (ASM)

$$ASM = \left| \sum_{i=1}^N \sqrt{|x_i|} \right| \quad (2)$$

Donde:

- x_i son los valores de la señal.
- N es el número total de muestras de la señal.
- $|x_i|$ es el valor absoluto de cada punto de la señal.

3.3.3 Energía (E)

$$E = \sum_{i=1}^N X_i \quad (3)$$

Donde:

- N es el número total de puntos de la señal X .
- X_i son los valores de la señal en cada punto.

3.3.4 Cardinalidad (CARD)

$$CARD = \sum_{i=1}^{N-1} [|Y_{i+1} - Y_i| > \theta] \quad (4)$$

Donde:

- Y es un vector ordenado.
- θ es el umbral definido.
- $[\cdot]$ es la función indicadora, que vale 1 si la condición es verdadera y 0 si es falsa.
- N es el número de elementos en Y .

3.3.5 Varianza de la Diferencia (DVARV)

$$DVARV = \frac{1}{N-2} \sum_{i=1}^{N-1} (X_{i+1} - X_i)^2 \quad (5)$$

Donde:

- X_i son los valores de la señal.
- $X_{i+1} - X_i$ es la diferencia entre dos muestras consecutivas.

3.3.6 Rango Intercuartil (IQR)

$$IQR = Q3 - Q1 \quad (6)$$

Donde:

- $Q1$ es el primer cuartil, el valor que marca el 25% de los datos.
- $Q3$ es el tercer cuartil, el valor que marca el 75% de los datos.

3.3.7 Desviación Absoluta Media (MAD)

$$MAD = \frac{1}{N} \sum_{i=1}^N |X_i - \mu| \quad (7)$$

Donde:

- X_i es el valor de la señal en el tiempo i .
- μ es el valor medio de la señal.
- N es el número total de elementos de la señal.

3.3.8 Raíz Cuadrada Media (RMS)

$$RMS = \sqrt{\frac{1}{N} \sum_{i=1}^N X(i)^2} \quad (8)$$

Donde:

- $X(i)$ son los valores de la señal.
- N es el número de elementos de la señal.

3.3.9 Desviación Estándar (SD)

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (X_i - \mu)^2} \quad (9)$$

Donde:

- N es el número total de elementos de la señal.
- X_i son los valores individuales de la señal X .
- μ es la media de la señal X .
- El factor $(N-1)$ es el grado de libertad de la muestra, se usa para corregir el sesgo en la estimación de la varianza muestral.

3.3.10 Amplitud de Willison (WA)

$$WA = \sum_{i=1}^{N-1} \delta(|X(i+1) - X(i)| > \theta) \quad (10)$$

Donde:

- $X(i)$ son los valores de la señal.



- N es el número total de muestras.
- θ es el umbral de activación.
- $\delta(\cdot)$ es una función indicadora que vale 1 si la condición es verdadera, y 0 en caso contrario.

3.3.11 Cruce por Cero (ZC)

$$ZC = \sum_{n=1}^{N-1} \delta(n) \quad (11)$$

Donde:

- N es la longitud de la señal.
- $\delta(n)$ indica si hay un cruce por cero significativo.

3.3.12 Longitud de la Forma de Onda (WL)

$$WL = \sum_{i=2}^N |X(i) - x(i-1)| \quad (12)$$

Donde:

- $X(i)$ es el valor de la señal en el punto i .
- N es la longitud total de la señal.

3.4 Filtrado de la señal

Existen diferentes formas de filtrar la señal, una de ellas es empleando filtros analógicos, y la otra es empleando filtros digitales, en la investigación la señal se filtró utilizando un filtro digital pasa bajos con una frecuencia de corte de 450 Hz y un filtro pasa altos con una frecuencia de corte de 20 Hz, el objetivo de los filtros fue eliminar los ruidos de baja frecuencia que surgen debido al movimiento de cables, y eliminar los ruidos de alta frecuencia que no pertenecen a la señal electromiográfica. La figura 3 muestra la señal electromiográfica filtrada.

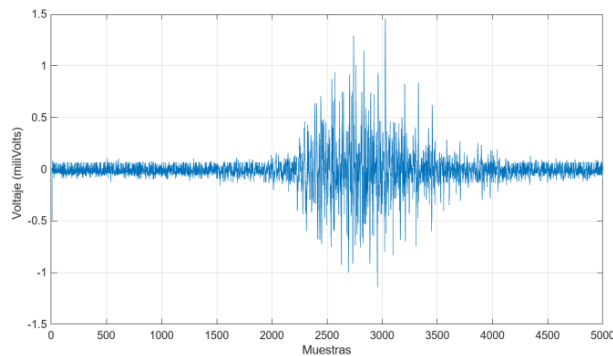


Figura 3. Señal electromiográfica filtrada.

3.5 Evaluación de la señal

En la actualidad el empleo de la descomposición de la señal a través de wavelets es ampliamente utilizado (Shariatzadeh et al., 2023). La transformada de wavelet permite la descomposición de las señales en diferentes componentes para seleccionar solo el rango de frecuencia en el que se encuentran las señales electromiográficas. La transformada de wavelet es ampliamente utilizada en señales que varían con el tiempo.

La descomposición de las wavelets permite eliminar las frecuencias que son consideradas como ruido en la señal, inducida por campos electromagnéticos exteriores. El estudio empleó la transformada de wavelet discreta (DWT). La transformada de wavelet permite ver cómo se comporta la señal con respecto a la frecuencia en el tiempo. La transformada de wavelet consiste en aplicar una serie de filtros a la señal para obtener los coeficientes de detalle y de aproximación.

La transformada de wavelet se descompuso en 5 niveles de detalle los niveles de descomposición fueron D1 (500 a 250 Hz), D2 (125 a 250 Hz), D3 (62.5 a 125 Hz), D4 (31.25 a 62.5 Hz) y D5 (de 0 a 31.25 Hz). Una vez obtenidos los coeficientes de detalle, se combinan los coeficientes que contienen las frecuencias electromiográficas deseadas. Para combinar los diferentes coeficientes de wavelet se puede reconstruir la señal si se emplea la transformada de wavelet inversa, la cual permite combinar las señales a diferentes frecuencias en una nueva señal que incluya las frecuencias deseadas. El trabajo utilizó la wavelet Daubechies wavelet (db7) de la cual se obtuvieron los coeficientes de detalle D2, D3, D4 los que se combinaron para obtener la señal, después se calcularon los parámetros de la sección 3.3 los cuales se introdujeron en el método de análisis discriminante lineal (LDA) y el método de análisis discriminante cuadrático (QDA) para la clasificación de las señales electromiográficas.

4. RESULTADOS

Para el análisis de la señal electromiográfica se emplearon los métodos análisis discriminante lineal y análisis discriminante cuadrático, se utilizó el método de validación cruzada (k fold cross validation) con un coeficiente de 10 con cada uno de los dos métodos utilizados, este método permite validar el modelo para reducir el sobreajuste, dividiendo los datos en partes iguales, y entrenando el modelo el número de veces que ha sido dividido proporcionando una clasificación más robusta del modelo. La figura 4 muestra la matriz de confusión de la clasificación empleando el método discriminante cuadrático del músculo vasto medial, donde se observan las distintas clases, siendo cero la clase en reposo, 1 la clase de movimiento sin peso, y 5 la clase de movimiento con peso, esta matriz de confusión permite ver que tan preciso fue diferenciar entre las distintas clases. En el caso de la clasificación del vasto medial se puede observar que la clase en reposo tuvo un 100% de precisión, mientras que durante el movimiento de la pierna sin peso el clasificador logró predecir el 65.6% de los datos y el 34.4% de los datos los interpreto como movimiento con peso, además durante el movimiento con peso el clasificador logró predecir el 63.3% de los datos mientras que el 35.6% los predijo como movimiento sin peso, también en la figura 5 se muestra la curva ROC la cual es una herramienta para predecir la capacidad discriminativa basada en falsos positivos, siendo 1 un rango de precisión más alto.

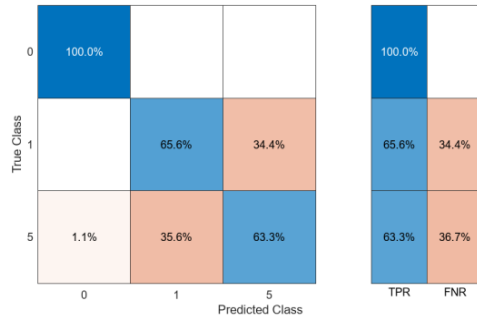


Figura 4. Matriz de confusión para el músculo vasto medial.

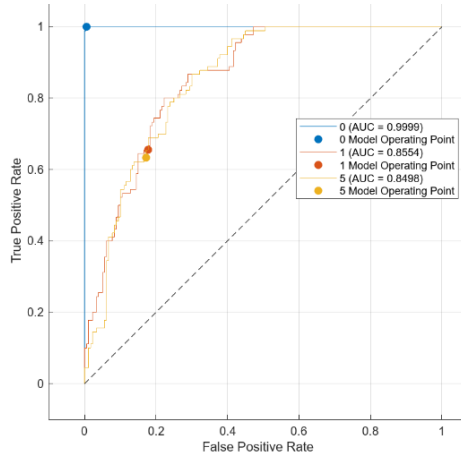


Figura 5. Curva ROC para el músculo vasto medial.

La figura 6 muestra la matriz de confusión para el músculo recto femoral donde se pueden observar los porcentajes de precisión empleando el método discriminante cuadrático, en el cual el clasificador fue capaz de predecir el 100% de los datos en reposo, mientras que para el movimiento sin peso se predijeron el 71.1% de los datos y el 28.9% de los datos fueron predichos como clase 5, así mismo para el movimiento con peso el clasificador predijo el 67.8% de los datos mientras del 32.2% de los datos los predijo como clase 1, en la figura 7 se muestra la curva ROC, donde se obtuvo un área bajo la curva (AUC) de .9946 en reposo, un AUC de .8741 con movimiento sin peso y un AUC de .8851 con movimiento agregando peso.

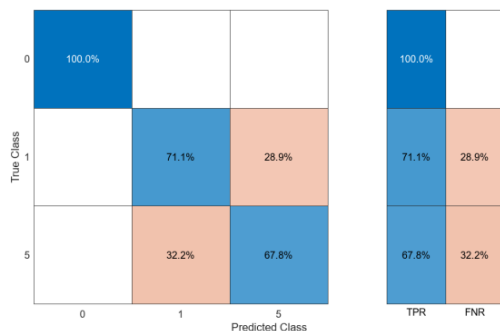


Figura 6. Matriz de confusión para el músculo recto femoral.

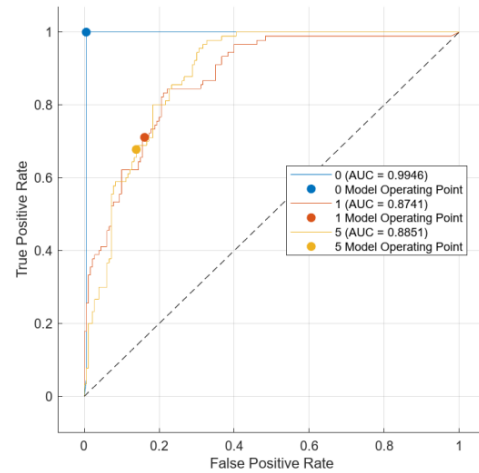


Figura 7. Curva ROC para el músculo recto femoral.

La figura 8 muestra los porcentajes de clasificación del músculo vasto medial, la matriz de confusión muestra que el algoritmo de clasificación fue capaz de identificar el 100% de los datos en reposo, durante el movimiento sin peso el algoritmo pudo identificar el 72.2% de los datos mientras que el 27.8% de los datos los interpreto como con peso, y durante el movimiento con peso el algoritmo detecto el 64.4% de los datos mientras que el 35.6 % de los datos los interpreto como movimiento sin peso. La figura 9 muestra la curva ROC para el músculo vasto medial donde se puede observar un AUC de 1 para el músculo en reposo, un AUC .8842 para el músculo con movimiento sin peso, y un AUC de .8842 para el músculo con movimiento agregando peso.

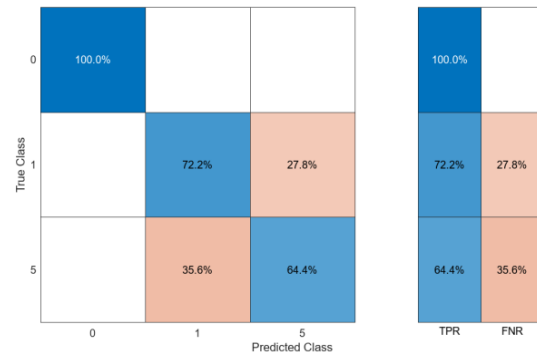


Figura 8. Matriz de confusión para el músculo vasto lateral.

En la tabla 1 se resumen los porcentajes precisión de la clasificación empleando el método análisis discriminante lineal y análisis discriminante cuadrático para cada uno de los músculos analizados, donde se puede observar que se obtuvieron mejores resultados empleando el método discriminante cuadrático con un porcentaje de clasificación de 80% para el músculo vasto medial, 79.73% para el recto femoral, y 80.74% para el músculo vasto lateral.

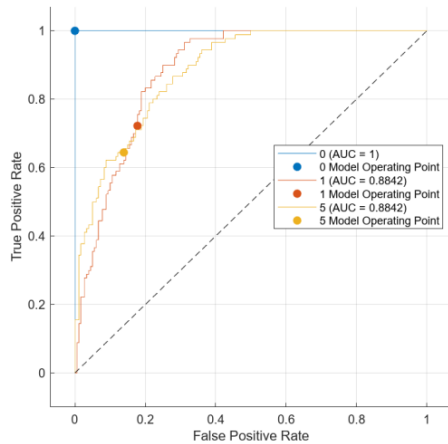


Figura 9. Curva ROC para el músculo vasto lateral.

Tabla 1. Clasificación entre diferentes músculos empleando los métodos LDA y QDA.

	LDA	QDA
Vasto medial	76.30%	80.00%
Recto femoral	77.04%	79.73%
Vasto lateral	78.15%	80.74%

En la tabla 2 se muestran los resultados de los porcentajes de los valores AUC para cada uno de los músculos bajo las diferentes condiciones de contracción del músculo.

Tabla 2. Porcentajes AUC entre diferentes músculos bajo diferentes condiciones de esfuerzo.

AUC	Vasto medial	Recto femoral	Vasto lateral
Reposo	.9999	.9946	1
Movimiento sin peso	.8554	.8741	.8842
Movimiento con peso	.8498	.8851	.8842

Se realizó la clasificación de cada uno de los sujetos empleando el método análisis discriminante cuadrático de manera individual en lugar de grupal como se hizo anteriormente, en la tabla 3 se resumen los porcentajes de clasificación obtenidos, la clasificación de sujetos de manera individual tuvo mejores resultados que la clasificación de forma grupal lo cual indica que fue posible distinguir entre las diferentes condiciones de carga para cada uno de los músculos.

Tabla 3. Clasificación por sujetos empleando método QDA.

	Vasto medial	Recto femoral	Vasto lateral
Sujeto1	93.33%	96.66%	76.67%
Sujeto2	90.00%	86.67%	73.33%
Sujeto 3	86.67%	93.33%	70.00%
Sujeto 4	83.33%	90.00%	76.67%
Sujeto 5	86.67%	93.33%	73.33%
Sujeto 6	86.67%	96.67%	83.33%
Sujeto 7	90.00%	83.33%	83.33%
Sujeto 8	90.00%	90.00%	73.33%
Sujeto 9	93.33%	100%	86.67%

La figura 10 muestra la matriz de confusión para el sujeto 1 del músculo vasto medial empleando el método de validación cruzada (k fold cross validation) con un coeficiente de 10, los resultados muestran el porcentaje de clasificación donde se puede observar que el porcentaje de precisión de los datos en reposo fue del 100%, mientras que el porcentaje predicho para el movimiento de la pierna sin peso fue de 90%, y el porcentaje de precisión para el movimiento de la pierna con peso fue de 90%, así mismo se muestra en la figura 11 el porcentaje de precisión AUC para cada una de las diferentes condiciones de carga, siendo el porcentaje AUC más cercano a 1 más preciso, obteniendo un porcentaje de AUC = 1 para cuando se encontraba en reposo, un AUC de .925 cuando la pierna se encontraba en movimiento sin peso y un AUC de .925 cuando la pierna se encontraba en movimiento agregando peso.

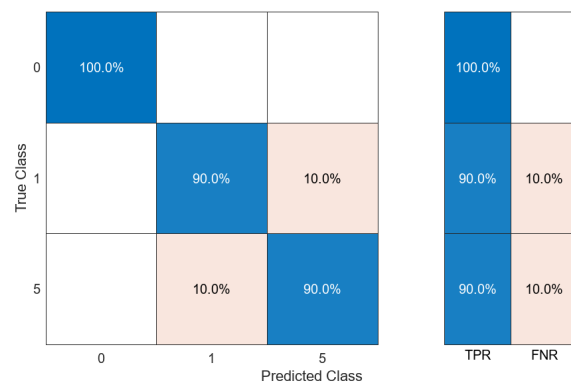


Figura 10. Matriz de confusión para el músculo vasto medial del sujeto 1.

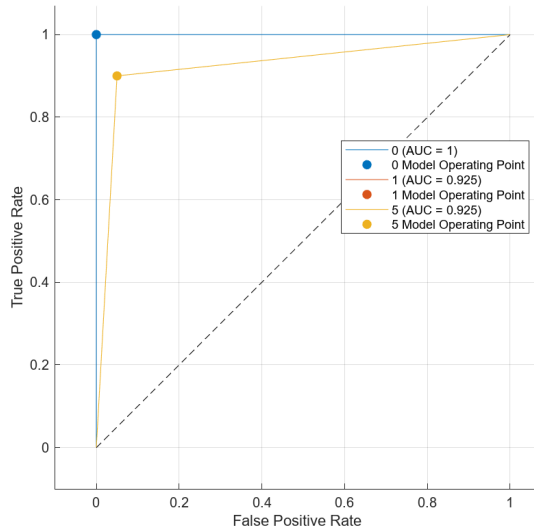


Figura 11. Curva ROC para el músculo vasto medial del sujeto 1.

5. CONCLUSIONES

Se utilizó el método análisis discriminante lineal y análisis discriminante cuadrático para clasificar los músculos vasto medial, recto femoral y vasto lateral de 9 sujetos, el procedimiento para analizar los tres músculos consistió en medir las señales electromiográficas bajo diferentes condiciones de contracción del músculo, en reposo, con movimiento sin peso y movimiento con peso, durante el análisis de cada músculo se utilizó el mismo peso para los 3 músculos evaluados, se clasificaron los músculos de manera individual y de manera grupal. La clasificación de manera grupal consistió en obtener el porcentaje de clasificación de todos los músculos de los pacientes, mientras que en la clasificación de manera individual se evaluó cada uno de los pacientes y se determinó el porcentaje de clasificación, el método análisis discriminante cuadrático obtuvo mejores resultados que el método análisis discriminante lineal y se observó que la clasificación de manera individual obtuvo mayores porcentajes de clasificación que la clasificación de manera grupal.

Como trabajo futuro, se propone incorporar nuevos métodos como redes neuronales y máquinas de vectores de soporte para analizar los porcentajes de clasificación, Asimismo, los resultados obtenidos podrían contribuir al desarrollo de sistemas de rehabilitación, prótesis inteligentes y tecnologías de asistencia motora.

Agradecimientos:

Los autores agradecen a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el apoyo otorgado para realizar este trabajo.

REFERENCIAS

Abdel-Maboud, N.F. *et al.* (2022). 'Comparative study of machine learning techniques based on TQWT for EMG signal classification', in *2022 5th International Conference on Computing and Informatics (ICCI)*. IEEE, pp. 374–377.

Adzkia, M., Setiawan, A.W. and Arland, F. (2023). 'Comparison Classification of EMG Signals in the Time Domain and Time-Frequency Domain', in *2023 International Conference on Electrical Engineering and Informatics (ICEEI)*. IEEE, pp. 1–5.

Alcan, V. and Zinnuroğlu, M. (2023). 'Current developments in surface electromyography', *Turkish journal of medical sciences*, 53(5), pp. 1019–1031.

Almaghout, K. *et al.* (2020). 'Design and control of a lower limb rehabilitation robot considering undesirable torques of the patient's limb', *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 234(12), pp. 1457–1471.

Amezquita-Garcia, J. *et al.* (2022). 'Applying machine learning to finger movements using electromyography and visualization in opensim', *Sensors*, 22(10), p. 3737.

Awad, S.F. and Kadhim, F.M. (2022). 'Compare EMG signals by using Myo-Ware muscle sensor and Myo-Trace device for measuring the electrical activity of the muscles', in *AIP Conference Proceedings*. AIP Publishing LLC, p. 40009.

Barański, R. *et al.* (2024). 'Towards hand grip force assessment by using EMG estimators', *Measurement*, 226, p. 114137.

Bawa, A. and Banitsas, K. (2022). 'Design validation of a low-cost EMG sensor compared to a commercial-based system for measuring muscle activity and fatigue', *Sensors*, 22(15), p. 5799.

Garouche, M. and Thamsuwan, O. (2023). 'Development of a low-cost portable EMG for measuring the muscular activity of workers in the field', *Sensors*, 23(18), p. 7873.

Guo, L. *et al.* (2020). 'Systematic review of textile-based electrodes for long-term and continuous surface electromyography recording', *Textile research journal*, 90(2), pp. 227–244.

Hurtado, E.F.A. *et al.* (2024). 'Movements classification system for transhumeral amputees using myoelectric signals', *Biomedical Signal Processing and Control*, 98, p. 106776.

Jochumsen, M. *et al.* (2020). 'Decoding attempted hand movements in stroke patients using surface electromyography', *Sensors*, 20(23), p. 6763.

Keleş, A.D., Türksoy, R.T. and Yucesoy, C.A. (2023). 'The use of nonnormalized surface EMG and feature inputs for LSTM-based powered ankle prosthesis control algorithm development', *Frontiers in Neuroscience*, 17, p. 1158280.

Kinugasa, R. and Kubo, S. (2023). 'Development of consumer-friendly surface electromyography system for muscle fatigue detection', *Ieee Access*, 11, pp. 6394–6403.



- Kok, C.L. *et al.* (2024). 'Machine learning-based feature extraction and classification of EMG signals for intuitive prosthetic control', *Applied sciences*, 14(13), p. 5784.
- Li, G. *et al.* (2023). 'A computational model of surface electromyography signal alterations after spinal cord injury', *Journal of Neural Engineering*, 20(6), p. 66020.
- Li, Z. *et al.* (2025). 'Surface Electromyography Monitoring of Muscle Changes in Male Basketball Players During Isotonic Training', *Sensors*, 25(5), p. 1355.
- Lin, L. *et al.* (2025). 'Low-computational EMG gesture recognition for prosthetic control via handcrafted features and lightweight MLP', *Results in Engineering*, p. 106602.
- Malesevic, N. *et al.* (2022). 'Evaluation of simple algorithms for proportional control of prosthetic hands using intramuscular electromyography', *Sensors*, 22(13), p. 5054.
- Mohapatra, N.R., Sahoo, J.P. and Sahoo, G.K. (2024). 'Hand gesture classification using surface electromyography signals based on fusion of time domain features', in *2024 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI)*. IEEE, pp. 1–6.
- Di Nardo, F. *et al.* (2022). 'Wavelet-based assessment of the muscle-activation frequency range by EMG analysis', *IEEE Access*, 10, pp. 9793–9805.
- Onay, F. and Mert, A. (2020). 'Phasor represented EMG feature extraction against varying contraction level of prosthetic control', *Biomedical Signal Processing and Control*, 59, p. 101881.
- Saranya, S., Poonguzhali, S. and Karunakaran, S. (2020). 'Gaussian mixture model based clustering of manual muscle testing grades using surface electromyogram signals', *Physical and Engineering sciences in Medicine*, 43(3), pp. 837–847.
- Sarwat, H. *et al.* (2024). 'Post-stroke hand gesture recognition via one-shot transfer learning using prototypical networks', *Journal of NeuroEngineering and Rehabilitation*, 21(1), p. 100.
- Shariatzadeh, M. *et al.* (2023). 'Predicting muscle fatigue during dynamic contractions using wavelet analysis of surface electromyography signal', *Biocybernetics and Biomedical Engineering*, 43(2), pp. 428–441.
- Sudharsan, R.R. *et al.* (2020). 'An analysis of different biopotential electrodes used for electromyography'.
- Tepe, C. and Demir, M.C. (2020). 'Detection and classification of muscle activation in EMG data acquired by Myo armband', *Avrupa Bilim ve Teknoloji Dergisi*, pp. 178–183.
- Too, J., Abdullah, A.R. and Saad, N.M. (2019). 'Classification of hand movements based on discrete wavelet transform and enhanced feature extraction', *International Journal of Advanced Computer Science and Applications*, 10(6).
- Turgunov, A. *et al.* (2020). 'Defining the features of EMG signals on the forearm of the hand using SVM, RF, k-NN classification algorithms', in *2020 Information Communication Technologies Conference (ICTC)*. IEEE, pp. 260–264.
- Verma, A.R. and Gupta, B. (2020). 'Detecting neuromuscular disorders using EMG signals based on TQWT features', *Augmented Human Research*, 5(1), p. 8.
- Wadud, A. and Showrov, M.I.H. (2020). 'EMG signal classification with effective features for diagnosis', in *International Conference on Image Processing and Capsule Networks*. Springer, pp. 629–637.
- Yang, J. *et al.* (2025). 'High-density electromyography for effective gesture-based control of physically assistive mobile manipulators', *npj Robotics*, 3(1), p. 2.

Valorization of rice by-products, a sustainable approach for a circular economy: Literature review

Loeza de la Cruz, Cynthia*, Hernández Barba, Nayeli del Milagro 2 *, Guerrero Ramírez, Gerardo Vicente*

**Tecnológico Nacional de México / Centro Nacional de Investigación y Desarrollo Tecnológico, Apatzingán 212, Palmira, 62490 Cuernavaca, Mor.*

(e-mail: m25ce028@cenidet.tecnm.mx, nayeli.hb@cenidet.tecnm.mx, gerardo.gr@cenidet.tecnm.mx)

Abstract: This study presents a literature review aimed at analyzing the current scientific knowledge on the valorization of rice by-products, specifically rice husk, rice straw, rice bran, and rice flour, within the framework of circular economy and sustainable agricultural systems. The main objective is to identify, classify, and synthesize the technological, environmental, economic, and social applications of these agricultural residues, as well as to evaluate their potential for integration into sustainable production systems, with special attention to artisanal rice production contexts such as those found in the state of Morelos, Mexico. The methodology follows a review approach based on established reporting guidelines, focusing on peer-reviewed scientific literature published between 2020 and 2025. A total of 40 scientific articles were selected, primarily retrieved from the Web of Science database, complemented by relevant institutional and academic sources. The selected studies were analyzed through qualitative synthesis and thematic categorization according to type of by-product and application sector, including energy production, agricultural soil improvement, construction materials, food innovation, and biotechnological processes. The results reveal a wide range of technological applications with high potential for resource recovery and value generation. Rice husk and rice straw are mainly associated with renewable energy production and soil enhancement, while rice bran and rice flour show strong applications in food systems and biodegradable materials. However, despite significant technological advances, a persistent gap remains between scientific development and practical implementation in local agricultural systems. The study highlights that in regions such as Morelos, traditional and artisanal production practices limit the adoption of circular economy strategies, resulting in the underutilization of valuable agricultural residues. This research emphasizes the importance of strengthening knowledge transfer, technological adaptation, and policy integration to promote sustainable and circular agricultural systems.

Keywords: circular economy agriculture, sustainable agriculture, agro-industrial waste.

1. INTRODUCCIÓN

El arroz constituye uno de los cereales básicos en la alimentación mundial y un componente estratégico para la seguridad alimentaria. En México, su relevancia es tanto nutricional como económica y cultural, formando parte esencial de la dieta cotidiana. No obstante, el país presenta una alta dependencia del mercado internacional, de acuerdo con el Servicio de Información Agroalimentaria Y Pesquera (SIAP), la producción nacional ha oscilado en años recientes entre 250 mil y 300 mil toneladas anuales, cifra insuficiente para cubrir la demanda interna, lo que ha derivado en importaciones superiores al 70 % del consumo nacional (SIAP, 2023; Secretaría de Agricultura y Desarrollo Rural (SADER), 2022).

En este contexto, el estado de Morelos destaca como una región emblemática en la producción arroceras, particularmente por la calidad diferenciada y conocida como "arroz del estado de Morelos", por su denominación de origen y reconocimiento a nivel nacional por sus características físicas y culinarias (Instituto Mexicano de la Propiedad Industrial (IMPI), 2012). La producción de este cereal en esta entidad, aunque limitada

en volumen, representa una actividad económica relevante para las familias morelenses, ya que contribuye a la generación de empleo y al fortalecimiento de la identidad agroalimentaria del estado (Gobierno del Estado de Morelos, 2019).

1.1 Desaprovechamiento de subproductos en el sector arroceras morelense

El procesamiento del arroz genera una cantidad significativa de subproductos, entre los que destacan la cascarilla, la paja, el salvado y la harina. Sin embargo, en el estado de Morelos, el aprovechamiento de estos materiales es aún limitado, en muchos casos, los residuos agrícolas son quemados a cielo abierto o utilizados de manera rudimentaria como forraje, sin considerar alternativas de valorización tecnológica o económica (Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT), 2020).

Esta práctica, no solo genera impactos ambientales negativos, como emisiones de gases de efecto invernadero y contaminación del aire, sino que también representa una pérdida de oportunidades económicas para las personas



productoras. De conformidad con la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO), el aprovechamiento y la valorización de subproductos agrícolas puede mejorar de manera significativa los ingresos en las cadenas de valor en el sector agroindustrial, como sería, por ejemplo, el aprovechamiento del uso de la biomasa residual (FAO, 2019). En ese sentido, el desaprovechamiento de los subproductos del arroz en el estado de Morelos, evidencia la desarticulación existente entre la producción agrícola y los principios de sostenibilidad.

1.2 Economía circular y transición hacia una agricultura sostenible

La economía circular se considera como una estrategia fundamental para cambiar los modelos de producción actuales, fomentando el aprovechamiento eficaz de los recursos y la reducción de desechos. En el sector agrícola, esto implica el cierre de ciclos productivos mediante la reincorporación de subproductos al sistema económico, lo que permite generar valor añadido y al mismo tiempo disminuir impactos ambientales (SEMARNAT, 2020).

En el caso del cultivo del arroz, la adopción de los principios de economía circular, representa una oportunidad para mejorar la sostenibilidad del sector, aumentar la eficiencia productiva y fortalecer la resiliencia económica de las personas productoras. Además, contribuye al cumplimiento de los compromisos internacionales en materia de desarrollo sostenible, tal como se menciona en el siguiente apartado.

1.3 Marco normativo y pertinencia del estudio

Existen diversos instrumentos normativos que respaldan la transición hacia modelos de producción sostenibles, a nivel internacional, por ejemplo, la Agenda 2030 de las Naciones Unidas para el Desarrollo Sostenible, define objetivos fundamentales relacionados con la producción responsable, la acción climática y la seguridad alimentaria (Organización de las Naciones Unidas (ONU), 2015).

En México, esta perspectiva se alinea con la Ley General de Economía Circular (LGEC), promulgada el 19 de enero de 2026, la cual promueve un modelo productivo y de consumo sostenible que reduzca la generación de desechos y permita aprovecharlos al máximo, a través de la reutilización, el reciclaje y la valorización (Diario Oficial de la Federación, 2026).

En relación con el arroz del estado de Morelos, la LGEC impulsa el aprovechamiento de los subproductos, tal como la cascarilla, la paja, el salvado y la harina, promoviendo el uso de los residuos como materias primas secundarias en nuevos procesos productivos; impulsando cadenas de valor circulares, en donde los subproductos agrícolas pueden reutilizarse en otras industrias como la energética, de biomateriales y alimentaria, además de priorizar la valorización antes que la disposición final, e incentivando su transformación en nuevos productos.

Es decir, este respaldo normativo establece el marco legal para que los residuos del arroz dejen de ser desechos y se conviertan en insumos económicos aprovechables.

Además, la LGEC se armonizó con la Ley General para la Prevención y Gestión Integral de los Residuos, que promueve la valorización de residuos, así como con la Ley de Desarrollo Rural Sustentable, que impulsa prácticas agrícolas sostenibles (Cámara de Diputados, 2021a;2021b). Así mismo, el Plan Nacional de Desarrollo 2025-2030 y los programas sectoriales de la SADER enfatizan la importancia de fortalecer la autosuficiencia alimentaria y el desarrollo sostenible del sector agroalimentario (Gobierno de México, 2019).

A nivel estatal, el Plan Estatal de Desarrollo 2025-2030, establece como una de sus prioridades el fortalecimiento sostenible del sector agropecuario, esto mediante la implementación de políticas públicas que se orienten a la innovación tecnológica, la capacitación y la adopción de prácticas agrícolas sostenibles. En este marco, destacan algunas iniciativas, como el programa *Morelos agroecológico* y el fortalecimiento al cultivo de arroz, iniciativas que promueven una producción más eficiente y resiliente, alineada con los principios de sostenibilidad y que da pauta para incorporar estrategias de economía circular en el aprovechamiento integral de los recursos agrícolas (Gobierno del Estado de Morelos, 2025).

Estos marcos normativos evidencian la pertinencia y urgencia de implementar modelos de economía circular en los sectores agrícolas del campo mexicano, y en este caso, en el campo morelense, particularmente en el sector arrocero.

1.4 Justificación del artículo de revisión

Pese al creciente reconocimiento del potencial de los subproductos agrícolas, existe una brecha significativa y persistente entre el conocimiento científico disponible y su aplicación en contextos locales. En el estado de Morelos, especialmente en el caso del arroz del estado de Morelos, esta brecha restringe la formulación de estrategias que permitan aprovechar de manera integral los subproductos de este cereal.

Por lo anterior, el presente artículo de revisión tiene como objetivo analizar la literatura existente sobre las aplicaciones de la cascarilla, paja, salvado y harina de arroz, con el objetivo de identificar oportunidades para su implementación en el contexto del campo morelense, con el fin de contribuir a la transición hacia una agricultura más sostenible que incorpore los principios de la economía circular.

1.5 Aplicaciones potenciales de los subproductos del arroz

En el contexto internacional, los subproductos del arroz han sido ampliamente estudiados y aprovechados en diferentes sectores productivos, por lo que, mediante un análisis de la literatura, se identificaron múltiples líneas de valorización de los subproductos de este cereal, como cascarilla, paja, harina y salvado, agrupadas en aplicaciones energéticas, agrícolas, industriales, alimentarias y biotecnológicas. Los hallazgos evidencian una transición hacia modelos de economía circular, donde los residuos agroindustriales son reincorporados como insumos de alto valor.



1.5.1 Aprovechamiento de la cascarilla de arroz

La cascarilla de arroz es uno de los subproductos más estudiados debido a su alta disponibilidad, bajo costo y propiedades fisicoquímicas.

La literatura refiere su uso en sectores diversos como se muestran en la Tabla 1.

Tabla 1. Aplicaciones de la cascarilla de arroz

Aplicación	Hallazgo	Autor	País
Energía (gasificación)	Viabilidad técnica para producir syngas en contextos rurales	Agudelo <i>et al.</i> (2022)	Colombia
Bioenergía	Obtención de hidrochar y gases con potencial energético	Ojeda-Rodríguez <i>et al.</i> (2024)	México
Electricidad	Alta eficiencia energética (~96%) y viabilidad económica	González <i>et al.</i> (2022)	España
Mejora del suelo	Incremento de fertilidad	Asadi <i>et al.</i> (2021)	China
Biofiltros	99 % remoción de color	Jácome & Escobar (2022)	Ecuador
Construcción	Mejora resistencia de concreto	Oblitas-Gastelo <i>et al.</i> (2022)	Colombia
Biomateriales	Sílice funcional para tecnologías	Rajan <i>et al.</i> (2024)	Malasia
Biotecnología	Producción de enzimas	Mora-López <i>et al.</i> (2023)	Colombia

Energía renovable

Diversos estudios han demostrado el potencial energético de la cascarilla de arroz, por ejemplo, la gasificación y combustión directa, permiten generar electricidad y biocombustibles, con aplicaciones exitosas en Asia y América Latina (Ortega & Quispe, 2021; Agudelo *et al.*, 2022). Asimismo, procesos como la carbonización hidrotérmica permiten obtener hidrochar y gas de síntesis (Ojeda-Rodríguez *et al.*, 2024).

Aplicaciones agrícolas

El biocarbón derivado de cascarilla de arroz, ha demostrado mejorar las propiedades del suelo como el pH, la retención de agua y la disponibilidad de nutrientes (Asadi *et al.*, 2021). Además, contribuye a reducir emisiones de gases como N₂O y NH₃ (Lee *et al.*, 2025).

También se ha demostrado su utilidad en producción de plántulas (Valdés-Zayas *et al.*, 2023) y sustitución de turba en sustratos agrícolas (Pérez-Cabrera *et al.*, 2021); se demostró que el compost con cascarilla de arroz mejora la retención de agua y el crecimiento vegetal (Leconte *et al.*, 2025); también se evidenció que las nanopartículas de sílice aumentan la tolerancia a la sequía en cultivos (Abdo *et al.*, 2024); Además se confirmó que la sílice derivada de cascarilla favorece el desarrollo vegetal (Sekifuji *et al.*, 2019).

Remediación ambiental

La cascarilla posee alta capacidad adsorbente, lo que permite su uso en el tratamiento de aguas; la literatura reporta eficiencias superiores al 95 % en la remoción de contaminantes (Jácome & Escobar, 2022; Li *et al.*, 2021).

Construcción sostenible

La ceniza de cascarilla de arroz se utiliza como sustituto parcial del cemento, mejorando propiedades mecánicas del concreto y también se ha incorporado en la fabricación de ladrillos ecológicos, morteros y biocompuestos (Oblitas-Gastelo *et al.*, 2022). De igual manera, se han desarrollado cementos alternativos a partir de ceniza de cascarilla de arroz (Espinoza *et al.*, 2023); asimismo, se han mejorado las propiedades mecánicas de morteros mediante la incorporación de cascarilla y nanopartículas (Ankit *et al.*, 2025). También se ha logrado la integración de cascarilla y residuos de vidrio en concretos sostenibles (Rúa-Suárez *et al.*, 2022). Además, se ha evaluado su uso como sustituto parcial de arena en losas, obteniendo resultados favorables (Caicedo *et al.*, 2024); de igual forma, se han desarrollado ladrillos ecológicos con ceniza de cascarilla (Frapanti *et al.*, 2024). Finalmente, se ha analizado la incorporación de fibras de cascarilla en bloques de tierra para construcción, con resultados alentadores (Bazuhair, 2023).

Biomateriales y biotecnología

Se ha identificado el uso de cascarilla en la producción de sílice de alta pureza (Park *et al.*, 2023), de nanopartículas (Rajan *et al.*, 2024) y de enzimas industriales (Mora-López *et al.*, 2023; Ceballos *et al.*, 2020). También, se han revisado biocompuestos reforzados con cascarilla, evidenciando mejoras en sus propiedades mecánicas y térmicas (Haruna *et al.*, 2025); asimismo, se han elaborado envases biodegradables a partir de cascarilla, con buenas propiedades de resistencia y baja humedad (Copari-Totora, 2024). De igual forma, se han optimizado bandejas biodegradables mediante mezclas de cascarilla y almidones, mejorando su desempeño estructural (Luna *et al.*, 2024). Además, se han evaluado sus propiedades emulsificantes con potencial para aplicaciones alimentarias e industriales (Arroyo-Arboleda *et al.*, 2022).

1.5.2 Aprovechamiento de la paja de arroz

La paja de arroz se posiciona como un recurso clave en construcción, energía y química verde, como se muestra en la Tabla 2.



Tabla 2. Aprovechamiento de la paja de arroz

Aplicación	Hallazgo	Autor	País
Construcción	Mejora resistencia estructural	Ramos-Rodríguez & Viera-Arroba (2023)	Cuba
Paneles ecológicos	Sustituto del yeso	Hurtado-Figueroa <i>et al.</i> , (2025).	México
Energía	Alto poder calorífico	Ponce <i>et al.</i> , (2020).	México
Química verde	Extracción de antioxidantes	Jorda (2020)	España

Construcción

La incorporación de paja de arroz en morteros y paneles mejora propiedades mecánicas y reduce el impacto ambiental (Ramos-Rodríguez & Viera-Arroba, 2023).

Energía

La producción de pellets con paja de arroz presenta altos valores caloríficos (~29 MJ/kg), lo que confirma su viabilidad como biocombustible (Ponce *et al.*, 2020).

Aplicaciones químicas

La paja es una fuente de compuestos fenólicos con propiedades antioxidantes y antimicrobianas (Jorda, 2020).

1.5.3 Aprovechamiento de la harina de arroz

La harina de arroz destaca en la industria alimentaria por su capacidad de mejorar textura y estabilidad en productos fermentados y libres de gluten. Asimismo, se ha incorporado en biocompuestos biodegradables con más del 70 % de degradación del suelo (Díaz-Herrera *et al.*, 2022). En la Tabla 3 se muestran los principales hallazgos de las aplicaciones de este subproducto.

Tabla 3. Aplicaciones de la harina de arroz

Aplicación	Hallazgo	Autor	País
Bebidas fermentadas	Alta aceptación sensorial (~90 %) y buena estabilidad fisicoquímica.	Colominas-Aspuro <i>et al.</i> , 2023; Colominas <i>et al.</i> , 2021.	Cuba
Panificación	Sustituto viable libre de gluten.	Hernández-Uribe <i>et al.</i> , 2025.	México
Biopolímeros	Alta biodegradabilidad del material.	Díaz-Herrera <i>et al.</i> , 2022.	México

1.5.2 Aprovechamiento del salvado de arroz

El salvado de arroz presenta un alto valor nutricional y funcional. Su incorporación en alimentos mejora la calidad nutricional y la aceptación sensorial (Gómez & Ramírez, 2020), además, permite desarrollar productos innovadores con valor agregado; sus principales aplicaciones se observan en la Tabla 4.

Tabla 4. Aplicaciones del salvado de arroz

Aplicación	Hallazgo	Autor	País
Alimentos	Mejora nutricional	Gómez & Ramírez (2020)	Venezuela
Grasas saludables	Alternativa sin grasas trans	Callejas, (2020)	Uruguay
Bebidas fermentadas	Vida útil hasta por 32 días	Goyes, (2020).	Ecuador
Comercialización de alimentos saludables	Competitividad como alimento saludable con apoyo de marketing	Benaque (2020)	México

2. METODOLOGÍA

El presente estudio se desarrolló bajo un enfoque de revisión de la literatura, con el objetivo de identificar, analizar y sintetizar el conocimiento científico existente sobre el aprovechamiento de los subproductos del arroz, en el marco de la economía circular y la sostenibilidad.

La revisión se llevó a cabo siguiendo los lineamientos del método PRISMA (*Preferred Reporting Items for Systematic Reviews and Meta-Analyses*), lo que permitió estructurar de manera transparente y reproducible el proceso de búsqueda, selección y análisis de la información científica.

2.1 Fuentes de información

El trabajo de búsqueda bibliográfica se realizó principalmente en la base de datos *Web of Science Group*, reconocida por su alta calidad y rigor en la indexación de literatura científica de alto impacto. De manera complementaria, se revisaron fuentes gubernamentales y documentos institucionales relevantes para reforzar el análisis y contextualizar aspectos normativos relacionados con la producción de arroz y la gestión sostenible de los residuos.

La búsqueda se desarrolló durante el periodo comprendido de agosto de 2025 a febrero de 2026.

2.2 Estrategia de búsqueda

La estrategia de búsqueda diseñada se basó en la utilización de palabras clave en inglés y español relacionadas con el tema de estudio y combinadas mediante operadores booleanos. Entre los principales términos empleados se encuentran: *rice husk valorization*, *rice straw utilization*, *rice bran*, *rice flour*,



circular economy sustainability, agricultura, agrícola, subproductos de arroz, entre otros.

Se utilizó, por ejemplo, la siguiente ecuación de búsqueda:

("rice husk" OR "rice straw" OR "rice bran" OR "rice flour") AND ("circular economy" OR "valorization" OR "sustainability")

La búsqueda priorizó filtros por tipo de documento (artículos científicos) y periodo de publicación (2020 - 2025).

2.3 Criterios de inclusión y exclusión

Con el fin de asegurar la calidad y la relevancia de la información analizada, se fijaron los siguientes criterios de selección:

2.3.1 Criterios de inclusión:

- Artículos científicos publicados entre 2020 y 2025
- Investigaciones relacionadas con la valorización y/o el aprovechamiento de subproductos del arroz
- Investigaciones con enfoque en sostenibilidad, bioeconomía o economía circular
- Publicaciones con acceso a texto completo

2.3.2 Criterios de exclusión:

- Documentos duplicados
- Publicaciones no relacionadas directamente con aplicaciones de subproductos del arroz
- Estudios sin claridad metodológica
- Estudios sin acceso a texto completo
- Literatura no científica (excepto documentos institucionales relevantes)

2.4 Proceso de selección de estudios

El proceso de selección de los estudios se realizó en cuatro etapas, de conformidad con el protocolo PRISMA:

1. *Identificación*: recopilación inicial de artículos en bases de datos científicas
2. *Depuración*: eliminación de registros duplicados
3. *Cribado*: revisión de títulos y resúmenes para evaluar su pertinencia
4. *Elegibilidad*: análisis a texto completo para determinar su inclusión final

Como resultado de este proceso, se seleccionaron 40 artículos científicos que cumplen con los criterios establecidos y que constituyen la base del análisis presentado en este estudio.

2.5 Análisis de la información

Los artículos seleccionados fueron analizados mediante un enfoque cualitativo y comparativo, clasificando la información en categorías temáticas según el tipo de subproducto (cascarilla, paja, harina y salvado) y sus aplicaciones (energéticas, agrícolas, industriales, alimentarias y biotecnológicas). Este enfoque permitió identificar tendencias de investigación, avances tecnológicos y oportunidades de implementación en el contexto de la economía circular aplicada al sector arrocerero; contribuyendo así, a una comprensión integral del potencial de estos subproductos

como recursos estratégicos dentro de modelos productivos circulares.

3. RESULTADOS Y/O DISCUSIÓN

El análisis de los estudios seleccionados permitió identificar patrones, tendencias y contrastes en torno al aprovechamiento de los subproductos del arroz, así como sus implicaciones en contextos productivos reales. Si bien la literatura revisada documenta avances significativos en el desarrollo de aplicaciones tecnológicas para la valorización de estos materiales, su integración en sistemas agrícolas locales presenta importantes limitaciones.

En el caso del estado de Morelos, estos hallazgos adquieren especial relevancia, ya que permiten contrastar el potencial teórico y experimental reportado a nivel internacional, con las prácticas predominantes en un sistema productivo tradicional.

3.1 Brecha entre conocimiento científico y prácticas locales

En Morelos, la producción de arroz se caracteriza por un fuerte arraigo cultural y prácticas heredadas, lo que ha permitido conservar la calidad del producto, pero también ha limitado la adopción de innovaciones tecnológicas. En este contexto, los subproductos del arroz continúan siendo subutilizados, destinándose principalmente a la quema a cielo abierto o a su uso como forraje, sin procesos de transformación que generen valor agregado.

Esta situación contrasta con la evidencia internacional, donde la cascarilla y la paja han sido integradas en procesos industriales como la generación de bioenergía, producción de materiales de construcción y desarrollo de biomateriales. La ausencia de prácticas circulares en Morelos, refleja no solo una falta de conocimiento, sino una desarticulación entre la investigación científica, los mecanismos de transferencia tecnológica y las capacidades reales del sistema productivo local.

Esta brecha refiere la existencia de limitaciones institucionales, como son, una débil vinculación entre la academia y las personas productoras, así como las barreras económicas y culturales que dificultan la adopción de innovaciones, especialmente en contextos con fuerte arraigo a los usos y las costumbres locales, así como a una lógica de producción tradicional que se inclina preponderantemente por lo conocido, para minimizar riesgos.

3.2 Limitaciones tecnológicas en un contexto artesanal

El carácter artesanal del sector arrocerero morelense representa tanto una fortaleza cultural como una limitante tecnológica, pues la adopción de tecnologías avanzadas como la gasificación de biomasa o la producción de sílice de alta pureza, resulta poco viable en el corto plazo debido a los altos costos de inversión inicial, la falta de infraestructura industrial, la limitada capacitación técnica y la baja escala de producción. No obstante, este mismo contexto abre oportunidades relevantes para el desarrollo de tecnologías apropiadas, adaptadas a las condiciones locales, como:

- Producción de biocarbón a pequeña escala



- Compostaje de paja y cascarilla
- Uso de subproductos como sustratos agrícolas
- Sistemas comunitarios de aprovechamiento de biomasa

En ese sentido, la innovación tecnológica en Morelos debe orientarse hacia soluciones accesibles, de bajo costo y fácil implementación, más que hacia tecnologías altamente sofisticadas.

3.3 Oportunidades económicas: de residuo a recurso

El desaprovechamiento de los subproductos del arroz en Morelos, implica una pérdida directa de oportunidades económicas. La evidencia analizada muestra que estos materiales pueden integrarse en cadenas de valor que generen ingresos adicionales, como:

- Producción de insumos agrícolas (como fertilizantes o sustratos)
- Desarrollo de materiales de construcción
- Generación de energía para autoconsumo
- Elaboración de productos alimenticios con valor agregado

Sin embargo, actualmente no existen mecanismos consolidados que permitan la comercialización o valorización de estos subproductos, lo que limita su incorporación en la economía local. Esto evidencia la necesidad de promover modelos de negocio circulares, esquemas de economía social y solidaria e incentivos para la innovación rural.

3.4 Implicaciones ambientales en el contexto local

Desde el punto de vista ambiental, las prácticas actuales de manejo de residuos en el sector arrocerero morelense generan impactos negativos, particularmente por la quema de paja y cascarilla, lo que contribuye a la emisión de contaminantes y gases de efecto invernadero. En contraste, la valorización de estos subproductos permitiría:

- Reducir emisiones contaminantes
- Mejorar la calidad del suelo
- Disminuir la presión sobre recursos naturales
- Sustituir materiales convencionales altamente contaminantes

Por tanto, el aprovechamiento de subproductos no solo representa una oportunidad económica, sino también una estrategia clave para avanzar hacia una agricultura responsable en la entidad.

3.5 Dimensión social

La transición hacia modelos circulares en el sector arrocerero del estado de Morelos, tiene implicaciones sociales profundas. La implementación de estrategias de valorización de subproductos podría generar empleo local, fortalecer capacidades técnicas en comunidades rurales, diversificar las actividades económicas y reducir la vulnerabilidad económica de las personas productoras. No obstante, este proceso requiere una transformación que va más allá de lo tecnológico, implicando capacitación y acompañamiento técnico, fortalecimiento de la organización entre personas productoras

y articulación institucional. Por lo tanto, la economía circular debe considerarse como un proceso de innovación social, más allá de una innovación tecnológica.

3.6 Hacia un modelo circular en la producción de arroz del estado de Morelos

Desde la perspectiva de la economía circular, el aprovechamiento de los subproductos del arroz, no debe entenderse sólo como una forma de gestionar residuos, sino como una oportunidad para transformar el modelo de producción, lo cual implica el aprovechamiento de los recursos y materiales dentro de la cadena de valor, para reducir desperdicios, cerrar ciclos, mejorar la salud del suelo, generar valor económico adicional y con ello, fortalecer la sostenibilidad del sistema.

Para el caso particular del estado de Morelos, este enfoque resulta interesante, dado que permite pensar en una transición hacia modelos de producción más sostenibles, en donde aspectos como la innovación tecnológica puedan integrarse de manera gradual, sin abandonar por completo el conocimiento tradicional ni perder la identidad cultural característica del arroz del estado de Morelos; esto considerando que el carácter artesanal de este cereal no debe concebirse como una limitación para la modernización, sino como un activo territorial que puede fortalecer modelos de economía circular basados en identidad cultural, calidad diferenciada y producción sostenible; la clave radica en el diseño de procesos circulares que sean compatibles con la preservación del patrimonio cultural en cuanto a las prácticas tradicionales de cultivo del arroz.

En ese sentido, los hallazgos indican que el sector arrocerero del estado de Morelos, puede transitar hacia un modelo de economía circular basado en el aprovechamiento integral de los subproductos, la creación de valor a lo largo de toda la cadena de valor, la disminución de residuos y la adopción de prácticas sostenibles. Sin embargo, esta transición enfrenta desafíos estructurales que demandan una sinergia coordinada entre gobierno, academia, sector productivo y sociedad.

4. CONCLUSIONES

La presente revisión de la literatura, permite evidenciar que los subproductos del arroz tienen un alto potencial de valorización en diversos sectores, con aplicaciones tecnológicas que permiten generar valor económico, reducir impactos medioambientales y avanzar hacia la transición a modelos de economía circular. Sin embargo, en contexto local, el estado de Morelos alberga una brecha importante entre el conocimiento científico disponible y su implantación. Pese al gran valor cultural e importancia productiva del cultivo del arroz del estado de Morelos, prevalecen prácticas artesanales bajo esquemas de producción lineal (extraer-producir-desechar) que limitan el aprovechamiento de los subproductos de este cereal, los cuales, con frecuencia son desaprovechados o subutilizados.

En este sentido, uno de los principales aportes de este artículo de revisión, es dar a conocer el conocimiento existente, poniendo en evidencia la amplia diversidad de aplicaciones y tecnologías disponibles para la valorización de los



subproductos del arroz. La integración de estos hallazgos, pretende no solo reconocer los avances científicos, sino también ampliar el abanico de posibilidades para su implementación en contextos locales, como el estado de Morelos; donde el mayor reto no está en la falta de soluciones, sino en las condiciones estructurales que obstaculizan su adopción, como son, limitada infraestructura, dificultad de acceso a financiamiento y desarticulación entre las instancias del sector.

De conformidad con los datos, se puede concluir que el aprovechamiento integral de los subproductos del arroz en el estado de Morelos, es una estrategia viable y necesaria para avanzar hacia un sistema agroalimentario más sostenible, competitivo y resiliente. Este trabajo pretende ofrecer una base de conocimiento que sirva como punto de partida para el desarrollo de futuras investigaciones, el diseño de políticas públicas y el desarrollo de iniciativas que impulsen la innovación y la circularidad en el sector arrocero.

5. AGRADECIMIENTOS

Se expresa un sincero agradecimiento a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por el valioso apoyo financiero brindado a través del programa de becas para estudios de posgrado, el cual ha sido fundamental para el desarrollo de esta investigación. De igual manera, se reconoce con aprecio al Tecnológico Nacional de México (TecNM) y especialmente, al Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), por la calidad de sus programas educativos y su compromiso con la comunidad estudiantil.

REFERENCIAS

- Instituto Mexicano de la Propiedad Industrial (2012) Declaración de denominación de origen "Arroz del Estado de Morelos". Disponible en: <https://www.gob.mx/impi>
- Gobierno del Estado de Morelos (2025). Plan Estatal de Desarrollo del Estado de Morelos 2025–2030. Disponible en: <https://mixtli.morelos.gob.mx/index.php/s/X8Q6TeGgM4rHrNy>
- Secretaría de Medio Ambiente y Recursos Naturales (2020) Informe sobre gestión de residuos agrícolas en México. Disponible en: <https://www.gob.mx/semarnat>
- Organización de las Naciones Unidas para la Alimentación y la Agricultura (2019) Valorización de residuos agrícolas y su impacto en cadenas agroindustriales. Disponible en: <https://www.fao.org>
- Organización de las Naciones Unidas (2015) Agenda 2030 para el Desarrollo Sostenible. Disponible en: <https://www.un.org/sustainabledevelopment/es/>
- Diario Oficial de la Federación (2026) Ley General de Economía Circular. Disponible en: <https://www.dof.gob.mx>
- Cámara de Diputados (2021a) Ley General para la Prevención y Gestión Integral de los Residuos. Disponible en: <https://www.diputados.gob.mx>
- Cámara de Diputados (2021b) Ley de Desarrollo Rural Sustentable. Disponible en: <https://www.diputados.gob.mx>
- Gobierno de México (2019) Plan Nacional de Desarrollo 2019–2024. Disponible en: <https://www.gob.mx/pnd>
- Ortega, A. & Quispe, I. (2021) Alternativas del uso de la cascarilla de arroz como fuente energética. *Revista Fuentes el Reventón Energético*, 19. <https://doi.org/10.18273/revfue.v19n2-2021005>
- Abdo, R. et al. (2024) Efficacy of nano-silicon extracted from rice husk to modulate the physio-biochemical constituents of wheat for ameliorating drought tolerance without causing cytotoxicity. *Beni-Suef University Journal of Basic and Applied Sciences*, 13. <https://doi.org/10.1186/s43088-024-00529-2>
- Agudelo, Ó. et al. (2022) Diseño de un gasificador de lecho fijo en equicorriente de núcleo abierto para operar con cascarilla de arroz. <https://doi.org/10.26507/paper.2501>
- Ankit, A. et al. (2025) Compressive strength evaluation and predictive modeling of mortar incorporating zirconia and rice husk ash using ensemble machine learning. *Journal of Infrastructure Preservation and Resilience*, 6. <https://doi.org/10.1186/s43065-025-00149-9>
- Asadi, H. et al. (2021) Application of Rice Husk Biochar for Achieving Sustainable Agriculture and Environment- A Review. *Rice Science*, 28, pp. 325–343. <https://doi.org/10.1016/j.rsci.2021.05.004>
- Ojeda-Rodríguez, V. et al. (2024) Cascarilla de arroz, un residuo agrícola sin ser aprovechado en México. *Tendencias en energías renovables y sustentabilidad*, 3, pp. 105–110. <https://doi.org/10.56845/terys.v3i1.319>
- González, M. et al. (2022) Estudio de viabilidad energético-económico de un sistema de gasificación para el aprovechamiento de cascarilla de arroz. *XV Congreso Iberoamericano de Ingeniería Mecánica*. <https://e-spacio.uned.es/entities/publication/634106db-e16d-47c2-8751-8a218566690a>
- Leconte, M. C. et al. (2025) Valor agrícola de compost producidos con aserrín, cascarilla de arroz y estiércol de gallina. *Ciencia del Suelo*, 43(1), pp. 187–196. <https://doi.org/10.64132/cds.v43i1.887>
- Lee, B.-R. et al. (2025) Comparative effectiveness of zeolite and rice husk biochar in mitigating NH₃ and N₂O emissions. *Chemical and Biological Technologies in Agriculture*, 12. <https://doi.org/10.1186/s40538-025-00859-y>
- Valdés-Zayas, D. et al. (2023) Combined effect of burned rice husks with controlled release fertilizer on *Coffea arabica* L. *Temas Agrarios*, 28, pp. 82–94.



- Sekifuji, R. et al. (2019) Investigation of negative effects of rice husk silica on komatsuna growth. *International Journal of Recycling of Organic Waste in Agriculture*, 8. <https://doi.org/10.1007/s40093-019-00303-w>
- Pérez-Cabrera, C. et al. (2021) Biocarbón de cascarilla de arroz como sustrato en el crecimiento de plántulas de pepino. *Revista Chapingo Serie Horticultura*, 27, pp. 171–183. <https://doi.org/10.5154/r.rchsh.2021.01.002>
- Jácome, X. & Escobar, A. (2022) Biofiltración de aguas residuales... mediante cascarilla de arroz. *Revista Iberoamericana Ambiente & Sustentabilidad*, 5. <https://doi.org/10.46380/rias.vol5.e271>
- Li, Y. et al. (2021) Rice husk hydrochars... as biosorbents of organics. *Bioresources and Bioprocessing*, 8, 99. <https://doi.org/10.1186/s40643-021-00451-w>
- Oblitas-Gastelo, B. E. et al. (2023) Influence of rice husk ash in concrete. *ITECKNE*, 20(1), pp. 58–67. <https://doi.org/10.15332/iteckne.v20i1.2920>
- Espinoza, L. et al. (2023) Cemento a base de ceniza de cascarilla de arroz. *Nexo Revista Científica*, 36, pp. 19–25. <https://doi.org/10.5377/nexo.v36i05.17275>
- Rúa-Suárez, A. et al. (2022) Hormigón verde con ceniza de cascarilla de arroz. *Revista ION*, 35. <https://doi.org/10.18273/revion.v35n2-2022008>
- Caicedo, R. et al. (2024) Uso de cascarilla de arroz en bloques para losas. *MQRInvestigar*, 8(2), pp. 2718–2735. <https://doi.org/10.56048/MQR20225.8.2.2024.2718-2735>
- Frapanti, S. et al. (2024) Rice husk ash bricks properties. *Journal of Engineering and Applied Science*, 71. <https://doi.org/10.1186/s44147-024-00479-7>
- Bazuhair, R. (2023) Effect of rice husk fiber on soil properties. *Journal of Umm Al-Qura University*, 14. <https://doi.org/10.1007/s43995-023-00027-z>
- Haruna, B. et al. (2025) Hybrid rice husk polymer biocomposites review. *Journal of Umm Al-Qura University*, 14. <https://doi.org/10.1007/s43995-025-00167-4>
- Copari-Totora, N. V. (2024) Envases biodegradables con cascarilla de arroz. *Revista Veritas*, 5(3), pp. 1691–1710. <https://doi.org/10.61616/rvdc.v5i3.306>
- Luna, J. et al. (2024) Bandejas biodegradables con cascarilla de arroz. ISBN 978-628-96613-0-9.
- Rajan, R. et al. (2024) Silica nanoparticles from rice husk for forensics. *Egyptian Journal of Forensic Sciences*, 14(1), 24. <https://doi.org/10.1186/s41935-024-00398-x>
- Park, J. et al. (2023) Biogenic silica extraction from rice husk. *Chemical and Biological Technologies in Agriculture*, 10. <https://doi.org/10.1186/s40538-023-00479-4>
- Arroyo-Arboleda, M. et al. (2022) Emulsifying properties of rice husk flour. *Acta Agronómica*, 70. <https://doi.org/10.15446/acag.v70n4.78743>
- Mora-López, M. et al. (2023) Producción de celulosas con cascarilla de arroz. *Revista Ingeniería y Aplicaciones*, 20. <https://doi.org/10.24050/reia.v20i40.1671>
- Ceballos, R. et al. (2020) α -amilasas con cascarilla de arroz. *Revista Científica*, 29, pp. 40–47. <https://doi.org/10.54495/Rev.Científica.v29i2.39>
- Ramos-Rodríguez, H. & Viera-Arroba, P. (2023) Morteros con paja de arroz. *Hábitat Sustentable*, 13(2), pp. 76–91. <https://doi.org/10.22320/07190700.2023.13.02.06>
- Hurtado-Figueroa, O. et al. (2025) Paneles con paja de arroz. *Encuentro Internacional de Educación en Ingeniería*. <https://doi.org/10.26507/paper.4260>
- Ponce, W. et al. (2020) Pellets de paja de arroz. ISSN 2477-8982.
- Colominas, A. et al. (2021) Bebida fermentada con harina de arroz. *Ciencia y Tecnología de Alimentos*, 31(1), pp. 1–6.
- Colominas-Aspuro, A. M. et al. (2023) Bebida funcional con harina de arroz. *Agronomía Mesoamericana*, 34(2), 51970. <https://doi.org/10.15517/am.v34i2.51970>
- Hernández-Urbe, J. P. et al. (2025) Mantecadas libres de gluten con harina de arroz. *Boletín de Ciencias Agropecuarias*, 11(21), pp. 45–51. <https://doi.org/10.29057/icap.v11i21.12755>
- Díaz-Herrera, Y. et al. (2022) Biocompuestos con harina de arroz. *Biotecnia*, 24(1), pp. 23–29. <https://doi.org/10.18633/biotecnia.v24i1.1468>
- Gómez, M. & Ramírez, L. (2020) Arepas con salvado de arroz. *Revista Alfa*, 4(10), pp. 57–70. <https://doi.org/10.33996/revistaalfa.v4i10.71>
- Callejas, A. (2020) Grasas saludables de salvado de arroz. Tesis doctoral, Universidad de la República de Uruguay. Disponible en: <https://www.colibri.udelar.edu.uy>
- Goyes, P. (2020) Bebida fermentada con salvado de arroz. Tesis, Universidad Técnica de Ambato. Disponible en: <https://repositorio.uta.edu.ec>
- Lluggla, J. (2020) Bebida fermentada con salvado de arroz. Tesis, Universidad Técnica de Ambato. Disponible en: <https://repositorio.uta.edu.ec>
- Benaque, C. (2024) Estrategias de marketing para productos con salvado de arroz. Tesis, UNAM. Disponible en: <https://ru.dgb.unam.mx>
- Jorda, P. (2020) *Extracción de antioxidantes de la paja de arroz. Aplicación de ultrasonidos con tratamiento térmico*. Trabajo de Fin de Grado, Universitat Politècnica de València. <https://riunet.upv.es/server/api/core/bitstreams/a903cf74-7d2c-466a-83bc-2c4420cd0d13/content>

Early Forest Fire Detection using the Internet of Things and Artificial Intelligence

Julio Víctor Sánchez-Hernández*, Hugo Estrada-Esquivel*, Alicia Martínez-Rebollar*, Pedro Enrique Albertí-Villalobos, Patricia Verónica Maldonado-Cardenas**, Fernando Pech-May*****

**Tecnológico Nacional de México / Centro Nacional de Investigación y Desarrollo Tecnológico (CENIDET), Interior Internado Palmira S/N, Col. Palmira, C.P. 62490 Cuernavaca, Morelos, México (E-mail: d23ce084@cenidet.tecnm.mx, hugo.ee@cenidet.tecnm.mx, alicia.mr@cenidet.tecnm.mx)*

***Universidad de Magallanes (UMAG), Av. Pdte. Manuel Bulnes 01855, C.P. 6210427, Punta Arenas, Magallanes y la Antártica Chilena, Chile (E-mail: pedro.alberti@umag.cl, patricia.maldonado@umag.cl)*

****Tecnológico Nacional de México / Instituto Tecnológico Superior de los Ríos (ITSR), Carretera Balancán - Villahermosa KM 3, C.P. 86930, Balancán, Tabasco, México (E-mail: fernando.pech@cinvestav.mx)*

Abstract: Forest fires pose an ever-growing climate threat and highlight the limitations of conventional monitoring systems, whose high latency reduces their effectiveness. This paper evaluates the integration of the Internet of Things and artificial intelligence as a basis for the early and proactive detection of these phenomena. To this end, a literature review was conducted following Brereton's protocol, with an emphasis on studies published between 2020 and 2026. The results show a trend toward edge computing architectures, the use of deep learning algorithms, and the adoption of LoRaWAN networks. By integrating chemical and meteorological variables, these approaches achieve accuracies exceeding 98%. Taken together, the evidence indicates that distributed and edge intelligence enables near-immediate responses, providing a robust and scalable solution to protect ecosystems from the limitations of traditional methods.

Keywords: Forest fires, early detection, Internet of Things, Artificial Intelligence, Wireless Sensor Networks, literature review.

1. INTRODUCCIÓN

El cambio climático es el principal motor de los desastres naturales en el mundo. El aumento constante de la temperatura altera los ciclos del agua y seca la vegetación con rapidez. Estas condiciones generan incendios más intensos y veloces (United Nations Environment Programme, 2022). Eventos recientes, como los incendios ocurridos en Chile a principios de 2026, muestran la vulnerabilidad de los ecosistemas y confirman que los mecanismos tradicionales de respuesta ya no son suficientes (Center for Disaster Management and Risk Reduction Technology, 2026). Ante esta situación, los sistemas tradicionales de detección, que dependen de la vigilancia humana o satelital, presentan graves limitaciones. Estos métodos buscan visualmente el humo o las llamas, lo que retrasa significativamente la emisión de alertas. Esta demora reduce las posibilidades de contener el fuego, ya que los equipos de emergencia suelen recibir la advertencia cuando el incendio ya está fuera de control (Sharobiddinov *et al.*, 2025).

Para resolver este problema, la integración del Internet de las Cosas y la Inteligencia Artificial permite transformar los sensores convencionales en dispositivos autónomos. El Internet de las Cosas monitorea constantemente las condiciones climáticas en el terreno, mientras que la Inteligencia Artificial utiliza la computación en el borde para analizar la información directamente dentro de cada sensor. Este enfoque evita el envío de volúmenes masivos de datos a la nube, reduce el tiempo de respuesta a milisegundos y hace

posible una detección casi en tiempo real (Rupanetti & Kaabouch, 2024).

En este contexto, este artículo presenta una revisión de la literatura publicada entre 2020 y 2026 sobre el uso conjunto de arquitecturas del Internet de las Cosas y algoritmos de Inteligencia Artificial para detectar incendios forestales de forma temprana. En particular, el artículo busca responder tres preguntas clave: ¿qué algoritmos de Inteligencia Artificial pueden implementarse en nodos del Internet de las Cosas para detectar incendios forestales en etapas tempranas? ¿qué variables ambientales resultan críticas para identificar estos eventos a tiempo? y ¿qué componentes de hardware predominan en el diseño de estos sistemas? El artículo se estructura de la siguiente manera. La Sección 2 expone el marco teórico, y la Sección 3 detalla la metodología de la revisión sistemática. Posteriormente, la Sección 4 analiza los resultados empíricos para responder a las preguntas de investigación. Finalmente, la Sección 5 presenta las conclusiones y los hallazgos técnicos del trabajo.

2. MARCO TEÓRICO

Esta sección presenta los fundamentos teóricos de la investigación. Primero, analiza la dinámica de los incendios forestales, enfocándose en su fase superficial para destacar la importancia de la detección temprana. Luego, describe la arquitectura del Internet de las Cosas como el sistema principal para adquirir datos en el terreno. Finalmente, examina la Inteligencia Artificial y la Computación en el Borde para

explicar cómo convergen el monitoreo físico y el procesamiento de la información.

2.1 Incendios forestales

Los incendios forestales son fuegos descontrolados que consumen la vegetación rural (Academia Nacional de Bomberos de Chile, 2020) (Corporación Nacional Forestal, 2012). El comportamiento de estos siniestros depende de la topografía, el clima y el combustible disponible, factores conocidos como el triángulo del comportamiento del fuego. Además, la combustión requiere calor, oxígeno, combustible y una reacción en cadena, elementos que forman el tetraedro del fuego (Carrasco-Ruano & Núñez-Moreno, 2019) (Ministerio de Ambiente de Panamá, 2015). En su etapa inicial, el fuego avanza de manera superficial y horizontal, con llamas de hasta 1.5 metros que queman hojarasca y matorrales. Esta propagación se alimenta principalmente de combustibles finos, con diámetros menores a 6 milímetros, que se secan rápido y arden con facilidad ante el calor (Comisión Nacional Forestal, 2010) (Ministerio de Ambiente de Panamá, 2015). Por ello, los sistemas de detección temprana deben vigilar este estrato superficial. Alertar sobre el fuego en esta fase evita el efecto escalera, un fenómeno mediante el cual las llamas ascienden hasta las copas de los árboles y provocan incendios mucho más destructivos y difíciles de apagar (ver Fig. 1).



Figura 1. Tipos de incendios forestales según su propagación, mostrando su transición e intensidad creciente

2.2 Internet de las Cosas

El Internet de las Cosas funciona como un puente entre el entorno físico y los sistemas digitales. Sus sensores y actuadores transforman variables físicas, como los cambios de temperatura, en datos útiles que permiten automatizar diferentes procesos (Dauda *et al.*, 2024) (Abdulhussain *et al.*, 2025). El Internet de las Cosas estructura el flujo de información a través de capas interconectadas. La capa de percepción recolecta los datos del ambiente, y la capa de red los envía en tiempo real a servidores locales o a la nube. Posteriormente, las capas de procesamiento y aplicación analizan la información para ejecutar acciones automáticas. Este diseño convierte a los objetos físicos en fuentes constantes de datos e impulsa avances en áreas como la industria, la gestión urbana y la prevención de desastres (Sánchez-Hernández *et al.*, 2025).

2.3 Inteligencia Artificial

La Inteligencia Artificial complementa esta infraestructura tecnológica porque aporta capacidades de aprendizaje y toma de decisiones ante escenarios inciertos. A través del aprendizaje automático y profundo, la Inteligencia Artificial identifica patrones en grandes conjuntos de datos y construye modelos predictivos (Siam *et al.*, 2024). Dentro de este ecosistema, la Computación en el Borde mejora el proceso analítico al trasladar el procesamiento de los datos desde la nube hacia los dispositivos locales. Esta estrategia disminuye los tiempos de respuesta y ahorra ancho de banda, algo fundamental para los sistemas que exigen actuar de inmediato (Shi *et al.*, 2016). La unión de estas tecnologías genera sistemas ciberfísicos eficientes. El Internet de las Cosas suministra la red para recolectar información en campo y la Inteligencia Artificial, al operar directamente en el borde, procesa los datos para tomar decisiones de forma autónoma y oportuna (Siam *et al.*, 2024) (Shi *et al.*, 2016).

3. METODOLOGÍA

Para garantizar un análisis riguroso, objetivo y reproducible, esta investigación aplica el protocolo de Brereton *et al.* (2007) (ver Fig. 2). Este método organiza el trabajo en tres etapas: planificación, ejecución y presentación de resultados; para reducir los posibles sesgos del investigador. Además, el artículo utiliza la declaración PRISMA para guiar el flujo documental y transparentar la selección de las fuentes.



Figura 2. Metodología implementada en la revisión de la literatura basada en Brereton *et al.* (2007)

3.1 Planificación de la revisión

Durante la fase de planificación, definimos el alcance de la revisión a partir de un problema operativo crítico. La alta latencia y la baja eficacia de los sistemas tradicionales para detectar incendios forestales superficiales. Para abordar estas deficiencias, centramos el análisis en la integración del



Internet de las Cosas y la Inteligencia Artificial. Bajo esta premisa, formulamos preguntas de investigación dirigidas a identificar los algoritmos más eficientes, las variables ambientales clave y el hardware predominante en la literatura reciente.

3.2 Realización de la revisión

En la etapa de ejecución, realizamos una búsqueda exhaustiva en bases de datos como IEEE Xplore, ScienceDirect, SpringerLink, Wiley, Nature y MDPI mediante el uso de cadenas booleanas específicas (ver Tabla 1). Esta exploración inicial arrojó 73 artículos, de los cuales descartamos 17 por estar duplicados. Luego, filtramos los 56 documentos restantes bajo criterios estrictos. Solo incluimos artículos científicos revisados por pares, escritos en inglés y publicados entre 2020 y 2026. También exigimos que los trabajos presentaran una validación experimental comprobable, integraran ambas tecnologías de forma explícita y detallaran sus parámetros técnicos. Por otro lado, excluimos estudios puramente teóricos, literatura gris, revisiones sistemáticas anteriores y sistemas que solo dependían de imágenes convencionales o multispectrales.

3.3 Presentación de los resultados

Por último, en la fase de presentación de resultados, organizamos la evidencia en una matriz de extracción de datos. Esta herramienta nos permitió sistematizar los detalles técnicos esenciales de cada artículo, como el hardware empleado, los protocolos de comunicación, los modelos analíticos y las limitaciones identificadas en la investigación. Esta estructura facilita un análisis temático profundo, el cual responde de manera directa y fundamentada a las preguntas planteadas al comienzo de la investigación.

Tabla 1. Cadenas de búsqueda utilizadas para la identificación de literatura sobre detección temprana de incendios forestales

Cadenas de búsqueda booleanas	
Cadena 1	(wildfire OR forest fire) AND (early detection OR early warning OR real-time detection) AND (Internet of Things OR wireless sensor network) AND (artificial intelligence OR machine learning)
Cadena 2	(wildfire OR forest fire) AND (early detection OR early warning OR real-time detection) AND (Internet of Things OR IoT OR wireless sensor network OR WSN) AND (artificial intelligence OR AI OR machine learning OR deep learning)

4. RESULTADOS

Esta sección presenta los resultados de la revisión de la literatura y organiza el análisis técnico, desde los datos bibliométricos generales hasta las respuestas concretas a las preguntas de investigación.

Tras aplicar los criterios de inclusión y exclusión, seleccionamos 20 artículos primarios para realizar el análisis cuantitativo y cualitativo. La producción científica muestra un creciente interés por diseñar soluciones tecnológicas contra el cambio climático, ya que el 55% de los artículos se publicó en los últimos tres años (ver Fig. 3). En cuanto al rigor metodológico, predomina la validación empírica en entornos reales. De hecho, el 60% de los trabajos evaluó prototipos físicos directamente en bosques y zonas rurales. De manera complementaria, otras investigaciones utilizan quemaduras controladas en laboratorios o simulaciones computacionales para medir la escalabilidad y la eficiencia de las redes en áreas de difícil acceso.

El análisis de estos estudios evidencia una importante evolución técnica en los sistemas de detección temprana (ver Tabla 2). Las arquitecturas pasaron de ser modelos reactivos, que emitían alertas al superar umbrales ambientales fijos, a sistemas proactivos basados en la Inteligencia Artificial y la Computación en el Borde. Este traslado del procesamiento hacia los dispositivos locales reduce el tiempo de respuesta y optimiza el consumo de energía, pues permite filtrar los datos irrelevantes antes de enviarlos a la nube. Al mismo tiempo, los protocolos de comunicación de bajo consumo y largo alcance, como LoRa y LoRaWAN, reemplazaron gradualmente a tecnologías tradicionales como WiFi o ZigBee en los entornos forestales. Esta madurez técnica se complementa con sistemas de alimentación basados en paneles solares y baterías avanzadas, lo que garantiza la operación autónoma de los equipos en zonas aisladas.

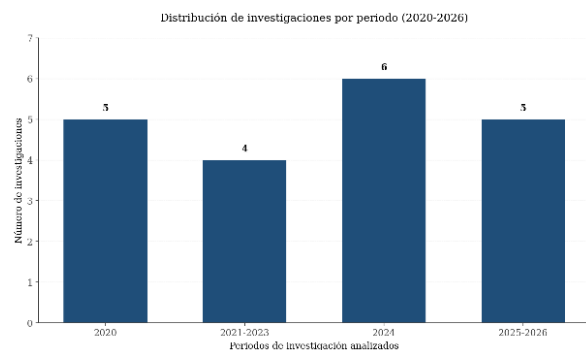


Figura 3. Número de investigaciones primarias seleccionadas para la detección y prevención de incendios forestales

En cuanto a la primera pregunta de investigación, centrada en los algoritmos de Inteligencia Artificial utilizados en los nodos, la literatura describe diversas opciones adaptadas a la capacidad de procesamiento del hardware (ver Fig. 4). Las redes neuronales profundas sobresalen al modelar relaciones complejas entre múltiples sensores, mientras que las redes recurrentes, como LSTM y GRU, analizan secuencias temporales para detectar patrones ambientales previos al fuego. Para procesar imágenes o firmas térmicas, las redes convolucionales ofrecen una alta precisión visual. En el caso de los dispositivos con recursos limitados, los algoritmos de aprendizaje supervisado tradicional, como las máquinas de vectores de soporte, los bosques aleatorios y los modelos de regresión, mantienen una gran utilidad. Por su parte, los sistemas que exigen respuestas inmediatas y bajo gasto

computacional recurren a la lógica difusa, modelos estadísticos como ARIMA y técnicas de agrupamiento para optimizar las mediciones y evitar las falsas alarmas.

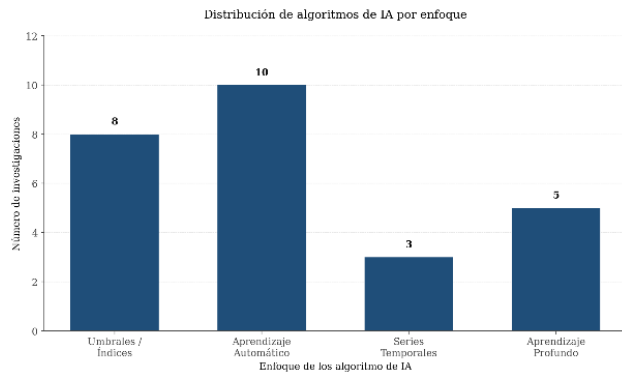


Figura 4. Clasificación de los algoritmos empleados en las investigaciones seleccionadas

Para responder a la segunda pregunta, relacionada con las variables ambientales críticas, observamos que todos los sistemas monitorean como base la temperatura del aire y la humedad relativa. Sin embargo, la detección en etapas tempranas requiere analizar la composición química del entorno mediante la medición de monóxido de carbono, dióxido de carbono y gases combustibles. Esto permite identificar la combustión antes de que aparezcan las llamas. Para confirmar la emergencia y diferenciar variaciones térmicas normales de un fuego real, los sistemas evalúan la concentración de material particulado y la densidad del humo. Además, la humedad del suelo funciona como un indicador predictivo sobre la vulnerabilidad del terreno, mientras que variables meteorológicas dinámicas, como la velocidad del viento, la presión atmosférica y la radiación solar, resultan esenciales para calcular cómo se propagará el incendio. Los modelos más avanzados también incorporan compuestos orgánicos volátiles, dióxido de nitrógeno y sensores acústicos

o infrarrojos. Finalmente, para gestionar el estado operativo de la red, los dispositivos miden parámetros de diagnóstico como la altitud, las coordenadas GPS, la calidad de la señal y la batería restante.

Respecto a la tercera pregunta, enfocada en el hardware predominante, la evidencia muestra una clara preferencia por sistemas modulares y de bajo costo que facilitan su instalación masiva. Para adquirir los datos, destacan los microcontroladores de la familia Arduino, así como las placas ESP32 y NodeMCU, preferidas por su conectividad integrada. En las pasarelas de red o nodos intermedios, las microcomputadoras como Raspberry Pi asumen el procesamiento central debido a su capacidad para ejecutar modelos analíticos complejos. El equipamiento de medición combina sensores climáticos, detectores de gases de la serie MQ, módulos multivariables de alta precisión y lectores específicos para material particulado (ver Fig. 5). Para transmitir la información, la infraestructura prioriza los módulos LoRa y las placas diseñadas para enlaces de largo alcance, aunque mantiene la conectividad celular o ZigBee como opciones de respaldo.

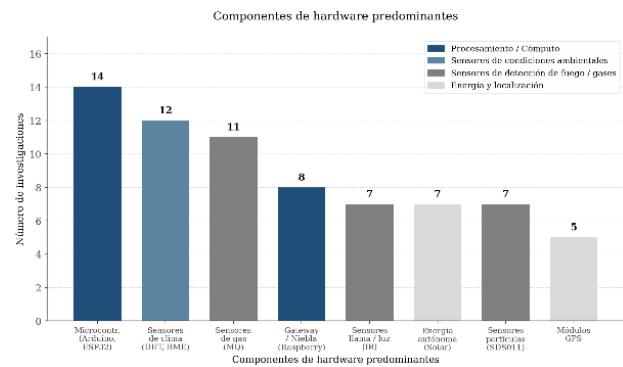


Figura 5. Componentes de hardware predominantes en las investigaciones analizadas

Tabla 2. Matriz de análisis de investigaciones seleccionados sobre detección temprana de incendios forestales basada en el Internet de las Cosas e inteligencia artificial.

Título de la investigación	Año de publicación	Algoritmos utilizados	Variables ambientales	Componentes de hardware	Tecnologías de comunicación	Limitaciones identificadas
IoT Sensor and Deep Neural Network based Wildfire Prediction System - (Zope <i>et al.</i> , 2020)	2020	Red Neuronal Densa (DNN), Regresión Logística, Modelo Funcional de Keras	Temperatura, humedad, humedad del suelo, presión atmosférica, altitud	Sensor DHT11, YL-69, BMP-280, GPS, dispositivo Bolt IoT, Arduino Uno	Protocolo I2C, Bolt API, Google Firebase, aplicación móvil Android (notificaciones push)	El sistema requiere una conexión a internet activa para obtener datos en tiempo real; los sensores utilizados son de baja potencia relativa.
Wireless Sensor Network for Ignitions Detection: An IoT Approach - (Brito <i>et al.</i> , 2020)	2020	Índice de Clima de Incendios (FWI), comparación de umbrales	Temperatura, humedad, radiación solar, presión, ignición de llama (IR), velocidad y dirección del	Placa CubeCell (LoRaWAN), módulo ATmega328P, sensores de llama IR, DHT11, anemómetro	LoRaWAN, pasarela Lora One, servidor The Things Network (TTN), visualización en ThingSpeak	Necesidad de mejorar los módulos con componentes de grado industrial y carcassas impermeables



			viento, humedad del suelo	ATMOS22, estación ATMOS41		(IP65) para despliegue final.
LPWAN Based IoT Surveillance System for Outdoor Fire Detection - (Roque & Sánchez-Padilla, 2020)	2020	Lógica basada en reglas (umbrales), lógica difusa	Temperatura, monóxido de carbono (CO), gas de hulla, gas licuado de petróleo	Arduino UNO, módulo Thinxtra Xkit (Sigfox), sensor LM35, sensor MQ9, panel solar, baterías Li-Po, escudo GSM	Sigfox, plataforma Ubidots, alertas vía Email, SMS y llamadas de voz (GSM)	Riesgo de problemas de escalabilidad al aumentar nodos; exclusión de protocolos propietarios por costos y configuración; requiere evaluar opciones de conexión.
Early Forest Fire Detection System using Wireless Sensor Network and Deep Learning - (Benzekri <i>et al.</i> , 2020)	2020	Redes Neuronales (RNN, GRU, LSTM), Índice Canadiense de Clima de Incendios (FWI)	Temperatura, humedad, presión, CO, CO2, PM2.5, PM10, velocidad del viento, ráfagas, radiación solar	Raspberry Pi, Lora32u4 (ATmega32u4), BME280, SDS011, MH-Z14A, ZE07, estación VISIOGREEN	LoRa, red LoRaWAN, protocolo MQTT, alertas vía Telegram e interfaz ThingsBoard	Carencia de estandarización global en capas de software; dificultad técnica para gestionar múltiples aplicaciones de forma simultánea.
An Alertness-Adjustable Cloud/Fog IoT Solution for Timely Environmental Monitoring Based on Wildfire Risk Forecasting - (Tsipis <i>et al.</i> , 2020)	2020	Índice de quema de Chandler (CBI)	Temperatura, humedad relativa	Nodos Arduino Uno, nodo sumidero Arduino Mega, escudo SD, módulo XBee ZigBee, sensor DHT22, Raspberry Pi (Niebla)	ZigBee, comunicación serie, 3G, alertas por correo electrónico/notificaciones web	Falta de despliegue real a gran escala; vulnerabilidad ante conectividad pobre, límites de energía y condiciones ambientales hostiles.
Intelligent Framework Using IoT-Based WSNs for Wildfire Detection - (Verma <i>et al.</i> , 2021)	2021	SEOF (Sleep scheduling-based Energy Optimized Framework), algoritmo TSA para cabeceras de clúster	Energía residual, proximidad de nodos, distancia a sumideros, retardo	Nodos sensores IoT heterogéneos, sumideros de datos (DCS) móviles y estacionarios	WSN basadas en IoT, comunicación de un solo salto a sumideros DCS, simulación MATLAB	Los resultados se basan en suposiciones de simulación; en la práctica, los obstáculos físicos y el calor extremo degradarían la operación.
Forest Fire Prediction Using IoT and Deep Learning - (Ananthi <i>et al.</i> , 2021)	2021	LBFFPS basado en Aprendizaje Profundo, SVM, Random Forest	Presencia de humo, temperatura del aire, humedad relativa	Controlador NodeMCU (WiFi), sensor MQ2, sensor DHT11, cámara digital	WiFi, servidor en la nube remoto para alertas	Experimentación realizada con una muestra pequeña; requiere validación en bosques a gran escala y por periodos prolongados.
Forest Fire Detection System Using Wireless Sensor Networks and Machine Learning - (Dampage <i>et al.</i> , 2022)	2022	Regresión lineal múltiple, análisis de relación de umbral (RTH), K-Means	Temperatura, humedad relativa, intensidad de luz, monóxido de carbono (CO)	Arduino Nano, módulo nrf24L01, sensores DHT22, LDR, MQ9, baterías 18650, paneles solares	Topología de árbol, módulo SIM800L para alertas vía SMS	Limitaciones de energía y capacidad de cómputo en los nodos finales para ejecutar algoritmos de ML altamente complejos.



Real-Time Forest Fire Detection, Monitoring, And Alert System Using Arduino - (Anuar <i>et al.</i> , 2023)	2023	Lógica de comparación de umbrales en tiempo real	Temperatura, humedad, CO ₂ , presencia de llama	Arduino UNO, shield Xbee, sensor DHT11, sensor MQ135, sensor de llama, módulo GPS	Zigbee, GPRS/WiFi, aplicación Blynk 2.0 (notificaciones push)	Falta de validación en entornos reales complejos y optimización para diferentes densidades forestales.
Real Time Smart Forest Fire Detection and Monitoring System Using IoT and Machine Learning - (Khushal <i>et al.</i> , 2024)	2024	CNN, algoritmos de umbrales, SVM, BRT, MARS	Temperatura, humedad, humo, vibración, llama	Sensores de gas, llama (IR), vibración (ADXL345), ESP32, Arduino Uno, módulo GPS, módulo GSM, IR Blaster, cámara	Módulo GSM (SMS), GPS para coordenadas	Interferencia de señal por follaje denso; necesidad de mantenimiento regular y vulnerabilidad de sensores ante fauna silvestre.
Detection And Reporting of Wildfires in Forests Using a Wireless Sensor Network - (Reddy <i>et al.</i> , 2024)	2024	Red Neuronal Artificial (ANN), CNN, Árboles de Decisión (DT), Random Forest	Temperatura, presión, humo, humedad, CO	Termopares, sensores de presión piezorresistivos, sensores de humo, sensores de humedad capacitivos, sensores de gas, MCU	Red de sensores inalámbricos (WSN), activación de alarmas y bombas de agua	Carácter semiaplicado del estudio; limitaciones en la evaluación de cobertura y tiempo de respuesta en escenarios extensos.
Integrated Internet of Things (IoT) Solutions for Early Fire Detection in Smart Agriculture - (Morchid <i>et al.</i> , 2024)	2024	Comparación de umbrales (SmokeThreshold y FlameThreshold), visualización con MATLAB	Concentración de humo, niveles de llama	Raspberry Pi 3 B+, sensores de humo y llama, conexión de red	Protocolo MQTT, plataforma en la nube ThingSpeak, alertas por correo electrónico	Falta de capacitación técnica para agricultores; dependencia de internet estable y necesidad de mantenimiento preventivo regular.
FDPA Internet Of Things System for Forest Fire Detection, Prediction and Behaviour Analysis - (Alkhatib & Jaber, 2024)	2024	Chandler Burning Index (CBI), algoritmos de detección, predicción y análisis de comportamiento	Temperatura, humedad relativa, índice CBI	Nodos Waspote (Libelium), transceptor XBee Pro (802.15.4), sensor MCP9700A, sensor 808H5V6SM, pasarela Meshlimum	WSN, ZigBee, alertas visuales en pantalla de monitoreo y base de datos MySQL/PHP	Los umbrales requieren ajuste por la alta combustibilidad del bosque; falta de permiso para quemas de propagación extensas en pruebas.
A System for Efficient Detection of Forest Fires through Low Power Environmental Data Monitoring and AI - (Uremek <i>et al.</i> , 2024)	2024	Modelo ARIMA para series temporales, Algoritmo de Árbol de Decisión	Temperatura, humedad, presión atmosférica, SNR, RSSI	Nodos LoRa Dragino, sensor BME280, pasarela Multitech LoRa Gateway	Tecnología LoRa, protocolo MQTT, red The Things Network (TTN), integración con Met Éireann	Discrepancia en la frecuencia de datos (horaria en servicios meteorológicos respecto a los minutos en sensores); validación de detección de fuego pendiente.
Enhancing Wildfire Risk Management Through Sensor-Based AI Integration in Social IoT	2024	Perceptrón Multicapa con funciones Sigmoide, ReLU y TANH	Temperatura, humedad, presión atmosférica, altitud, VOC, CO, CO ₂ , NO ₂ ,	Arduino UNO (Rev. 3 y 4), escudo LoRa (SX1278), sensores Bosch BME688, MQ-2, CJMCU4541,	LoRa (868 MHz), LoRaWAN, 3G/4G/5G/Ether net/WiFi para la pasarela,	Sensibilidad de la señal LoRa ante la orografía; los sistemas individuales (WSN o video) no son



Frameworks - (Putzu <i>et al.</i> , 2024)			infrarrojos e índice UV	panel solar, baterías 18650, cargador bq24074, pasarela LoRa	plataforma SIoT Lysis	suficientes por sí solos.
Smart Wireless Sensor Networks with Virtual Sensors for Forest Fire Evolution Prediction Using Machine Learning - (Haque & Soliman, 2025)	2025	Perceptrón Multicapa, regresión polinómica para sensores virtuales	Temperatura, humedad, intensidad de luz, humo, sonido	Red de sensores inalámbricos (WSN) con 15 sensores físicos, sensores virtuales emulados	Red de sensores inalámbricos inteligente y segura (SSWSN), estructura de supervectores híbridos (HSPV)	Rendimiento inconsistente de sensores virtuales en categorías específicas; el proceso de generación virtual requiere optimización.
Towards Smart Wildfire Prevention: Development of a LoRa-Based IoT Node for Environmental Hazard Detection - (Pires <i>et al.</i> , 2025)	2025	Lógica basada en reglas (regla 30-30-30)	Temperatura, humedad, monóxido de carbono (CO), calidad del aire (PM2.5 y PM10), velocidad del viento	Arduino Mega, módulo LoRa RFM95, sensor DHT11, sensor SDS011 (PM), sensor MQ7 (CO), anemómetro, Raspberry Pi 3, pantalla TFT, panel solar, batería LiPo	LoRa (868 MHz), protocolo MQTT, Node-RED (panel de control)	Desviación y ruido en sensores económicos; materiales de carcasa (madera) no aptos para despliegue permanente; falta medir dirección del viento.
Optimized Wireless Sensor Network Architecture for AI-Based Wildfire Detection in Remote Areas - (Almarri <i>et al.</i> , 2025)	2025	Regresión logística multiclasa	Temperatura, humedad, humo	Nodos Arduino Uno, sensores DHT11, módulo XBee ZigBee, módulo GSM SIM900D	ZigBee (IEEE 802.15.4), GSM (vía SMS para alertas remotas)	La tasa de pérdida de paquetes se incrementa significativamente con el diámetro de la red y la complejidad del terreno.
Intelligent Fire Detection in Agriculture Using Machine Learning and Embedded Systems for Risk Prevention and Improved Sustainability - (Morchid <i>et al.</i> , 2026)	2026	Random Forest, Regresión Lineal, Regresión Logística	Nivel de llama, nivel de humo	Raspberry Pi 3 B+, sensores de humo (MQ-135) y llama, convertidor ADC MCP3208, pantalla LCD, zumbador audible	Funcionamiento autónomo, alertas visuales por LCD y sonoras por zumbador	Entorno de prueba controlado que no replica variaciones climáticas extremas; sensibilidad a falsas alarmas por polvo o niebla.
Holistic IoT and Cloud-based Telemetry Architecture for Proactive Fire Monitoring in Smart Agriculture - (Morchid <i>et al.</i> , 2026)	2026	Algoritmo de agregación personalizado	Llama, humo, uso de CPU y memoria del sistema	Raspberry Pi 3 Model B+, sensores de llama y humo, convertidor ADC MCP3208, pasarela de red	Protocolo MQTT (formato JSON), plataforma en la nube ThingsBoard, alertas por correo electrónico	Alta dependencia de internet en zonas de bajo ancho de banda; el factor humano es crítico (el usuario debe interpretar las alertas).

5. CONCLUSIONES

La revisión de la literatura confirma que integrar el Internet de las Cosas, el Aprendizaje Profundo y las arquitecturas distribuidas transforma la prevención de incendios

forestales en una labor proactiva. La combinación de la computación en el borde con el procesamiento intermedio reduce el consumo de energía y acorta la latencia de respuesta respecto a la nube, lo que permite ajustar la frecuencia de medición de los sensores en tiempo real ante



condiciones de alto riesgo. Para asegurar la comunicación en áreas remotas, la tecnología LoRa proporciona un alcance de hasta 20 kilómetros sin requerir infraestructura previa. Además, los nodos operan de forma autónoma a largo plazo mediante el uso de paneles solares y algoritmos de eficiencia energética. En el ámbito analítico, los modelos de aprendizaje automático optimizados se ejecutan directamente en los dispositivos locales con una precisión superior al 99%. Las redes neuronales GRU lideran el procesamiento intermedio, mientras que el uso de sensores virtuales estima mediciones en zonas sin hardware físico para abaratar los costos de instalación. Estos ecosistemas logran anticipar el fuego antes de que exista humo visible mediante el monitoreo de variables químicas, como el monóxido de carbono, y aplican reglas operativas directas, como la norma 30-30-30, para detonar alertas inmediatas. A pesar de este sólido desarrollo tecnológico, el sector aún carece de despliegues masivos y prolongados en escenarios reales, donde el clima y la naturaleza desgastan los equipos y afectan la conectividad. Por consiguiente, los próximos avances deben integrar mecanismos capaces de detectar anomalías y filtrar las falsas alarmas derivadas de fallos físicos, asegurando así la fiabilidad operativa de estos sistemas en el terreno.

REFERENCIAS

- Abdilhussain, S., Mahmmud, B., Alwhelat, A., Shehada, D., & Shihab, Z. (2025). A Comprehensive Review of Sensor Technologies in IoT: Technical Aspects, Challenges, and Future Directions. *Computers*, 14(8), 342 - 354. <https://doi.org/10.3390/computers14080342>
- Academia Nacional de Bomberos de Chile. (2020). Origen y comportamiento del fuego. In M. Alfaro-Bermúdez, F. Muñoz-Jélvez, R. Romo-Silva, D. Fariás-Salazar, M. Chavarriga-Guerrero, & P. Fregonara-Bermúdez, *Control de Incendios Forestales para Bomberos* (pp. 13 - 25). Academia Nacional de Bomberos de Chile. Retrieved from https://www.bomberos.cl/images/mini_sitio_sno/forestal/Curso%20IIFB.pdf
- Alkhatib, A., & Jaber, K. (2024). FDPA Internet Of Things System for Forest Fire Detection, Prediction and Behaviour Analysis. *IET Wireless Sensor Systems*, 14(3), 56 - 71. <https://doi.org/10.1049/wss2.12076>
- Almarri, S., Safwan, H., Al-Qisoom, S., Gdaim, S., & Zitouni, A. (2025). Optimized Wireless Sensor Network Architecture for AI-Based Wildfire Detection in Remote Areas. *Fire*, 8(7), 245 - 267. <https://doi.org/10.3390/fire8070245>
- Ananthi, J., Sengottaiyan, N., Anbukaruppusamy, S., Upreti, K., & Dubey, A. (2021). Forest Fire Prediction Using IoT and Deep Learning. *International Journal of Advanced Technology and Engineering Exploration*, 9(87), 246 - 256. <https://doi.org/10.19101/IJATEE.2021.87464>
- Anuar, A., Mohamad, R., Morkom, A., & Concepcion II, R. (2023). Real-Time Forest Fire Detection, Monitoring, And Alert System Using Arduino. *Indonesian Journal of Electrical Engineering and Computer Science*, 33(2), 942 - 950. <https://doi.org/10.11591/ijeecs.v33.i2.pp942-950>
- Benzekri, W., Moussati, A., Moussaoui, O., & Berrajaa, M. (2020). Early Forest Fire Detection System using Wireless Sensor Network and Deep Learning. *International Journal of Advanced Computer Science and Applications (IJACSA)*, 11(5), 1 - 11. <https://doi.org/10.14569/IJACSA.2020.0110564>
- Brereton, P., Kitchenham, B., Budgen, D., Turner, M., & Khalil, M. (2007). Lessons from Applying the Systematic Literature Review Process within the Software Engineering Domain. *Journal of Systems and Software*, 80(4), 571 - 583. <https://doi.org/10.1016/j.jss.2006.07.009>
- Brito, T., Pereira, A., Lima, J., & Valente, A. (2020). Wireless Sensor Network for Ignitions Detection: An IoT Approach. *Electronics*, 9(6), 893 - 906. <https://doi.org/10.3390/electronics9060893>
- Carrasco-Ruano, Y., & Núñez-Moreno, M. (2019). Los Incendios Forestales y su Impacto Ambiental en las Reservas Ecológicas. *Explorador Digital*, 3(3), 104 - 113. <https://doi.org/10.33262/exploradordigital.v2i1.324>
- Center for Disaster Management and Risk Reduction Technology. (2026). Wildfires January 2026: South-Central Chile. *Risk Layer*, 1(1), 1 - 12. <https://doi.org/10.5445/IR/1000190174>
- Comisión Nacional Forestal. (2010). *Incendios Forestales: Guía Práctica para Comunicadores* (Primera ed.). Comisión Nacional Forestal. Retrieved from <http://www.conafor.gob.mx:8080/documentos/docs/10/236Gu%C3%ADa%20pr%C3%A1ctica%20para%20comunicadores%20-%20Incendios%20Forestales.pdf>
- Corporación Nacional Forestal. (2012). ¿Qué es un incendio forestal? In C. N. Forestal, *Manual con Medidas para la Prevención de Incendios Forestales* (pp. 9 - 17). Corporación Nacional Forestal. Retrieved from https://avisoenlinea.conaf.cl/documentos/doc_tecnicos/Manual_OHiggins.pdf
- Dampage, U., Bandaranayake, L., Wanasinghe, R., Kottahachchi, K., & Jayasanka, B. (2022). Forest Fire Detection System Using Wireless Sensor Networks and Machine Learning. *Scientific Reports*, 12(46), 1 - 12. <https://doi.org/10.1038/s41598-021-03882-9>



- Dauda, A., Flauzac, O., & Nolot, F. (2024). A Survey on IoT Application Architectures. *Sensors*, 24(16), 5320 - 5332. <https://doi.org/10.3390/s24165320>
- Haque, A., & Soliman, H. (2025). Smart Wireless Sensor Networks with Virtual Sensors for Forest Fire Evolution Prediction Using Machine Learning. *Electronics*, 14(2), 223 - 236. <https://doi.org/10.3390/electronics14020223>
- Khushal, A., Kishore, A., Tejas, N., Mohana, M., & Badarinath, K. (2024). Real Time Smart Forest Fire Detection and Monitoring System Using IoT and Machine Learning. *2024 8th International Conference on Computational System and Information Technology for Sustainable Solutions (CSITSS)*, 8(1), 1 - 5. <https://doi.org/10.1109/CSITSS64042.2024.10816855>
- Ministerio de Ambiente de Panamá. (2015). *Metodología de Investigación de Causas que Provocan los Incendios Forestales* (Primera ed.). Metodología de Investigación de Causas que Provocan los Incendios Forestales. Retrieved from https://www.itto.int/files/itto_project_db_input/2902/technical/INVESTIGACION%20DE%20CAUSAS%20DE%20OCURRENCIA%20DE%20INCENDIOS%20FORESTALES.pdf?v=1709399164
- Morchid, A., Elbasri, A., Qjidaa, H., El-Alami, R., Bossoufi, B., & Mobayen, S. (2026). Intelligent Fire Detection in Agriculture Using Machine Learning and Embedded Systems for Risk Prevention and Improved Sustainability. *Scientific Reports*, 16(1), 1 - 18. <https://doi.org/10.1038/s41598-026-40378-w>
- Morchid, A., Oughannou, Z., El-Alami, R., Qjidaa, H., Jamil, M., & Khalid, H. (2024). Integrated Internet of Things (IoT) Solutions for Early Fire Detection in Smart Agriculture. *Results in Engineering*, 24(2024), 1 - 8. <https://doi.org/10.1016/j.rineng.2024.103392>
- Morchid, A., Salami, A., Khalid, H., & Said, Z. (2026). Holistic IoT and Cloud-based Telemetry Architecture for Proactive Fire Monitoring in Smart Agriculture. *Scientific Reports*, 16(1), 8669 - 8687. <https://doi.org/10.1038/s41598-026-43538-0>
- Pires, L., Fialho, V., Pécurto, T., & Madeira, A. (2025). Towards Smart Wildfire Prevention: Development of a LoRa-Based IoT Node for Environmental Hazard Detection. *Designs*, 9(4), 91 - 127. <https://doi.org/10.3390/designs9040091>
- Putzu, M., Loru, D., Carta, F., Ledda, A., Chirigu, A., Sole, M., . . . Giusto, D. (2024). Enhancing Wildfire Risk Management Through Sensor-Based AI Integration in Social IoT Frameworks. *Engineering Proceedings*, 78(1), 4 - 15. <https://doi.org/10.3390/engproc2024078004>
- Reddy, R., Anandh, R., Lonhde, G., Vignesh, G., Purohit, P., & Rao, V. (2024). Detection And Reporting of Wildfires in Forests Using a Wireless Sensor Network. *Proceedings of the 2nd International Conference on Challenges in Information, Communication, and Computing Technology (ICCICCT 2024)*, 1(1), 1 - 6. <https://doi.org/10.1201/9781003559085>
- Roque, G., & Sánchez-Padilla, V. (2020). LPWAN Based IoT Surveillance System for Outdoor Fire Detection. *IEEE Access*, 8(1), 114900 - 114909. <https://doi.org/10.1109/ACCESS.2020.3003848>
- Rupanetti, D., & Kaabouch, N. (2024). Combining Edge Computing-Assisted Internet of Things Security with Artificial Intelligence: Applications, Challenges, and Opportunities. *Applied Sciences*, 14(16), 7104 - 7113. <https://doi.org/10.3390/app14167104>
- Sánchez-Hernández, J., Estrada-Esquivel, H., Martínez-Rebollar, A., & Pech-May, F. (2025). Marco de evaluación para la óptima selección de componentes de hardware para sistemas del Internet de las Cosas. *Computación y Sistemas*, 29(3), 1391 - 1404. <https://doi.org/10.13053/CyS-29-3-5901>
- Sharobiddinov, D., Siddiqui, H., Saleem, A., Méndez, G., Ramírez, D., & De la Torre, I. (2025). Edge-Based Autonomous Fire and Smoke Detection Using MobileNetV2. *Sensors*, 25(20), 6419 - 6433. <https://doi.org/10.3390/s25206419>
- Shi, W., Cao, J., Zhang, Q., Li, Y., & Xu, L. (2016). Edge Computing: Vision and Challenges. *IEEE Internet of Things Journal*, 3(5), 637 - 646. <https://doi.org/10.1109/JIOT.2016.2579198>
- Siam, S., Ahn, H., Liu, L., Alam, S., Shen, H., Cao, Z., & Shroff, N. (2024). Artificial Intelligence of Things: A Survey. *ACM Transactions on Sensor Networks*, 21(1), 1 - 75. <https://doi.org/10.1145/3690639>
- Tsipis, A., Papamichail, A., Angelis, I., Koufoudakis, G., Tsoumanis, G., & Oikonomou, K. (2020). An Alertness-Adjustable Cloud/Fog IoT Solution for Timely Environmental Monitoring Based on Wildfire Risk Forecasting. *Energies*, 13(14), 3693 - 3705. <https://doi.org/10.3390/en13143693>
- United Nations Environment Programme. (2022). *UN Environment Programme*. Retrieved from Spreading like Wildfire: The Rising Threat of



Extraordinary Landscape Fires:
<https://www.unep.org/resources/report/spreading-wildfire-rising-threat-extraordinary-landscape-fires>

Uremek, I., Leahy, P., & Popovici, E. (2024). A System for Efficient Detection of Forest Fires through Low Power Environmental Data Monitoring and AI. *Engineering Proceedings*, 68(1), 38 - 49. <https://doi.org/10.3390/engproc2024068038>

Verma, S., Kaur, S., Rawat, D., Xi, C., Alex, L., & Jhanji, N. (2021). Intelligent Framework Using IoT-Based WSNs for Wildfire Detection. *IEEE Access*, 9(1), 48185 - 48196. <https://doi.org/10.1109/ACCESS.2021.3060549>

Zope, V., Dadlani, T., Matai, A., Tembhurnikar, P., & Kalani, R. (2020). IoT Sensor and Deep Neural Network based Wildfire Prediction System. *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 1(1), 205 - 205. <https://doi.org/10.1109/ICICCS48265.2020.9120949>

Artificial Intelligence in Companies: PRISMA Bibliometric Review in Scopus and Web of Science

de la Serna-Tuya, Andrés Saúl*, Jasso-Romero, Evangelina**, B.Flores, Angélica***, Gallardo-Martínez, Yurixhi****, Sánchez-Román, Guillermina*****,

* Universidad Popular Autónoma del Estado de Puebla, 17 Sur 901, Colonia Santiago, C.P. 72160, Puebla, MX. (Tel: 222 229 94 00; e-mail: asdelaserna@gmail.com ORCID: <https://orcid.org/0000-0002-0409-7685>)

**Centro Universitario de la Costa Sur/ Universidad de Guadalajara, Centro Universitario de la Costa Sur, Av. Independencia Nacional 151, Centro, C.P. 48900 Autlán de Navarro, Jalisco, MX. (e-mail: evangelina.jasso@academicos.udg.mx ORCID: <https://orcid.org/0000-0003-2079-7486>).

*** Universidad de las Américas Puebla, Ex-Hacienda Santa Catarina Mártir S/N C.P. 72810 San Andrés Cholula, Puebla, MX. (e-mail: angelica.flores@udlap.mx ORCID: <https://orcid.org/0000-0003-1843-6497>)

**** Instituto de Humanidades, Universidad Panamericana, campus Guadalajara. Calzada Álvaro del Portillo No. 49, Ciudad Granja, C.P. 45010 Zapopan, Jalisco, MX. (e-mail: ygallard@up.edu.mx ORCID: <https://orcid.org/0000-0003-1365-0179>)

***** Facultad de Ciencias de la Computación, Benemérita universidad Autónoma de Puebla, Av. San Claudio y 14 Sur, Colonia San Manuel, Edificio CCO4-211, Ciudad Universitaria, C.P. 72570, Puebla, Puebla, MX. (e-mail: guille.sroman@correo.buap.mx ORCID: <https://orcid.org/0000-0002-8819-4019>) }

Abstract: Artificial intelligence (AI) currently plays a relevant role in companies. This research aims to examine the scientific production related to AI in this field, through a bibliometric review in the Scopus and Web of Science (WoS) databases. Methodology: A bibliometric analysis was conducted following the PRISMA protocol, analyzing 379 documents published in journals indexed in the Scopus and WoS databases. Results: A total of 1,141 documents were identified in the Scopus database and 700 in WoS. India and the United States stand out as the leading countries in terms of scientific production in this field. The journals with the highest number of publications in Scopus are Applied Sciences Switzerland, IEEE Transactions on Engineering Management, and IEEE Access, and in WoS are Waste Management and ACM Transactions on Interactive Intelligent Systems. Most of the studies are concentrated in the area of computer science. Regarding institutional affiliation, the authors come mainly from institutions in Switzerland, the United States, and the Netherlands. Conclusions: Further research is needed on AI and its implications for businesses, given the relevance of the SDGs, which affect the business sector, workers, their training, and interdisciplinary innovation and improving technological capacity.

Keywords: Artificial Intelligence (AI); Bibliometric Analysis; Scopus; Web of Science; Professional Development.

1. INTRODUCCIÓN

El uso de la inteligencia artificial (IA) en las empresas se ha convertido en una tendencia creciente en los últimos años, impulsada por la capacidad de estas tecnologías para transformar y optimizar diversos procesos organizacionales.

Por medio del uso de la IA, en los diferentes espacios empresariales, se pueden reducir costos y mejorar la toma de decisiones.

La IA utiliza programación avanzada, y por ello implica una revolución en sectores, como la atención al cliente, personalización de servicios (Sneider-Calle García et al.,

2024), lo que permite a las empresas ser mas competitivas a nivel local y globalmente (Davenport y Ronanki, 2018).

Claro esta que existen obstáculos dentro de las organizaciones como por ejemplo la cualificación del personal y reestructuración del mismo esto genera situaciones éticas (Fjeld et al., 2020; UNESCO, 2023). Sus aplicaciones clave incluyen automatización vía chatbots y robots de software, elevando la eficiencia empresarial (Chen et al., 2012; Waller y Fawcett, 2013).

Esta investigación se sustenta en la revisión de la producción científica sobre la IA en el mundo empresarial, se aplica la metodología PRISMA en bases de datos Scopus y WoS.



2. METODOLOGIA

El diseño metodológico de esta investigación adoptó un enfoque descriptivo retrospectivo (Corona-Martínez y Fonseca-Hernández, 2021). Para el análisis bibliométrico, se seleccionaron publicaciones científicas sobre aplicaciones de IA en empresas, utilizando las bases Scopus y WoS (Baas et al., 2020; Chaparro-Martínez et al., 2016).

Siguiendo el protocolo PRISMA, se estructuró rigurosa y transparentemente la identificación, evaluación y selección de documentos (Carrera-Rivera et al., 2022; Page, McKenzie, et al., 2021).

El proceso de selección y depuración de documentos lo que permitió garantizar la transparencia, trazabilidad y reproducibilidad del procedimiento (Page, Moher, et al., 2021; PRISMA, 2024). Se implementaron las diez etapas del protocolo en cuatro bloques:

Diseño del protocolo de revisión

1. Determinación del propósito de la revisión
2. Define criterios de elegibilidad
3. Determina fuentes de información
4. Define estrategia de búsqueda

Búsqueda de literatura y Selección de Estudios

5. Identificación
6. Cribado
7. Elegibilidad

Extracción de datos

8. Lectura en profundidad

Análisis de datos

9. Análisis e interpretación
10. Resultados

Durante estas fases, se registraron el número de documentos encontrados, los motivos de exclusión y los textos finalmente seleccionados para el análisis.

2.1 Recopilación de datos

En paralelo a la descripción de las etapas se va a describir los resultados.

Etapas 1: Determinación del propósito de la revisión

La fase preliminar de la revisión de la literatura consistió en definir su objetivo principal, el cual se centró en identificar el uso de las IAs en las empresas. Para guiar esta revisión, se formuló una pregunta de investigación: "¿Cuál es la producción científica sobre inteligencia artificial en los procesos de gestión organizacional de las compañías, según las bases de datos Scopus y WoS?", considerando el nivel de capacitación del personal, la eficiencia, innovación, productividad, procesos, procesos administrativos, toma de decisiones.

Etapas 2: Define criterios de elegibilidad y **Etapas 3:** Determina fuentes de información

El siguiente paso fue seleccionar las fuentes de información adecuadas, la estrategia de búsqueda contempló la inclusión de

artículos publicados y debidamente indexados en dos bases de datos científicas especializadas Scopus y WoS, las dos son bases de datos de publicaciones científicas caracterizada por su amplia cobertura temática y reconocida por la rigurosidad de sus procesos editoriales. Estas bases de datos se distinguen por su extensa cobertura temática y por albergar un número considerable de revistas con alto factor de impacto. Y como ejemplo Scopus se destaca por ofrecer herramientas analíticas que resultan especialmente valiosas durante las fases iniciales del proceso de revisión.

Etapas 4: Define estrategia de búsqueda

Con el objetivo de dar respuesta a la pregunta de investigación, se construyó una cadena de palabras claves utilizando los operadores booleanos AND y OR, lo que permitió formular la correspondiente ecuación de búsqueda ("artificial intelligence" OR AI) AND ("companies" OR "firms" OR "organizations" OR "business enterprises") AND ("management" OR "administration" OR "organizational management") AND ("training" OR "professional development" OR "staff development" OR "employee training").

Para la organización de toda la información y la elaboración del documento se utilizaron los siguientes software, Microsoft Word, Microsoft Excel, Mendeley, VOSviewer y bases de datos, en la que se incluyó los siguientes datos, Nombres completos del autor, Título, Año, Título de la fuente, Volumen, Número, Núm. de art., Inicio de página, Fin de página, Número de páginas, Citado por, DOI, Enlace, Afiliaciones, Autores con afiliaciones, Índice Palabras clave, Dirección de correspondencia, Editores, Editorial, ISSN, ISBN, CODEN, PubMed ID, Idioma del documento original, Título abreviado de la fuente, Tipo de documento, Etapa de publicación, Acceso abierto, Fuente, EID.

Etapas 5: Búsqueda de literatura y selección (Identificación).

Esta etapa 5, junto con la 6 y 7 forman un mismo bloque de Búsqueda de literatura y selección, por medio de dos procesos característicos del método PRISMA: identificación, cribado y elegibilidad.

En esta etapa inicial, se realiza la búsqueda exhaustiva de estudios relevantes en diversas fuentes de información, como bases de datos científicas. El objetivo es reunir el mayor número posible de documentos relacionados con la temática de investigación, sin aplicar aún criterios de inclusión o exclusión estrictos. Se registran todos los resultados encontrados y se eliminan los duplicados para evitar sesgos.

En primavera del 2025 se llevó a cabo una búsqueda utilizando una estrategia específica que permitió recuperar 1141 documentos para Scopus y 700 para WoS. La primera fase de exclusión se contempló respecto a la temporalidad de los estudios; es decir, solo aquellos entre el 2015 y 2024. Se redujeron los datos a 905 documentos para Scopus y 570 para WoS. Los sistemas de bases de datos muestran las mismas tendencias al ser una temática relativamente nueva se ve un crecimiento exponencial en desde el 2018-2019.

No obstante, la cantidad de publicaciones anuales se mantiene relativamente reducida hasta el año 2021, donde se multiplica



entre 2015 al 2024 por 16.2 en Scopus y 20.3 en WoS, siendo el promedio de crecimiento de las dos bases de datos de 18.1, estos resultados indican que, es un tema en de tendencia y crecimiento y claro es un campo por explorar.

Seguido, se delimitaron las áreas temáticas, Informática, Ingeniería, Negocios, Administración y Contabilidad, Ciencias Sociales, Ciencias Ambientales, Multidisciplinar. Luego se delimitaron solo aquellos artículos en fase de publicación final. Por otro lado, se excluyeron aquellos con metadatos incompletos (DOI y que no haya pasado por revisión de pares), y los duplicados de las bases de datos eliminando 9 documentos, lo que resultó en un total de 1,475 artículos incluidos entre las dos bases de datos.

En el contexto global actual, IA ha emergido como una de las tecnologías más disruptivas e influyentes en los distintos ámbitos de la sociedad, siendo la educación uno de los sectores donde su impacto ha sido más evidente.

La Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura (UNESCO) (UNESCO, 2022; Vázquez-Pita, 2022) hace hincapié en remarcar que se necesitan espacios para desarrollar el uso de las Tecnologías de la Información y la Comunicación (TIC).

Desde hace años las competencias digitales (CD) definidas como la capacidad ética para emplear tecnologías de comunicación y participar en la sociedad del conocimiento (Jasso-Romero y De La Serna-Tuya, 2018; Junta de Andalucía, 2024) son promovidas por los diferentes organismos locales e internacionales (Centro de Investigaciones Comunes de la Comisión Europea, 2022).

Diferentes autores empatizan en ventajas y dificultades a la hora de su uso:

- Uso interdisciplinario por el personal (Vázquez-Pita, 2022).
- Regulación y control ético (Stahl et al., 2023).
- Transformación de la educación por medio del aprendizaje personalizado (Birenbaum et al., 2015).
- Creación de contenidos (Chan y Hu, 2023).
- Ámbito legal (Castro-Granados & Artavia-Díaz, 2020).

Al ser un fenómeno que afecta a múltiples áreas, como tal es un fenómeno social, pero también de transformación empresarial.

Etapas 6: Búsqueda de literatura y selección (cribado).

En la fase de cribado, se examinan preliminarmente títulos y resúmenes de documentos identificados.

Aquí se aplican los primeros criterios de inclusión y exclusión para descartar aquellos estudios que no guardan relación directa con el objetivo de la revisión. El cribado permite reducir el volumen de documentos, manteniendo únicamente aquellos con potencial relevancia.

Criterios de inclusión

1. Estudios publicados entre 2015 y 2024, para asegurar actualidad en relación con la IA.
2. Publicaciones que analicen el impacto de la IA en la gestión organizacional (eficiencia, innovación, productividad, procesos, procesos administrativos, toma de decisiones, eficiencia).
3. Estudios que consideren el nivel de capacitación o desarrollo del personal como factor de análisis o moderador.
4. Artículos escritos en inglés o español y con acceso al texto completo.
5. Estudios que analicen aplicaciones de IA en contextos empresariales del sector privado o público.

Criterios de exclusión

1. Artículos centrados exclusivamente en desarrollos técnicos o algorítmicos de IA sin aplicación organizacional.
2. Estudios realizados en los sectores no empresariales como (escuelas, universidades), sin vinculación con compañías o gestión empresarial.
3. Publicaciones orientadas a consumidores finales o interacción cliente-IA, sin enfoque en la estructura organizacional.
4. Estudios teóricos sin evidencia empírica o sin revisión sistemática (como ensayos o editoriales).
5. Duplicados o publicaciones que no cuenten con información suficiente para evaluación metodológica.
6. Investigaciones con enfoque exclusivo en ética, sin análisis de gestión o capacitación.
7. Estudios publicados en blogs, revistas no arbitradas o sin revisión por pares.

La búsqueda inicial arrojó un total de 1,475 (Scopus=905; WoS=570). Con el fin de garantizar la adecuación de la muestra para el análisis subsiguiente, se determinó una muestra probabilística representativa conformada por 379 documentos, considerando un nivel de confianza del 95 % y un margen de error del 5%.

La muestra n fue calculada aplicando la fórmula adecuada, en la cual N indica el tamaño total del conjunto original de documentos, Z es el nivel de confianza, E el margen de error y p es la proporción de elementos que poseen las características

Siendo como resultado $n = (1416.59) / (4.64) \approx 304.9$

Se obtuvo una necesidad de muestra de 305 documentos.

Etapas 7: Búsqueda de literatura y selección (Elegibilidad).

Para finalizar con el bloque de Búsqueda de literatura y selección nos encontramos con la tercera parte, Elegibilidad, en la cual, se lleva a cabo una revisión del texto de los estudios que superaron el cribado.

Se aplican de manera rigurosa los criterios de inclusión y exclusión previamente establecidos (como diseño metodológico, población de estudio, idioma, fecha de publicación, etc.) para determinar la elegibilidad final de cada documento.



Como tal el propósito es confirmar que los productos científicos sean los más pertinentes, siendo la pregunta de investigación.

Se obteniendo un total de (n=379) que cumplen los criterios de la muestra.

Etapas 8: Extracción de datos.

La extracción de los datos se utilizó por medio de la revisión de los artículos seleccionados, utilizando una tabla en Excel que facilitó un análisis mixto, combinando técnicas cualitativas y cuantitativas.

La fase cualitativa se categorizó y realizó una codificación abierta por medio de una codificación axial para interrelacionar categorías y conceptos.

En fase cuantitativa se hizo un cálculo de frecuencias de las categorías y términos claves, revelando las tendencias temáticas.

Etapas 9: Análisis de datos.

Para esta última parte se utilizó el modelo IMRaD (Introducción, Métodos, Resultados y Discusión), aplicando un análisis mixto para detectar patrones y hallazgos clave.

Un análisis de frecuencias determinó la distribución y relaciones de categorías conceptuales, incluyendo tipo de documento, país e institución de origen, revistas especializadas, áreas temáticas, autores y publicaciones más citadas.

En la etapa 5 (Búsqueda de literatura y selección) se identificaron estos documentos académicos, en Scopus, se recuperaron principalmente artículos científicos (67.02%), seguidos por capítulos de libro (22.16%) y en menor proporción artículos de revisión (4.75%).

En WoS la mayoría de las publicaciones eran artículos científicos 92.34%, y solo eran de revisión un 5.11% y en Scopus se sigue la misma tendencia.

Por países existe una hegemonía de potencias asiáticas y norteamericanas, en Scopus la India lidera (34.75%), seguida de EE.UU. (22.55%) y China (16.98%); en WoS, EE.UU. prevalece (30%), con China e India como segunda y tercera, reflejando hegemonía de potencias asiáticas y norteamericanas, otros países con menor producción son Reino Unido, Italia y Alemania.

El análisis institucional revela que la Symbiosis International (Deemed University) de India lidera la producción en la base de datos Scopus con 9 publicaciones. Le siguen tres instituciones con 6 documentos cada una: Georgia Institute of Technology (Estados Unidos), Graphic Era Deemed to be University (India) y SR University (India), lo que confirma una importante presencia de instituciones indias en esta línea de investigación.

Entre las instituciones con cinco o más publicaciones destacan The University of Sheffield, Christ University y Chinese Academy of Sciences, evidenciando el interés de

universidades de Asia y Europa por el desarrollo de la IA en contextos organizacionales. Asimismo, instituciones como la University of Cambridge aparecen de manera recurrente en ambas bases de datos, reflejando una participación constante en la difusión científica. La presencia predominante de universidades de India, Reino Unido, Estados Unidos y China coincide con la concentración geográfica de la producción académica identificada previamente.

La revista *Applied Sciences Switzerland* destaca en Scopus como la de mayor artículos publicados, seguidos de la revista *IEEE Transactions on Engineering Management e IEEE Access*, ambas clasificadas en Q1.

A nivel de áreas dominantes se destacan en Scopus:

- Computer Science (39.78%),
- Engineering (25.27%) y
- Business, Management and Accounting (16.39%),

Lo que refleja tanto el origen tecnológico de la IA como su expansión hacia aplicaciones organizacionales y estratégicas.

Se puede observar que se da cada vez más un enfoque multidisciplinario.

Como publicación destacada, el artículo más citado en Scopus fue "So what if ChatGPT wrote it?" de Dwivedi et al. (2023), con más de 2000 citas,

Estos hallazgos evidencian que las investigaciones más influyentes combinan enfoques tecnológicos y empresariales, reforzando la importancia estratégica de la IA en las organizaciones modernas.

3. RESULTADOS

Etapas 10: Resultados.

El análisis sistemático de las publicaciones sobre IA en el ámbito empresarial, considerando las bases de datos Scopus y WoS, revela un crecimiento sostenido y multidisciplinario en esta línea de investigación.

La revisión por medio de análisis bibliométrico de publicaciones científicas sobre IA en entornos empresariales las bases de datos en Scopus y WoS, se destaca un crecimiento exponencial y multidisciplinario. Predominan artículos arbitrados (67.02% en Scopus; 92.34% en WoS), con capítulos de libro destacando en Scopus (22.16%), indicando tendencias colaborativas.

Por países existe una hegemonía de potencias asiáticas y norteamericanas, en Scopus la India lidera (34.75%), seguida de EE.UU. (22.55%) y China (16.98%); en WoS, EE.UU. prevalece (30%), con China e India como segunda y tercera, reflejando hegemonía de potencias asiáticas y norteamericanas.

En conjunto, los hallazgos evidencian que la inteligencia artificial en el entorno empresarial se ha consolidado como una línea de investigación estratégica, global e interdisciplinaria, con fuerte presencia de países asiáticos, publicaciones en



revistas indexadas de alta calidad, y abordajes diversos desde lo técnico hasta lo ético y organizacional.

A nivel de áreas dominantes se destacan:

- Ciencias de la Computación,
- Ingeniería,
- Administración y Ciencias Sociales,
- Interdisciplinario.

La IA en entornos empresariales emerge como línea estratégica, global e interdisciplinaria.

4. CONCLUSIONES

Los resultados evidencian un creciente interés académico por estudiar la implementación de la IA en contextos organizacionales, con la India y Estados Unidos destacan como los principales países liderando la producción científica. Las áreas temáticas predominantes, como las ciencias de la computación y la gestión organizacional, revelan una clara orientación hacia la innovación tecnológica aplicada. Este panorama se conecta de manera directa con los Objetivos de Desarrollo Sostenible (ODS), especialmente con el ODS 4 (Educación de calidad), el ODS 8 (Trabajo decente y crecimiento económico) y el ODS 9 (Industria, innovación e infraestructura), al reflejar cómo la IA impulsa procesos de capacitación, desarrollo profesional, y transformación estructural en el tejido empresarial.

La IA afecta a los entornos productivos empresariales y profesionales, pues modifica el impacto, tanto en la gestión y producción, representa un nuevo fenómeno social que reconfigura el quehacer laboral.

La revisión bibliométrica bajo PRISMA analizó 1.141 documentos en Scopus y 700 en WoS sobre IA empresarial, estos criterios rigurosos de selección y búsquedas sistemáticas aseguraron validez y transparencia metodológica.

Resultados revelan interés académico creciente, con India y EE.UU. liderando producción científica; disciplinas clave como ciencias de la computación y gestión organizacional enfatizan innovación aplicada a procesos empresariales.

Finalmente, urge investigación interdisciplinaria para elucidar efectos actuales y prever implicaciones futuras, abordando desafíos en capacitación continua, adaptación organizacional y políticas para desarrollo sostenible, inclusivo y ético.

REFERENCIAS

- Baas, J., Schotten, M., Plume, A., Côté, G., & Karimi, R. (2020). *Scopus as a curated, high-quality bibliometric data source for academic research in quantitative science studies*. https://doi.org/10.1162/qss_a_00019
- Birenbaum, M., DeLuca, C., Earl, L., Heritage, M., Klenowski, V., Looney, A., Smith, K., Timperley, H., Volante, L., & Wyatt-Smith, C. (2015). International trends in the implementation of assessment for learning: Implications for policy and practice. *Policy Futures in Education*, 13(1), 117-140. <https://doi.org/10.1177/1478210314566733>
- Carrera-Rivera, A., Ochoa, W., Larrinaga, F., & Lasa, G. (2022). How-to conduct a systematic literature review: A quick guide for computer science research. *MethodsX*, 9, 101895. <https://doi.org/10.1016/J.MEX.2022.101895>
- Castro-Granados, A., & Artavia-Díaz, K. Y. (2020). Competencias digitales docentes: un acercamiento inicial. *Revista Electrónica Calidad en la Educación Superior*, 11(1), 47-80. <https://doi.org/10.22458/caes.v11i1.2932>
- Centro de Investigaciones Comunes de la Comisión Europea. (2022). *DigComp 2.2 Marco de Competencias Digitales para la Ciudadanía – Asociación*. Asociación Somos Digital. <https://somos-digital.org/digcomp/>
- Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education*, 20(1). <https://doi.org/10.1186/s41239-023-00411-8>
- Chaparro-Martínez, E. I., Álvarez-Muñoz, P., & D'Armas-Regnault, M. (2016). *Gestión de la información: uso de las bases de datos Scopus y Web of Science con fines académicos*. https://www.researchgate.net/publication/312384415_Gestion_de_la_informacion_uso_de_las_bases_de_datos_Scopus_y_Web_of_Science_con_fines_academicos
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly: Management Information Systems*, 36(4), 1165-1188. <https://doi.org/10.2307/41703503>
- Corona-Martínez, L. A., & Fonseca-Hernández, M. (2021). *Acerca del carácter retrospectivo o prospectivo en la investigación científica*. <https://www.redalyc.org/journal/1800/180068639021/html/>
- Davenport, T. H., & Ronanki, R. (2018). Artificial intelligence for the real world. *Harvard Business Review Article*. <https://hbr.org/2018/01/artificial-intelligence-for-the-real-world>
- Fjeld, J., Achten, N., Hilligoss, H., Nagy, A. C., & Srikumar, M. (2020). *Principled Artificial Intelligence: Mapping Consensus in Ethical and Rights-based Approaches to Principles for AI*. <https://ssrn.com/abstract=3518482>
- Jasso-Romero, E., & De La Serna-Tuya, A. S. (2018). El impacto de las TIC en el desarrollo de Competencias investigativas: estudio de caso licenciatura en administración del CUCSUR de la UDG. *AMMACI*, 1.
- Junta de Andalucía. (2024). Herramientas digitales basadas en metodologías activas para el ámbito de la Formación Profesional. En *Revista Andalucía Educativa*. <https://www.juntadeandalucia.es/educacion/portals/web>



- /revista-andalucia-educativa/contenidos/-
/contenidos/detalle/herramientas-digitales-basadas-en-
metodologias-activas-para-el-ambito-de-la-formacion-
profesional
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Systematic Reviews*, 10(1), 1-11. <https://doi.org/10.1186/S13643-021-01626-4/FIGURES/1>
- Page, M. J., Moher, D., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... McKenzie, J. E. (2021). PRISMA 2020 explanation and elaboration: Updated guidance and exemplars for reporting systematic reviews. *The BMJ*, 372. <https://doi.org/10.1136/BMJ.N160>
- PRISMA. (2024). *PRISMA statement*. <https://www.prisma-statement.org/>
- Sneider-Calle García, J., Sotaminga-Andi, A. S., Garay-Arias, G. N., & Villavicencio-Tuarez, R. R. (2024). INTELLIGENCE AND ITS CONTRIBUTION TO INNOVATION IN COMPANIES. *Ciencia y Desarrollo. Universidad Alas Peruanas*. <http://revistas.uap.edu.pe/ojs/index.php/CYD/index>
- Stahl, B. C., Antoniou, J., Bhalla, N., Brooks, L., Jansen, P., Lindqvist, B., Kirichenko, A., Marchal, S., Rodrigues, R., Santiago, N., Warsø, Z., & Wright, D. (2023). A systematic review of artificial intelligence impact assessments. En *Artificial Intelligence Review* (Vol. 56, Número 11). Springer Netherlands. <https://doi.org/10.1007/s10462-023-10420-8>
- UNESCO. (2022). *Aplicación de la recomendación sobre la ética de la inteligencia artificial*.
- UNESCO. (2023). *Ethical impact assessment. A tool of the Recommendation on the Ethics of Artificial Intelligence*. <https://doi.org/10.54678/YTSA7796>
- Vázquez-Pita, E. (2022). La UNESCO y la gobernanza de la inteligencia artificial en un mundo globalizado. La necesidad de una nueva arquitectura legal. *Anuario de la Facultad de Derecho. Universidad de Extremadura*, (37), 273-302. <https://doi.org/10.17398/2695-7728.37.273>
- Waller, M. A., & Fawcett, S. E. (2013). Data Science, Predictive Analytics, and Big Data: A Revolution That

Will Transform Supply Chain Design and Management. *Journal of Business Logistics*, 34(2), 77-84. <https://doi.org/10.1111/JBL.12010>

Energy efficiency of power supply cables for 1 MW temporary generators under the NOM-001-SEDE standard and international standards and thermo-electrical optimization under real operating conditions

Ruiz Flores Luis Ivan, Herrera Gutierrez Hugo

Facultad de Ciencias Químicas e Ingeniería, Universidad Autónoma del Estado de Morelos, Av. Universidad 1001, Colonia Chamilpa, C.P. 62209, Cuernavaca, Morelos, México. (Tel: +52.7772884896; e-mail: luis.ruizf@fcqi.uaem.mx; hugo.herrera1@fcqi.uaem.mx).

Abstract: This article develops a comprehensive methodology for the sizing, validation and selection of electric feeders in temporary generation systems of the order of 1 MW, considering real operating conditions in industrial power systems. From the technical analysis of a calculation memory applied to a 480 V three-phase generator, a systematic approach is built that integrates regulatory, electrical and thermal criteria to guarantee the safety, reliability and regulatory compliance of the system. The study addresses in a structured way the fundamental elements of electrical design: calculation of nominal current under continuous load conditions, application of demand factors in accordance with regulations, selection of conductors by effective ampacity, evaluation of voltage drop, thermal analysis in the event of short circuit events and coordination of electrical protections. Likewise, the channeling analysis is incorporated by means of ventilated cable trays, considering thermal dissipation, physical occupancy and geometric distribution of conductors in parallel. The Mexican standard NOM-001-SEDE is used as the main reference, complemented by international standards such as NEC (NFPA 70), IEEE Std 141, IEEE Std 399 and IEC recommendations for the design of power systems. The article highlights the importance of considering real installation conditions, such as ambient temperature, physical arrangement of conductors and insulation characteristics, factors that directly influence the effective ampacity of the system. A relevant contribution of this work is the formalization of a technical guide for calculation memory applicable to temporary generators, particularly in backup scenarios, construction, events or critical systems. This guide establishes reproducible criteria that allow the validation of installations before verification bodies such as UVIE, minimizing electrical risks and ensuring continuous performance. The results show that a correct selection of conductors in parallel, combined with a comprehensive analysis of protections and channeling, allows simultaneous compliance with the criteria of ampacity, voltage drop (<2%) and interruptive capacity, guaranteeing safe operation even under fault conditions. Finally, the article positions the design of feeders for temporary generation as a multidisciplinary problem that requires integration of electrical theory, regulations and engineering practice, proposing a solid basis for future research and applications in modern power systems.

Keywords: voltage drop, sizing, efficiency, protections, power systems.

1. INTRODUCCIÓN

En la actualidad, se presentan diversos casos de metodología integral para el diseño, validación y justificación de ingeniería de los alimentadores para generadores temporales con una potencia de 1 MW. Con base en una memoria de cálculo real, la metodología integra modelado eléctrico, análisis térmico y cumplimiento normativo bajo las normas NOM-001-SEDE, NEC e IEEE. El estudio y trabajo propone una guía de ingeniería estructurada para el cálculo del desarrollo de memorias, que apoya la validación por parte de organismos de inspección y permite la reproducibilidad en aplicaciones industriales de acuerdo con el modelo de red de la Figura 1 promoviendo una eficiencia energética de cables alimentadores en generadores de 1 MW similares (NFPA, 2023; IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012).

Para fines de publicación técnica, este tipo de metodología adquiere valor cuando la memoria de cálculo no se limita a presentar resultados finales, sino que muestra la trazabilidad

entre datos de placa, criterios normativos, verificación térmica y decisión de canalización, permitiendo auditoría y repetibilidad del diseño (IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012).

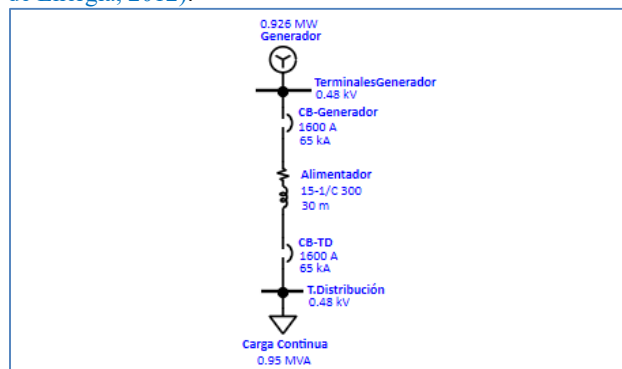


Figura 1. Modelo de red conceptual para obtener eficiencia energética en cable alimentador de un generador de 1 MW.



La integración de los resultados de la simulación, incluyendo el análisis de flujo de potencia y cortocircuitos, confirma la fiabilidad y robustez del sistema bajo condiciones transitorias y de estado estable para obtener una eficiencia energética en el desempeño del cable y del generador eléctrico (IEEE, 1993; IEEE, 1997).

El artículo menciona un análisis explícito de eficiencia termoeléctrica, estimación de pérdidas de conductores, validación térmica por cortocircuito, aclaración de metodología de software, justificación normativa actualizada y criterios comparativos de ingeniería para la selección de alimentadores en condiciones reales de funcionamiento.

1.1 Contexto y Problemática

El crecimiento de aplicaciones industriales que requieren generación temporal —tales como respaldo crítico, eventos de alta demanda, obras de infraestructura y contingencias energéticas— ha incrementado la necesidad de metodologías robustas para el diseño de alimentadores eléctricos confiables. En este contexto, la correcta selección de conductores eléctricos representa un factor determinante para garantizar la continuidad operativa, la seguridad del sistema y el cumplimiento normativo (NFPA, 2023; IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012).

Los generadores de aproximadamente 1 MW constituyen un rango crítico en sistemas de potencia, debido a que operan en niveles de corriente elevados (superiores a 1 kA), lo cual implica desafíos significativos en términos de disipación térmica, coordinación de protecciones y manejo de corrientes de cortocircuito. La ingeniería asociada a estos sistemas requiere una integración rigurosa de criterios eléctricos, mecánicos y normativos (IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012; IEC, 2023).

1.2 Objetivo

Desarrollar una metodología integral para el dimensionamiento de alimentadores eléctricos en sistemas de generación temporal de 1 MW, garantizando cumplimiento normativo, seguridad operativa y confiabilidad (Secretaría de Energía, 2012).

1.3 Glosario.

Término	Definición
Ampacidad	Capacidad máxima de conducción sin sobrecalentamiento
Generador temporal	Fuente eléctrica móvil o provisional
Conductores en paralelo	Uso de múltiples conductores para una fase
UVIE	Unidad de Verificación de Instalaciones Eléctricas

1.4 Marco teórico.

El comportamiento térmico del conductor está gobernado por: (Smith et al., 2020; IEC, 2023)

$$I^2R = Q$$

Este principio establece que la generación de calor es proporcional al cuadrado de la corriente, lo que hace crítico el

análisis en sistemas de alta potencia (Smith et al., 2020; IEC, 2023). Desde la teoría de cables, la ampacidad está asociada al equilibrio térmico entre pérdidas por efecto Joule y disipación hacia el medio circundante; por ello, la instalación en charola, la temperatura ambiente y el número de conductores por capa modifican directamente la capacidad real de conducción (Smith et al., 2020).

1.5 Normatividad internacional

- NFPA 70 (NEC)
- IEEE Std 141 (*Red Book*)
- IEEE Std 399 (*Brown Book*)
- IEC 60364
- IEC 60287
- IEC 60949

Estas normas proporcionan criterios para diseño seguro, análisis de fallas y selección de conductores. (NFPA, 2023; IEEE, 1993; IEEE, 1997; IEC, 2019; Smith et al., 2020; IEC, 2023). Asimismo, para la evaluación de ampacidad en estado estacionario y para la verificación térmica de conductores sometidos a corrientes de cortocircuito, ya que ambos documentos fortalecen el sustento analítico del artículo (Smith et al., 2020; IEC, 1988).

1.6 Normatividad nacional

NOM-001-SEDE (Artículos: 110, 215, 240, 250, 300, 310, 392, 445) (Secretaría de Energía, 2012). Define: i) ampacidad, ii) canalización, iii) puesta a tierra y iv) protección.

1.7 Guía para memoria de cálculo de cables en generadores de 1 MW.

A continuación, se seguirán los pasos recomendados para validar y verificar la memoria de cálculo del cable para conectar 1 generador de 1 MW (Secretaría de Energía, 2012; Smith et al., 2020; IEC, 2023)

- 1) Definir carga y corriente
- 2) Aplicar factor 125%
- 3) Seleccionar conductores por ampacidad
- 4) Verificar caída de tensión (<2%)
- 5) Validar cortocircuito
- 6) Seleccionar protección
- 7) Diseñar canalización
- 8) Validar con software

2. DESARROLLO DEL ARTÍCULO

2.1 Datos de diseño del generador

Para fines de diseño y del ejercicio con el artículo técnico contamos con un generador que funciona a 950 kVA y 480 V. La corriente se calcula usando ecuaciones de potencia trifásicas, resultando en aproximadamente 1,143 A. Aplicando el criterio de carga continua del 125 por ciento, la ampacidad requerida supera los 1400 A (Secretaría de Energía, 2012; Smith et al., 2020). En la Tabla 1 se muestran los datos del modelo de red de la Figura 1.



El uso de la potencia continua de operación, en lugar de emplear únicamente la potencia máxima de placa, evita subestimar la exigencia térmica del alimentador cuando el generador opera durante periodos prolongados o en servicio base temporal (IEEE, 1993; Secretaría de Energía, 2012).

Usando la potencia de operación continua, se calcula la correspondiente corriente usando la fórmula:

$$I = \frac{S}{\sqrt{3} \times V}$$

Donde S es la potencia aparente en kVA y V es el voltaje nominal de línea a línea en kV según (Secretaría de Energía, 2012).

$$I_{cont} = \frac{950 \text{ kVA}}{\sqrt{3} \times 0.48 \text{ kV}} = 1,143 \text{ A}$$

La corriente de operación continua del generador es de (según (Secretaría de Energía, 2012)) $I_{cont} = 1,143 \text{ A}$

Tabla 1. Datos de diseño para dimensionar el cable y corriente que alimentará un generador de 1 MW

Parámetro	Valor
Potencia nominal	1158 kVA
Potencia de operación continua	950 kVA
Tensión nominal	480 V _{L-L}
Sistema	Trifásico
FP	0.8
Distancia del circuito	30 m
Temperatura ambiente	30 °C
Reactancia subtransitoria	13.6%

2.2 Factor de demanda (carga continua)

Aplicando el criterio “No menor al de la carga continua”, según (Secretaría de Energía, 2012), la ampacidad requerida en el conductor es de: 125%

$$I_{cond} = 1.25 \times I_{cont}$$
$$I_{cond} = 1.25 \times 1,143 \text{ A} = 1,429 \text{ A}$$

2.3 Selección de conductores por ampacidad

Se considera cable monoconductor de cobre, con aislamiento THW-LS/THHW-LS 90 °C, canalizados en charola portables tipo escalera en una sola capa sin separación (Secretaría de Energía, 2012; Smith et al., 2020).

Tabla 2. Conductor seleccionado.

Calibre	Ampacidad 310-15(b)(17)
152 mm2 (300 kcmil)	289.3 A *

* Se toma en cuenta el 65% de la ampacidad (a 75 °C por el criterio de terminales) de conformidad con 392-80(a)(2) acorde a (Secretaría de Energía, 2012).

Conductores en paralelo:

$$\frac{1,429 \text{ A}}{289.3 \text{ A}} = 4.9 \approx 5$$

Se requieren 5 conductores por fase de calibre 152 mm2 (300 kcmil) acorde a (Secretaría de Energía, 2012; Smith et al., 2020), tal como se muestra en la Tabla 2.

Ampacidad total:

$$289.3 \text{ A} \times 5 = 1,446.5 \text{ A}$$

Nota 1. No se realizan ajustes a la ampacidad por temperatura debido a que se considera una temperatura ambiente de 30 °C y se utiliza la tabla 310-15(b)(17) según (Secretaría de Energía, 2012; Smith et al., 2020).

2.4 Caída de tensión

Para el cálculo de la caída de tensión se utiliza la siguiente fórmula para sistemas trifásicos acorde a (Secretaría de Energía, 2012):

$$\Delta V = \sqrt{3} \times I \times (R \cos \theta + X \sin \theta) \times L$$

De la Tabla 9 de la NOM-001-SEDE, para conductores de calibre 152 mm2 (300 kcmil) de cobre instalados en charola portables tipo escalera (aluminio) se tiene:

$$R = 0.000161 \Omega/\text{m}$$

$$X_L = 0.000135 \Omega/\text{m}$$

Para 5 conductores en paralelo:

$$R_{eq} = \frac{0.000161 \Omega/\text{m}}{5} = 0.0000322 \Omega/\text{m}$$
$$X_L = \frac{0.000135 \Omega/\text{m}}{5} = 0.000027 \Omega/\text{m}$$

Para un factor de potencia de 0.8 se tiene:

$$\cos \theta = 0.8$$

$$\sin \theta = 0.6$$

Considerando una distancia del circuito de 30 m, la caída de tensión calculada es:

$$\Delta V = \sqrt{3} \times 1,143 \text{ A} \times [(0.0000322 \Omega/\text{m} \times 0.8) + (0.000027 \Omega/\text{m} \times 0.6)] \times 30 \text{ m}$$
$$\Delta V = 2.49 \text{ V}$$
$$\%VD = \frac{2.23 \text{ V}}{480 \text{ V}} \times 100 = 0.52\%$$

Cumple NOM (<2% para alimentadores) según (Secretaría de Energía, 2012).

En la Figura 2, se muestra el modelo de red con el análisis en el software comercial ETAP mostrando la caída de tensión.

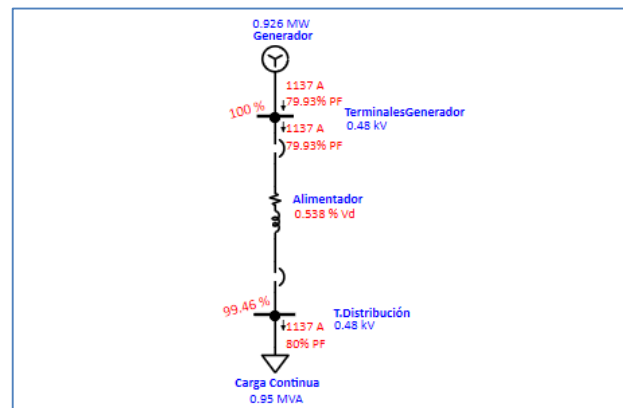


Figura 2. Modelo de red con el análisis de flujos de carga usando software ETAP.

2.5 Cálculo y análisis de corto circuito

La corriente de corto circuito subtransitoria en terminales del generador según (Secretaría de Energía, 2012; IEC, 1988) se calcula a partir de su corriente nominal. En términos prácticos, la verificación debe considerar tanto la magnitud inicial de la falla como el tiempo efectivo de despeje de la protección, dado que la integridad del conductor depende de la energía térmica dejada pasar y no solo del valor instantáneo de corriente acorde a (IEEE, 1997; IEC, 1988).

$$I_{nom} = \frac{S}{\sqrt{3} \times V_n} = \frac{1,158 \text{ kVA}}{\sqrt{3} \times 0.48 \text{ kV}} = 1,393 \text{ A}$$

$$I_{cc} = \frac{I_{nom}}{X''_{p.u.}}$$

$$I_{cc} = \frac{1,393 \text{ A}}{0.136} = 10,243 \text{ A} \approx 10.2 \text{ kA}$$

El tiempo en segundos en que se alcanza el límite térmico de corto circuito para 5 conductores por fase de cobre calibre 152 mm² (300 kcmil) se calcula según (IEC, 1988), usando la siguiente ecuación:

$$t = \frac{0.0297 \log_{10} \left[\frac{(T_2 + 234)}{(T_1 + 234)} \right]}{\frac{I^2}{A^2}}$$

Donde I es la corriente de corto circuito en amperes, A es el área de la sección transversal de los conductores en circular mil, T_1 es la temperatura inicial (nominal) y T_2 la temperatura final del conductor en grados Celsius; para termoplásticos, $T_2 = 150^\circ\text{C}$.

$$t = \frac{0.0297 \log_{10} \left[\frac{(150 + 234)}{(90 + 234)} \right]}{\frac{(10,243)^2}{(1,500,000)^2}} = 47 \text{ s}$$

Los conductores seleccionados soportan ampliamente esta corriente por los tiempos típicos de despeje de las protecciones termomagnéticas (<1 s) (Secretaría de Energía, 2012; IEC, 1988).

En la Figura 3, se muestra el modelo de red con el análisis de corto circuito en el software comercial ETAP.

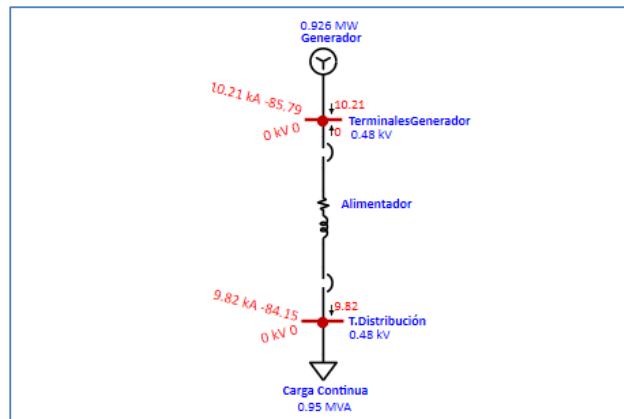


Figura 3. Modelo de red con el análisis de corto circuito usando software.

2.6 Selección de protección

Esta protección es tomada de acuerdo con lo indicado en la ficha técnica de la planta donde se instalará el generador (datos

proporcionados por el cliente del servicio), como se muestra en la Tabla 3.

Tabla 3. Datos de la protección seleccionada según (Secretaría de Energía, 2012).

Parámetro	Valor
Tipo	Termomagnético
Marca	Schneider Electric
Modelo	NW16H14FCC3AUL
Corriente	1,600 A
Ajuste (It)	1,440 A
Capacidad interruptiva	65 kA

La protección seleccionada cumple, al tener un ajuste nominal (1,440 A) mayor que la corriente de operación continua (1,143 A) y menor que la ampacidad del alimentador propuesto (1,446.5 A); además, la capacidad interruptiva es mayor que la corriente de corto circuito calculada en las terminales del generador (IEEE, 1997; Secretaría de Energía, 2012).

En la Figura 4, se muestra la curva de tiempo corriente para la protección del alimentador ante corto circuito.

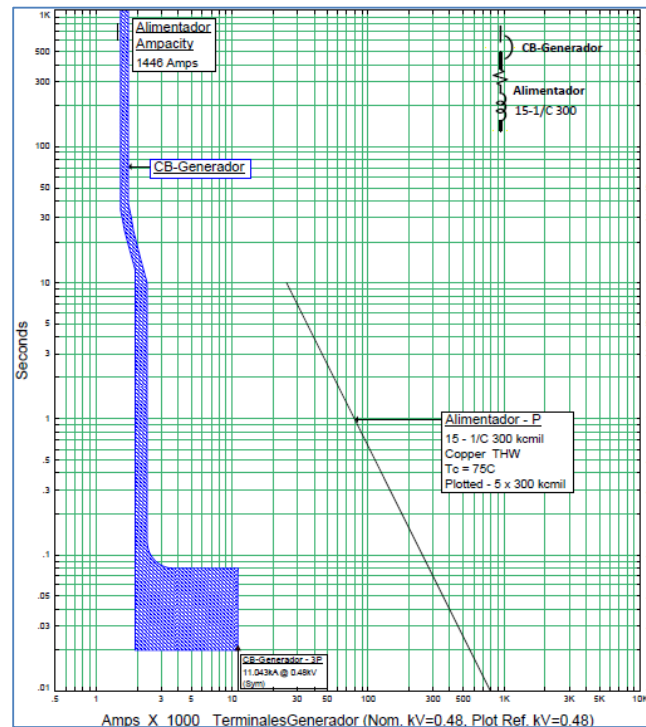


Figura 4. Protección del alimentador ante corto circuito usando software.

2.7 Conductor de puesta a tierra

El conductor de puesta a tierra se selecciona conforme a la Tabla 250-122 de la NOM-001-SEDE; tomando en cuenta que el ajuste nominal de la protección es de 1,440 A, el conductor de puesta a tierra seleccionados es lo que se muestra en la Tabla 4:

Tabla 4. Conductor de puesta a tierra seleccionado según (Secretaría de Energía, 2012).

Valor nominal o ajuste de la Protección	Conductor de Puesta a Tierra requerida
1,440 A	107 mm ² (4/0 AWG) de cobre

2.8 Cálculo y selección de canalización. Ver Tablas 5, 6 y 7.

2.8.1. Tipo de canalización

Se selecciona charola portables tipo escalera (ventilada), conforme al Art. 392 de la NOM-001-SEDE, adecuada para:

- ✓ Sistemas de potencia de alta corriente
- ✓ Conductores en paralelo
- ✓ Requerimientos de disipación térmica
- ✓ Facilidad de mantenimiento

Tabla 5. Cantidad de conductores.

Tipo	Cantidad
Fase 152 mm ² (300 kcmil), Cu	15
Neutro 152 mm ² (300 kcmil), Cu	5
Tierra 107 mm ² (4/0 AWG), Cu	1
Total	21

Tabla 6. Dimensiones de los conductores

Cable	Diámetro aproximado
152 mm ² (300 kcmil)	20.34 mm
107 mm ² (4/0 AWG)	17.06 mm

2.8.2. Cálculo de área ocupada

Área por conductor:

- 152 mm² (300 kcmil): 324.9 mm²
- 107 mm² (4/0 AWG): 228.6 mm²

Área total:

- $A = (20 \times 324.9 \text{ mm}^2) + (1 \times 228.6 \text{ mm}^2)$
- $A = 6,726.6 \text{ mm}^2$

2.8.3. Área de ocupación máxima permisible

De acuerdo con la Tabla 392-22(b)(1), la charola portables de 30 cm de ancho permite una ocupación máxima de 8,400 mm² (Secretaría de Energía, 2012).

2.8.4 Cálculo dimensional

El diámetro exterior de un conductor de 152 mm² (300 kcmil) es de 20.34 mm, mientras que el diámetro exterior de un conductor de 107 mm² (4/0 AWG) es de 17.06. Para realizar el acomodo de los conductores en una sola capa, sin separación se requiere un ancho interior de charola de:

Ancho interior de charola

$$= (20.34 \text{ mm} \times 20) + 17.06 \text{ mm}$$

$$= 423.86 \text{ mm}$$

Por lo que se requiere una charola portables con un ancho interior comercial mínimo de 450 mm.

2.8.5 Verificación de arreglo físico

Para evitar sobrecalentamiento y cumplir con Art. 392 según (Secretaría de Energía, 2012; Smith et al., 2020):

Se requiere la instalación de los conductores en una sola capa, sin separación. Ver la Figura 5, para la instalación recomendada de cables

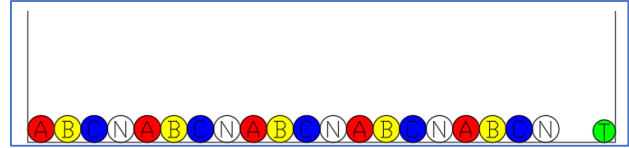


Figura 5. Configuración recomendada.

Tabla 7. Selección final de la charola portables.

Parámetro	Valor
Ancho	45 cm (18 pulgadas)
Altura	10 cm
Material	Aluminio
Tipo	Ventilada

2.8.6 Justificación técnica (según (Secretaría de Energía, 2012)(6)

La selección cumple con:]

- NOM-001-SEDE Art. 392
- Disipación térmica adecuada
- Distribución correcta de conductores en paralelo
- Facilidad de instalación y mantenimiento

2.8.7 Recomendaciones de instalación

- Mantener agrupamiento por fases equilibrado
- Identificación por colores
- Sujeción conforme a norma
- Continuidad eléctrica de la charola (puesta a tierra)
- Evitar cruces innecesarios

3. ANÁLISIS Y RESULTADOS DE LOS CÁLCULOS

3.1 Modelado eléctrico con memoria de cálculo

El generador funciona a 950 kVA y 480 V. La corriente se calcula usando ecuaciones de potencia trifásicas, resultando en aproximadamente 1143 A. Aplicando el criterio de carga continua del 125 por ciento, la ampacidad requerida supera los 1400 A (IEEE, 1993; Secretaría de Energía, 2012; Smith et al., 2020).

Para cumplir con este requisito, se seleccionan cinco conductores de cobre paralelo de 300 kcmil por fase. Cada director aporta aproximadamente 289 A, lo que resulta en una ampacidad total de aproximadamente 1446 A. Esta configuración satisface tanto las restricciones regulatorias como térmicas (Secretaría de Energía, 2012; Smith et al., 2020).

La caída de tensión se evalúa utilizando parámetros de impedancia derivados de tablas estándar. El valor calculado permanece por debajo del 2 por ciento, cumpliendo con las recomendaciones de la NOM-001-SEDE. El análisis de cortocircuitos indica una corriente subtransitoria de aproximadamente 1393 A, que está dentro de la capacidad de resistencia térmica de los conductores seleccionados (Secretaría de Energía, 2012; Smith et al., 2020).

El sistema de protección consta de un interruptor automático de 1600 A con una capacidad de interrupción de 65 kA, que garantiza la coordinación con los parámetros del sistema y la interrupción segura ante fallos (IEEE, 1997; Secretaría de Energía, 2012).



3.2 Análisis de ingeniería

El análisis de ingeniería del sistema demuestra que la selección de conductores debe abordarse como un problema multidimensional que implica consideraciones eléctricas, térmicas y espaciales. El uso de conductores paralelos es esencial para aplicaciones de alta corriente, ya que distribuye el flujo de corriente y reduce los efectos de calentamiento localizados (IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012; Smith et al., 2020; IEC, 1988).

Esta visión integrada es especialmente útil en instalaciones temporales, donde la solución final debe ser técnicamente robusta, físicamente montable y documentalmente defendible frente a inspección, comisionamiento y operación continua (IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012).

La disposición física de los conductores en una bandeja de cables tipo escalera es un factor clave de diseño. La configuración seleccionada garantiza que todos los conductores estén instalados en una sola capa, minimizando el calentamiento mutuo y permitiendo una circulación eficiente del aire. Esta disposición afecta directamente al rendimiento térmico del sistema y previene la degradación del aislamiento (Secretaría de Energía, 2012; Smith et al., 2020).

Los resultados de la simulación derivados de software especializado validan los cálculos analíticos. El análisis de flujo de potencia confirma que la caída de tensión se mantiene dentro de los límites aceptables, mientras que las simulaciones de cortocircuito verifican que el sistema puede soportar condiciones de fallo transitorio sin exceder los límites térmicos. Las curvas de coordinación de protección demuestran que el interruptor seleccionado opera dentro de las características tiempo-corriente requeridas, asegurando la selectividad del sistema (IEEE, 1993; IEEE, 1997; IEC, 1988).

El diseño de la bandeja de cables, con un ancho de 45 cm, acomoda a todos los conductores manteniendo el cumplimiento de los límites de ocupación. La distribución geométrica de los cables garantiza una carga de fase equilibrada y facilita las actividades de mantenimiento (Secretaría de Energía, 2012; Smith et al., 2020).

En general, el análisis confirma que la integración de cálculos teóricos, validación de simulaciones y criterios prácticos de instalación da como resultado un diseño de sistema robusto y fiable. Este enfoque está alineado con las mejores prácticas en ingeniería de sistemas eléctricos y proporciona un marco para futuras aplicaciones en escenarios similares (IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012; Smith et al., 2020; IEC, 1988).

3.3 Evaluación de interpretaciones en análisis del uso del cable para generador de 1 MW

3.3.1 Análisis adicional de eficiencia: La evaluación de eficiencia del alimentador se complementó estimando las pérdidas en julios en los conductores bajo condiciones de carga continua. Basándose en la configuración de conductores paralelos seleccionada, se compararon las pérdidas térmicas con disposiciones alternativas de conductores para validar el criterio técnico de selección. El análisis confirma que la distribución de conductores afecta directamente a la acumulación térmica, la eficiencia operativa y la regulación de voltaje en sistemas de generación temporal (IEC, 2023; IEEE, 1993).

3.3.2 Aclaración de la metodología de software: Las simulaciones eléctricas presentadas en las Figuras 2–4 se realizaron utilizando software comercial ETAP de análisis de sistemas eléctricos bajo módulos de estado estacionario y de cortocircuitos equivalentes a entornos de análisis de sistemas de potencia. Las simulaciones incluyeron la evaluación de flujo de carga, caída de tensión, coordinación de protección y cortocircuitos transitorios para validar cálculos analíticos y criterios de resistencia térmica (IEEE, 1997; IEC, 1988).

3.3.3 Aclaración sobre la interpretación de cortocircuitos: El artículo ahora distingue explícitamente entre la magnitud instantánea de la corriente de cortocircuito y la energía de resistencia térmica. La evaluación térmica fue validada teniendo en cuenta el tiempo de limpieza y el calentamiento permitido de conductores según la metodología IEC 60949, reduciendo la ambigüedad en la interpretación de los valores calculados de corriente de fallo (IEC, 1988; IEEE, 1997). En la *Tabla 8*, se muestra una comparativa de rendimiento térmico y caída de tensión en %.

Tabla 8. Comparativa de configuración de acuerdo a calibre del cable

Configuración	Ampacidad (A)	Pérdidas (%)	Caída de tensión (%)	Rendimiento térmico
4 x 300 kcmil	1,157	8.7	2.4	Marginal
5 x 300 kcmil	1,446	5.2	1.8	Óptima
6 x 300 kcmil	1,735	4.8	1.5	Sobredimensionado

3.3.4 Aclaración normativa: Aunque la NOM-001-SEDE-2012 sigue aplicándose ampliamente en la práctica industrial mexicana y en los procesos de validación de la UVIE, este trabajo reconoce la evolución continua de las normas internacionales y, por tanto, complementa la metodología utilizando referencias actualizadas de la IEC e IEEE para reforzar la justificación técnica y la reproducibilidad analítica.

La transición entre la NOM-001-SEDE-2012 y la NOM-001-SEDE-2018 representa un avance significativo en la modernización de los criterios de diseño aplicables a sistemas eléctricos industriales y, particularmente, a alimentadores para generación temporal de alta capacidad. Aunque ambas versiones mantienen la estructura fundamental basada en principios de seguridad eléctrica, ampacidad y coordinación de protecciones, la edición 2018 introduce una alineación técnica más cercana con versiones recientes del NEC y con prácticas internacionales de ingeniería eléctrica. En el contexto del presente artículo, la NOM-001-SEDE-2018 fortalece la interpretación de criterios asociados al cálculo de conductores en paralelo, selección de canalizaciones y verificación térmica



bajo condiciones reales de operación, aspectos críticos en sistemas de generación temporal de 1 MW. Mientras que la versión 2012 establece lineamientos ampliamente utilizados en procesos de validación UVIE y en instalaciones industriales existentes, la edición 2018 incorpora aclaraciones adicionales relacionadas con capacidad de conducción de corriente, métodos de instalación y evaluación de condiciones ambientales que impactan directamente la disipación térmica del conductor. Asimismo, la actualización normativa promueve una integración más rigurosa entre diseño eléctrico y desempeño operativo, enfatizando la necesidad de considerar factores reales de carga continua, agrupamiento y comportamiento térmico dinámico. Desde el punto de vista metodológico, la NOM-001-SEDE-2018 proporciona un marco más robusto para justificar técnicamente la selección de alimentadores en sistemas temporales, permitiendo reducir ambigüedades en interpretación normativa y fortaleciendo la trazabilidad de las memorias de cálculo utilizadas en proyectos de ingeniería eléctrica avanzada.

Desde una perspectiva especializada, la transición entre la NOM-001-SEDE-2012 y la NOM-001-SEDE-2018 no representa únicamente una actualización editorial, sino una evolución hacia un enfoque más compatible con metodologías contemporáneas de ingeniería eléctrica aplicada. En sistemas de generación temporal de alta capacidad, como el analizado en este artículo, la edición 2018 permite una mejor correlación entre cálculos normativos, simulación eléctrica y validación térmica, especialmente en escenarios donde existen conductores en paralelo, variaciones dinámicas de carga y restricciones de disipación térmica. Esta compatibilidad técnica resulta fundamental para sustentar memorias de cálculo avanzadas frente a procesos de validación UVIE, auditorías energéticas y aplicaciones industriales de potencia elevada. En la [Tabla 9](#), se muestra una comparativa de normatividad 2012 y 2018 para el tema de dimensionamiento de cables para generadores temporales de 1 MW.

Tabla 9. Comparativa técnica entre NOM-001-SEDE-2012 y NOM-001-SEDE-2018 aplicada al dimensionamiento de alimentadores para generadores temporales de 1 MW

Criterio técnico	NOM-001-SEDE-2012	NOM-001-SEDE-2018	Impacto técnico en el artículo
Enfoque general de diseño	Basado principalmente en criterios prescriptivos derivados del NEC anterior.	Integra criterios más alineados con desempeño operativo y condiciones reales de instalación.	La versión 2018 permite justificar con mayor precisión el comportamiento térmico y operativo de alimentadores de alta corriente en generación temporal.
Conductores en paralelo	Define requisitos básicos de igualdad de longitud, calibre y material.	Refuerza criterios de simetría física, trayectorias equivalentes y distribución electromagnética.	Mejora la validación técnica del arreglo de 5 conductores por fase utilizado en el artículo, reduciendo riesgos de desbalance térmico.
Ampacidad en charolas portables	Aplicación conservadora basada en tablas de ocupación y factores generales.	Introduce interpretaciones más claras sobre disipación térmica y condiciones de agrupamiento.	Favorece el análisis termoelectrónico avanzado del sistema propuesto en la investigación.
Carga continua en alimentadores	Mantiene el criterio mínimo de 125% para carga continua.	Conserva el criterio, pero enfatiza interpretación bajo condiciones reales de operación.	Fortalece la justificación del cálculo de 1,143 A y la selección del alimentador de 1,446 A.
Coordinación de protecciones	Enfoque centrado en capacidad interruptiva y protección básica.	Integra interpretación más rigurosa sobre coordinación y selectividad.	Permite validar de manera más robusta el interruptor termomagnético de 1,600 A empleado en el estudio.
Evaluación térmica de conductores	Referencias limitadas a tablas normativas.	Mayor compatibilidad con metodologías IEC y análisis térmicos modernos.	Facilita complementar el artículo con análisis de pérdidas Joule y acumulación térmica.
Canalización en charolas	Criterios dimensionales y de ocupación tradicionales.	Mayor énfasis en ventilación, accesibilidad y desempeño térmico.	Sustenta técnicamente la selección de charola ventilada de 45 cm propuesta en el artículo.
Validación mediante software	No se aborda explícitamente el uso de simulación avanzada.	Compatible con metodologías modernas de simulación y análisis de ingeniería.	Permite justificar el uso de software tipo ETAP para validación de flujo de potencia y cortocircuito.
Interpretación de cortocircuito	Predomina enfoque de corriente instantánea.	Facilita integración de criterios térmicos y energía dejada pasar.	Refuerza el análisis térmico de corto circuito desarrollado con IEC 60949.
Aplicación en sistemas temporales	Aplicable, pero sin enfoque explícito en generación temporal de alta capacidad.	Más adaptable a configuraciones industriales dinámicas y sistemas críticos.	Mejora la pertinencia normativa del artículo para aplicaciones de generación temporal de 1 MW.

Fuente: Elaboración propia, acorde a Secretaría de Energía (2012 y Secretaría de Energía (2018)

6. CONCLUSIONES

El dimensionamiento de alimentadores eléctricos en sistemas de generación temporal de 1 MW representa un problema de ingeniería compleja que requiere la integración de múltiples disciplinas. A lo largo de este trabajo se ha demostrado que el uso exclusivo de criterios normativos resulta insuficiente para garantizar un diseño óptimo, siendo necesario incorporar análisis térmico, validación eléctrica y condiciones reales de instalación (NFPA, 2023; IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012; Smith et al., 2020; IEC, 1988).

Uno de los hallazgos principales es la relevancia del uso de conductores en paralelo como solución técnica para manejar corrientes elevadas. Sin embargo, esta configuración introduce desafíos adicionales relacionados con el balance de corrientes, la distribución geométrica y la disipación térmica, los cuales deben ser cuidadosamente analizados para evitar fallas operativas (Secretaría de Energía, 2012; Smith et al., 2020).

Asimismo, se confirmó que la caída de tensión y la capacidad térmica ante cortocircuito son parámetros críticos que deben ser evaluados de forma simultánea, ya que un diseño que cumple con ampacidad puede no ser adecuado en condiciones de falla (Secretaría de Energía, 2012; IEC, 1988).

El análisis de canalización evidenció la importancia del uso de charolas ventiladas en sistemas de alta corriente, ya que permiten una adecuada disipación térmica y facilitan el mantenimiento del sistema (Secretaría de Energía, 2012; Smith et al., 2020).

Finalmente, la integración de normativas nacionales e internacionales proporciona una base sólida para el diseño, pero es la interpretación ingenieril la que permite optimizar el sistema y garantizar su operación segura (NFPA, 2023; IEEE, 1993; IEEE, 1997; IEC, 2019; Secretaría de Energía, 2012).

Como aporte metodológico, el artículo transforma un caso real de ingeniería en una guía reutilizable para futuras memorias de cálculo de generadores temporales de potencia similar, con énfasis en trazabilidad técnica y verificabilidad (IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012).

De los resultados obtenidos, se seleccionaron para el alimentador del generador 5 monoconductores por fase y de neutro de calibre 152 mm² (300 kcmil) y un conductor de puesta a tierra de 107 mm² (4/0 AWG), proponiéndose como medio de canalización una charola portacables tipo escalera con un ancho interior de 45 cm. El sistema está diseñado para operar de forma segura, continua y confiable, cumpliendo con los requerimientos normativos aplicables para su validación por UVIE (Secretaría de Energía, 2012).

Esta guía en este ejemplo se puede utilizar en otros cálculos o memorias de cálculo usando fórmulas descritas en el presente artículo, así como usando software comercial ETAP. Este ejemplo en México es común para ponerlo en práctica por la comunidad de ingeniería eléctrica que desarrolla estudios eléctricos en estado estable similares para alimentadores de generadores de capacidades de 1 MW (IEEE, 1993; IEEE, 1997; Secretaría de Energía, 2012).

REFERENCIAS

- NFPA (2023) NFPA 70: National Electrical Code. DOI: 10.6028/NFPA.70.2023.
- IEEE (1993) IEEE Std 141-1993: Electric Power Distribution for Industrial Plants. DOI: 10.1109/IEEEESTD.1994.121641.
- IEEE (1997) IEEE Std 399-1997: Industrial Power Systems Analysis. DOI: 10.1109/IEEEESTD.1998.89423.
- IEC (2019) IEC 60364 Electrical Installations of Buildings.
- Secretaría de Energía (2012) NOM-001-SEDE-2012 Instalaciones eléctricas, México.
- Smith, J. et al. (2020) 'Cable Thermal Analysis in Power Systems', IEEE Transactions on Power Delivery. DOI: 10.1109/TPWRD.2020.297.
- International Electrotechnical Commission (2023) IEC 60287-1-1:2023, Electric cables - Calculation of the current rating - Part 1-1: Current rating equations (100 % load factor) and calculation of losses - General. Geneva, Switzerland.
- International Electrotechnical Commission (1988) IEC 60949:1988, Calculation of thermally permissible short-circuit currents, taking into account non-adiabatic heating effects. Geneva, Switzerland.
- Artritt, R.F. y Dugan, R.C. (2011) 'Análisis de sistemas de distribución y la red inteligente del futuro', IEEE Transactions on Industry Applications, 47(6), pp. 2343–2350. DOI: 10.1109/TIA.2011.2168596.
- Gönen, T. (2014) Ingeniería de distribución de energía eléctrica. 3ª edición. CRC Press, Boca Ratón.
- Kersting, W.H. (2018) Modelado y análisis de sistemas de distribución. 4ª edición. CRC Press.
- IEEE Std 242-2001 (Buff Book), Protección y coordinación de sistemas eléctricos industriales y comerciales. DOI: 10.1109/IEEEESTD.2001.93397.
- Secretaría de Energía (2018) NOM-001-SEDE-2018, Instalaciones Eléctricas (Utilización). Diario Oficial de la Federación, México. Disponible en: https://www.dof.gob.mx/nota_detalle.php?codigo=5544135&fecha=06/11/2018

A comparative analysis of PEM and alkaline electrolyzers for H₂ production in areas with high solar

Karina Delgado Osorio*, Rodolfo Amalio Vargas Méndez*, Gloria Lilia Osorio Gordillo*, Felipe Ramírez Rasgado*

*Tecnológico Nacional de México/CENIDET, Cuernavaca, Mor, México (e-mail: d23ce120@cenidet.tecnm.mx, rodolfo.vm@cenidet.tecnm.mx, gloria.og@cenidet.tecnm.mx).

Abstract: The production of green hydrogen via solar-powered water electrolysis represents one of the most promising strategies for the decarbonisation of the global energy system. This article presents a comprehensive review of scientific literature on alkaline (AEL) and proton exchange membrane (PEM) electrolyzers, focusing on their technical performance, energy efficiency, costs and dynamic behaviour under high solar irradiance conditions. A mathematical modelling framework and a techno-economic analysis based on data reported in the literature are integrated. Furthermore, a specific analysis for the case of Mexico is included, highlighting its high solar potential and its impact on the competitiveness of green hydrogen.

Keywords: PEM electrolysis, alkaline electrolysis, green hydrogen, solar radiation, renewable energy.

1. INTRODUCCIÓN

El hidrógeno verde se ha consolidado como un vector energético clave en la transición hacia sistemas energéticos sostenibles, especialmente en el contexto de reducción de emisiones de CO₂ y almacenamiento de energía renovable (IEA, 2021). Actualmente, la electrólisis del agua es una de las principales tecnologías para su producción, particularmente cuando se acopla con fuentes renovables como la energía solar.

Las tecnologías predominantes son los electrolizadores alcalinos (AEL) y los electrolizadores PEM. Estas tecnologías presentan diferencias significativas en eficiencia, costo, flexibilidad operativa y madurez tecnológica. Diversos estudios han señalado que la selección entre ambas depende fuertemente del tipo de fuente energética y del perfil de operación (Buttler and Spliethoff, 2022).

En regiones con alta radiación solar, la variabilidad de la generación fotovoltaica introduce desafíos adicionales que afectan directamente el desempeño de los electrolizadores, lo cual hace necesario un análisis comparativo profundo sustentado en literatura científica.

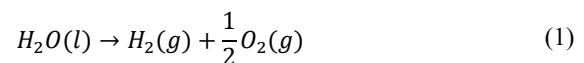
En este contexto, la integración de sistemas de electrólisis con generación fotovoltaica ha cobrado especial relevancia en regiones con alta irradiancia solar, donde el recurso renovable presenta un elevado potencial energético, pero también una marcada variabilidad temporal. Esta intermitencia impone retos técnicos significativos en la operación de los electrolizadores, afectando su eficiencia, vida útil y costos asociados. Por ello, la selección adecuada de la tecnología de electrólisis resulta un factor clave para maximizar el aprovechamiento del recurso solar y garantizar la viabilidad técnica y económica de la producción de hidrógeno verde.

El objetivo de este artículo es realizar un análisis comparativo entre electrolizadores AEL y PEM para la producción de

hidrógeno a partir de sistemas fotovoltaicos en zonas de alta irradiancia solar. El estudio se basa en una revisión de literatura técnica y científica, evaluando parámetros de desempeño como eficiencia, producción anual de hidrógeno, respuesta dinámica y costos, con el fin de identificar las ventajas, limitaciones y condiciones de aplicación más favorables para cada tecnología.

2. FUNDAMENTOS DE LA ELECTRÓLISIS DEL AGUA

La electrólisis del agua se basa en la descomposición de la molécula del agua mediante energía eléctrica:



El voltaje teórico mínimo requerido es de 1.23 V, aunque en condiciones reales este valor se incrementa debido a pérdidas electroquímicas:

$$V = V_{rev} + \eta_{act} + \eta_{ohm} + \eta_{conc} \quad (2)$$

Donde V_{rev} es el voltaje reversible, η_{act} es el sobrepotencial de activación, η_{ohm} son las pérdidas resistivas y η_{conc} son las pérdidas por concentración.

Estas pérdidas determinan el consumo energético específico, el cual típicamente se encuentra entre 45 y 55 kWh/kg de H₂ en sistemas comerciales (Arunachalam and Han, 2024).

Además de las pérdidas electroquímicas mencionadas, es importante considerar que el rendimiento del proceso de electrólisis también depende de variables operativas como la temperatura, la presión y la densidad de corriente. Diversos estudios han demostrado que un incremento en la temperatura de operación reduce las pérdidas por activación, mejorando la

cinética de reacción, aunque puede acelerar procesos de degradación de los materiales (Chatenet et al., 2022, Yassin et al., 2025, Carlson et al., 2023).

Asimismo, la presión de operación influye directamente en la pureza del hidrógeno producido y en la eficiencia global del sistema. Operaciones a alta presión permiten reducir o eliminar etapas de compresión posteriores, lo cual impacta favorablemente el balance energético del sistema.

Por otro lado, la densidad de corriente representa un parámetro crítico, ya que valores elevados incrementan la producción de hidrógeno, pero también aumentan las pérdidas óhmicas y térmicas. Por ello, existe un punto óptimo de operación que depende del tipo de electrolizador y de su diseño específico.

3. ELECTROLIZADORES ALCALINOS

Los electrolizadores alcalinos son la tecnología más madura y ampliamente utilizada. Operan con soluciones de hidróxido de potasio (KOH) al 20-40 % a temperaturas entre 60 y 80 °C, ver Figura 1. (Ajeeb et al., 2024).

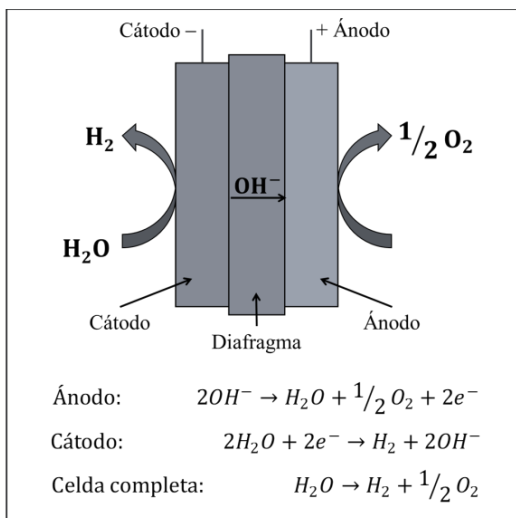


Figura 1. Ilustración esquemática de la electrólisis del agua alcalina.

Diversos estudios reportan las siguientes características:

- Eficiencia: 60-70 %
- Consumo energético: 50-55 kWh/kg
- Densidad de corriente: hasta 0.5-1 A/cm²

Su principal ventaja es el bajo costo debido a la ausencia de catalizadores nobles. Sin embargo, presentan limitaciones importantes tales como:

- Baja respuesta dinámica
- Operación menos eficiente a cargas parciales
- Mayor tamaño físico del sistema

Una revisión reciente destaca que los modelos de AEL se han utilizado ampliamente en optimización de sistemas energéticos debido a su estabilidad operativa (Huang et al., 2025).

Desde el punto de vista del modelado, los electrolizadores alcalinos suelen representarse mediante circuitos equivalentes que consideran resistencias óhmicas del electrolito y la separación de gases en el diafragma. La conductividad iónica del electrolito (KOH) depende fuertemente de la temperatura y la concentración, lo cual impacta directamente en las pérdidas óhmicas (Rolo et al., 2024).

Asimismo, la formación de burbujas de gas en los electrodos introduce resistencias adicionales al transporte de masa, reduciendo el área activa efectiva. Este fenómeno se vuelve más significativo a altas densidades de corriente, limitando el desempeño del sistema.

Modelos recientes incorporan ecuaciones de transporte multifásico para describir estos efectos, mejorando la precisión en la predicción del comportamiento del electrolizador bajo condiciones reales de operación (Lake et al., 2022, Vedrtan et al., 2025).

4. ELECTROLIZADORES PEM

Los electrolizadores PEM han ganado relevancia debido a su alta eficiencia y flexibilidad operativa, ver Figura 2 (Boretti, 2024, Wang et al., 2025).

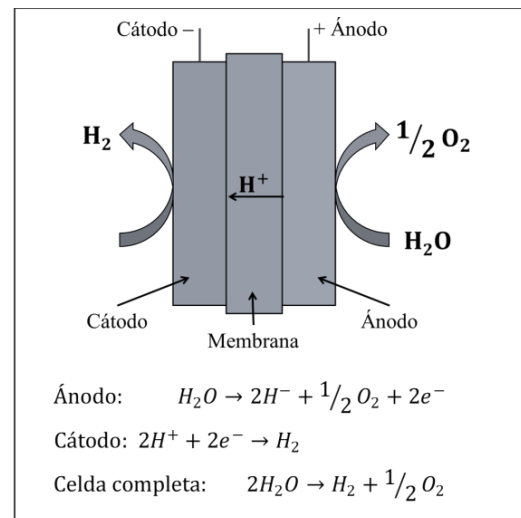


Figura 2. Ilustración esquemática de la electrólisis del agua PEM.

Características principales:

- Eficiencia: 65-80 %
- Consumo: 45-50 kWh/kg
- Densidad de corriente: 1-2 A/cm²

Estudios experimentales reportan consumos tan bajos como 43.46 kWh/kg con eficiencias superiores al 90 % en condiciones óptimas (Arunachalam and Han, 2024).

No obstante, presentan desventajas (Maoulida et al., 2026):

- Alto CAPEX
- Uso de materiales críticos: Iridio (Ir) y Platino (Pt)
- Mayor sensibilidad a degradación



4.1 Degradación y vida útil en electrolizadores PEM

Uno de los principales retos asociados a los electrolizadores PEM es su degradación a lo largo del tiempo, especialmente bajo condiciones de operación dinámica. La degradación se manifiesta principalmente en la membrana de intercambio protonico y en los catalizadores, afectando directamente la eficiencia del sistema.

Estudios indican que la degradación puede estar asociada a fenómenos como:

- Disolución del catalizador (Ir, Pt)
- Formación de radicales libres en la membrana
- Ciclos de arranque/parada frecuentes

La tasa de degradación típicamente se encuentra en el rango de 1-5 % por cada 1000 horas de operación, dependiendo de las condiciones operativas (Maoulida et al., 2026).

Este comportamiento es particularmente relevante en sistemas acoplados a energía solar, donde la intermitencia provoca variaciones constantes en la carga, acelerando los mecanismos de envejecimiento. Por ello, el diseño de estrategias de control y operación resulta fundamental para extender la vida útil del sistema (Arun et al., 2025, Makhsoos et al., 2025).

5. COMPARACIÓN TÉCNICA ENTRE TECNOLOGÍAS

Los resultados de la literatura coinciden en que PEM presenta ventajas operativas, mientras que AEL es más competitivo económicamente, ver Tabla 1 (Shanian and Savadogo, 2024, Wang et al., 2025, Wang et al., 2024).

Tabla 1. Comparación de parámetros operativos clave de electrolizadores Alcalinos y PEM

Parámetro	Alcalino	PEM
Eficiencia (%)	60-70	65-85
Consumo (kWh/kg)	50-55	45-50
Densidad de corriente (A/cm ²)	0.2-1	1-2
Tiempo de respuesta	Minutos	Segundos
CAPEX (USD/kW)	500-1000	1200-1800
Pureza del H ₂ (%)	99.5	>99.99

6. INTEGRACIÓN CON ENERGÍA SOLAR

Los sistemas fotovoltaicos presentan alta variabilidad temporal. En este contexto:

- PEM: mejor adaptación a variabilidad

- AEL: requiere operación más estable

Un estudio reciente basado en simulaciones TRNSYS.MATLAB reporta una producción anual de hasta 455 kg de H₂ en sistemas PV-PEM, con LCOH mínimo de 5.6 USD/kg (Lai et al., 2026a).

Además, se ha demostrado que la producción de hidrógeno aumenta significativamente con las horas pico de sol, confirmando la fuerte dependencia con la irradiancia solar.

Adicionalmente, la integración de sistemas de almacenamiento energético, como baterías o sistemas híbridos, ha sido propuesta como una solución para mitigar la variabilidad de la energía solar. Estos sistemas permiten suavizar las fluctuaciones de potencia, mejorando la operación del electrolizador y aumentando su eficiencia global (Rezaei et al., 2022).

En particular, configuraciones PV-batería-electrolizador han demostrado mejoras significativas en el factor de capacidad del sistema, lo que se traduce en una mayor producción anual de hidrógeno y una reducción del costo nivelado (LCOH) (Zhang and Xu, 2024, Arcos-Presedo and Arcos-Vargas, 2025).

La Tabla 2 presenta una comparativa de estudios sobre el desempeño de sistemas fotovoltaicos acoplados a los electrolizadores bajo condiciones de alta irradiancia solar. Los resultados muestran que las configuraciones PV-PEM presentan mayores niveles de producción anual de hidrógeno y eficiencias globales en comparación con sistemas PV-AEL. Asimismo, la integración de almacenamiento mediante baterías mejora el desempeño de ambas tecnologías, reduciendo el LCOH al favorecer una operación más estable del electrolizador.

Tabla 2. Desempeño de sistemas PV-electrolizador bajo alta irradiancia

Estudio	Configuración	Producción H ₂ (kg/año)	Eficiencia global (%)	LCOH (USD/kg)
(Lai et al., 2026b)	PV-PEM	455	72	5.6
(Hassan et al., 2023)	PV-AEL	380	65	6.8
(Zhang and Xu, 2024)	PV-PEM-Batería	520	75	5.2
(Guinot et al., 2015)	PV-AEL-Batería	415	69	6.1



El modelado dinámico de sistemas electrolíticos acoplados a energía solar requiere la integración de perfiles temporales de irradiancia, curvas de eficiencia dependientes de la carga y restricciones operativas del electrolizador. A diferencia de modelos estacionarios, los enfoques dinámicos permiten capturar los efectos de la intermitencia solar sobre la producción real de hidrógeno.

En la literatura, estos modelos suelen implementarse mediante simulaciones horarias o sub-horarias, donde la potencia fotovoltaica disponible determina el punto de operación instantáneo del electrolizador. Este enfoque resulta fundamental para evaluar pérdidas por curtailment, factores de capacidad y degradación acumulada (Park et al., 2023, Al-Mandhari and Ghosh, 2025).

Diversos estudios coinciden en que la omisión del comportamiento dinámico conduce a sobreestimaciones significativas de la producción anual de hidrógeno, particularmente en sistemas sin almacenamiento (Virah-Sawmy et al., 2024).

Sin embargo, la inclusión de almacenamiento implica un incremento en el CAPEX, por lo que su implementación debe evaluarse cuidadosamente mediante análisis tecno-económico.

7. MODELO MATEMÁTICO DEL SISTEMA

La potencia fotovoltaica se expresa como:

$$P_{PV}(t) = GHI(t) \cdot A \cdot \eta_{PV} \quad (3)$$

Donde $GHI(t)$ es la irradiancia solar global horizontal (W/m^2), A es el área del arreglo fotovoltaico (m^2) y η_{PV} es la eficiencia del panel fotovoltaico (%).

La producción instantánea de hidrógeno:

$$\dot{m}_{H_2}(t) = \frac{P_{PV}(t) \cdot \eta_{el}}{HHV} \quad (4)$$

Donde $P_{PV}(t)$ es la potencia disponible (kW), η_{el} es la eficiencia del electrolizador y HHV es el poder calorífico superior del hidrógeno (~ 39.4 kWh/kg).

La producción anual:

$$H_{2,anual} = \int_0^{8760} \dot{m}_{H_2}(t) dt \quad (5)$$

Donde $\dot{m}_{H_2}(t)$ es la producción instantánea y **8760** son el número de horas en un año.

La eficiencia eléctrica del proceso de electrólisis se define como la relación entre la energía química contenida en el hidrógeno producido y la energía eléctrica consumida por el electrolizador. Para ambos tipos de tecnología (AEL y PEM), la eficiencia puede expresarse como:

$$\eta_{el} = \frac{\dot{m}_{H_2} \cdot HHV_{H_2}}{P_{el}} \quad (6)$$

Donde HHV_{H_2} es el poder calorífico superior del hidrógeno (39.4 kWh/kg) y P_{el} es la potencia eléctrica consumida por el electrolizador (kW).

Adicionalmente, la eficiencia puede expresarse en función de los voltajes del sistema electroquímico:

$$\eta_{el} = \frac{V_{rev}}{V_{cell}} \quad (7)$$

Donde V_{rev} es el voltaje reversible teórico de la electrólisis del agua (1.23 V) y V_{cell} es el voltaje de operación real de la celda, que incluye las pérdidas por activación, resistivas y de concentración, que se expresa en la Ec. 2.

En los electrolizadores alcalinos, las pérdidas óhmicas están dominadas por la resistencia del electrolito y del diafragma, mientras que en electrolizadores PEM estas pérdidas están asociadas principalmente a la resistencia de la membrana polimérica. Asimismo, los electrolizadores PEM presentan menores sobrepotenciales de activación y mayor densidad de corriente, lo que se traduce en eficiencias superiores bajo operación dinámica.

Este modelo es ampliamente utilizado en estudios de integración renovable-electrólisis, la Figura 3 muestra un diagrama esquemático de un sistema de producción de hidrógeno verde acoplado a energía solar fotovoltaica. Se muestran las principales etapas de conversión energética, desde la radiación solar hasta el almacenamiento de hidrógeno, incluyendo las eficiencias asociadas a cada componente.

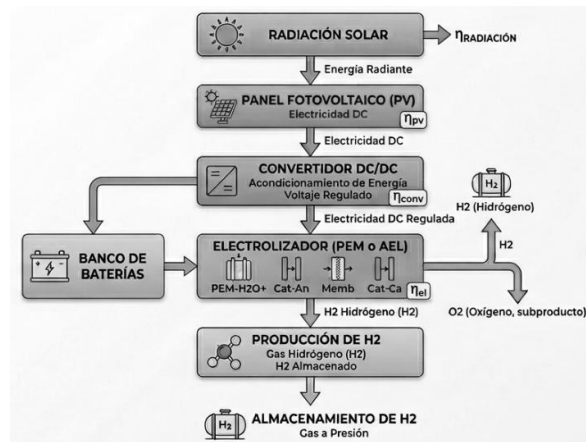


Figura 3. Diagrama esquemático de un sistema de producción de hidrógeno verde.

8. ANÁLISIS TECNO-ECONÓMICO

El costo nivelado del hidrógeno se define como:

$$LCOH = \frac{CAPEX \cdot CRF + OPEX + C_{elec}}{H_{2,anual}} \quad Ec. 8$$

Donde $CAPEX$ es la inversión inicial, CRF es el factor de recuperación de capital, $OPEX$ son los costos de operación, C_{elec} es el costo de electricidad y $H_{2,anual}$ es la producción anual de hidrógeno verde.



Resultados típicos:

- PEM: 5-15 USD/kg
- AEL: 3-10 USD/kg

El costo está fuertemente influenciado por:

- Precio de electricidad
- Factor de capacidad
- CAPEX

8.1 Análisis de sensibilidad económica

El costo nivelado del hidrógeno es altamente sensible a diversas variables, lo que hace necesario realizar un análisis de sensibilidad para identificar los parámetros más influyentes.

Entre los factores más relevantes se encuentran:

- Costos de la electricidad
- Factor de capacidad del sistema
- CAPEX
- Vida útil del electrolizador

Diversos estudios coinciden en que el costo de la electricidad es el parámetro más crítico, representando hasta el 60-70 % del LCOH en sistemas basados en energías renovables (Park et al., 2023, Acharya, 2023).

Asimismo, un aumento en el factor de capacidad, derivado de mejores condiciones de irradiancia o estrategias de almacenamiento, puede reducir significativamente el costo del hidrógeno (Acharya, 2023).

Además de los parámetros económicos clásicos, estudios recientes han demostrado que la degradación del electrolizador tiene un impacto no despreciable en el LCOH, especialmente en tecnologías PEM operando bajo condiciones dinámicas (Maoulida et al., 2026, He et al., 2025, Schofield et al., 2025). La pérdida progresiva de eficiencia incrementa el consumo eléctrico específico a lo largo del tiempo, elevando los costos operativos.

En este contexto, la inclusión de modelos de degradación en el análisis tecno-económico permite obtener estimaciones más realistas del costo del hidrógeno a largo plazo. Estrategias de operación que limiten los ciclos de arranque y parada, así como el uso de almacenamiento intermedio, pueden mitigar estos efectos, mejorando la viabilidad económica del sistema.

Este tipo de análisis es fundamental para la toma de decisiones en proyectos de hidrógeno verde, especialmente en países en desarrollo donde las condiciones económicas y energéticas son altamente variables.

9. CASO DE ESTUDIO: MÉXICO

México presenta condiciones excepcionales para la producción de hidrógeno verde (Jiménez, 2026):

- Irradiancia: 5-7 kWh/m²/día
- Alta disponibilidad de terreno
- Potencial fotovoltaico elevado

Comparado con Europa (~3 kWh/m²/día), México puede duplicar la producción anual de hidrógeno (Becerra-Lopez et al., 2025, Ashraf et al., 2024).

Esto impacta directamente el LCOH:

- Mayor irradiancia → Mayor producción → menor costo

Sin embargo, existen retos:

- Falta de infraestructura
- Costos iniciales elevados
- Marco regulatorio en desarrollo

En este contexto, la literatura sugiere que:

- PEM es más adecuado para integración directa con solar
- AEL es viable para sistemas híbridos o industriales

Adicionalmente, México cuenta con ventajas geográficas estratégicas para la exportación de hidrógeno verde, particularmente hacia mercados como Estados Unidos y Europa. La cercanía con centros industriales y la posibilidad de desarrollar corredores energéticos representan una oportunidad clave para posicionar al país como un actor relevante en la economía del hidrógeno.

No obstante, para materializar este potencial, es necesario fortalecer aspectos como:

- Políticas públicas y regulación energética
- Incentivos a la inversión privada
- Desarrollo de infraestructura de transporte y almacenamiento

Asimismo, la formación de capital humano especializado y el impulso a la investigación científica serán factores determinantes para acelerar la adopción de estas tecnologías en el país.

10. DISCUSIÓN

Además de los aspectos técnicos y económicos, es importante considerar factores ambientales y de sostenibilidad en la selección de la tecnología. La huella de carbono del hidrógeno producido depende no solo de la fuente eléctrica, sino también del ciclo de vida de los materiales utilizados en los electrolizadores.

En este sentido, tecnologías que utilizan metales críticos como PEM, presentan desafíos adicionales en términos de sostenibilidad y disponibilidad de recurso, lo cual debe ser considerado en evaluaciones a largo plazo.

El análisis integrado muestra que la selección tecnológica no depende únicamente del costo, sino del perfil de operación.

En escenarios de alta radiación solar, la flexibilidad operativa se vuelve crítica, favoreciendo el desempeño de los electrolizadores PEM, que maximizan la producción bajo condiciones intermitentes, mientras que los sistemas AEL presentan pérdidas de eficiencia en operación variable.



Esto sugiere que, en países como México, PEM puede ofrecer ventajas estratégicas a pesar de su mayor costo inicial.

11. CONCLUSIÓN

Este estudio confirma la viabilidad técnica de las tecnologías de electrólisis analizadas, destacando AEL por su madurez y rentabilidad económica, mientras que la PEM sobresale por su desempeño superior en sistemas solares. Se identificó que la irradiancia es el factor crítico que determina el costo final del hidrógeno, posicionando a México como un actor con alto potencial estratégico para liderar la producción de hidrógeno verde a nivel global.

AGRADECIMIENTOS

Agradecimientos y financiamiento: A la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI), por su compromiso con la investigación. Así como al TecNM por su contribución económica otorgada en la Convocatoria Proyectos de Investigación Científica, Humanística, de Desarrollo Tecnológico e Innovación 2026.

REFERENCIAS

- ACHARYA, A. 2023. A Review of Renewable Electricity Cost and Capacity Factor Impact on Green Hydrogen Levelized Cost in Off-grid Configuration. *Journal of Sustainable Development*, 16, 106-118.
- AJEEB, W., BAPTISTA, P. & NETO, R. C. 2024. Life cycle analysis of hydrogen production by different alkaline electrolyser technologies sourced with renewable energy. *Energy Conversion and Management*, 316, 118840.
- AL-MANDHARI, M. & GHOSH, A. 2025. Modelling Solar Intermittency Effects on PEM Electrolyser Performance & Degradation: A Comparison of Oman and UK. *Energies*, 18, 6131.
- ARCOS-PRESEDO, M. C. & ARCOS-VARGAS, A. 2025. Optimizing Green Hydrogen Cost with PV Energy and Storage. *Energies* [Online], 18.
- ARUN, M., GIDDEY, S., JOSEPH, P. & DHAWALE, D. S. 2025. Challenges and mitigation strategies for general failure and degradation in polymer electrolyte membrane-based fuel cells and electrolyzers. *Journal of Materials Chemistry A*, 13, 11236-11263.
- ARUNACHALAM, M. & HAN, D. S. 2024. Efficient solar-powered PEM electrolysis for sustainable hydrogen production: an integrated approach. *Emergent Materials*, 7, 1401-1415.
- ASHRAF, U., MORELLI, T. L., SMITH, A. B. & HERNANDEZ, R. R. 2024. Climate-Smart Siting for renewable energy expansion. *iScience*, 27, 110666.
- BECERRA-LOPEZ, H. R., DE SILVA-MUNOZ, L., FLORES-HERNANDEZ, J. R., LIRA-ARGUELLO, R. & ROMERO-CASTANON, T. 2025. Forecasting regional green hydrogen demand in Mexico by 2035. *International Journal of Hydrogen Energy*, 141, 1039-1049.
- BORETTI, A. 2024. Estimating the efficiency of a PEM electrolyzer fed by a PV plant in NEOM City. *Solar Energy Advances*, 4, 100072.
- BUTTLER, A. & SPLIETHOFF, H. 2022. Current status of water electrolysis for energy storage, grid balancing and sector coupling via power-to-gas and power-to-liquids: A review. *Renewable and Sustainable Energy Reviews*, 82, 2440-2454.
- CARLSON, E. L., PICKFORD, K. & NYGA-LUKASZEWSKA, H. 2023. Green hydrogen and an evolving concept of energy security: Challenges and comparisons. *Renewable energy*, 219, 119410.
- CHATENET, M., POLLET, B. G., DEKEL, D. R., DIONIGI, F., DESEURE, J., MILLET, P., BRAATZ, R. D., BAZANT, M. Z., EIKERLING, M., STAFFELL, I., BALCOMBE, P., SHAO-HORN, Y. & SCHÄFER, H. 2022. Water electrolysis: from textbook knowledge to the latest scientific strategies and industrial developments. *Chem Soc Rev*, 51, 4583-4762.
- GUINOT, B., CHAMPEL, B., MONTIGNAC, F., LEMAIRE, E., VANNUCCI, D., SAILLER, S. & BULTEL, Y. 2015. Techno-economic study of a PV-hydrogen-battery hybrid system for off-grid power supply: Impact of performances' ageing on optimal system sizing and competitiveness. *International journal of hydrogen energy*, 40, 623-632.
- HASSAN, Q., ZUHAIR, A., SALMAN, H. & JASZCZUR, M. 2023. Large-scale green hydrogen production using alkaline water electrolysis based on seasonal solar radiation. *Energy Harvesting and Systems*, 11.
- HE, D., CHEN, K., CHEN, W., LUO, Z., XIONG, Z., ZOU, G., LI, G. & CHEN, B. 2025. Experimental study on the degradation of proton exchange membrane water electrolyzers under different dynamic loads and flow field configurations. *Journal of Power Sources*, 641, 236844.
- HUANG, C., TORRES, J. L. R., ZONG, Y., YOU, S. & JIN, X. 2025. A review of alkaline electrolyzer technology modeling and applications for decision-making optimization in energy systems. *Renewable and Sustainable Energy Reviews*, 224, 116005.
- IEA 2021. Global Hydrogen Review 2021. IEA.
- JIMÉNEZ, A. N. 2026. Scenario-based mapping of state-level green hydrogen potential in Mexico: Policy-relevant insights from a multi-technology framework. *Energy and Climate Change*, 7, 100240.
- LAI, F., WANG, M., ZHANG, T., SONG, J. & ZHOU, J. 2026a. Techno-economic evaluation of solar-driven PEM water electrolysis systems for hydrogen production: Comparison and regional insights. *Renewable Energy*, 264, 125558.
- LAI, F., WANG, M., ZHANG, T., SONG, J. & ZHOU, J. 2026b. Techno-economic evaluation of solar-driven PEM water electrolysis systems for hydrogen production: Comparison and regional insights. *Renewable Energy*, 125558.
- LAKE, J. R., SOTO, Á. M. & VARANASI, K. K. 2022. Impact of Bubbles on Electrochemically Active



- Surface Area of Microtextured Gas-Evolving Electrodes. *Langmuir*, 38, 3276-3283.
- MAKHSOOS, A., KANDIDAYENI, M., POLLET, B. G. & BOULON, L. 2025. Proton exchange membrane water electrolyzers degradation models review: implications for power allocation and energy management. *Journal of Power Sources*, 655, 238003.
- MAOULIDA, F., GUILBERT, D., CAMARA, M.-B. & DAKYO, B. 2026. Dynamic electrical degradation of PEM electrolyzers under renewable energy Intermittency: Mechanisms, diagnostics, and mitigation strategies—A comprehensive review. *Renewable and Sustainable Energy Reviews*, 225, 116170.
- PARK, J., HWAN RYU, K., KIM, C.-H., CHUL CHO, W., KIM, M., HUN LEE, J., CHO, H.-S. & LEE, J. H. 2023. Green hydrogen to tackle the power curtailment: Meteorological data-based capacity factor and techno-economic analysis. *Applied Energy*, 340, 121016.
- REZAEI, M., AKIMOV, A. & GRAY, E. M. 2022. Economics of solar-based hydrogen production: Sensitivity to financial and technical factors. *International Journal of Hydrogen Energy*, 47, 27930-27943.
- ROLO, I., COSTA, V. A. F. & BRITO, F. P. 2024. Hydrogen-Based Energy Systems: Current Technology Development Status, Opportunities and Challenges. *Energies* [Online], 17.
- SCHOFIELD, L., PAREN, B., MACDONALD, R., SHAO-HORN, Y. & MALLAPRAGADA, D. 2025. Dynamic optimization of proton exchange membrane water electrolyzers considering usage-based degradation. *AIChE Journal*, 71, e18635.
- SHANIAN, S. & SAVADOGO, O. 2024. Techno economic analysis of electrolytic hydrogen production by alkaline and PEM electrolyzers using MCDM methods. *Discover Energy*, 4, 23.
- VEDRTNAM, A., KALAUNI, K. & PAHWA, R. 2025. Water electrolysis technologies and their modeling approaches: a comprehensive review. *Eng*, 6, 81.
- VIRAH-SAWMY, D., BECK, F. J. & STURMBERG, B. 2024. Ignore variability, overestimate hydrogen production—Quantifying the effects of electrolyzer efficiency curves on hydrogen production from renewable energy sources. *International Journal of Hydrogen Energy*, 72, 49-59.
- WANG, J., WEN, J., WANG, J., YANG, B. & JIANG, L. 2024. Water electrolyzer operation scheduling for green hydrogen production: A review. *Renewable and Sustainable Energy Reviews*, 203, 114779.
- WANG, J., YANG, J., FENG, Y., HUA, J., CHEN, Z., LIAO, M., ZHANG, J. & QIN, J. 2025. Comparative experimental study of alkaline and proton exchange membrane water electrolysis for green hydrogen production. *Applied Energy*, 379, 124936.
- YASSIN, K., ATTIAS, R., TSUR, Y. & DEKEL, D. R. 2025. Identifying and Quantifying Loss Sources in Anion-Exchange Membrane Water Electrolyzers. *ACS Electrochemistry*, 1, 655-666.
- ZHANG, W. & XU, C. 2024. Capacity configuration optimization of photovoltaic-battery-electrolysis hybrid system for hydrogen generation considering dynamic efficiency and cost learning. *Energy Conversion and Economics*, 5, 78-92.

Thermodynamic simulation of sustainable syngas production with Reverse Water-Gas Shift using captured carbon dioxide and green hydrogen.

Marisol Coba-Martínez*, Marisol Cervantes-Bobadilla**, Jarniel García-Morales**

Alicia Martínez-Rebollar,**, Noe Alejandro Castro-Sánchez**, Hugo Estrada-Esquivel**

*Posgrado Tecnológico Nacional de México/CENIDET, Interior Internado Palmira S/N, Col. Palmira, Cuernavaca, Mor. México, 62490. (e-mail: d24ce042@cenidet.tecnm.mx).

**Tecnológico Nacional de México/CENIDET, Interior Internado Palmira S/N, Col. Palmira, Cuernavaca, Mor. México, 62490. (e-mail: marisol.cb@cenidet.tecnm.mx, jarniel.gm@cenidet.tecnm.mx, alicia.mr@cenidet.tecnm.mx, noe.cs@cenidet.tecnm.mx, hugo.ee@cenidet.tecnm.mx).

Abstract: This study presents a thermodynamic analysis and optimization of syngas production via the Reverse Water-Gas Shift (RWGS) reaction using captured carbon dioxide (CO₂) and green hydrogen. The process was simulated in Aspen Plus using a Gibbs equilibrium reactor (RGibbs) and the Peng–Robinson equation of state to evaluate the effects of temperature and H₂/CO₂ molar ratio on system performance. A parametric analysis was conducted over a temperature range of 500–900 °C and molar ratios of 1:1, 2:1, and 3:1. Results indicate that high temperatures significantly enhance CO₂ conversion and suppress methane formation. A maximum CO₂ conversion of 92.20% was achieved at 900 °C with an H₂/CO₂ ratio of 2:1, yielding an H₂/CO ratio close to the optimal value required for downstream Fischer–Tropsch synthesis. Lower temperatures favored methanation, reducing syngas quality. The findings demonstrate that RWGS is a technically viable and scalable pathway for sustainable syngas production, supporting the integration of carbon capture and utilization (CCU) with green hydrogen systems for low-carbon fuel synthesis. This work provides a robust thermodynamic basis for the design and optimization of integrated CO₂ conversion processes in future sustainable energy systems.

Keywords: Syngas production, Carbon dioxide utilization, Green hydrogen, Reverse water-gas shift, Sustainable fuel synthesis.

1. INTRODUCCIÓN

El incremento sostenido de las emisiones de dióxido de carbono (CO₂), derivado de la actividad industrial y del uso intensivo de combustibles fósiles, ha sido identificado como uno de los principales impulsores del cambio climático global. En 2022, las emisiones alcanzaron 36.1 Gt y las concentraciones atmosféricas superaron las 414 ppm, evidenciando la necesidad urgente de desarrollar soluciones tecnológicas efectivas para mitigar este fenómeno (Triviño et al., 2024).

En este escenario, la Captura y Utilización de Carbono (CCU) ha emergido como una estrategia clave para transformar el CO₂ capturado en productos químicos y combustibles de valor agregado. Entre las rutas más prometedoras destaca la reacción inversa de desplazamiento gas-agua (Reverse Water-Gas Shift, RWGS), una reacción endotérmica que convierte CO₂ e hidrógeno (H₂) en monóxido de carbono (CO) y agua, generando gas de síntesis (syngas), un intermediario fundamental en la industria química y energética (Zhu et al., 2020; Triviño et al., 2024).

El syngas, compuesto principalmente por H₂ y CO, constituye la base de procesos industriales ampliamente utilizados, como la síntesis de metanol, la producción de combustibles

sintéticos mediante Fischer–Tropsch y la obtención de diversos productos químicos intermedios (Bahmanpour et al., 2021; Ruggiero et al., 2025). Tradicionalmente, su producción ha dependido de fuentes fósiles; sin embargo, su obtención mediante RWGS utilizando hidrógeno verde y CO₂ capturado representa una alternativa sostenible alineada con la transición hacia sistemas energéticos bajos en carbono (Zhou et al., 2023).

A pesar de su potencial, el proceso RWGS enfrenta importantes limitaciones termodinámicas. A bajas temperaturas, la reacción compite con la metanación del CO₂, favoreciendo la formación de metano (CH₄) y reduciendo la selectividad hacia CO (Lykaki et al., 2024; Memon et al., 2023). Por el contrario, a temperaturas elevadas (>700 °C), el equilibrio se desplaza hacia la formación de CO, mejorando la conversión y la calidad del syngas. Desde el punto de vista práctico, el desarrollo de catalizadores activos y selectivos, así como el control de las condiciones operativas, resulta fundamental para optimizar el desempeño del proceso (Chen et al., 2020; Ebrahimi et al., 2022). En años recientes, se han reportado avances significativos en catalizadores basados en metales nobles (Pt, Pd, Au) y no nobles (Ni, Cu, Fe), además de soportes como CeO₂, ZrO₂ y Al₂O₃, que incrementan la

estabilidad térmica y la dispersión activa (Choi et al., 2024; Jaubert et al., 2022; Paes et al., 2024).

Además de los avances catalíticos, se han propuesto estrategias como el chemical looping (Abbas et al., 2024) y la integración del RWGS con procesos Fischer–Tropsch, con el objetivo de maximizar el aprovechamiento del CO producido y mejorar la eficiencia global del sistema (Corrao et al., 2023; Buthelezi et al., 2024; Molefe et al., 2024). Desde la ingeniería de procesos, diversos estudios han demostrado la utilidad de Aspen Plus y de modelos termodinámicos como la ecuación de estado de Peng–Robinson para analizar sistemas RWGS bajo condiciones de equilibrio, permitiendo predecir la composición del syngas y evaluar el efecto de variables operativas sobre la conversión de CO₂ y la selectividad del proceso (Ruggiero et al., 2025; Hiraga & Ushiki, 2024; Khaghani et al., 2025). Además, el uso de reactores de equilibrio tipo RGibbs ha mostrado ser una herramienta efectiva para el análisis preliminar y la optimización conceptual de procesos de valorización de CO₂ mediante hidrógeno verde.

Aún existe la necesidad de estudios que no solo evalúen la conversión de CO₂, sino que también optimicen simultáneamente la calidad del syngas en términos de la relación H₂/CO, parámetro crítico para su integración con procesos posteriores como la síntesis Fischer–Tropsch o la producción de metanol. Desde esta perspectiva, el presente trabajo analiza de manera conjunta el efecto de la temperatura y de la relación molar H₂/CO₂ sobre la conversión de CO₂, la relación H₂/CO y la formación de CH₄ durante el proceso RWGS, con el objetivo de identificar condiciones favorables para la producción sostenible de syngas.

Finalmente, este trabajo establece un marco de simulación basado en Aspen Plus y la ecuación de estado de Peng–Robinson, aplicable al diseño conceptual y escalamiento de sistemas integrados de valorización de CO₂ mediante hidrógeno verde, contribuyendo al desarrollo de rutas tecnológicas sostenibles para la producción de combustibles sintéticos.

2. METODOLOGÍA

Este estudio se desarrolló bajo un enfoque cuantitativo, no experimental y de simulación computacional, orientado al análisis termodinámico del proceso Reverse Water-Gas Shift (RWGS) para la producción de gas de síntesis (syngas) a partir de CO₂ capturado e hidrógeno verde. La metodología se basó en la implementación de un modelo de equilibrio en el software Aspen Plus, empleando la ecuación de estado Peng–Robinson (PR-EoS) para la predicción del comportamiento de fases y composición de productos bajo diferentes condiciones de operación.

2.1 Herramientas y modelo termodinámico

La simulación del sistema se llevó a cabo utilizando Aspen Plus V14. Para representar el proceso RWGS se empleó un reactor de equilibrio tipo RGibbs, mientras que la ecuación de estado Peng–Robinson (PR-EoS) fue seleccionada para describir el comportamiento termodinámico de la mezcla gaseosa compuesta por CO₂, H₂, CO, CH₄ y H₂O bajo

condiciones de alta temperatura. Esta configuración permitió evaluar el efecto de la temperatura y de la relación molar H₂/CO₂ sobre la conversión de CO₂ y la composición del gas de síntesis generado.

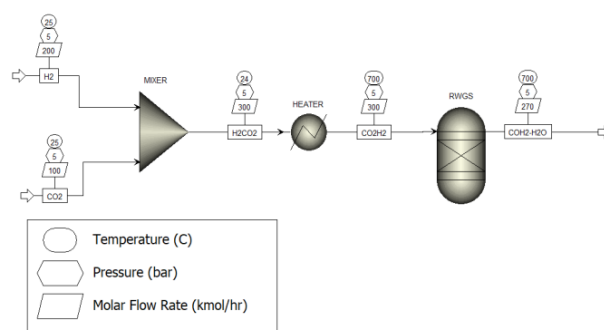


Figura 1 Diagrama de flujo del proceso RWGS simulado en Aspen Plus.

La Figura 1 muestra el diagrama de flujo del proceso de producción de gas de síntesis (syngas) mediante la reacción inversa de desplazamiento gas-agua (RWGS), modelado en Aspen Plus. El sistema inicia con la alimentación de hidrógeno (H₂) y dióxido de carbono (CO₂) en condiciones ambientales (25 °C y 5 bar), con flujos molares de 200 kmol/h y 100 kmol/h, respectivamente, estableciendo una relación molar H₂/CO₂ de 2:1.

Como se detalla en la Tabla 1, ambas corrientes son conducidas a un mezclador (MIXER), donde se combinan mediante un sistema de mezcla estática en línea, garantizando una distribución homogénea de los reactivos. A la salida de esta etapa se obtiene una corriente unificada con un flujo total de 300 kmol/h, manteniendo prácticamente las mismas condiciones de presión y temperatura.

Posteriormente, la mezcla es sometida a un proceso de calentamiento (HEATER), en el cual la temperatura se incrementa hasta 700 °C, condición necesaria para impulsar la reacción endotérmica RWGS. En la simulación, esta etapa se representa mediante un bloque tipo HEATER; sin embargo, a nivel industrial, se implementa mediante una combinación de recuperación de calor en intercambiadores tipo coraza y tubos y un sistema de calentamiento de alta temperatura, como un horno tubular (fired heater), para alcanzar las condiciones requeridas de operación.

La corriente calentada es alimentada a un reactor de equilibrio tipo RGibbs, donde se lleva a cabo la conversión de CO₂ e H₂ en monóxido de carbono (CO) y agua (H₂O), considerando simultáneamente reacciones secundarias como la metanación. Este modelo permite estimar la composición de salida mediante la minimización de la energía libre de Gibbs. En términos industriales, este comportamiento se aproxima mediante un reactor de lecho fijo (Packed Bed Reactor), donde la mezcla gaseosa fluye a través de un catalizador sólido.

Como resultado del proceso, se obtiene una corriente de salida a 700 °C y 5 bar, con un flujo molar de 270 kmol/h,



correspondiente al gas de síntesis generado. La composición de esta corriente depende directamente de las condiciones de operación, particularmente la temperatura y la relación molar de alimentación, tal como se analiza en las secciones posteriores.

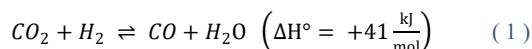
Cabe señalar que la reducción del flujo molar total en la corriente de salida (de 300 a 270 kmol/h) se debe principalmente a la ocurrencia de reacciones secundarias como la metanación, las cuales implican una disminución en el número total de moles gaseosos. Este comportamiento es consistente con el modelo de equilibrio termodinámico (RGibbs), el cual considera la contribución simultánea de todas las reacciones presentes en el sistema.

Tabla 1 Materias primas, etapas de proceso de la simulación en Aspen Plus.

Entradas de materia prima	
H ₂ (hidrógeno) Temperatura: 25 °C Presión: 5 bar Flujo molar: 200 kmol/h	CO ₂ (dióxido de carbono) Temperatura: 25 °C Presión: 5 bar Flujo molar: 100 kmol/h
Etapas del proceso	
Mezclador (MIXER)	
Temperatura: 24 °C Presión: 5 bar Flujo molar combinado: 300 kmol/h	
Intercambiador de calor (HEATER)	
Temperatura: 700 °C Presión: 5 bar Flujo molar: 300 kmol/h	
Reactor RWGS (RGibbs)	
Temperatura: 700 °C Presión: 5 bar Flujo molar: 270 kmol/h	

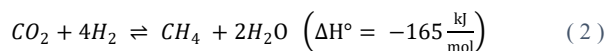
2.2 Reacción modelada

La reacción principal modelada en este estudio es la reacción inversa del desplazamiento gas-agua, conocida como Reverse Water-Gas Shift (RWGS, por sus siglas en inglés):



Esta reacción es endotérmica, lo que significa que requiere el aporte de energía térmica para desplazarse hacia los productos. Desde el punto de vista termodinámico, la formación de monóxido de carbono (CO) es favorecida a temperaturas elevadas, ya que el incremento de temperatura desplaza el equilibrio hacia la derecha según el principio de Le Châtelier.

Uno de los principales desafíos del proceso RWGS es evitar reacciones secundarias como la metanación del CO₂:



Esta reacción es exotérmica y se ve favorecida a temperaturas más bajas (<500 °C), lo que puede disminuir la selectividad

hacia CO y afectar la calidad del gas de síntesis obtenido. Por tanto, la elección adecuada de las condiciones de operación (temperatura y relación molar H₂/CO₂) es fundamental para maximizar la conversión deseada y minimizar subproductos indeseables como el metano (Memon et al., 2023; Choi et al., 2024).

En esta investigación se consideró principalmente la metanación del CO₂ como reacción secundaria debido a su mayor relevancia termodinámica bajo las condiciones evaluadas y a su impacto directo sobre la selectividad hacia CO y el consumo de hidrógeno. Otras reacciones secundarias, como la metanación de CO, la reacción de Boudouard o la formación de coque, no fueron incluidas explícitamente en el análisis debido a que su contribución resulta menos significativa en sistemas ricos en hidrógeno y a temperaturas elevadas, además de requerir consideraciones cinéticas y catalíticas específicas que se encuentran fuera del alcance del presente estudio termodinámico de equilibrio (Lykaki et al., 2024; Choi et al., 2024).

El modelo de equilibrio implementado en Aspen Plus permitió analizar el efecto de las variables operativas sobre la composición del sistema RWGS, identificando tendencias en la conversión de CO₂, la relación H₂/CO y la formación de CH₄ bajo diferentes condiciones de operación.

Este enfoque permitió identificar tendencias operativas relevantes y establecer condiciones favorables para la producción de syngas, proporcionando una base preliminar para futuras etapas de optimización cinética y evaluación energética del proceso. Además, su aplicación en sistemas que integran hidrógeno verde y captura de CO₂ contribuye al análisis preliminar de alternativas sostenibles para la producción de syngas y su posterior integración en procesos como Fischer-Tropsch o síntesis de metanol.

2.3 Condiciones de simulación

Se establecieron las siguientes condiciones operativas:

Tabla 2 Condiciones del caso base

Parámetros	Rangos
Presión:	5 bar
Temperatura:	700 °C
Relación molar H ₂ /CO ₂ :	2:1
Entrada de reactivos:	100 kmol/h de CO ₂ y 200 kmol/h de H ₂

Estos rangos fueron seleccionados con base en estudios experimentales y de simulación reportados en la literatura para maximizar la conversión de CO₂ y obtener una relación H₂/CO adecuada para su uso posterior en procesos como Fischer-Tropsch o metanación (Zhu et al., 2020; Choi et al., 2024; Zhou et al., 2023).

2.4 Supuestos del modelo

Para el desarrollo de la simulación en Aspen Plus, se adoptaron los siguientes supuestos con el fin de simplificar el análisis y



centrarse en el comportamiento termodinámico del sistema RWGS:

- Estado estacionario: se modeló el proceso bajo condiciones de operación estacionarias, es decir, sin considerar transitorios ni fluctuaciones en el tiempo. Esto permite evaluar el rendimiento en régimen permanente, como es común en estudios conceptuales.
- Mezcla ideal de entrada: se asumió que los reactivos (CO_2 y H_2) ingresan en forma pura y perfectamente mezclados, sin la presencia de impurezas (como N_2 , O_2 o vapor de agua) ni compuestos que puedan actuar como inhibidores catalíticos. Esta simplificación permite aislar el efecto de la temperatura y la relación molar sobre el equilibrio químico.
- Ausencia de efectos cinéticos y de transporte: el modelo emplea un reactor tipo RGibbs, el cual calcula la composición en equilibrio mediante la minimización de la energía libre de Gibbs. Por tanto, no se considera la influencia de la velocidad de reacción, ni fenómenos de transporte de masa o calor. Esta aproximación es adecuada para una primera etapa de simulación orientada a evaluar condiciones óptimas de operación (Ruggiero et al., 2025).
- Sistema cerrado, sin reciclaje ni purgas: se considera una sola pasada de los reactivos, sin recirculación de productos ni corrientes de purga. Esto simplifica el modelo y representa un escenario base que puede ampliarse en estudios posteriores con integración de etapas downstream (etapas posteriores que utilizan el gas de síntesis como insumo) o recuperación de reactivos no reaccionados.
- Flujos molares fijos: el flujo de CO_2 se mantiene constante en 100 kmol/h, mientras que el de H_2 varía según la relación molar H_2/CO_2 definida en cada escenario. Este enfoque permite comparar escenarios de sensibilidad térmica bajo condiciones normalizadas.
- Sin generación o consumo de energía fuera del sistema: no se modela el consumo energético de calentamiento, compresión o integración térmica con otros procesos. El objetivo es estudiar el comportamiento químico del sistema idealizado.

Estos supuestos son comunes en la literatura técnica cuando se busca analizar el potencial termodinámico de un proceso sin complicaciones operativas ni detalles de diseño de planta. Sin embargo, deben ser tenidos en cuenta al extrapolar los resultados a condiciones industriales reales.

3. RESULTADOS Y DISCUSIÓN

3.1 Caso base

El caso base fue evaluado a una temperatura de 700 °C y una presión de operación de 5 bar, bajo una relación molar de alimentación H_2/CO_2 de 2:1. En estas condiciones, se obtuvo una conversión de CO_2 de 76.27 %, lo que sugiere un desempeño favorable del proceso bajo estas condiciones operativas. La relación H_2/CO obtenida fue de 3.09, valor que,

aunque superior al óptimo requerido para procesos como Fischer–Tropsch, muestra una tendencia adecuada hacia la producción de gas de síntesis. De igual forma, la fracción molar de CH_4 fue de 0.0199, indicando que la formación de metano es considerablemente baja a esta temperatura.

Este comportamiento coincide con la naturaleza endotérmica de la reacción RWGS ($\Delta H^\circ = +41$ kJ/mol), ya que temperaturas elevadas contribuyen a la formación de CO y H_2O , al tiempo que limitan las reacciones secundarias de metanación. El uso del modelo termodinámico de Peng–Robinson, junto con un reactor RGibbs, permitió representar adecuadamente el equilibrio del sistema, siendo una aproximación válida para estudios preliminares.

En conjunto, el caso base refleja un balance adecuado entre conversión, selectividad y calidad del syngas. Pese a ello, la relación H_2/CO aún se encuentra por encima del valor ideal, lo que sugiere que el sistema puede optimizarse mediante el ajuste de condiciones operativas, particularmente la temperatura, para mejorar el aprovechamiento del hidrógeno.

3.2 Análisis de sensibilidad térmica

Se realizó un análisis de sensibilidad térmica variando la temperatura de operación del reactor entre 500 °C y 900 °C, con el objetivo de evaluar su impacto en tres variables clave: la conversión de CO_2 , la relación H_2/CO del gas de síntesis y la formación de CH_4 como subproducto no deseado. Dado su carácter endotérmico, se evaluaron temperaturas elevadas para promover la formación de CO y minimizar la producción de CH_4 (Lykaki et al., 2024).

Este análisis se llevó a cabo considerando tres relaciones molares de alimentación H_2/CO_2 : 1:1, 2:1 y 3:1, lo cual permite estudiar no solo la influencia térmica, sino también el impacto del exceso de hidrógeno sobre el equilibrio del sistema. La Tabla 3 describe las condiciones de operación consideradas.

Tabla 3 Condiciones de operación evaluadas para la simulación del proceso RWGS

Parámetro	Rangos evaluados	Descripción / Justificación
Temperatura (°C)	500, 600, 700, 800, 900	La reacción RWGS es endotérmica y favorecida a alta temperatura (Zhu et al., 2020).
Relación H_2/CO_2 (mol/mol)	1:1, 2:1, 3:1	Se exploran diferentes relaciones para ajustar la proporción H_2/CO del syngas (Choi et al., 2024).
Modelo termodinámico	Peng–Robinson (PR–EoS)	Ampliamente utilizado para gases reales con buen ajuste en sistemas $\text{CO}_2/\text{CO}/\text{H}_2$ (Khaghani et al., 2025).
Tipo de reactor	RGibbs (mín. energía libre de Gibbs)	Permite cálculo del equilibrio termodinámico sin necesidad de datos cinéticos.



Alimentación de CO ₂	100 kmol/h	Base de referencia para estandarizar condiciones de entrada.
Hidrógeno verde (H ₂)	Variable (según relación molar)	Se ajusta para cumplir con las relaciones H ₂ /CO ₂ definidas.

Esta tabla presenta los parámetros clave de simulación empleados en el modelo termodinámico, incluyendo las variables evaluadas y su justificación técnica, para representar adecuadamente el comportamiento del proceso RWGS en condiciones representativas de operación.

Los resultados se presentan en las Tablas 4, 5 y 6, correspondientes a cada relación molar evaluada:

Tabla 4 Resultados del análisis de sensibilidad 1:1

Temperatura (°C)	Conversión CO ₂ (%)	Relación H ₂ /CO	Fracción molar CH ₄
500.0	24.86	3.13	0.13033412
600.0	30.37	1.71	0.07965244
700.0	41.05	1.25	0.02237608
800.0	48.53	1.04	0.00230433
900.0	52.84	0.89	0.00022101

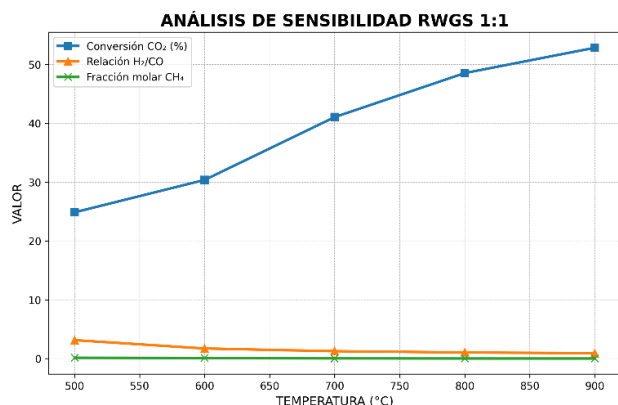


Figura 2 Efecto de la temperatura sobre la conversión de CO₂, relación H₂/CO y fracción molar de CH₄ en el proceso RWGS 1:1

La Tabla 4 y la Figura 2 muestran el comportamiento del sistema para una relación H₂/CO₂ de 1:1. La línea de conversión de CO₂ (%) (azul) aumenta con la temperatura, de ~25 % a ~53 %, confirmando el carácter endotérmico de la reacción; sin embargo, la conversión sigue siendo limitada por la baja disponibilidad de hidrógeno. La línea de relación H₂/CO (morado) disminuye progresivamente, acercándose a 1 a altas temperaturas, lo que refleja un mayor consumo de H₂, aunque este valor no es ideal para procesos como Fischer-Tropsch. Por su parte, la línea de fracción molar de CH₄ (naranja) decrece significativamente hasta valores cercanos a cero, indicando la supresión de la metanación a altas temperaturas. En conjunto, esta relación minimiza la formación de CH₄, pero limita la conversión de CO₂ y la

calidad del syngas, evidenciando la necesidad de mayor proporción de hidrógeno para optimizar el proceso.

Tabla 5 Resultados del análisis de sensibilidad 2:1

Temperatura (°C)	Conversión CO ₂ (%)	Relación H ₂ /CO	Fracción molar CH ₄
500.0	55.75	4.02	0.084170395
600.0	66.26	3.48	0.044381519
700.0	76.27	3.09	0.019955227
800.0	85.36	2.77	0.004372834
900.0	92.20	2.55	0.000544583

En la relación 2:1, mostrada en la Tabla 5 y Figura 3, se alcanza la mayor conversión de CO₂ con una relación H₂/CO que se acerca a la ideal (~2). La fracción molar de CH₄ también disminuye de manera significativa, demostrando que esta relación molar es óptima para eficiencia y selectividad.

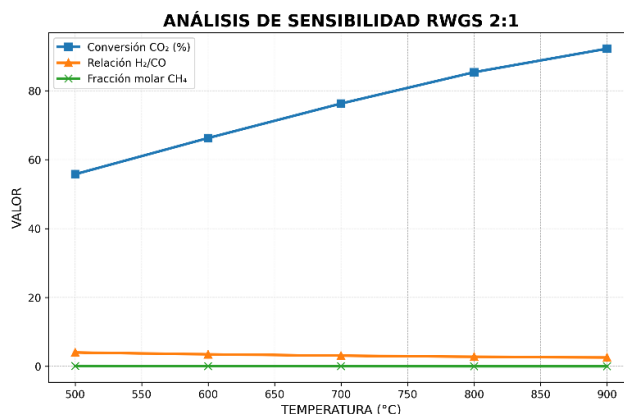


Figura 3 Efecto de la temperatura sobre la conversión de CO₂, relación H₂/CO y fracción molar de CH₄ en el proceso RWGS 2:1

La Tabla 5 y la Figura 3 presentan el comportamiento del sistema para una relación H₂/CO₂ de 2:1. La conversión de CO₂ (línea azul) aumenta de manera sostenida con la temperatura, pasando de 55.75 % a 500 °C hasta 92.20 % a 900 °C, lo que indica una mejora importante en la conversión del sistema conforme aumenta la temperatura. La relación H₂/CO (línea morada) disminuye progresivamente de 4.02 a 2.55, acercándose al valor óptimo (~2) requerido para procesos como Fischer-Tropsch, lo que indica un mejor aprovechamiento del hidrógeno. Por su parte, la fracción molar de CH₄ (línea naranja) decrece significativamente hasta valores prácticamente nulos, reflejando la supresión de la metanación a altas temperaturas. Estos resultados indican que la relación 2:1 permite obtener un equilibrio adecuado entre conversión de CO₂, selectividad y calidad del syngas.

Tabla 6 Resultados del análisis de sensibilidad 3:1

Temperatura (°C)	Conversión CO ₂ (%)	Relación H ₂ /CO	Fracción molar CH ₄
500.0	65.26	19.03	0.229020088
600.0	60.89	7.53	0.159789949
700.0	64.26	4.13	0.075864704



800.0	77.94	2.85	0.002747109
900.0	72.42	3.16	0.018461261

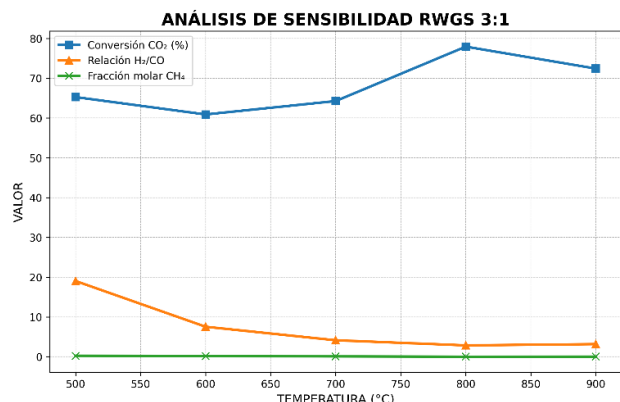


Figura 4 Efecto de la temperatura sobre la conversión de CO₂, relación H₂/CO y fracción molar de CH₄ en el proceso RWGS 3:1

La Tabla 6 y la Figura 4 muestran el comportamiento del sistema para una relación H₂/CO₂ de 3:1, donde se evidencia un exceso significativo de hidrógeno en la alimentación. La conversión de CO₂ (línea azul) presenta un comportamiento no monótono, aumentando con la temperatura hasta un máximo cercano al 78 % a 800 °C y disminuyendo ligeramente a 900 °C, lo que indica efectos de equilibrio y dilución por exceso de H₂. La relación H₂/CO (línea morada) inicia en valores extremadamente altos (≈ 19) a 500 °C y disminuye conforme aumenta la temperatura, aunque se mantiene por encima del rango óptimo, reflejando un uso ineficiente del hidrógeno. Por su parte, la fracción molar de CH₄ (línea naranja) es elevada a bajas temperaturas (≈ 0.23), evidenciando una fuerte presencia de reacciones de metanación, y disminuye considerablemente a temperaturas superiores, aunque no se elimina completamente.

Sin embargo, a 900 °C se observa una ligera disminución en la conversión de CO₂, acompañada de un incremento en la fracción molar de CH₄ respecto a 800 °C. Este comportamiento puede atribuirse al exceso de hidrógeno presente en la alimentación 3:1, el cual modifica el equilibrio global del sistema y aumenta la redistribución de especies químicas bajo condiciones extremas de temperatura. En el modelo de equilibrio RGibbs, la composición final se determina mediante la minimización de la energía libre de Gibbs considerando simultáneamente todas las especies presentes, por lo que el exceso de H₂ puede promover reacciones secundarias y alterar la selectividad hacia CO. Estos resultados evidencian que un incremento adicional de hidrógeno no necesariamente mejora el desempeño del proceso RWGS y puede generar desviaciones respecto al comportamiento esperado a altas temperaturas.

3.3 Comparación general de condiciones

El análisis de sensibilidad realizado permitió identificar el efecto de la temperatura y de la relación molar H₂/CO₂ sobre la conversión de CO₂, la relación H₂/CO y la formación de CH₄

en el proceso RWGS. Se observó que la temperatura es el parámetro más influyente, ya que su incremento contribuye al desplazamiento del equilibrio hacia la formación de CO y reduce significativamente la generación de metano, en concordancia con la naturaleza endotérmica de la reacción.

La relación molar 1:1 presentó las conversiones más bajas debido a la limitada disponibilidad de hidrógeno, restringiendo el avance de la reacción hacia la producción de syngas. Por otro lado, la relación 3:1 generó un exceso de H₂, provocando relaciones H₂/CO desbalanceadas y una mayor formación de CH₄ a bajas temperaturas, lo que disminuye la eficiencia global del proceso y el aprovechamiento del hidrógeno alimentado.

Por el contrario, la relación H₂/CO₂ de 2:1 mostró el comportamiento más favorable, al combinar altas conversiones de CO₂, baja formación de metano y relaciones H₂/CO cercanas al intervalo requerido para procesos downstream como Fischer–Tropsch y síntesis de metanol. Bajo estas condiciones, operar entre 800 y 900 °C permitió alcanzar conversiones de hasta 92.20 %, además de mejorar la calidad del gas de síntesis producido. Por otra parte, la disminución progresiva de la fracción molar de CH₄ conforme aumenta la temperatura evidencia la supresión de reacciones de metanación a altas condiciones térmicas.

A pesar de los beneficios observados, operar a temperaturas elevadas implica un importante requerimiento energético para el calentamiento de las corrientes de alimentación, por lo que resulta necesario considerar estrategias de integración térmica y recuperación de calor para mejorar la viabilidad industrial del proceso. En este sentido, los resultados obtenidos sirven como referencia preliminar para futuras etapas de evaluación y optimización termodinámica de sistemas RWGS orientados a la valorización de CO₂ mediante hidrógeno verde.

3.4 Comparación con trabajos similares

Las tendencias observadas en este estudio coinciden con diversos trabajos recientes enfocados en la producción de syngas mediante la reacción RWGS. Zhu et al. (2020) reportaron que temperaturas superiores a 700 °C mejoran significativamente la formación de CO y reducen la producción de CH₄, comportamiento que coincide con las tendencias observadas en este trabajo. De manera similar, Lykaki et al. (2024) señalaron que el incremento de temperatura desplaza el equilibrio hacia la formación de productos, aumentando la selectividad hacia CO y disminuyendo la contribución de reacciones de metanación.

En cuanto a la relación molar H₂/CO₂, Choi et al. (2024) y Ruggiero et al. (2025) identificaron relaciones H₂/CO cercanas a 2 como adecuadas para procesos downstream orientados a la producción de combustibles sintéticos. En el presente estudio, la relación 2:1 permitió obtener valores cercanos a dicho intervalo, además de alcanzar altas conversiones de CO₂ y baja formación de subproductos, lo que confirma la viabilidad de esta condición para aplicaciones de integración energética y síntesis Fischer–Tropsch.



También, Triviño et al. (2024) destacaron la utilidad de Aspen Plus y del modelo termodinámico Peng–Robinson para evaluar sistemas de valorización de CO₂ bajo condiciones de equilibrio. Los resultados obtenidos en esta investigación coinciden con las tendencias reportadas con dichos trabajos, particularmente en el incremento de conversión conforme aumenta la temperatura y en la reducción progresiva de la formación de metano a altas condiciones térmicas.

Sin embargo, diversos autores indican que los sistemas reales pueden presentar conversiones inferiores a las predichas por modelos de equilibrio debido a limitaciones cinéticas, fenómenos de transferencia de calor y masa, así como efectos asociados a la actividad y estabilidad catalítica (Memon et al., 2023; Khaghani et al., 2025). Por ello, los resultados presentados representan un límite termodinámico ideal del proceso RWGS y constituyen una referencia útil para etapas preliminares de diseño y optimización de sistemas sostenibles de producción de syngas a partir de CO₂ capturado e hidrógeno verde.

4. CONCLUSIONES

Este estudio confirma la viabilidad termodinámica de la producción de gas de síntesis a partir de CO₂ capturado e hidrógeno verde mediante la reacción Reverse Water-Gas Shift (RWGS), empleando un enfoque de simulación en Aspen Plus basado en un reactor de equilibrio RGibbs y la ecuación de estado de Peng–Robinson. Los resultados demuestran que la temperatura es el parámetro dominante, ya que su incremento hacia rangos de 800–900 °C favorece significativamente la conversión de CO₂ y reduce la formación de metano, en concordancia con el carácter endotérmico de la reacción.

Entre las condiciones evaluadas, la relación molar H₂/CO₂ de 2:1 fue la condición de operación más favorable, al combinar altas conversiones de CO₂ (hasta 92.20 %), baja formación de subproductos y una relación H₂/CO cercana al valor óptimo para procesos downstream como la síntesis Fischer–Tropsch o la producción de metanol. En cambio, relaciones más bajas limitaron la conversión debido a la insuficiencia de hidrógeno, mientras que un exceso de H₂ (3:1) generó desbalances en la composición del syngas y favoreció la formación de metano a bajas temperaturas, reduciendo la eficiencia del proceso.

El enfoque termodinámico empleado demostró ser adecuado para la evaluación preliminar del sistema, lo que permitió identificar tendencias en la conversión, selectividad y calidad del gas de síntesis bajo distintas condiciones de operación. En términos generales, los resultados obtenidos permiten identificar condiciones operativas favorables para futuras etapas de diseño y optimización del proceso de valorización de CO₂, confirmando el potencial del proceso RWGS para futuras aplicaciones en valorización de CO₂ y producción de combustibles sostenibles.

Aun así, las condiciones óptimas identificadas implican un elevado requerimiento energético, por lo que su implementación a escala industrial requiere la integración de estrategias de recuperación de calor y optimización energética. Como trabajo futuro, se recomienda la incorporación de

modelos cinéticos, análisis de transferencia de calor y masa, así como simulaciones dinámicas, con el fin de evaluar el comportamiento del sistema bajo condiciones representativas de operación y avanzar hacia su escalamiento industrial.

5. AGRADECIMIENTOS

Los autores agradecen al Tecnológico Nacional de México / Centro Nacional de Investigación y Desarrollo Tecnológico (TecNM/CENIDET) por su apoyo, así como a la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) por la asignación de una beca de doctorado para Marisol Coba Martínez.

REFERENCIAS

- Abbas, S.Z., De Leeuwe, C., Amieiro, A., Poulston, S. and Spallina, V. (2024). Experimental assessment of reverse water gas shift integrated with chemical looping for low-carbon fuels. *Journal of CO₂ Utilization*, 83, 102775. <https://doi.org/10.1016/j.jcou.2024.102775>
- Bahmanpour, A.M., Signorile, M. and Kröcher, O. (2021). Recent progress in syngas production via catalytic CO₂ hydrogenation reaction. *Applied Catalysis B: Environment and Energy*, 295, 120319. <https://doi.org/10.1016/j.apcatb.2021.120319>
- Buthlezi, A.S., Tucker, C., Heeres, H.J., Shoji, M.L., Van de Bovenkamp, H.H. and Ntola, P. (2024). Fischer-Tropsch synthesis using promoted, unsupported, supported, bimetallic and spray-dried iron catalysts: A review. *Results in Chemistry*, 9, 101623. <https://doi.org/10.1016/j.rechem.2024.101623>
- Chen, X., Chen, Y., Song, C., Ji, P., Wang, N., Wang, W. and Cui, L. (2020). Recent advances in supported metal catalysts and oxide catalysts for the reverse water-gas shift reaction. *Frontiers in Chemistry*, 8. <https://doi.org/10.3389/fchem.2020.00709>
- Choi, Y. et al. (2024). Copper catalysts for CO₂ hydrogenation to CO through reverse water–gas shift reaction for e-fuel production: Fundamentals, recent advances, and prospects. *Chemical Engineering Journal*, 492, 152283. <https://doi.org/10.1016/j.cej.2024.152283>
- Corrao, E. et al. (2023). CO₂ conversion into hydrocarbons via modified Fischer-Tropsch synthesis by using bulk iron catalysts combined with zeolites. *Process Safety and Environmental Protection*, 197, 449–465. <https://doi.org/10.1016/j.cherd.2023.07.052>
- Ebrahimi, P., Kumar, A. and Khraisheh, M. (2022). A review of CeO₂ supported catalysts for CO₂ reduction to CO through the reverse water gas shift reaction. *Catalysts*, 12(10), 1101. <https://doi.org/10.3390/catal12101101>
- Hiraga, Y. and Ushiki, I. (2024). Prediction of ionic liquid solubilities in supercritical CO₂ + co-solvent systems using Peng–Robinson equation of state with accurate critical temperature. *Journal of Molecular Liquids*, 398, 124324. <https://doi.org/10.1016/j.molliq.2024.124324>
- Jaubert, J., Qian, J., Lasala, S. and Privat, R. (2022). The impressive impact of including enthalpy and heat capacity



- of mixing data when parameterising equations of state. Application to the development of the E-PPR78 (Enhanced-Predictive-Peng-Robinson-78) model. *Fluid Phase Equilibria*, 560, 113456. <https://doi.org/10.1016/j.fluid.2022.113456>
- Khaghani, A., SaeiMoghaddam, M. and Kheshti, M.F. (2025). Investigation of the liquid-vapor equilibrium in the mixture of carbon dioxide and normal alkanes in binary systems using the Peng-Robinson equation of state modified with mixing rules. *Chemical Thermodynamics and Thermal Analysis*, 100161. <https://doi.org/10.1016/j.ctta.2025.100161>
- Lykaki, M. et al. (2024). State-of-the-art thermocatalytic systems for CH₄ and CO production via CO₂ hydrogenation: Critical comparison, mechanistic considerations and structure-performance insights. *Discover Chemical Engineering*, 4(1). <https://doi.org/10.1007/s43938-024-00048-7>
- Memon, M. et al. (2023). Heterogeneous catalysts for carbon dioxide methanation: A view on catalytic performance. *Catalysts*, 13(12), 1514. <https://doi.org/10.3390/catal13121514>
- Molefe, T., Masilo-Kumi, J., Forbes, R.P. and Coville, N.J. (2024). Fischer-Tropsch synthesis: Platinum promoted Co@HCS catalysts. *Applied Catalysis A: General*, 685, 119863. <https://doi.org/10.1016/j.apcata.2024.119863>
- Paes, F., Privat, R., Jaubert, J. and Sirjean, B. (2024). Predicting solvation energies of free radicals and their mixtures: A robust approach coupling the Peng-Robinson and COSMO-RS models. *Journal of Molecular Liquids*, 401, 124641. <https://doi.org/10.1016/j.molliq.2024.124641>
- Ruggiero, R., Coppola, A., Urciuolo, M. and Scala, F. (2025). Process modeling of the production of synthetic natural gas from biomass-derived syngas: Focus on tar cleaning and fuel synthesis stages. *Fuel*, 393, 134900. <https://doi.org/10.1016/j.fuel.2025.134900>
- Tawalbeh, M. et al. (2023). Unlocking the potential of CO₂ hydrogenation into valuable products using noble metal catalysts: A comprehensive review. *Environmental Technology & Innovation*, 31, 103217. <https://doi.org/10.1016/j.eti.2023.103217>
- Triviño, M.L.T., Arriola, N.C., Kang, Y.S. and Seo, J.G. (2024). Transforming CO₂ to valuable feedstocks: Emerging catalytic and technological advances for the reverse water gas shift reaction. *Chemical Engineering Journal*, 487, 150369. <https://doi.org/10.1016/j.cej.2024.150369>
- Zhou, C., Zhang, J., Fu, Y. and Dai, H. (2023). Recent advances in the reverse water-gas conversion reaction. *Molecules*, 28(22), 7657. <https://doi.org/10.3390/molecules28227657>
- Zhu, M., Ge, Q. and Zhu, X. (2020). Catalytic reduction of CO₂ to CO via reverse water gas shift reaction: Recent advances in the design of active and selective supported metal catalysts. *Transactions of Tianjin University*, 26(3), 172-187. <https://doi.org/10.1007/s12209-020-00246-8>

The Importance of Applying an Adequate Distribution in Wind Speeds

Alfredo, Gutiérrez Alarcón* Emmanuel Olmos Barba* Luisana Claudio Pachecano*

* Centro Nacional de Investigación y Desarrollo Tecnológico,
62490. m24ce090@cenidet.tecnm.mx,
emmanuel.ob@cenidet.tecnm.mx, luisana.cp@cenidet.tecnm.mx

Abstract: Accurate characterization of wind resources is essential not only for estimating energy production but also for ensuring the structural reliability and operational safety of wind turbines. The Weibull distribution has traditionally been adopted as the standard probabilistic model for wind speed analysis due to its simplicity and its satisfactory performance in representing average wind conditions. However, in regions characterized by complex atmospheric behavior and high-intensity wind regimes, such as the Isthmus of Tehuantepec, the Weibull model may fail to adequately capture extreme wind events. This study evaluates the importance of selecting an appropriate probabilistic distribution for wind resource assessment by comparing the conventional Weibull model with approaches derived from Extreme Value Theory (EVT), specifically the Generalized Extreme Value (GEV) distribution and the Peaks over Threshold (POT) method. Wind speed data collected from four wind farms, La Ventosa I, La Ventosa II, La Venta, and Ciudad Ixtepec, were analyzed to assess the capability of each model to represent the upper tail of the distribution, corresponding to wind speeds exceeding the third quartile. Model performance was evaluated using statistical error metrics, including the Mean Absolute Percentage Error (MAPE) and the Mean Squared Error (MSE). The results indicate that the use of inadequate probabilistic models for extreme wind characterization may lead to significant technical risks, including the underestimation of critical aerodynamic loads and the accelerated fatigue of turbine components. Among the evaluated methods, the POT approach demonstrated superior performance in modeling extreme wind behavior. Consequently, this study recommends a hybrid modeling strategy in which Weibull-based models are employed for operational and energy production analyses, while EVT-based methods, particularly POT, are applied for structural reliability and extreme load assessment. This combined approach contributes to improving both the technical robustness and the financial viability of wind energy projects developed in regions with extreme wind conditions.

Keywords: Extreme Value Theory, wind speeds, wind energy, Probability distribution.

1. INTRODUCCIÓN

En el campo de la energía eólica, el Análisis de Valores Extremos es usado para predecir valores máximos de la velocidad del viento llamadas comúnmente como ráfagas de viento, las cuales son conocidas por causar daño a la infraestructura de aerogeneradores, Cheng & Yeung (2002).

En México la capacidad instalada para la generación de energía eólica incrementó hasta en 6238 MW en el 2019. El Istmo de Tehuantepec, ubicado en Oaxaca México, cuenta con el 62% de esta capacidad, gracias a su abundante recurso eólico. Esta región se caracteriza por sus fuertes ráfagas de viento que llegan a volcar vehículos pesados en temporadas de huracanes o de frentes fríos. En el caso de los aerogeneradores, la subestimación de la velocidad de viento puede implicar ignorar esfuerzos y deformaciones que lleguen a comprometer la seguridad y vida útil de los mismos (Lorenzo et al., 2016).

El análisis del impacto de las velocidades del viento sobre las estructuras que conforman los aerogeneradores requiere un estudio detallado de los datos disponibles, para ello se desarrolló el siguiente trabajo, con el fin de caracterizar las

velocidades extremas del viento mediante distribuciones basadas en la teoría de los valores extremos.

2. ESTADO DEL ARTE Y FUNDAMENTOS TEÓRICOS

Un estudio más con enfoque estadístico que busca estimar el potencial eólico en una región específica es el de Prado & Quisirumbay (2024), se centró en el estudio de la variabilidad del recurso eólico en la región de la Central Eólica Villonaco, ubicada en Loja, Ecuador. Su objetivo principal fue analizar la velocidad del viento, sus patrones estacionales y tendencias a lo largo del tiempo, utilizando datos históricos recopilados durante cuatro años (2020-2023) a una altura de 60 metros. Para completar datos faltantes, y un análisis estadístico detallado para cuantificar las características del viento, utilizaron la función de distribución de Weibull para identificar patrones y aplicaron técnicas de descomposición de series temporales para distinguir componentes de tendencia y estacionalidad. Dentro de los resultados que obtuvieron fueron que se identificaron patrones de variabilidad en la velocidad del viento a través de análisis diarios, semanales y mensuales. Además, se calculó la media, mediana y desviación estándar



de las velocidades del viento, lo que permitió evaluar la energía eólica disponible y su variabilidad a lo largo del tiempo.

Así mismo, estudiando el comportamiento del viento desde un enfoque estadístico y probabilístico usado para evaluar el recurso eólico, Yancachajlla Tito & Vilca Huayta (2022) analizaron el potencial de la energía eólica en Laraqueri, un área rural del sur de Perú que carece de acceso a la red eléctrica. Se utilizaron datos de viento de 2020 a una altura de 10 metros, se aplicaron dos métodos numéricos para estimar los parámetros de la función de distribución de Weibull. Los resultados mostraron que, aunque la densidad de potencia promedio es baja (3.5 m/s y 61.89 W/m²), hay meses con potencial más alto, como agosto, donde se registraron valores de hasta 123.40 W/m². Esto sugiere que, aunque no es adecuado para turbinas rápidas ya que estas son turbinas operan en condiciones de viento extremadamente muy fuerte, por lo tanto, el viento puede ser aprovechado para turbinas de eje vertical ya que estas turbinas son menos sensibles a la dirección del viento, especialmente para aplicaciones como la extracción de agua. El estudio resalta la importancia de las energías renovables en la búsqueda de soluciones sostenibles para el suministro energético en áreas aisladas.

Utilizando la distribución Weibull para el estudio del viento, particularmente en el análisis de la velocidad del viento y la modelización de la energía eólica, (Palavecino et al., 2023) analizaron la intensidad del viento en la estación SAEZ (Ezeiza) durante un periodo de 11 años, desde 2009 hasta 2019, utilizando datos de vientos.

Por último, el trabajo de Castillo, E. (2012) tuvo como objetivo evaluar y caracterizar el recurso eólico en la Universidad Francisco de Paula Santander (UFPS) en Cúcuta, Colombia, utilizando los datos meteorológicos que se recopilaban durante un periodo de diez años. Además, se analiza la viabilidad del aprovechamiento energético del viento en la región de Norte de Santander. Para ello se utilizaron registros en horarios de variables meteorológicas como temperatura, presión, velocidad y dirección del viento, todas estas obtenidas de la estación meteorológica de la UFPS. Los resultados indicaron que la viabilidad de un emplazamiento eólico depende de la velocidad media del viento, se analiza la importancia de la variabilidad horaria y mensual, así como la aplicación de la función de densidad de probabilidad de Weibull para calcular el potencial energético de los aerogeneradores. Como conclusión en el estudio mencionan que en algunos lugares la velocidad media del viento puede ser baja, existen momentos en el día donde se pueden registrar incrementos significativos, lo que puede hacer viable la instalación de aerogeneradores.

3. METODOLOGÍA Y DESARROLLO

La metodología para el análisis estadístico de las 4 granjas consistió en un preprocesamiento de datos descrito en la

sección 3.1 en donde finalmente se obtienen 4 datos datos set para cada estación.

Se hace una caracterización estadística descriptiva de cada granja. En la sección 3.2 se muestran los datos representados a través del tiempo (fig. 2.) y los histogramas, en los que revelan las ocurrencias de distintos intervalos de velocidades (fig. 3).

A partir de estas características se procederá a analizar distintos tipos de distribuciones provenientes de la Teoría de los valores extremos (TVE) para caracterizar las velocidades más altas. En la sección 3.3, Como primera instancia, se modela con una distribución de Weibull las 4 granjas como referencia de comparación.

En la sección 3.4 y 3.5 se hace una comparación del modelado de las distribuciones de velocidad con las distribuciones Generalizada de Valores Extremos (GEV) y Distribución Picos por encima del umbral (POT) que tienen su origen, al igual que la distribución de Weibull en la teoría de valores extremos.

3.1 Localización de cada granja y creación de cada data-set.

En la figura 1 se muestra como están distribuidas las granjas eólicas y las distancias aproximadas entre ellas. Como podemos observar la distancia ubicada al oeste (izquierda) y las granjas centrales es de 14 km, y la distancia entre la granja central y la granja eólica ubicada al este (derecha) es de 16 km, la distancia total entre la granja eólica del oeste y la granja del este es de 30 km. Sin embargo, se analizaron sus diferencias estadísticas entre cada una de ellas.

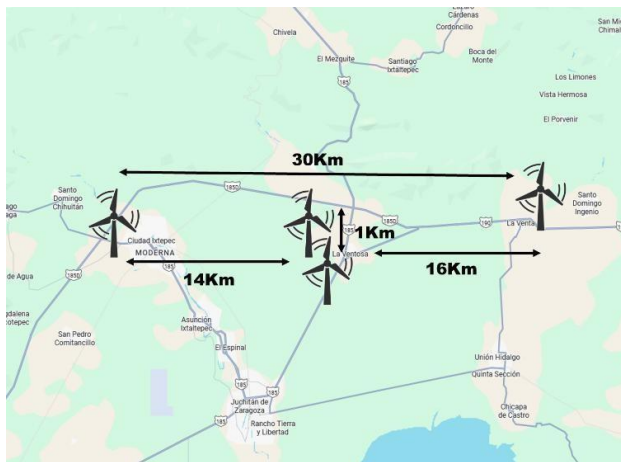


Fig. 1. Ubicación y distancia entre granjas: De izquierda a derecha: Se encuentra la estación Ixtépec, después las estaciones Ventosa 1 y Ventosa 2 las cuales tienen una distancia aproximada de un 1km. La última estación es La Venta.

Previo a la creación a la data set, se hizo una evaluación de los datos, en los cuales se encontraron algunos datos faltantes dando, parte de este proceso tenemos que la estación de la venta 612 datos faltantes, la estación de la ventosa uno: 377 datos faltantes, la ventosa dos: 277 datos



faltantes y por último la estación Cd Ixtepec: 949 datos faltantes. Para continuar con la creación de la data set y tomando en cuenta el análisis previo se determinó que no afectarían a los data set porque cada estación no supera el 5% de los datos totales registrados.

Además, mencionar que cada granja cuenta con datos generales, parte de estos datos son: estado, región, nombre de la estación, año en que fueron tomados y alturas de medición. Esto se aplicó a las cuatro granjas eólicas de esta región.

3.2 Generación de series temporales e histogramas.

En la figura 3 se muestran las series temporales de las 4 granjas eólicas en donde se marcó en línea roja el promedio, los cuales son: para la ventosa uno 8.169, la ventosa dos 8.627, Cd Ixtepec 6.238 y por último la venta con un promedio de 9.130.

Se pueda notar que considerando sólo el promedio de cada granja no representa la variabilidad de las series a través de todo el año. Una caracterización mensual de los datos se puede apreciar en la figura 4, en donde se grafican los histogramas de cada granja. Los histogramas muestran de forma ordenada las frecuencias de ocurrencias en los intervalos de velocidades. Los colores representan en qué mes ocurrieron dichas velocidades. Se muestran en línea punteada roja el primer cuartil en el cual representa que el 25% de los datos se encuentran hacia la izquierda del histograma. La línea azul punteada representa el inicio del tercer cuartil el 75% de los datos hacia la izquierda del histograma. A partir de esa línea hacia la derecha se muestran las velocidades menos probables de ocurrencia, lo que se puede observar es que está línea sirve de referencia para describir la cola de cada distribución que pudiera modelar estos datos.

Día	Hora	Rapidez del nivel inferior	Rapidez máxima del nivel inferior
1	1	10	14.05
2	1	20	13.48
3	1	30	13.19
4	1	40	13.38
5	1	50	13.93
6	1	100	14.31
7	1	110	14.01
8	1	120	14.04
9	1	130	13.70
10	1	140	13.50
11	1	150	12.71
12	1	200	12.73
13	1	210	12.85
14	1	220	13.24

Fig. 2 Ejemplo de un data set para la generación de gráficas y análisis estadístico, Variables de la data set.

3.3 Caracterización de la distribución de Weibull

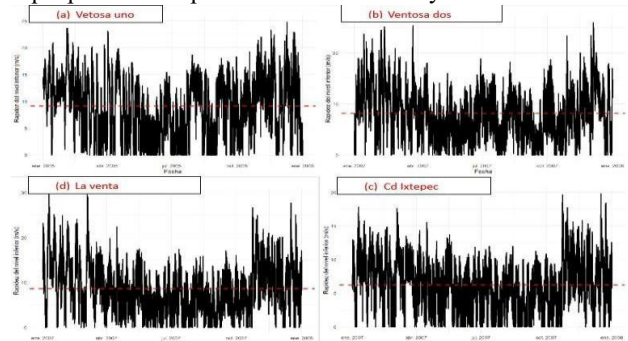
La función de densidad de distribución de Weibull está en función de dos parámetros: Escala y forma.

$$f(x) = \frac{k}{\lambda} \left(\frac{x}{\lambda}\right)^{k-1} \exp \left[-\left(\frac{x}{\lambda}\right)^k \right], \quad x \geq 0$$

Como observamos en la formula, x representa la velocidad del viento, $f(x)$ nos indica la probabilidad de que el viento tenga esa velocidad, λ (lambda) es el parámetro escala y nos indica que tan fuertes son los vientos en promedio, si lambda aumenta, generalmente las velocidades del viento son mayores, si lambda disminuye, los vientos suelen ser mas bajos. El parámetro de k (forma) controla la forma de la gráfica, continuando con la parte exponencial, (exp) esta parte hace que la probabilidad disminuya para velocidades muy grandes, porque esas velocidades ocurren menos veces. En conclusión, esta fórmula combina todos los elementos para crear una curva en los histogramas.

Para los parámetros de escala, y forma (Fig.5), y entenderlos más a detalle, tenemos que el parámetro k (forma) es el que le da forma a la curva de la densidad de distribución, este parámetro le da pico, cola y simetría, mientras que el parámetro λ (escala) es el que tira o comprime horizontalmente la curva.

En la Figura 6 se ilustra el ajuste de la distribución de Weibull sobre el histograma de frecuencias para cada estación. La curva roja representa el modelo propuesto para describir los datos empíricos ordenados como histograma. Los resultados muestran que este ajuste es relativamente aceptable para la Ventosa 2 con 16.98 % de error. Sin embargo, para las demás estaciones la distribución de Weibull no es representativa de los histogramas. Por ejemplo, para Cd Ixtepec, podemos notar que tiene una asimetría hacia la derecha, con un skewness de 0.14, en comparación de las demás granjas con un Skewness menor. La venta de igual forma presenta un skewness de -0.04 esto nos indica que está un poco más sesgado a la izquierda, lo que hace que la densidad de distribución Weibull no sea adecuada para esta granja, por último, tenemos la ventosa uno con un skewness de 0.43, aunque con este valor se ajusta más a la distribución, aun podemos observar en la (Fig. 6.) que presenta un poco de subestimación y sobrestimación en



histograma.

Fig. 3 Series temporales del nivel inferior. (a) La Ventosa uno, en esta granja las mediciones se hicieron a una altura de 20 metros. (b) La ventosa dos, en esta granja las mediciones se hicieron a una

altura de 40 metros. (c) Cd Ixtepec, en esta granja las mediciones se hicieron a una altura de 20 metros, (d) La venta, en esta granja las mediciones se hicieron a una altura de 15 metros.

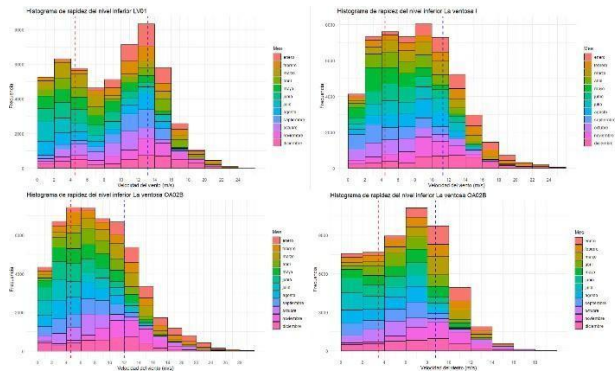


Fig. 4 Histogramas. (a) La Venta. El primer el cuartil se encuentra en (línea roja punteada) con un valor de 4.468. El tercer cuartil en (línea azul punteada) con valor de 13.221, (b) La ventosa dos. El primer el cuartil se encuentra en (línea roja punteada) con un valor de 4.597. El tercer cuartil en (línea azul punteada) con valor de 12.000. (c) Cd Ixtepec. El primer el cuartil se encuentra en (línea roja punteada) con un valor de 3.463. El tercer cuartil en (línea azul punteada) con valor de 8.770. (d) La venta. El primer el cuartil se encuentra en (línea roja punteada) con un valor de 4.414. El tercer cuartil en (línea azul punteada) con valor de 11.340.

El análisis realizado demuestra que la distribución de Weibull proporciona un ajuste estadístico adecuado solo en ocasiones cuando las gráficas son unimodales, pero observamos que para gráficas bimodales la distribución de Weibull no logra graficar adecuadamente.

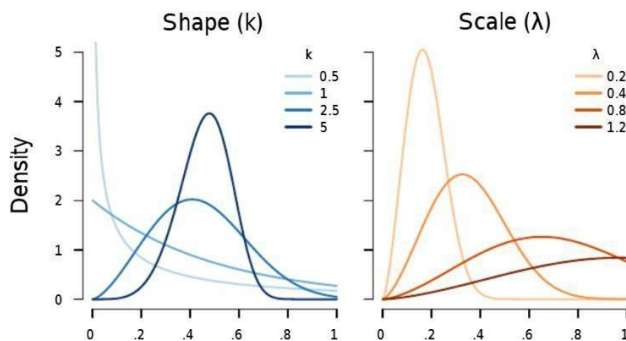
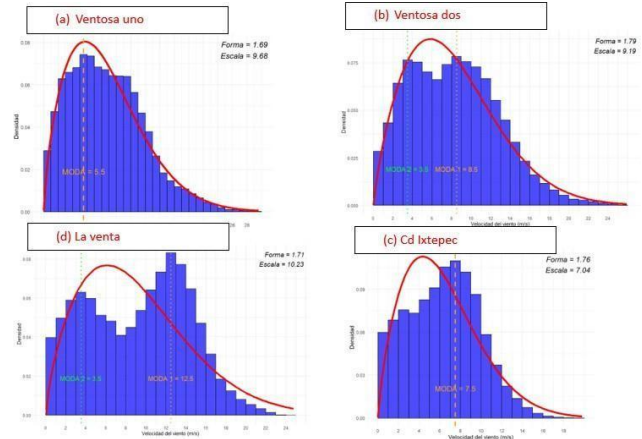


Fig.5 Parámetros forma y escala. (a) Forma, cuando tiene un valor de 0.5 es una distribución muy sesgada, cuando es valor es de 1 su distribución es exponencial, cuando el valor es de 2.5 comienza a tener moda, por último, cuando el valor es de 5 tiene un pico marcado y es parecida a la gaussiana. (b) Escala, cuando escala tiene valor de 0.2 tiene pico muy alto y angosto, cerca de 0, cuando escala tiene valor de 0.4 es un pico más bajo, pero aun a la izquierda, cuando escala tiene valor de 0.8 es una curva más ancha, y el pico se desplaza a la derecha, por último, cuando escala tiene valor de 1.2 es una curva extendida y baja.



A continuación, se presenta la distribución Generalizada de Valores Extremos (GEV) cuyos parámetros podrían ayudar a describir histogramas con asimetrías como la estación de Ixtepec.

Fig. 6 Distribución de Weibull. (a) El ajuste de Weibull para la ventosa dos en donde se encontró: parámetros de forma:1.69 y escala: 9.68 (b) la ventosa uno parámetros forma: 1.79 y escala: 9.19, (c) Cd Ixtepec, parámetros forma: 1.76 y escala: 7.04, (d) la venta, parámetros forma: 1.71 y escala: 10.23

3.4 Aplicación de la distribución Generalizada de Valores Extremos (GEV)

Se aplico la distribución generalizada de valores extremos (GEV) basada en la teoría de los valores extremos para poder modelar con más precisión los histogramas de frecuencias observados en cada una de las granjas.

La robustez de esta distribución reside en su formulación triparamétrica. Al igual que Weibull tiene los parámetros de forma y escala, sin embargo, se diferencia con un nuevo parámetro de localización

$$\frac{1}{\sigma} t(x)^{\xi+1} e^{-t(x)}$$

$$\text{donde } t(x) = \begin{cases} \left[1 + \xi \left(\frac{x-\mu}{\sigma} \right) \right]^{-1/\xi} & \text{if } \xi \neq 0 \\ \exp \left(-\frac{x-\mu}{\sigma} \right) & \text{if } \xi = 0 \end{cases}$$

Como observamos la fórmula de la distribución generalizada de valores extremos, x Es el dato que se está analizando, por ejemplo, la velocidad extrema del viento. μ (mu) representa el valor central de los datos extremos, σ (sigma) indica qué tan dispersos están los datos, si sigma es grande, los valores extremos varían mucho, por último, ξ (xi) Es el parámetro más importante porque define la forma de la distribución.

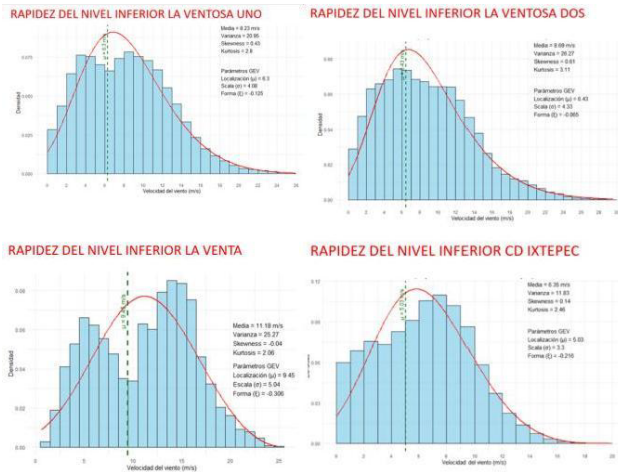


Figura 7. Distribución GEV y parámetro de localización (línea verde). (a) La ventosa uno (b) ventosa dos, (c) Cd Ixtepec (d) La venta

Continuando, tenemos que el parámetro de localización controla la posición de la distribución, causa un efecto en la gráfica, desplaza la curva a la izquierda o derecha dependiendo el valor. Por ejemplo, valores de μ con igual a 4 o menores de 4 la curva se desplaza a la izquierda y tiende a estar cercano a cero. Cuando μ es igual a 6, la curva tiende a desplazarse a la derecha, es una curva un poco más centrada y con un pico un poco más pronunciado. Por último, cuando μ es igual a 8, la curva tiende a desplazarse a la derecha, lo que hace que la curva se haga más larga y este pegada a los valores extremos. Esto permite una convergencia óptima hacia las familias de distribuciones Gumbel, Fréchet o Weibull.

En la tabla 1 podemos observar los valores de localización, forma y escala de cada granja eólica.

Tabla 1. Valores de parámetros de cada granja.

Parámetro	Ventosa uno	Ventosa dos	Cd Ixtepec	La venta
Localización	6.3	6.43	5.03	9.45
Forma	-0.125	4.33	3.3	5.04
Escala	4.08	-0.065	-0.216	-0.306

Como conclusión el parámetro de localización mejora en el caso de asimetrías con sesgo hacia la derecha, caso de Ixtepec, sin embargo aun presenta un porcentaje de error elevado (21% en GEV & 35% en Weibull), aunque también existe una mejora para el caso bimodal como lo es en la venta, aún tenemos un porcentaje de error elevado (27% en GEV y 22% en Weibull), esto nos indica que este modelo aun no logra ajustar con claridad la curva en los histogramas, por lo que se sugiere aplicar otra distribución basada en la teoría de los valores extremos.

3.5 Aplicación de la Distribución Picos por encima del umbral (POT)

El método Peaks Over Threshold (POT) es una distribución que pertenece a la Teoría de Valores Extremos, utilizada para analizar eventos extremos dentro de un conjunto de datos.

$$f_{\mu, \sigma, \xi}(x) = \begin{cases} \frac{1}{\sigma} \left(1 + \xi \frac{x - \mu}{\sigma} \right)^{-1 - 1/\xi} & \text{for } \xi \neq 0, \\ \frac{1}{\sigma} \exp\left(-\frac{x - \mu}{\sigma}\right) & \text{for } \xi = 0, \end{cases}$$

x , Es el valor que se quiere analizar, en este caso la velocidad del viento, μ (μ) es el parámetro de ubicación, indica el nivel central o promedio de los valores, σ (σ) Es el parámetro de escala, describe qué tan dispersos o variables son los datos, ξ (ξ), es el parámetro de forma, controla el comportamiento de los valores extremos y define la forma de la distribución, la fórmula cambia dependiendo del valor de ξ : Cuando $\xi \neq 0$, la distribución puede adaptarse a distintos tipos de eventos extremos, cuando $\xi = 0$, la fórmula se simplifica y toma una forma exponencial.

A diferencia de otros enfoques que consideran únicamente los máximos, el método POT se enfoca en todos los valores que superan el umbral previamente definido. En este trabajo se tomó como referencia el parámetro del 3 cuartil. Como se puede observar en la figura 4, a partir de este valor, observamos un decaimiento de la probabilidad de encontrar velocidades mayores. En la tabla 2, se muestran los valores de referencia para modelar con la distribución POT.

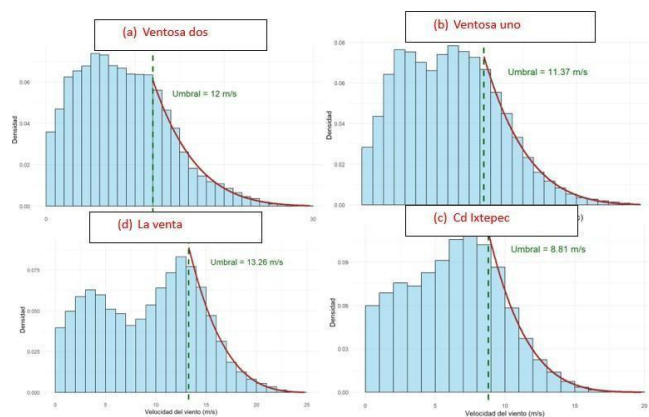


Figura 8 Distribución Picos por encima del umbral (POT), (a) Ventosa uno, (b) ventosa dos, (c) Cd Ixtepec, (d) la venta.



Tabla 2. Valores umbrales de cada estación.

Estación	Ventosa dos	Ventosa uno	Cd Ixtepec	La venta
Umbral	12 m/s	11.37 m/s	8.81 m/s	13.26 m/s

En la granja de Cd Ixtepec y la venta el ajuste de la distribución POT en comparación con Weibull y GEV (fig. 6 y 7), existe una mejora visualmente evidente. Esta mejora también queda registrada en las tablas 3, 4, 5, 6, en donde las métricas de error muestran una mejora.

3.6 Comparación de resultados de cada una de las distribuciones

Para comparar los resultados de cada una de las distribuciones aplicadas se procedió a aplicar tres indicadores de error, el primero es el indicador MAPE, seguido del indicador MAE y por último el indicador MSE.

La tabla 3 presenta una comparación del desempeño de los modelos Weibull, GEV y POT mediante tres métricas de error estadísticas: MAPE, MAE y MSE, evaluadas en el tercer cuartil (Q3). Estas métricas permiten determinar qué tan bien cada modelo representa los datos observados de velocidad del viento, especialmente en valores extremos

El modelo Weibull presenta los valores de error más elevados en las tres métricas analizadas. Su MAPE de 26.85 % indica que, en promedio, las estimaciones del modelo se desvían aproximadamente un 27 % de los valores reales, lo cual representa una precisión relativamente baja. Por otro lado, el modelo GEV (Generalized Extreme Value) presenta una mejora considerable respecto al modelo Weibull. El valor de MAPE disminuye hasta 11.97 %, indicando una representación mucho más cercana a los datos reales. El modelo POT (GPD) es el que presenta el mejor desempeño estadístico de toda la comparación, ya que obtiene los valores más bajos en las tres métricas de error. Su MAPE de 10.33 % indica el menor error porcentual promedio, demostrando una alta precisión en las estimaciones.

Tabla 3. Comparación de resultados La Ventosa Uno

Modelo	MAPE (Q3 %)	MAE Q3	MSE Q3
Weibull	26.85%	3.598×10^{-3}	2.5×10^{-5}
GEV	11.97%	2.497×10^{-3}	1.5×10^{-5}
POT (GPD)	10.33%	1.023×10^{-3}	1×10^{-6}

La tabla 6 de igual forma presenta una comparativa de los modelos GEV, POT y Weibull evaluados en el tercer cuartil (Q3), que es donde se encuentran las velocidades de viento más altas y críticas. El modelo POT es el ganador indiscutible en las tres métricas. tiene el MAPE más bajo (19.91%), lo que indica que es el modelo que menos se desvía porcentualmente de los datos reales, su MSE es extremadamente bajo (0.000004), lo que significa que el

modelo casi no presenta errores grandes. El modelo GEV ocupa el segundo lugar en precisión porcentual (MAPE de 22.61%). Sin embargo, su MSE (0.001016) es el más alto de los tres, lo que sugiere que, aunque en promedio no es tan malo, tiene puntos específicos donde la predicción se aleja considerablemente de los datos observados. Weibull presenta el MAPE más alto (27.64%) y el MAE más elevado. Esto confirma estadísticamente lo que vimos en las gráficas: la distribución Weibull suele ser demasiado general y no logra "estirarse" lo suficiente para capturar con precisión los vientos del tercer cuartil. Tras evaluar la precisión estadística de los tres modelos en el régimen de vientos del tercer cuartil (Q3), se concluye que el modelo POT (GPD) es el más robusto para este conjunto de datos. Presenta un error porcentual (MAPE) de apenas 19.91% y un MSE de 0.000004. Estos resultados validan el uso de POT para la estimación de valores extremos en el sitio analizado, superando la capacidad predictiva de los modelos GEV y Weibull.

Continuando con el análisis de los resultados, la tabla 4 evalúa el desempeño de los modelos en el tercer cuartil (Q3) de la granja la Ventosa Dos que corresponde al 25% de los datos con vientos más veloces. El modelo de Picos sobre el Umbral (POT), es el más preciso de los tres, tiene el error porcentual más bajo (12.30%) y el MSE más pequeño, lo que demuestra que es el modelo que mejor sigue la forma de la cola del histograma sin generar desviaciones grandes. El modelo GEV se mantiene como segunda opción con un error del 14.36%. Aunque es ligeramente menos preciso que POT en los extremos, su desempeño es significativamente mejor que el de Weibull. Como es común en sitios de vientos fuertes o complejos, la distribución Weibull es la que presenta mayores errores (16.98% en MAPE). Esto confirma que Weibull tiende a ser demasiado generalista y pierde precisión cuando se trata de modelar las velocidades más altas del viento. Tras el análisis estadístico de los errores en el tercer cuartil (Q3), se determina que el modelo POT es el que presenta el mejor ajuste para los datos observados, logrando una reducción del error porcentual (MAPE) de casi 5 puntos respecto a la distribución Weibull tradicional. El bajo valor en MSE (0.000002) asegura que este modelo es altamente confiable para predecir comportamientos de viento.

Tabla 4. Comparación de resultados La Ventosa Dos

Modelo	MAPE (Q3 %)	MAE Q3	MSE Q3
Weibull	16.98%	2.7166×10^{-3}	1.5×10^{-5}
GEV	14.36%	2.4749×10^{-3}	1.2×10^{-5}
POT (GPD)	12.30%	1.5151×10^{-3}	2×10^{-6}

Para continuar con la comparación de resultados, en la tabla 5 podemos observar un dominio significativo del método POT, es el único modelo con un error porcentual de un solo dígito (9.70%). Esto es una precisión excelente para datos meteorológicos extremos. Además, su MSE es casi despreciable (0.000002), lo que indica un ajuste casi perfecto a la cola de la distribución. Seguido del modelo



GEV donde observamos que, aunque la distribución de Valores Extremos Generalizada (GEV) es funcional, su error porcentual (21.17%) es más del doble que el de POT. Esto sugiere que, para este sitio, GEV no logra capturar la forma exacta de los vientos máximos tan bien como el método de umbral (POT).

Por último, tenemos la distribución Weibull donde presenta el error más alto en todas las categorías, con un MAPE del 35.27%. Un error superior al 30% en el tercer cuartil esto nos indica que Weibull es un modelo inadecuado para describir los vientos fuertes en esta ubicación, probablemente porque subestima drásticamente la frecuencia de las ráfagas extremas del viento.

Tabla 5. Comparación de resultados CD. Ixtepec

Modelo	MAPE (Q3 %)	MAE Q3	MSE Q3
Weibull	35.27%	7.122×10^{-3}	1.16×10^{-4}
GEV	21.17%	4.1381×10^{-3}	3×10^{-6}
POT (GPD)	9.70%	1.3491×10^{-3}	2×10^{-6}

Tabla 6. Comparación de resultados La Venta

Modelo	MAPE (Q3 %)	MAE Q3	MSE Q3
Weibull	27.64 %	7.113×10^{-3}	1.99×10^{-4}
GEV	22.61%	5.143×10^{-3}	1.6×10^{-5}
POT (GPD)	19.91%	1.848×10^{-3}	4×10^{-6}

En las cuatro tablas analizadas, el modelo POT obtuvo consistentemente el MAPE más bajo (entre el 9% y el 19%). Esto demuestra que, para el análisis de vientos en regiones ráfagas intensas, el enfoque de umbral es estadísticamente superior a los modelos que intentan ajustar todos los datos a la vez. Por otro lado, se observa que el modelo Weibull presenta sistemáticamente los errores más altos en el tercer cuartil llegando hasta un 35.27% en uno de los casos. Esto sugiere que el uso de Weibull para calcular ráfagas de diseño o cargas estructurales en estos sitios específicos introduciría un nivel de incertidumbre (riesgo) inaceptable, tendiendo a subestimar el recurso eólico extremo. Para finalizar El modelo GEV se posiciona como una opción intermedia. Aunque es superior a Weibull, no alcanza a superar el ajuste de POT. Su utilidad queda relegada a casos donde no se disponga de un umbral de datos claramente definido, pero para propósitos de precisión técnica, POT sigue siendo la recomendación principal.

1. RESULTADOS Y DISCUSIÓN

El análisis comparativo de las distribuciones aplicadas para la rapidez del viento en las granjas estudiadas nos permite establecer un marco de precisión basado en métricas de error. A continuación, se detallan los resultados obtenidos. determinó que la distribución Weibull, aunque es el estándar de la industria para la caracterización general del recurso eólico, presenta limitaciones críticas en sitios con regímenes de viento bimodales o de alta energía. Los resultados nos muestran que Weibull tiende a subestimar la frecuencia de

los vientos en el tercer cuartil, alcanzando errores porcentuales (MAPE) de hasta el 35.27%.

Por otra parte, la distribución de Valores Extremos Generalizada (GEV) demostró una mayor flexibilidad estructural que Weibull, logrando mejores ajustes en la zona de la moda; sin embargo, su capacidad predictiva se ve comprometida ante la presencia de valores atípicos extremos, reflejado en valores de MSE superiores a los del modelo POT. El modelo POT obtuvo un desempeño más sobresaliente en todas las granjas evaluadas enfocándose únicamente en los valores que superan el umbral, este método logra reducir el MAPE y un MSE garantizando el buen modelado de la distribución. Se concluye que el método POT es la herramienta estadística más robusta para el análisis de vientos extremos en las zonas de estudio. Su capacidad para modelar la cola de la distribución sin verse influenciada por los vientos de baja energía lo hace superior a los modelos de ajuste global, se sugiere que el uso exclusivo de la distribución Weibull en sitios como La Venta o Ciudad Ixtepec conlleva un riesgo técnico elevado, debido a su tendencia sistemática a subestimar las velocidades de viento extremas. Esto podría derivar en un diseño estructural de los aerogeneradores. Además, la presencia de distribuciones bimodales en los datos observados resalta la necesidad de implementar modelos híbridos o de valores extremos. La distribución GEV se presenta como una alternativa válida para la caracterización de vientos medios y altos, pero debe ser complementada con la distribución POT. La precisión ganada mediante el modelo POT que reduce el error porcentual en más de 20 puntos respecto a Weibull en algunos casos permite optimizar el cálculo de las cargas de diseño y la selección de tecnología de aerogeneradores, aumentando la viabilidad técnica y financiera de los proyectos eólicos en regiones con potencial extremo.

REFERENCIAS

- Castillo, E. (2012) *Extreme Value Theory in Engineering*. Amsterdam: Elsevier.
- Cheng, E. and Yeung, C.Y. (2002) 'Generalized extreme gust wind speeds distributions', *Journal of Wind Engineering*, pp. 1657–1669.
- Durañona, V. (2008) *Actualización de la estadística de vientos extremos para Uruguay*.
- Lassig, J.L., Palese, C. and Apcarian, A. (2011) *Vientos extremos en la provincia de Neuquén*, 36.
- Lorenzo, I.F., Beatriz, V. and Parnás, E. (2016) 'Análisis de métodos de vientos extremos para calcular las velocidades básicas', *Revista Cubana de Ingeniería*, 7(2).
- Palavecino, M.A., Godoy, A.A., Dillon, M.E. and de los Milagros Skansi, M. (2023) 'Characterization of wind intensity at the Ezeiza station', *Meteorologica*, 48(2).



- Prado León, W.A. and Quisirumbay Ludeña, R.S. (2024) Caracterización y análisis de la variabilidad del viento en la zona de influencia de la central eólica Villonaco. Bachelor's thesis.
- Wang, J., Qin, S., Zhou, Q. and Jiang, H. (2015) 'Medium-term wind speed forecasting utilizing hybrid models for three different sites in Xinjiang, China', *Renewable Energy*, 76, pp. 91–101.
- Yancachajlla Tito, U. and Vilca Huayta, O.A. (2022) 'Caracterización del viento con la función de Weibull para una zona altoandina, Laraqueri - Perú', *Revista de Investigaciones Altoandinas*, 24(3), pp. 190–198.
- Bortoli, D., Canavesio, M.E., Benitez, O. and Makuch, F. (n.d.) Análisis de velocidades de viento atmosféricas medias horaria anual extrema a través del análisis probabilístico de registros cortos.
- Natalini, B. and Bruno Natalini, M. (n.d.) 'Review of extreme wind speeds in three sites of southern Argentina', *Accidentes e Infraestructura Civil*, 13(2).

A Review of Electrical Stimulation Therapy (EST) as an Adjunct to Myofascial Release Therapy (MFRT) for Posterior Lower-Leg Muscle Pain.

R. T. Castellanos-Baltazar * H. R. Azcaray-Rivera *
V. Ortiz-Méndez * J. A. Velasco-Cruz * P. A. López Ramírez *

** Instituto Tecnológico de Oaxaca, Departamento de Ingeniería Eléctrica y Electrónica, Oaxaca de Juárez, Oax. 68030 Mexico (e-mail: roberto.cb@oaxaca.tecnm.mx, hector.ar@oaxaca.tecnm.mx, virginia.om@oaxaca.tecnm.mx, jose.vc@oaxaca.tecnm.mx, 21160375@itoaxaca.edu.mx).*

Abstract: This article compiles evidence related to Myofascial Release Therapy (MFRT) and, through its analysis, describes findings ranging from pain assessment records to current therapeutic approaches. In this context, the implementation of electrostimulation devices is presented as an auxiliary tool within improvement strategies for Myofascial Release Therapy (MFRT) applied to musculoskeletal structures, considering documented evidence from both patients with chronic pain and athletes across various high-performance disciplines. Due to the lack of conclusive information, the possibility arises of establishing a methodology for the design of measurement instruments that allows these findings to be documented and recorded.

Keywords: Myofascial Release, Chronic Pain, Electrical Stimulation, High Performance, Musculoskeletal System.

1. INTRODUCCIÓN

La creación de este artículo se centra en el análisis de los antecedentes del uso de la Terapia de Liberación Miofascial en atletas en procesos de competición, la cual se ha implementado como tratamiento de lesiones musculoesqueléticas, estas por lo general son asociadas al dolor, existe una discrepancia en el registro de los alcances de esta misma, con la efectividad y potencial terapéutico, ya sea por el escaso registro de efectividad durante su tratamiento o por la falta de un proceso metodológico para su medición. Ante la existencia de literatura carente de evidencia medible del desempeño y éxito de esta, es imprescindible, establecer un panorama claro de los avances y del futuro estudio para la aplicación de tecnología e innovación en la Terapia de Liberación Miofascial, abordamos la problemática primeramente desde el aspecto formal de liberación miofascial, en seguida abordando los diferentes aspectos que las terapias de electroestimulación bioeléctrica son aplicables en la actualidad, para finalmente llegar a las tecnologías que se utilizan para la liberación miofascial. Es importante resaltar que el aporte de este artículo, es la búsqueda de una metodología que permita crear nueva tecnología de bajo costo con los enfoques que se describen en el mismo, sentando las bases para que en futuras experimentaciones, se dirijan puntualmente a las estructuras musculoesqueléticas posteriores de las piernas.

La Terapia de Liberación Miofascial también ha mostrado resultados prometedores en el tratamiento de trastornos neurológicos, particularmente en la reducción de síntomas

como espasmos musculares, disfunción motora y problemas de equilibrio. Estudios han demostrado que la liberación miofascial puede mejorar los parámetros de la marcha y aumentar la función física general en pacientes en proceso de recuperación de condiciones neurológicas como el accidente cerebro-vascular. Esto sugiere que la liberación miofascial puede desempeñar un papel fundamental en la neuro-rehabilitación, ofreciendo un enfoque no invasivo para mejorar la calidad de vida de los pacientes con discapacidades neurológicas Lv and Yin (2024).

2. TERAPIA POR LIBERACIÓN MIOFASCIAL AUTOINDUCIDA

Para Martínez et al. (2024), el uso de Liberación Miofascial Autoinducida por parte de atletas profesionales y aficionados durante el calentamiento, la recuperación y los entrenamientos pueden tener efectos significativos en atributos de rendimiento físico, tales como el Rango de Movimiento y la fuerza. Los resultados sugieren que esta técnica parece tener efectos agudos positivos en la flexibilidad y el rango de movimiento, sin afectar el rendimiento muscular durante acciones de fuerza máxima y potencia, pero favoreciendo la percepción de recuperación y reduciendo el dolor muscular de aparición tardía. También se identificaron algunos efectos positivos en la agilidad y en acciones de alta velocidad de corta duración Martínez et al. (2024). La utilización de implementos, así como la terapia mecánica, esta recomendado para la movilización de las fibras y fascia muscular, tanto en la liberación, como en la recuperación de movilidad y rendimiento físico acompañado de fuerza,

de manera que en los márgenes se genere disminución de viscosidad en las fibras musculares y por tanto exista mayor rango de movilidad y flexibilidad. He aquí donde se presenta el primer reto, debido a la variación entre los diferentes métodos para la Liberación Miofascial, ya sea aplicado por un profesional o autoinducido, ya que los umbrales del dolor varían según la percepción del afectado.

2.1 Liberación post-competición

En evidencias, existen mediciones post-esfuerzo físico de Alto Rendimiento en el cual no mejora la recuperación en atletas post-competición, de manera mecánica y con aditamentos (foamrollers). Esto sugiere que, en un contexto de rehabilitación deportiva, la Liberación Miofascial Autoinducida puntual puede no ser suficiente para acelerar la recuperación tras partidos; posiblemente se requieran aplicaciones más frecuentes o complementarias a otras estrategias para obtener beneficios medibles en la recuperación de deportistas de alto nivel Oliveira et al. (2024). El desarrollo de estrategias para mejorar las diferentes capacidades mencionadas anteriormente es una necesidad en cualquier deporte. Aunque existen estudios en los que se utiliza la Terapia Manipulativa Osteopática, para mejorar variables asociadas al deporte, como el Rango de Movimiento o la actividad muscular en personas con patologías, pocos estudios han empleado esta intervención en atletas y ninguno ha evaluado sus efectos sobre el rendimiento deportivo.

Mientras que la fuerza y la potencia están enfocadas en el rendimiento, la flexibilidad, ya sea estática, medida a través del Rango de Movimiento, o dinámica, evaluada mediante pruebas como el "Sit and Reach" en V, ver Fig. 1, está más relacionada con la incidencia de lesiones, siendo los atletas con menor flexibilidad más susceptibles a ellas Afanador et al. (2023). A razón de las intensas rutinas de entrenamiento y competencia que mantienen los atletas universitarios, es común encontrar daño muscular crónico que afecta la flexibilidad, lo que lleva al uso frecuente de diferentes métodos de tratamiento, como el estiramiento, el masaje, los medicamentos antiinflamatorios, el ejercicio, la inmersión en agua fría y las intervenciones nutricionales, entre otros.



Fig. 1. Test "Sit and Reach en V"

2.2 Liberación Miofascial Autoinducida

La Liberación Miofascial autoinducida se utilizó como complemento de otros tratamientos, por lo que no se puede determinar su efecto específico. Además, debido a la heterogeneidad de los estudios, fue difícil establecer una aplicación óptima de la liberación miofascial autoinducida en atletas ver Fig.2. La Liberación Miofascial Autoinducida es una técnica efectiva para mejorar la flexibilidad y el Rango de Movimiento de forma aguda, pero sus efectos crónicos no son prometedores. Al compararla con otras intervenciones como el estiramiento dinámico, la Liberación Miofascial Autoinducida resultó ser menos efectiva para mejorar la flexibilidad en atletas Gunjan et al. (2025). Es fundamental continuar investigando los efectos específicos de la liberación miofascial autoinducida con rodillo de espuma en la población atlética. Dado que la presente revisión abarca estudios con una marcada heterogeneidad metodológica, resulta necesario realizar un nuevo análisis sistemático que aplique criterios de inclusión y exclusión más rigurosos. Esto permitiría minimizar la variabilidad entre estudios y ofrecer una evaluación más precisa y detallada sobre los efectos del rodillo de espuma en el rendimiento, la recuperación y la prevención de lesiones en atletas.



Fig. 2. Liberación Miofascial Autoinducida

2.3 Estimulación Eléctrica Neuromuscular

La Estimulación Eléctrica Neuromuscular es un método que utiliza corriente eléctrica pulsada de baja frecuencia para estimular directamente los nervios o músculos, con el fin de mejorar la función muscular o tratar trastornos neuromusculares y lesiones mediante la inducción pasiva de la contracción muscular. Ha sido ampliamente utilizada en neurología clínica, neuro-rehabilitación y otros campos como una tecnología de tratamiento indolora y no invasiva Nie et al. (2024). Existen evidencias de los beneficios fisiológicos asociados a la Estimulación Eléctrica Neuromuscular, incluyendo vasodilatación capilar, modificaciones en el metabolismo y optimización del reclutamiento neural. Estos efectos fisiológicos se traducen en una reducción del dolor, mejoría funcional y aumento en la activación muscular, especialmente en condiciones asociadas a edema. Además, se ha investigado el impacto inmediato de un protocolo específico de Estimulación Eléctrica Neuromuscular aplicado durante cuatro semanas sobre la modulación del

dolor y la activación neuromuscular en pacientes diagnosticados con el Síndrome De Dolor Patelofemoral.

La Estimulación Eléctrica Intramuscular, es una modalidad utilizada por los clínicos para tratar el dolor miofascial que ha mostrado resultados positivos para el tratamiento del dolor, sin embargo, aún faltan investigaciones que analicen los posibles mecanismos de acción de la Estimulación Eléctrica Intramuscular aplicada directamente sobre los puntos gatillo Perreault et al. (2022). Para zonas con menor cantidad de fibras musculares a diferencia de lo expuesto, se considera que, la terapia manual, el estiramiento, los ejercicios de coordinación y la estimulación eléctrica han demostrado ser beneficiosas para personas con dolor miofascial temporomandibular Espejo et al. (2024). La Terapia Manual consiste en la aplicación de estrategias orientadas al movimiento, que integran ejercicios y técnicas de movilización y/o manipulaciones manuales. Como en los trastornos temporomandibulares y de tejido blando en el cuello y los músculos masticatorios. Existen múltiples mecanismos potenciales por los cuales la Liberación Miofascial Autoinducida puede reducir la actividad eléctrica muscular. En el caso de los golfistas estudiados, es posible que la técnica haya modulado los husos musculares y reducido los espasmos musculares originados en los puntos gatillo Banaśkiewicz et al. (2024). Los Puntos Gatillo son nódulos contracturados localizados dentro de una banda tensa del tejido muscular. Estos nódulos, que suelen tener un diámetro de apenas unos milímetros, pueden encontrarse en diversas zonas dentro de los músculos y en el tejido miofascial, afectando la función muscular y generando molestias o dolor referido Zhu et al. (2024).

2.4 Aplicación de Liberación Miofascial en dolor Crónico

Las disfunciones lumbosacras y los síndromes de dolor asociados, como el Dolor Lumbar, representan uno de los problemas musculoesqueléticos más comunes a nivel mundial. Se ha sugerido que la disfunción de la Fascia Toracolumbar podría desempeñar un papel en algunos casos de dolor lumbar. Además, investigaciones han confirmado que las personas con Dolor Lumbar presentan un aumento en la actividad muscular en reposo en la región de la Fascia Toracolumbar Ozóg et al. (2021).

Es evidente que la electroestimulación para la Liberación Miofascial se ha convertido en una opción terapéutica para pacientes con dolor lumbar crónico, con el objetivo de mejorar la flexibilidad y el deslizamiento entre las capas de los tejidos blandos, reducir la actividad muscular, disminuir la intensidad del dolor y mejorar el rendimiento funcional.

2.5 Afecciones de Dolor Crónico atendida por la Liberación Miofascial

El dolor musculoesquelético crónico representa un desafío de salud persistente y debilitante, con implicaciones significativas tanto a nivel individual como social ver Fig.3. Afecciones como la fibromialgia, la osteoartritis y el síndrome de dolor miofascial afectan a millones de personas en todo el mundo, generando una carga considerable en términos de bienestar y calidad de vida. Más allá del

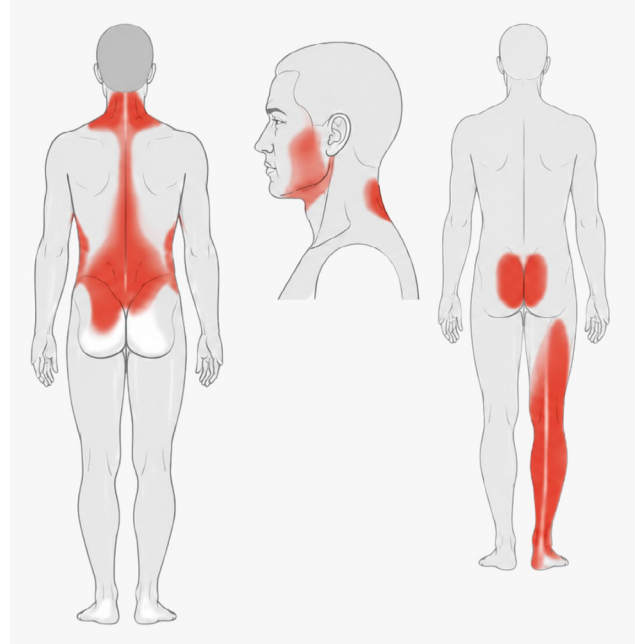


Fig. 3. Regiones musculares más comunes con Dolor Crónico para la Aplicación de la Liberación Miofascial

malestar físico, el dolor crónico puede provocar una disminución en la calidad de vida, problemas de salud mental y un alto costo económico tanto para los pacientes como para los sistemas de salud Sur et al. (2024).

Uno de los problemas más comunes que afectan la región del pie es la fascitis plantar, una afección caracterizada por la inflamación de la aponeurosis plantar en el punto donde se conecta con la tuberosidad calcánea. El dolor en el talón es el síntoma principal y suele manifestarse con mayor intensidad en las primeras horas de la mañana, disminuyendo gradualmente con el movimiento Nadeem et al. (2023). Esta afección presenta dolor en la planta del pie, especialmente en la parte inferior del talón. El cual resulta más intenso al dar los primeros pasos por la mañana y, aunque tiende a disminuir durante el día con el movimiento, puede volverse migratorio o generalizado inicialmente, hasta asentarse cerca de la tuberosidad medial del calcáneo, cabe señalar que la aplicación de electroestimulación es reducida en su intensidad, debido a la proporción del dolor y su intensidad en relación al área relativamente pequeña donde las fibras musculares son menores que las de la región posterior de la pierna. En los últimos años, la tecnología Estimulación Eléctrica Neuromuscular ha experimentado un avance significativo, tanto en el desarrollo de dispositivos como en una mejor comprensión de los parámetros de tratamiento, incluyendo frecuencia, intensidad y duración del pulso Ackermann et al. (2024). La eficacia de la ES en el tratamiento del Dolor Muscular De Aparición Tardía y su impacto en la recuperación muscular siguen sin estar claros, y se requiere más investigación para determinar su verdadero papel en la recuperación post-ejercicio.

La producción de Especies Reactivas De Oxígeno inducida por el ejercicio ha sido identificada como un mecanismo que contribuye al daño muscular. Estas moléculas son gen-

eradas continuamente en el organismo y, bajo condiciones normales, son neutralizadas de manera eficiente por los sistemas antioxidantes, las Especies Reactivas De Oxígeno desempeñan un papel fundamental en la producción de fuerza muscular, ya que en bajas concentraciones regulan diversos procesos de señalización celular Gussoni et al. (2023). Es importante señalar que cuando su concentración aumenta, el impacto se vuelve negativo, ya que reducen la capacidad de generación de fuerza de manera dependiente del tiempo, contribuyendo así a la aparición de fatiga muscular.

3. ESTIMULACIÓN BIOELÉCTRICA

La terapia de estimulación bioeléctrica ha comenzado a captar atención debido a su potencial para favorecer la recuperación muscular y aliviar el dolor Chen (2024). Actualmente, se presentan limitaciones en cuanto al diseño de estudio y evaluación de la eficacia, lo que dificulta generar la validación integral de su seguridad y efectividad clínica. Se ha demostrado claramente que la longitud del músculo influye en el daño causado por las contracciones musculares isométricas inducidas por Estimulación Eléctrica Neuromuscular. Cuando el músculo se encuentra en una posición corta, la producción de fuerza durante la aplicación de NMES es mayor en comparación con la posición alargada Fouré et al. (2020). El uso de Estimulación Eléctrica Neuromuscular en músculos acortados ayuda a minimizar la aparición y el alcance del daño muscular, en comparación con su aplicación en músculos más largos. Esto favorece la readaptación muscular, a pesar de que no observarse alteraciones significativas en los tejidos musculares. Los sistemas de estimulación eléctrica disponibles en el mercado generan señales de estimulación con amplitudes de voltaje-corriente no personalizadas, lo que puede provocar fatiga muscular o movimientos terapéuticos incompletos Orona et al. (2024). Para esto es necesario regular la señal de estimulación según las necesidades terapéuticas individuales según la afección.

3.1 Características Eléctricas de la Estimulación Eléctrica Neuromuscular

La Estimulación Eléctrica Neuromuscular convencional, que utiliza estimulación de pulso estrecho con una duración de 0.05 a 0.4 ms, el reclutamiento de unidades motoras ocurre en un orden no fisiológico. Este tipo de estimulación genera una despolarización directa de los axones motores, activando una gran cantidad de unidades motoras de contracción rápida y fatigable Alahmari et al. (2022). Esto genera fatiga muscular rápida, lo que limita su aplicación en la práctica clínica. La coestimulación electromecánica de las células es una herramienta prometedora en la ingeniería de tejidos, ya que permite influir en el desarrollo y la funcionalidad celular. Sin embargo, las plataformas actuales presentan limitaciones, principalmente debido a la baja durabilidad de los componentes y la dificultad para optimizar los parámetros de estimulación. Aunque se han explorado diversos sistemas de coestimulación electromecánica, la integración de materiales y componentes altamente duraderos sigue siendo un desafío Kim et al. (2024a).

3.2 Aplicación de Inteligencia Artificial para Estimulación Eléctrica Neuromuscular

Entre los desafíos más preocupantes para los pacientes con parálisis facial aguda o crónica acompañada de sincinesias, se encuentran la pérdida de la capacidad para sonreír y el cierre insuficiente del párpado. De acuerdo a Steiner et al. (2024), se propone un algoritmo activado auricularmente basado en Inteligencia Artificial, el cual clasifica los movimientos faciales que el paciente desea realizar. Esta clasificación se realiza a partir de grabaciones de Electromiografía De Superficie de los músculos auriculares extrínsecos, concretamente del músculo auricular anterior, superior y posterior del lado parético. Posteriormente, el sistema administra estimulación eléctrica superficial dirigida para contraer los músculos faciales correspondientes. La limitación que se propone es derivada de la unilateralidad de la lesión, aun cuando los hallazgos si resultan significativos, se observa que esta herramienta, sí mejora la calidad de vida de los pacientes por la mejora en la función facial aplicando puntualmente la Estimulación Eléctrica Neuromuscular en las zonas identificadas como dañadas.

3.3 Estimulación Electrotáctil

La Estimulación Electrotáctil se basa en la aplicación de Diferencias de Potencial Eléctrico sobre la piel, estimulando directamente las fibras nerviosas cutáneas ver Fig.4.

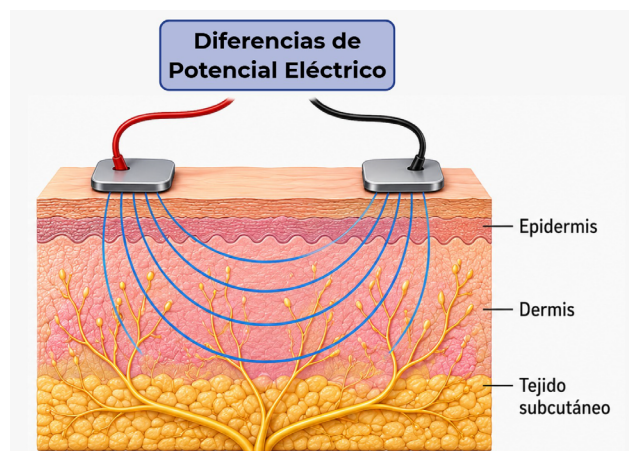


Fig. 4. Estimulación Electrotáctil

En el ámbito médico, la Estimulación Electrotáctil ha demostrado ser una herramienta útil en el desarrollo de prótesis, permitiendo que los pacientes recuperen la sensación táctil en sus extremidades faltantes, lo que les proporciona mayor control y precisión en la manipulación de objetos. Además, en aplicaciones de rehabilitación, se ha demostrado que la Estimulación Electrotáctil mejora la abducción y flexión del hombro en pacientes con accidente cerebrovascular crónico Doan et al. (2025). Al ser de bajo costo y poco invasiva, esta técnica ha cobrado mucho auge por su alcance comercial, convirtiéndose en una herramienta asequible.

3.4 Estimulación Eléctrica Aplicada

La regulación efectiva de la estimulación mecánica tiene un gran potencial para optimizar la funcionalidad celular,

lo que ha llevado al desarrollo de métodos in vitro para simular y controlar estos estímulos. Como resultado, la biomecánica se ha consolidado como un área de investigación clave, impulsando la creación de nuevos materiales, dispositivos y tecnologías destinados a proporcionar estimulación mecánica precisa a las células Gu et al. (2024). Para la biomecánica, estos materiales representan una nueva clase de materiales inteligentes y flexibles, capaces de modificar su forma o tamaño en respuesta a la excitación eléctrica. Debido a su comportamiento los Polímeros electroactivos también son conocidos como músculos artificiales, ver Fig. 5.

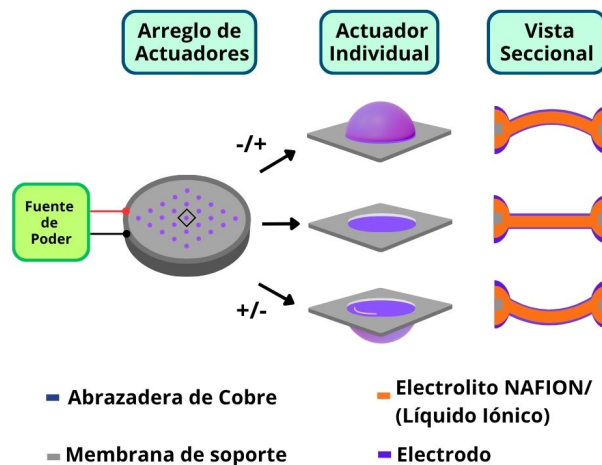


Fig. 5. Ilustraciones esquemáticas de la matriz de microactuadores basada en músculos artificiales Polímeros Electroactivos iónicos para la estimulación mecánica celular.

La activación háptica puede desencadenar una mayor actividad muscular, amplificando y clarificando la sensación en ciclos de estimulación sucesivos, esto ha revelado una interacción fascinante entre la actividad muscular y la percepción de señales hápticas electroestimuladas durante la marcha Kim et al. (2024b). Se ha observado que una mayor activación muscular puede estar relacionada con una mejor precisión en el reconocimiento de los patrones hápticos, lo que sugiere un vínculo con la propiocepción y la atención focalizada. Resaltan la relevancia de la estimulación háptica electroactiva en el ámbito médico y tecnológico, con aplicaciones clave en el fortalecimiento del control motor, la rehabilitación de la marcha y la mejora de la interacción entre el usuario y sistemas hápticos de alta precisión.

Quienes reciben estimulación real muestran señales cerebrales más intensas, asociadas con una mayor sensibilidad táctil, probablemente debido al efecto de la Estimulación Eléctrica Sensorial sobre el sistema nervioso central. Los hallazgos sugieren que la Estimulación Eléctrica Sensorial mejora la discriminación de texturas después del entrenamiento y tiene un impacto positivo en las áreas cerebrales relacionadas con el procesamiento sensorial Villar et al. (2024).

Se puede estimular la mejora en paciente con úlceras venosas que no soportan terapia de compresión, mediante la Estimulación Eléctrica Neuromuscular. Se ha revisado la evidencia científica sobre el uso de la activación de la Bomba Muscular mediante Estimulación Eléctrica Neuromuscular del nervio peroneo común, con el objetivo de mejorar la función de la bomba muscular de la pantorrilla. Mostrando que existen pruebas fisiológicas de que la Bomba Muscular mejora la función de la bomba muscular de la pantorrilla y el retorno venoso, tanto en sujetos sanos como en pacientes con enfermedad venosa Stacey et al. (2024). De igual manera, se encontró evidencia de que el tratamiento de los pacientes con Trastornos Temporomandibulares suele requerir un enfoque multimodal, involucrando no solo a odontólogos y especialistas en dolor orofacial, sino también a profesionales de la salud especializados en el manejo conservador del dolor.

Así mismo las señales Mecanomiográficas, pueden evaluar el rendimiento muscular, en entornos clínicos, se ha demostrado que la implementación de Estimulación Eléctrica Neuromuscular se ha tornado crucial para la restauración de funciones mecánicas que por si solo los pacientes no son capaces de activar. La caracterización de las propiedades contráctiles de los músculos mediante Mecanomiográficas a través de Estimulación Eléctrica Neuromuscular ha mostrado una fiabilidad excelente. Por ello, esta revisión analiza el uso de las señales Mecanomiográficas para la evaluación de los músculos esqueléticos bajo estimulación eléctrica. Los hallazgos indican que el uso de Mecanomiográficas a través de Estimulación Eléctrica Neuromuscular proporciona factores clave para la evaluación de la actividad muscular, destacando la importancia de la estandarización de parámetros de electroestimulación y Mecanomiográficas en diferentes protocolos experimentales Uwamahoro et al. (2021). El ejercicio de resistencia regula la expresión de una amplia variedad de genes en el músculo esquelético, aumentando o disminuyendo su actividad según la demanda Kusano et al. (2024). La tasa de variación de contracción muscular se midió bajo el régimen de 5 series de 10 contracciones de 3 segundos a 100Hz, mostrando una mejora a la fatiga después de 24 horas, y describiendo que el patrón sostenido de expresión y la magnitud de la respuesta, juega un papel clave en los mecanismos de adaptación al ejercicio de resistencia, haciendo factible la medición.

3.5 Técnicas de Electroestimulación

La electroestimulación representa una herramienta fundamental en rehabilitación neuromuscular y medicina deportiva debido a su capacidad para modular respuestas fisiológicas mediante parámetros eléctricos controlados.

Para Ackermann et al. (2024), las técnicas como Estimulación Nerviosa Eléctrica Transcutánea se enfocan principalmente en analgesia, mientras que Estimulación Eléctrica Neuromuscular y Estimulación Eléctrica Funcional buscan inducir contracciones funcionales para fortalecimiento muscular y restauración motora. Por su parte, Perreault et al. (2022) la Estimulación Intramuscular Eléctrica permite actuar directamente sobre puntos gatillo miofasciales, mejorando flujo sanguíneo y modulando mecanismos antinociceptivos ver Fig.6.



Fig. 6. Aplicación de la Estimulación Eléctrica Neuromuscular

La eficacia terapéutica depende principalmente de parámetros como frecuencia, intensidad y ancho de pulso. De acuerdo con Ackermann et al. (2024), las frecuencias bajas favorecen analgesia, mientras frecuencias intermedias entre 20 y 50 Hz producen contracciones funcionales útiles para rehabilitación y fortalecimiento muscular. En la actualidad, la electroestimulación evoluciona hacia sistemas inteligentes integrados con electromiografía e inteligencia artificial, capaces de personalizar terapias y optimizar el control motor y biomecánico ver Tabla1.

Técnica	Invasiva	Objetivo	Nivel de Contracción.	Precisión
Estimulación Nerviosa Eléctrica Transcutánea	No	Analgesia y modulación sensorial del dolor.	Baja	Moderada
Estimulación Eléctrica Neuromuscular	No	Fortalecimiento muscular y activación neuromuscular.	Media	Alta
Estimulación Eléctrica Funcional	No	Movimiento funcional asistido.	Alta	Muy alta
Estimulación Eléctrica Intramuscular	Sí	Tratamiento localizado de puntos gatillo miofasciales.	Variable	Muy alta
Estimulación Nerviosa Eléctrica por Microcorrientes	No	Reparación tisular mediante microcorrientes.	Muy baja	Moderada
Electromecánica avanzada	Variable	Control biomédico inteligente y adaptativo.	Adaptativa	Muy alta

Tabla 1. Comparación de modalidades de electroestimulación

4. TECNOLOGÍAS ASOCIADAS LA ESTIMULACIÓN ELÉCTRICA

La Estimulación Eléctrica de los nervios periféricos está emergiendo como una terapia antiinflamatoria innovadora, capaz de modular las señales bioeléctricas y los neurotransmisores para reducir la inflamación. Los picos de estimulación inadecuados pueden inducir daño nervioso y aumentar el riesgo de complicaciones ver Fig.7. Los estimuladores comerciales basados en semiconductores del tipo "Semiconductor Complementario de Óxido Metálico" generan pulsos de estimulación con amplitudes fijas y abruptas, lo que no permite programabilidad y no imita fielmente las señales bioeléctricas naturales Bao et al. (2024). Los progresos en semiconductores bidimensionales han impulsado avances significativos en la electrónica sensorial biónica y la ingeniería neuromórfica. Gracias a su estructura ultradelgada, estos materiales presentan características clave, como operación a bajo voltaje, eficiencia energética mejorada y una respuesta adaptable a estímulos físicos externos, lo que los hace ideales para aplicaciones en tecnologías emergentes.

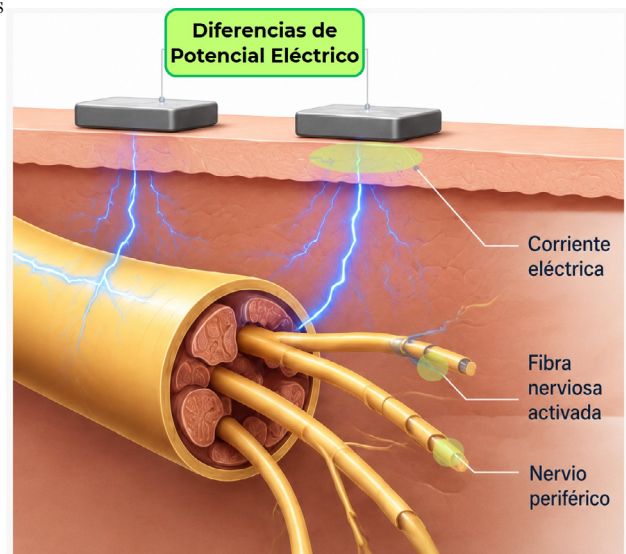


Fig. 7. Estimulación Eléctrica de los nervios periféricos

La Estimulación Eléctrica muscular de cuerpo utiliza electrodos colocados en lados opuestos del cuerpo para estimular simultáneamente los músculos agonistas y antagonistas. Puede cubrir una superficie muscular de hasta 2800 cm^2 Zarzeczny et al. (2024). Estos resultados subrayan la importancia de seguir explorando el uso de la Estimulación Eléctrica Neuromuscular como estrategia terapéutica, enfocándose en su capacidad para mejorar la función muscular y regular procesos biológicos fundamentales relacionados con el envejecimiento y la prevención de patologías asociadas. En la práctica clínica, se recomienda aumentar progresivamente la intensidad de la corriente para lograr una mayor adaptación de la fuerza y torque muscular inducido por Estimulación Eléctrica Neuromuscular .

Existen diferentes tipos de ondas de Estimulación Eléctrica Neuromuscular utilizadas para inducir contracciones musculares:



- (1) Corriente pulsada convencional: Consiste en pulsos eléctricos de duración de fase $<500 \mu s$; Se aplica en un rango de frecuencia de 20 a 100 Hz.
- (2) Corriente Alterna De Kilohertz: Opera en un rango de 1 kHz a 10 kHz, con una frecuencia de ráfaga entre 1 Hz y 100 Hz y un ciclo de trabajo del 10% o más Pinto Damo et al. (2021).

Actualmente, no se han realizado investigaciones que analicen de manera conjunta el torque muscular, tanto submáximo como máximo, en función de la percepción de incomodidad, considerando al mismo tiempo las distintas modalidades de corrientes utilizadas en la Estimulación Eléctrica Neuromuscular y las variaciones en la duración de fase. Esta ausencia evidencia una brecha en la literatura que limita la comprensión integral de los efectos combinados de estos parámetros sobre la eficacia y tolerabilidad del estímulo. Los efectos fisiológicos de la estimulación eléctrica externa pueden analizarse en distintos niveles de complejidad, desde las moléculas y células individuales hasta los órganos y el organismo completo Lee et al. (2024). En particular, los mecanismos moleculares y celulares que regulan el inicio de respuestas metabólicas, transcripcionales o mecánicas frente a estímulos del entorno siguen siendo en gran parte desconocidos. Para abordar estas limitaciones, resulta esencial avanzar en el conocimiento de las condiciones intracelulares y extracelulares que se generan como consecuencia de la estimulación eléctrica, lo cual abriría paso al desarrollo de aplicaciones más precisas y eficaces dentro del ámbito de la medicina bioeléctrica.

5. COMENTARIOS

- Desde su origen la Liberación Miofascial, ha tenido diferentes formas de aplicación, considerando a la mecánica como la primordial.
- Al no haber evidencia medible resulta prioritario en posteriores trabajos documentar la evidencia de los sujetos de prueba que deben ser hombres en etapa de competición preferentemente de Alto Rendimiento.
- Se tiene el hallazgo de que la subjetividad en la medición del dolor altera los resultados en la medición de la mejora de las afecciones desde la perspectiva de los pacientes.
- La medicina bioeléctrica está cobrando cada vez más relevancia en la investigación y la práctica médica actual.
- La implementación de la Estimulación Eléctrica Neuromuscular no tiene registros concluyentes de la reducción de dolor en los pacientes con dolor crónico.

6. CONCLUSION

Después de la revisión de todas las evidencias, es importante resaltar el área de oportunidad que este artículo demuestra, puesto que, al no haber información concluyente en rangos de medición, mejora y disminución de dolor en músculos posteriores de la pierna, tanto en pacientes con dolor crónico y/o deportistas de alto rendimiento, surge la posibilidad de establecer una metodología para el diseño de instrumentos de medición que permita documentar y registrar estos hallazgos.

REFERENCIAS

- Ackermann, P. et al. (2024). Unlocking the potential of neuromuscular electrical stimulation: achieving physical activity benefits for all abilities. *Frontiers in Sports and Active Living*, 6, 1–8. doi:10.3389/fspor.2024.1507402.
- Afanador, D. et al. (2023). Effects of myofascial release using finding-oriented manual therapy combined with foam roller on physical performance in university athletes: A randomized controlled study. *International Journal of Environmental Research and Public Health*, 20, 1–13. doi:10.3390/ijerph20021364.
- Alahmari, S.K. et al. (2022). Effects of three neuromuscular electrical stimulation methods on muscle force production and neuromuscular fatigue. *Scandinavian Journal of Medicine & Science in Sports*, 32(10), 1456–1463. doi:10.1111/sms.14210.
- Banaśkiewicz, J. et al. (2024). Self-myofascial release therapy moderately influence electromyography of longissimus dorsi muscle in golfers. *Biomedical Human Kinetics*, 16, 295–304. doi:10.21203/rs.3.rs-4489765/v1.
- Bao, R. et al. (2024). Neuromorphic electro-stimulation based on atomically thin semiconductor for damage-free inflammation inhibition. *Nature Communications*, 15, 1327. doi:10.1038/s41467-024-45590-8.
- Chen, J. (2024). Bioelectrical stimulation therapy for muscle injuries in aerobics athletes. *Molecular & Cellular Biomechanics*, 21(2), 1–19. doi:10.62617/mcb268.
- Doan, H. et al. (2025). Electro-tactile modulation of muscle activation and intermuscular coordination in the human upper extremity. *Scientific Reports*, 15, 2559. doi:10.1038/s41598-025-86342-y.
- Espejo, L. et al. (2024). Effects of adding electro-massage to manual therapy for the treatment of individuals with myofascial temporomandibular pain: a randomized controlled trial. *Journal of Applied Oral Science*, 32, 1–12. doi:10.1590/1678-7757-2024-0109.
- Fouré, A. et al. (2020). Muscle alterations induced by electrostimulation are lower at short quadriceps femoris length. *European Journal of Applied Physiology*, 325–335. doi:10.1007/s00421-019-04277-5.
- Gu, J. et al. (2024). A microactuator array based on ionic electroactive artificial muscles for cell mechanical stimulation. *Biomimetics*, 9, 281. doi:10.3390/biomimetics9050281.
- Gunjan, B. et al. (2025). Effect of self myofascial release on flexibility in athletes: A systematic review. *African Journal of Biomedical Research*, 28, 604–614. doi:10.53555/AJBR.v28i1S.6174.
- Gussoni, M. et al. (2023). Effects of electrical stimulation on delayed onset muscle soreness (doms): Evidences from laboratory and in-field studies. *Journal of Functional Morphology and Kinesiology*, 8, 146. doi:10.3390/jfmr8040146.
- Kim, A.R. et al. (2024a). Highly durable, stretchable multi-electrode array for electro-mechanical co-stimulation of cells. *Biomaterials Research*, 28, 30.
- Kim, Y. et al. (2024b). The interplay between muscular activity and pattern recognition of electro-stimulated haptic cues during normal walking: A pilot study. *Bioengineering*, 11(12), 1248. doi:10.3390/bioengineering11121248.
- Kusano, T. et al. (2024). Time-series transcriptomics reveals distinctive mrna expression dynamics associated

- with gene ontology specificity and protein expression in skeletal muscle after electrical stimulation-induced resistance exercise. *The FASEB Journal*, 38, 1–16. doi: 10.1096/fj.202401420RR.
- Lee, H. et al. (2024). Bioelectric medicine: unveiling the therapeutic potential of micro-current stimulation. *Biomedical Engineering Letters*, 14, 367–392. doi: 10.1007/s13534-024-00366-3.
- Lv, Y. and Yin, Y. (2024). A review of the application of myofascial release therapy in the treatment of diseases. *Journal of Multidisciplinary Healthcare*, 17, 4507–4517. doi:10.2147/jmdh.s481706.
- Martínez, L.M. et al. (2024). Effects of self-myofascial release on athletes physical performance: A systematic review. *Journal of Functional Morphology and Kinesiology*, 9, 20.
- Nadeem, W. et al. (2023). Effectiveness of myofascial release techniques in the management. *Alied Medical Research Journal*, 1(2), 161–175. doi: 10.59564/amrj/01.02/015.
- Nie, Q. et al. (2024). Effects of adding neuromuscular electrical stimulation to functional training on muscle recruitment, pain reduction, and knee joint function in patellofemoral pain syndrome patients. *Medicine*, 103, e36095. doi:10.1097/MD.00000000000036095.
- Oliveira, C. et al. (2024). Comparison of the effect of passive recovery and self-myofascial release in post-match recovery in female soccer players. *Journal of Human Sport and Exercise*, 19(2), 510–521. doi: 10.55860/j6dyvg90.
- Orona, L., Chairez, I., and Alfaro, M. (2024). Functional electrostimulation system for a prototype of a human hand prosthesis using electromyography signal classification by machine learning techniques. *Machines*, 12, 49. doi:10.3390/machines12010049.
- Ozóg, P. et al. (2021). Analysis of muscle activity following the application of myofascial release techniques for low-back pain: A randomized-controlled trial. *Journal of Clinical Medicine*, 10, 1–11.
- Perreault, T. et al. (2022). Intramuscular electrical stimulation to trigger points: Insights into mechanisms and clinical applications: A scoping review. *Journal of Clinical Medicine*, 11, 6039. doi:10.3390/jcm11206039.
- Pinto Damo, N.L. et al. (2021). Effects of different electrical stimulation currents and phase durations on submaximal and maximum torque, efficiency, and discomfort: a randomized crossover trial. *Brazilian Journal of Physical Therapy*, 25(5), 593–600. doi: 10.1016/j.bjpt.2021.03.001.
- Stacey, M. et al. (2024). Continuous muscle pump activation by neuromuscular electrical stimulation of the common peroneal nerve in the treatment of patients with venous leg ulcers: A position paper. *International Wound Journal*, 21, 1–10. doi:10.1111/iwj.70040.
- Steiner, K. et al. (2024). Electro-stimulation system with artificial-intelligence-based auricular-triggered algorithm to support facial movements in peripheral facial palsy: A simulation pilot study. *Diagnostics*, 14, 2158. doi:10.3390/diagnostics14192158.
- Sur, M. et al. (2024). Unlocking relief: Myofascial release in chronic musculoskeletal pain management. *International Journal of Orthopaedics and Physiotherapy*, 6(1), 17–20. doi:10.33545/26648989.2024.v6.i1a.20.
- Uwamahoro, R. et al. (2021). Assessment of muscle activity using electrical stimulation and mechanomyography: a systematic review. *Biomedical Engineering Online*, 20, 1–47. doi:10.1186/s12938-020-00840-w.
- Villar, E. et al. (2024). Enhancing touch sensibility with sensory electrical stimulation and sensory retraining. *Journal of NeuroEngineering and Rehabilitation*, 21, 79. doi:10.1186/s12984-024-01371-4.
- Zarzeczny, R. et al. (2024). Effects of 12 weeks of neuromuscular electrical stimulation of the kfac quadriceps muscles on the function and physio-biochemical traits in functionally fit female nursing-home residents aged 75+ years: a pilot study. *European Journal of Applied Physiology*, 124, 945–962. doi:10.1007/s00421-023-05321-1.
- Zhu, M. et al. (2024). Efficacy and factors of myofascial release therapy combined with electrical and magnetic stimulation in the treatment of chronic pelvic pain syndrome. *Open Medicine*, 19, 1–7. doi:10.1515/med-2024-0936.

Dynamic Analysis of a High-Order Buck–Boost Converter with Reduced Power Processing

Iván A. Reyes-Portillo* Jorge U. Muñoz-Minjares*
Saúl R. Méndez-Elizondo** Claudia A. Rivera-Romero*,***
Luis G. Carreto-Hernández**** Wendy G. Báez-García†

* Universidad Autónoma de Zacatecas, Unidad Académica de Ingeniería Eléctrica Plantel Jalpa, 99601, Jalpa, Zacatecas, México, (e-mail: ivan.reyes@upslp.edu.mx and ju.munoz@uaz.edu.mx)

** Universidad Autónoma de San Luis Potosí, 78210, San Luis Potosí, S.L.P., México, (e-mail: saul.mendez@uaslp.mx)

*** Instituto Nacional de Astrofísica, Óptica y Electrónica, 72840, Tonantzintla, Puebla, México (e-mail: c.a.riveraromero@inaoe.mx)

**** Universidad Tecnológica de Gutiérrez Zamora, 93556, Gutiérrez Zamora, Veracruz, México (e-mail: luis.carreto@utgz.edu.mx).

† Instituto Tecnológico Superior de Misanthla, 93821, Misanthla, Veracruz, México (e-mail: wgbaezg@itsm.edu.mx).

Abstract: High-order DC–DC converters have emerged as a suitable solution for applications requiring wide voltage conversion ranges and high efficiency, such as electric vehicles, photovoltaic systems and DC microgrids. This paper presents a dynamic analysis of a high-order Buck–Boost converter based on the *Reduced Redundant Power Processing* concept. A state-space averaged model is developed, along with its corresponding small-signal linearised model, with the aim of describing the system's dynamics. Based on these models, the poles, zeros and frequency response of the converter are analysed, highlighting the presence of non-minimum phase characteristics and their implications for controller design. Simulation results validate the proposed models, showing good agreement with the behaviour of the converter during switching. The results obtained demonstrate that the reduction in redundant power processing significantly influences the dynamic characteristics of high-order converters, providing relevant information for the design of control strategies in high-performance power conversion systems.

Keywords: DC–DC, Buck–Boost converter, High-order systems, R^2P^2 , Dynamic analysis.

1. INTRODUCTION

In recent years, the growth of power electronics-based applications, such as electric vehicles, photovoltaic systems, and DC microgrids, has driven the development of DC–DC converters capable of operating over wide voltage conversion ranges while maintaining high levels of efficiency and power density (Reyes-Portillo et al., 2022, 2026). In this context, high-order buck-boost converters have emerged as an attractive alternative due to their ability to achieve high conversion ratios without the use of coupled magnetic elements, enabling more compact and flexible designs (Gholizadeh et al., 2023; Diaz et al., 2026). However, an increase in system order leads to more complex dynamic behavior, which makes accurate modeling, stability analysis, and the design of control strategies more difficult (Xiong and He, 2025). In particular, high-order converters exhibit complex interactions between poles and zeros, which can lead to resonant effects and the emergence

of non-minimum phase characteristics, imposing fundamental limitations on the system's dynamic performance. Furthermore, many conventional high-gain topologies rely on cascaded configurations to achieve high levels of voltage conversion, which reduces efficiency due to the multiple stages of power processing (Chen et al., 2019). In these configurations, power is processed repeatedly before reaching the load, which increases losses and degrades the overall system performance (Anuraag et al., 2021).

To address these limitations, the concept of *Reduced Redundant Power Processing* (R^2P^2) has been proposed as an effective strategy for improving the efficiency of DC–DC converters. This approach allows a fraction of the input power to be transferred directly to the load through intermediate storage elements, thereby reducing redundant power processing and associated losses (Reyes-Portillo et al., 2025; Liu et al., 2025). However, although the advantages of R^2P^2 have been extensively studied from the perspective of efficiency and structural design, its impact on the dynamic behavior of high-order converters has not yet been fully characterized (Klein Faistel et al., 2019). In particular, the influence of partial power pro-

* We would like to thank the Universidad Autónoma de Zacatecas and the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) for their support and funding for this work.

cessing on system stability, the distribution of poles and zeros, and the emergence of non-minimum phase behavior remains an open research problem. Unlike previous studies based on R^2P^2 , which focus primarily on steady-state efficiency and component sizing—such as the high-voltage boost converter in Zogogianni et al. (2018) and the LED driver in Liu et al. (2025) the dynamic behavior of these topologies remains largely unexplored. Specifically, no previous study has systematically analyzed the small-signal transfer functions, pole-zero locations, or non-minimum phase characteristics of high-order buck-boost converters under reduced redundant power processing.

This paper presents a comprehensive dynamic analysis of a high-order Buck-Boost converter based on the R^2P^2 concept. First, a state-space averaged model is developed, followed by its corresponding small-signal representation. Based on these models, the poles, zeros, and frequency response of the converter are systematically analyzed, with special emphasis on the identification of non-minimum phase characteristics and their implications for controller design. Additionally, a set of generalized transfer functions is obtained that relate the state variables to the duty cycle, providing a deeper understanding of the system's internal dynamics. The proposed models are validated through frequency-domain analysis, comparing the linearized model with a switched-mode model obtained via simulation. The results show a high degree of agreement between the two approaches, confirming the accuracy of the developed models. Furthermore, the analysis demonstrates how the R^2P^2 structure influences the converter's dynamic response, providing relevant guidelines for the design of control strategies in high-performance power conversion systems.

2. PROPOSED BUCK-BOOST CONVERTER

The R^2P^2 concep aims to enhance converter efficiency by reducing the portion of power that undergoes full processing. Instead of relying on fully cascaded stages, this approach enables partial direct energy transfer from the source to the load, while an intermediate storage element (buffer) manages the remaining power exchange. As a result, only a fraction of the total energy is actively processed by the conversion stages, which leads to improved efficiency when compared to traditional cascaded configurations (Zogogianni et al., 2018). Based on this principle, a high conversion ratio buck-boost topology is developed using the I-II structural configuration of the R^2P^2 framework. The proposed converter is depicted in Fig. 1.

In the I-IIA configuration, the overall efficiency is determined by the interaction between the efficiencies of the individual conversion stages. Denoting η_1 and η_2 as the efficiencies of each stage, the global efficiency can be approximated as

$$\eta_T = \eta_1\eta_2 + (1 - k)\eta_2(1 - \eta_1) \quad (1)$$

where $k \in (0, 1)$ defines the proportion of power processed through the buffer element (Tse, 2003). This expression highlights that efficiency gains arise from reducing redundant power processing. However, it is important to note

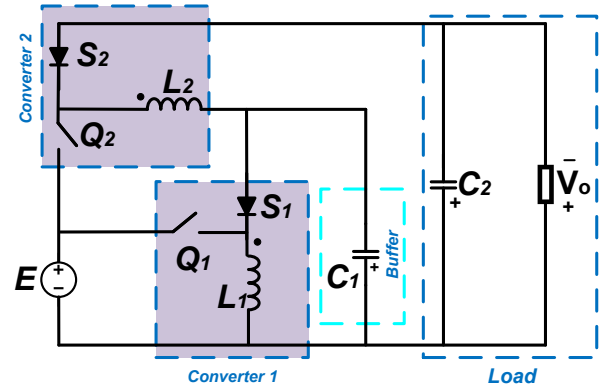


Fig. 1. Proposed high-gain buck-boost converter derived from the R^2P^2 structure.

that this analysis assumes ideal conditions. In practical implementations, the presence of parasitic elements—such as resistive losses, leakage inductances, and non-ideal capacitances—introduces additional losses that may limit the achievable efficiency improvement (Reyes-Portillo et al., 2023). The converter has two conduction modes, the Continuous Conduction Mode CCM where the current in the inductors never reaches zero and the Discontinuous Conduction Mode (DCM) where the current in the inductors is zero at an instant of the switching period. Through a steady state analysis it is possible to determine its equilibrium points of the Buck-Boost converter with R^2P^2 . This analysis is only valid for the CCM and the ideal elements thisis, without considering parasitic elements (resistors, capacitances and inductances). Expressions (2-5) represent the steady-state voltage and current nominals of capacitors and inductors, in terms of input voltage (E), duty ratio (D) and load (R).

$$V_{C1} = \frac{ED}{(1 - D)} \quad (2)$$

$$V_{C2} = \frac{ED(2 - D)}{(1 - D)^2} \quad (3)$$

$$I_{L1} = \frac{ED(2 - D)}{R(1 - D)^4} \quad (4)$$

$$I_{L2} = \frac{ED(2 - D)}{R(1 - D)^3} \quad (5)$$

The gain or transformation ratio (M) is given by:

$$M = \frac{V_o}{E} = \frac{(2 - D)D}{(1 - D)^2} \quad (6)$$

3. MODELING OF THE CONVERTER

The converter in CCM presents two operating states, the (ON state) consists of switches Q_1 and Q_2 being turned on while diodes S_1 and S_2 do not conduct current and the (OFF state) when switches Q_1 and Q_2 are opened while diodes S_1 and S_2 are turned on. The modeling of the converter was performed through state space, analyzing the different operating states. This modeling does not



consider the parasitic resistances present in the electronic devices. The following variables i_{L_1} , i_{L_2} , v_{C_1} , v_{C_2} were considered for modeling. The variables correspond to the currents in the inductors and voltages in the capacitors. Fig. 2 shows the electrical network corresponding to the ON state.

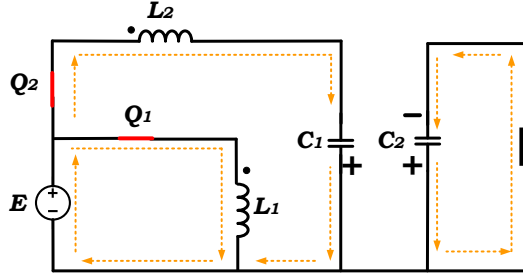


Fig. 2. Electrical network corresponding to the ON state.

The expressions (7-10) represent the dynamic equations of the converter in ON state.

$$\dot{i}_{L_1} = \frac{E}{L_1} \quad (7)$$

$$\dot{i}_{L_2} = \frac{E}{L_2} - \frac{v_{C_1}}{L_2} \quad (8)$$

$$\dot{v}_{C_1} = \frac{i_{L_2}}{C_1} \quad (9)$$

$$\dot{v}_{C_2} = -\frac{v_{C_2}}{RC_2} \quad (10)$$

Fig. 3 shows the electrical network corresponding to the OFF state.

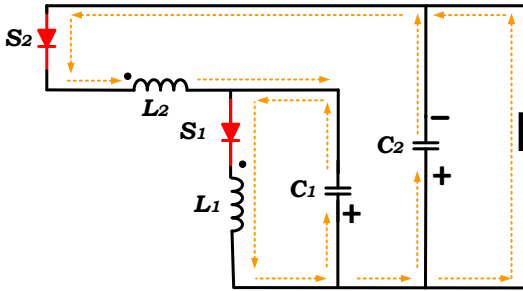


Fig. 3. Electrical network corresponding to the OFF state.

The expressions (11-14) represent the dynamic equations of the converter in "OFF state".

$$\dot{i}_{L_1} = \frac{v_{C_1}}{L_1} \quad (11)$$

$$\dot{i}_{L_2} = -\frac{v_{C_1}}{L_2} - \frac{v_{C_2}}{L_2} \quad (12)$$

$$\dot{v}_{C_2} = \frac{i_{L_2}}{C_1} - \frac{i_{L_1}}{C_1} \quad (13)$$

$$\dot{v}_{C_2} = \frac{i_{L_2}}{C_2} - \frac{v_{C_2}}{RC_2} \quad (14)$$

Once the expressions of a switching period have been obtained, it is possible to obtain an average model, using:

$$\dot{x} = A_1(D) + A_2(1-D) + B_1(D) + B_2(1-D) \quad (15)$$

The average model state by:

$$\dot{x} = \begin{bmatrix} 0 & 0 & \frac{(1-D)}{L_1} & 0 \\ 0 & 0 & -\frac{1}{L_2} & -\frac{(1-D)}{L_2} \\ -\frac{(1-D)}{C_1} & \frac{1}{C_1} & 0 & 0 \\ 0 & \frac{(1-D)}{C_2} & 0 & -\frac{1}{C_2 R} \end{bmatrix} \begin{bmatrix} i_{L_1} \\ i_{L_2} \\ v_{C_1} \\ v_{C_2} \end{bmatrix} + \begin{bmatrix} \frac{D}{L_1} \\ \frac{D}{L_2} \\ 0 \\ 0 \end{bmatrix} E \quad (16)$$

The linear model of the system at an operating point is obtained by perturbing the average model with small A.C. signals resulting in:

$$\tilde{\dot{x}} = A\tilde{x} + B\tilde{u} \quad (17)$$

where:

$$A\tilde{x} = \begin{bmatrix} 0 & 0 & \frac{(1-D)}{L_1} & 0 \\ 0 & 0 & -\frac{1}{L_2} & -\frac{(1-D)}{L_2} \\ -\frac{(1-D)}{C_1} & \frac{1}{C_1} & 0 & 0 \\ 0 & \frac{(1-D)}{C_2} & 0 & -\frac{1}{C_2 R} \end{bmatrix} \begin{bmatrix} \dot{i}_{L_1} \\ \dot{i}_{L_2} \\ \dot{v}_{C_1} \\ \dot{v}_{C_2} \end{bmatrix} \quad (18)$$

$$B\tilde{u} = \begin{bmatrix} \frac{D}{L_1} & \frac{E}{(1-D)L_1} \\ \frac{D}{L_2} & \frac{E}{(1-D)^2 L_2} \\ 0 & \frac{ED(2-D)}{C_1 R(1-D)^4} \\ 0 & \frac{ED(2-D)}{C_1 R(1-D)^3} \end{bmatrix} \begin{bmatrix} \tilde{e} \\ \tilde{d} \end{bmatrix} \quad (19)$$

The state-space system consists of the state vector $x = [i_{L_1}, i_{L_2}, v_{C_1}, v_{C_2}]^T$, the control input is the ratio $d(t)$, applied synchronously to switches Q_1 and Q_2 , and the output variable is the load voltage $v_o(t) = v_{C_2}(t)$. The resulting converter is given in a representation single-input single-output (SISO) under the condition that the signals for switches Q_1 and Q_2 are operated synchronously. The system is fourth order under the conditions analyzed.

4. DYNAMIC ANALYSIS OF THE CONVERTER

The Buck-Boost converter was designed under R^2P^2 operation to step up the voltage from a 24 V DC microgrid to 170 V, corresponding to the required DC-link voltage of a single-phase inverter (120 V_{rms} AC output). Table 1 summarizes the design parameters used to validate the switched, averaged, and linear models.

Equation (20) presents the generalized transfer function representation of the state variables $i_{L_1}(G_1)$, $i_{L_2}(G_2)$, $v_{C_1}(G_3)$, and $v_{C_2}(G_4)$ with respect to the duty ratio (d).

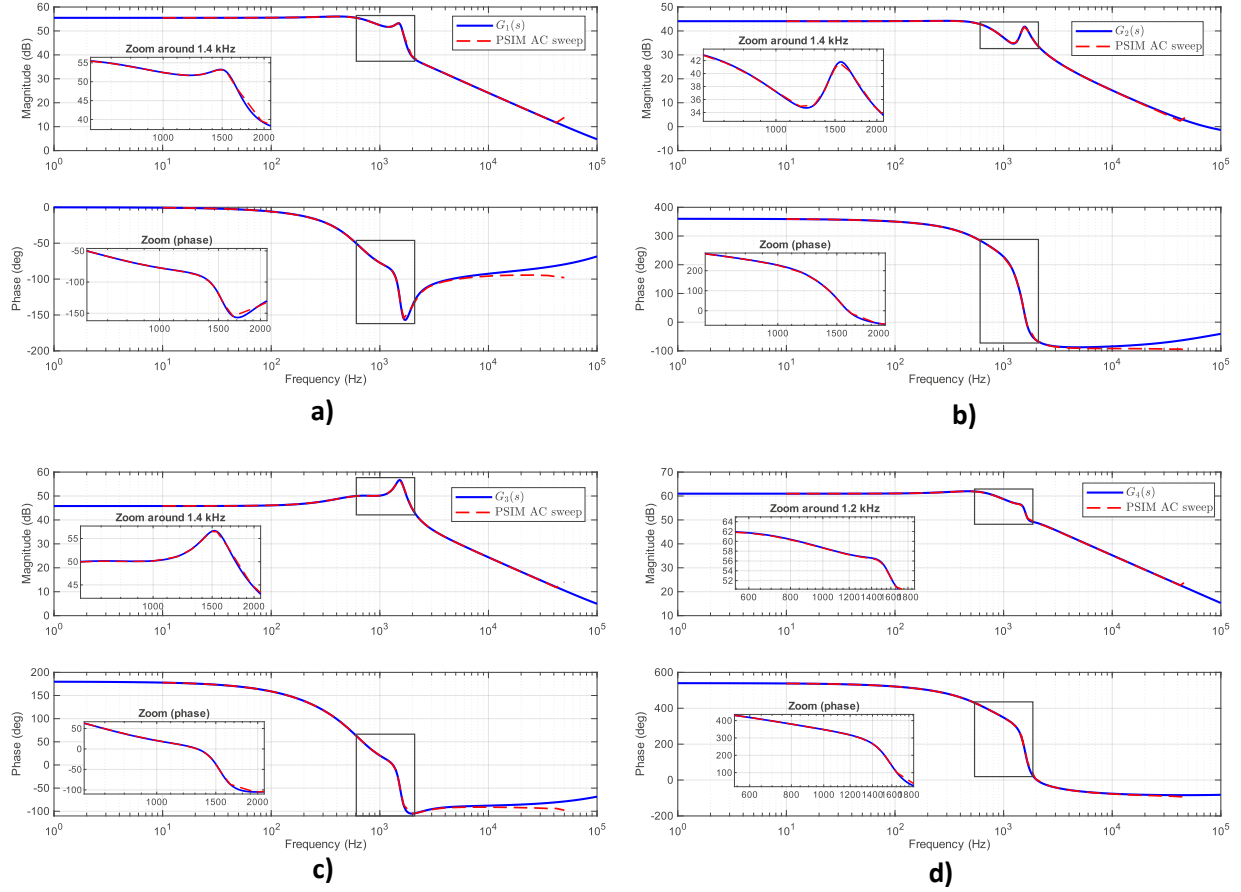


Fig. 4. Validation of the linear model and the switched model through the frequency response of the Buck-Boost converter state variables: a) $G_1(s)$, b) $G_2(s)$, c) $G_3(s)$, and d) $G_4(s)$.

Table 1. Design parameters.

Symbol	Parameter	Value
E	Input voltage	24 V
I_o	Output current	5.9 A
P_o	Output power	1000 W
V_o	Output voltage	170 V
f_s	Switching frequency	100 kHz
L_1, L_2	Inductors	68 μ H and 560 μ H
C_1, C_2	Capacitors	47 μ F and 4.7 μ F
D	Duty ratio	65%

$$G_{(1,2,3,4)}(s) = \frac{a_3 s^3 + a_2 s^2 + a_1 s + a_0}{D(s)} \quad (20)$$

The Table 2 lists the coefficients of the numerators associated with four transfer functions, $G_1(s)$, $G_2(s)$, $G_3(s)$, and $G_4(s)$, expressed as third-order polynomials in the complex variable s . The coefficients a_3 , a_2 , a_1 , and a_0 correspond to the terms s^3 , s^2 , s , and the constant term, respectively. It can be seen that all the transfer functions share a common denominator $D(s)$.

The system's dynamics are governed by four poles, which characterizes it as a fourth-order system. Furthermore, the lowest degree of the numerators confirms that the transfer functions are eigenfunctions. Fig.4 shows the frequency

Table 2. Transfer function coefficients.

FT	a_3	a_2	a_1	a_0
$G_1(s)$	1.008×10^6	1.274×10^{10}	1.627×10^{14}	1.059×10^{18}
$G_2(s)$	3.499×10^5	2.991×10^9	1.324×10^{13}	2.841×10^{17}
$G_3(s)$	1.033×10^6	7.541×10^9	9.569×10^{13}	-3.495×10^{17}
$G_4(s)$	3.616×10^6	-2.605×10^{10}	4.134×10^{14}	-1.997×10^{18}

$$D(s) = s^4 + 7.362 \times 10^3 s^3 + 1.229 \times 10^8 s^2 + 5.619 \times 10^{11} s + 1.784 \times 10^{15}$$

response of the four generalized transfer functions $G_1(s)$, $G_2(s)$, $G_3(s)$, and $G_4(s)$, comparing the linear model with the switched model obtained via AC sweep in PSIM. The solid line represents the linear model, while the dashed line corresponds to the switched model. There is a high degree of agreement between the two models across nearly the entire frequency spectrum, which confirms the accuracy of the linear model. In DC-DC converters, the most important variable is the output voltage, as it is the primary parameter required by the applications. Consequently, control strategies are designed based on its regulation. Therefore, it is essential to analyze the system dynamics by determining the locations of the poles and zeros with respect to this variable, which is related to G_4 . Fig. 5 shows the locus of the poles of the transfer function G_4 as the duty ratio D is varied from 10% to 90%.

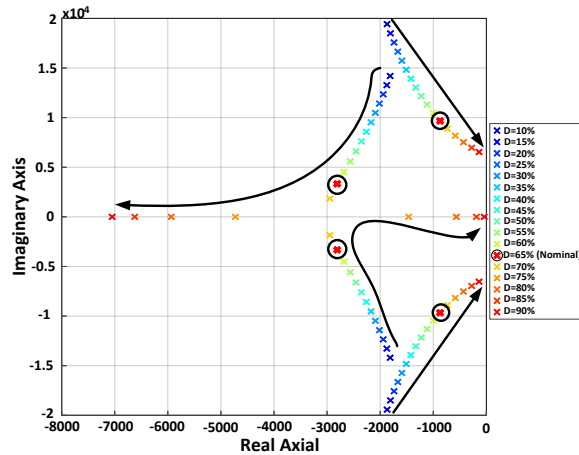


Fig. 5. Geometric locus of the poles of the transfer function G_4 as the duty ratio D (10%–90%) is swept across the complex plane.

It can be seen that, as D increases, the poles move in the complex plane along well-defined trajectories. For low values of D in buck mode, the poles are predominantly located in the left half-plane with a large negative real component, indicating fast and highly damped dynamics. As D increases in boost mode, the poles migrate to the right, reducing their damping and moving closer to the imaginary axis. At the nominal point ($D = 65\%$), highlighted in the Fig. 5, the system exhibits a trade-off between speed and damping. For larger values of D , some poles tend to concentrate in regions with a smaller negative real part, suggesting a deterioration in relative stability and an increase in system oscillations. Fig. 6 shows the behavior of the zeros of the transfer function G_4 in the complex plane as the duty ratio D varies.

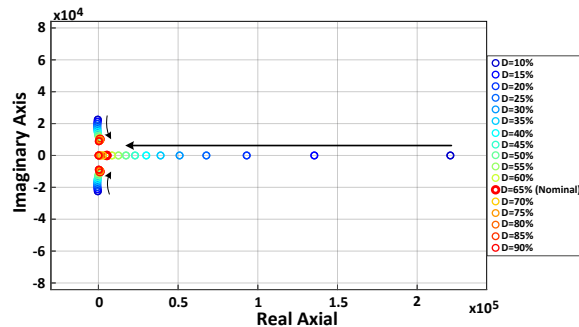


Fig. 6. Plot of the zeros of the transfer function G_4 in the complex plane as the duty ratio D varies.

It can be seen that, as D changes within the analyzed range, the zeros follow specific trajectories, highlighting the dependence of the system's dynamics on this parameter. The zeros are distributed primarily along the real axis, remaining in the left half-plane, which suggests that they do not introduce instabilities into the system. Furthermore, the nominal point ($D = 65\%$) stands out in the Fig. 6, making it possible to identify the location

of the zeros under typical operating conditions. Another parameter that plays an important role in the transfer function is the value of capacitor C_2 , as it is directly related to the voltage ripple at the system output. For this reason, this parameter is critical to the performance of both the converter and the controller. Fig. 7 shows the locus of the poles of the transfer function G_4 in the complex plane as the capacitor C_2 varies over a range of $\pm 50\%$ from its nominal value.

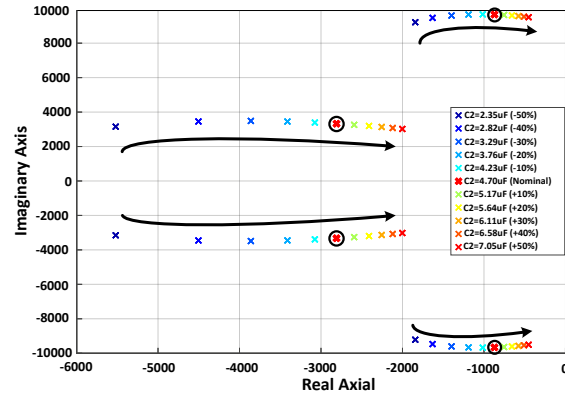


Fig. 7. Geometric locus of the poles of the transfer function G_4 as the capacitor C_2 ($\pm 50\%$) is swept across the complex plane.

The complex poles cause shifts in both the real and imaginary parts, altering the system's damping and natural frequency. As the value of C_2 increases, the poles tend to move toward regions with a smaller negative real part, implying a reduction in damping and a more oscillatory behavior. Conversely, as C_2 decreases, the poles shift to the left, indicating a faster and more damped response. The nominal point is highlighted in the Fig 7, allowing the design condition to be compared with extreme scenarios. The arrows indicate the direction of the poles' movement as C_2 increases. Fig. 8 shows the zeros as the value of capacitor C_2 varies; however, no significant changes in their offset relative to the nominal operating point are observed.

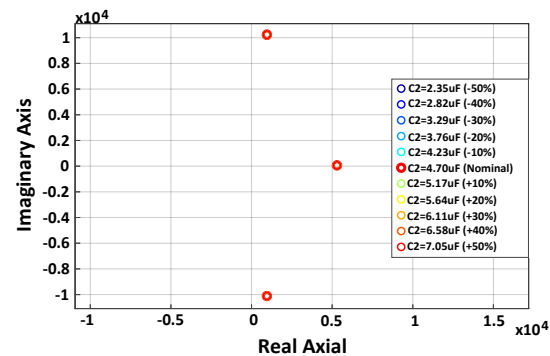


Fig. 8. Plot of the zeros of the transfer function G_4 in the complex plane as the capacitor C_2 varies.

Table 3. Quantitative dynamic results for three duty ratio values.

Parameter	D = 0.10	D = 0.65	D = 0.90
DC gain $G_4(s)$ (V)	2.41×10^1	1.35×10^3	2.98×10^4
Dominant poles (real part)	-8.20×10^3	-1.01×10^2	-1.20×10^1
Dominant poles (imag part)	$\pm 4.12 \times 10^3$	$\pm 7.86 \times 10^3$	$\pm 8.95 \times 10^3$
Damping ratio ξ	8.95×10^{-1}	1.30×10^{-2}	1.34×10^{-3}
Natural frequency ω_n (rad/s)	9.18×10^3	7.86×10^3	8.95×10^3
Real zero of $G_4(s)$ (RHP)	$+1.20 \times 10^4$	$+5.02 \times 10^3$	$+1.10 \times 10^3$
Complex zeros (LHP)	$-2.10 \times 10^3 \pm j1.85 \times 10^3$	$-1.10 \times 10^3 \pm j2.48 \times 10^3$	$-2.01 \times 10^2 \pm j3.10 \times 10^3$
Stability	stable	stable	stable (marginally)
Non-minimum phase	yes	yes	yes

Table 3 summarizes the quantitative results of the dynamic analysis for three representative duty cycle values: $D = 0.10$ (Buck mode), $D = 0.65$ (nominal operating point), and $D = 0.90$ (extreme Boost mode). The dominant poles shift toward the imaginary axis as D increases: its real part decreases from -8.20×10^3 to -1.20×10^1 , while the imaginary part increases slightly from $\pm 4.12 \times 10^3$ to $\pm 8.95 \times 10^3$. As a result, the damping ratio ξ decreases dramatically from 0.895 (well-damped system) to 0.00134 (nearly sustained oscillations). The natural frequency ω_n remains on the order of 10^3 rad/s.

5. CONCLUSION

This article presents a comprehensive dynamic analysis of a high-order Buck-Boost converter based on the concept of Reduced Redundant Power Processing. A state-space averaged model was developed along with its corresponding linearized representation, which allowed for an accurate description of the system's dynamic behavior. Based on these models, the poles, zeros, and frequency response of the converter were analyzed, identifying non-minimum phase characteristics that impose fundamental limitations on the design of control systems. The simulation results validated the proposed models, showing high agreement with the behavior of the converter during switching. It was demonstrated that reducing redundant power processing significantly influences the dynamic characteristics of high-order converters, particularly the location of poles and zeros in the complex plane as the duty cycle and output capacitor vary. These findings provide relevant information for the design of control strategies in high-efficiency power conversion systems, especially in applications such as DC microgrids, electric vehicles, and photovoltaic systems. In future work, control strategies will be designed to regulate the output while taking into account the system's dynamic characteristics. In addition, the converter will be experimentally implemented in order to validate the theoretical results by comparing them with experimental data.

ACKNOWLEDGEMENTS

The authors would like to thank the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) and the Universidad Autónoma de Zacatecas (UAZ) for their support in carrying out this work.

REFERENCES

Anuraag, B., Mahalakshmi, R., Likhtih, S., Bhargavi, P., and Mohanty, A. (2021). Design and comparative study of dc-dc quadratic buck-boost converter and cascaded

buck-boost converter. In *2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT)*, 642–646. IEEE. doi: 10.1109/RTEICT52294.2021.9573716.

Chen, X., Pise, A.A., Elmes, J., and Batarseh, I. (2019). Ultra-highly efficient low-power bidirectional cascaded buck-boost converter for portable pv-battery-devices applications. *IEEE Transactions on Industry Applications*, 55(4), 3989–4000. doi:10.1109/TIA.2019.2911566.

Diaz, V., Barrado, A., Zumel, P., and Lazaro, A. (2026). Optimized modulation for the four-switch buck-boost converter with reduced current ripple. *IEEE Open Journal of the Industrial Electronics Society*. doi: 10.1109/OJIES.2026.3653129.

Gholizadeh, H., Gorji, S.A., and Sera, D. (2023). A quadratic buck-boost converter with continuous input and output currents. *IEEE Access*, 11, 22376–22393. doi:10.1109/ACCESS.2023.3253102.

Klein Faistel, T.M., Illa Font, C.H., Santos Spencer Andrade, A.M., and Lúcio da Silva Martins, M. (2019). Modelling and control of dc-dc cuk converter with r^2p^2 and voltage multiplier cells for pv applications. *IET Power Electronics (Wiley-Blackwell)*, 12(9). doi: 10.1049/iet-pel.2019.0073.

Liu, X., Lei, K., Feng, Y., Wang, H., Bao, P., and Wan, Y. (2025). Reduced redundant power processing flicker-free led driver with adaptive peak voltage tracking control. *IEEE Journal of Emerging and Selected Topics in Power Electronics*. doi:10.1109/JESTPE.2025.3601631.

Reyes-Portillo, I.A., Claudio-Sánchez, A., Morales-Saldaña, J.A., Mina-Antonio, J.D., Netzahuatl-Huerta, E.M., Claudio-Pachecano, L., Ponce-Silva, M., and Reyes-Sánchez, E. (2022). Study of the effects of current imbalance in a multiphase buck converter for electric vehicles. *World Electric Vehicle Journal*, 13(5), 88. doi:https://doi.org/10.3390/wevj13050088.

Reyes-Portillo, I.A., Méndez-Elizondo, S.R., Morales-Saldaña, J.A., Muñoz-Minjares, J.U., Rivera-Romero, C.A., and Castro-López, D.L. (2026). Analysis and modeling of a boost converter with power processing reduction for pv applications. *IEEE Latin America Transactions*, 24(4), 412–421. doi:10.1109/TLA.2026.11435591.

Reyes-Portillo, I.A., Méndez-Elizondo, S.R., Morales-Saldaña, J.A., Rivera-Romero, C.A., and Castro-López, D.L. (2025). Design of a transformerless dc regulator with reduced redundant power processing for a residential dc microgrid. *IEEE Latin America Transactions*, 23(5), 427–436. doi:10.1109/TLA.2025.10974364.

Reyes-Portillo, I.A., Morales-Saldaña, J., Romero-Rivera, C., and Palacios-Hernández, E. (2023). Design and modeling of a high current ratio converter for pv ap-



- plications. *IEEE Latin America Transactions*, 21(10), 1144–1155. doi:10.1109/TLA.2023.10255452.
- Tse, C.K. (2003). Circuit theory of power factor correction in switching converters. *International journal of circuit theory and applications*, 31(2), 157–198. doi: <https://doi.org/10.1002/cta.216>.
- Xiong, Z. and He, L. (2025). An online neural network based model-free control method for high-order dc-dc converter. *IEEE Journal of Emerging and Selected Topics in Power Electronics*. doi: 10.1109/JESTPE.2025.3598871.
- Zogogianni, C.G., Tatakis, E.C., and Porobic, V. (2018). Investigation of a non-isolated reduced redundant power processing dc/dc converter for high-power high step-up applications. *IEEE Transactions on Power Electronics*, 34(6), 5229–5242. doi:10.1109/TPEL.2018.2868165.

Observer-Based Fault Diagnosis From Sampled Data^{*}

E. Alcorta-García^{*} D. A. Díaz-Romero^{*}
M. A. Platas-Garza^{*}

^{} Facultad de Ingeniería Mecánica y Eléctrica, Universidad Autónoma de Nuevo León, Mexico (e-mail: efrain.alcortagr@uanl.edu.mx; david.diazrr@uanl.edu.mx; miguel.platasgrz@uanl.edu.mx).*

Abstract: Many industrial processes are characterized by continuous-time signals, while most measurement devices operate digitally. Consequently, only sampled measurements of these signals are available. Although a discrete-time equivalent of the system can be derived, the presence of multiple faults may lead to coupling effects among them. In this work, a procedure to implement a structured observer-based residual generator for fault isolation using sampled data is proposed. An advantage of the proposed approach is the simplicity of its implementation, while preserving the properties of the continuous-time approach. An example is provided to illustrate the effectiveness of the proposed method.

Keywords: filter, observers, fault isolation, continuous time systems, discretization.

1. INTRODUCTION

Fault diagnosis has become a highly active research area due to increasing industrial demands for safety, reliability and performance. Beyond hardware redundancy, model-based and model-inspired approaches have been extensively studied, as evidenced in Ding (2021), Kong et al. (2024), Aldrini et al. (2024), and Girard et al. (2023).

Many industrial processes are inherently described by continuous-time signals, whereas most measurement devices operate in discrete time. As a result, only sampled data are available for monitoring and fault diagnosis purposes.

Several approaches to fault diagnosis in sampled-data systems have been reported in the literature. Early work addressing multi-rate sampled-data systems was presented in Viswanadham and Minto (1990). In Zhang et al. (2001), an operator-based framework is introduced to characterize the mapping between continuous-time signals over a finite time horizon and their discrete-time representations, yielding an exact description at the expense of requiring a finite observation window. Alternatively, Zhang et al. (2003) proposes a frequency-domain formulation; however, such approaches may face limitations associated with bandwidth constraints induced by sampling.

An optimal framework for fast-rate fault detection is proposed in Izadi et al. (2005). In addition, an H_∞ -based methodology for sampled-data systems is developed in Izadi et al. (2006), with further refinements discussed in Izadi et al. (2007).

From the reviewed literature, a key limitation can be identified: most existing approaches rely on an equivalent

discrete-time model either exact or approximate of the underlying continuous-time system. This transformation typically introduces coupling among faults defined in the continuous-time domain and may increase the complexity of the fault isolation task. In Niemann and Poulsen (2015), a different approach based on so-called active fault diagnosis (i.e., the injection of an additional signal into the system) is proposed. Two schemes are presented that can handle fault coupling; in fact, the presence of fault coupling is also demonstrated.

In this work, a novel approach is proposed for fault detection and isolation directly at the continuous-time level using sampled data. The method is based on structured residuals Alcorta García and Frank (1999) and incorporates concepts from continuous-time system identification using sampled measurements Garnier and Wang (2008). The proposed approach avoids the need for an equivalent discrete-time model, requires only an observer of the same order as the system, and leads to a simpler structure for fault isolation. A numerical example is provided to illustrate the effectiveness of the proposed method.

2. PRELIMINARIES

2.1 Eigenstructure of Linear Systems

In linear systems theory, the *eigenstructure* of a system extends the notion of eigenvalues by incorporating the associated eigenvectors (and, when relevant, generalized eigenvectors). While eigenvalues determine the system's modal dynamics—such as stability, natural frequencies, and damping—the eigenstructure provides a more complete characterization by also describing the directional behavior of the state trajectories.

Consider a continuous-time linear time-invariant (LTI) system:

^{*} Sponsor and financial support acknowledgment goes here. Paper titles should be written in uppercase and lowercase letters, not all uppercase.



$$\dot{x}(t) = Ax(t) \quad (1)$$

where $A \in \mathbb{R}^{n \times n}$. The eigenstructure of A consists of pairs (λ_i, v_i) , where λ_i are the eigenvalues and v_i the corresponding right eigenvectors satisfying:

$$Av_i = \lambda_i v_i \quad (2)$$

The solution to the system can be expressed as a linear combination of modal components:

$$x(t) = \sum_{i=1}^n \alpha_i e^{\lambda_i t} v_i \quad (3)$$

where the coefficients α_i depend on the initial condition.

In observer design, consider the Luenberger observer:

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + L(y(t) - C\hat{x}(t)) \quad (4)$$

which leads to the estimation error dynamics:

$$\dot{e}(t) = (A - LC)e(t) \quad (5)$$

The observer gain L can be used to assign the eigenvalues of $(A - LC)$ if (A, C) is observable. However, eigenvector assignment is constrained.

Result (Eigenvector Assignability in Observers).

Let (A, C) be an observable pair with $C \in \mathbb{R}^{p \times n}$. Then, for a given set of n distinct eigenvalues, at most p independent eigenvectors of $(A - LC)$ can be arbitrarily assigned. The remaining $n - p$ eigenvectors are structurally constrained.

Procedure for Eigenvector Assignment.

Let $\{\lambda_i\}_{i=1}^n$ be the desired eigenvalues. For each eigenvalue λ_i , the associated eigenvector v_i of $(A - LC)$ must satisfy:

$$(A - \lambda_i I)v_i = LCv_i \quad (6)$$

A constructive procedure is as follows:

- (1) Select up to p eigenvectors $\{v_i\}$ to be assigned freely, ensuring that they are linearly independent and satisfy $Cv_i \neq 0$.
- (2) For each selected pair (λ_i, v_i) , compute:

$$w_i := (A - \lambda_i I)v_i \quad (7)$$

- (3) Enforce the compatibility condition:

$$w_i = L(Cv_i) \quad (8)$$

which can be rewritten as a linear constraint on L :

$$Lz_i = w_i, \quad \text{with } z_i := Cv_i \quad (9)$$

- (4) Stack the constraints for all assignable eigenvectors:

$$LZ = W \quad (10)$$

where $Z = [z_1 \cdots z_r] \in \mathbb{R}^{p \times r}$ and $W = [w_1 \cdots w_r] \in \mathbb{R}^{n \times r}$, with $r \leq p$.

- (5) If Z has full row rank, compute L as:

$$L = WZ^\dagger \quad (11)$$

where $Z^\dagger$ denotes the Moore–Penrose pseudoinverse. Additional degrees of freedom can be used to assign the remaining eigenvalues.

The remaining eigenvectors are implicitly determined by the choice of L and cannot be independently specified. This procedure highlights the trade-off between eigenvalue placement and eigenvector shaping.

2.2 Model-Based Fault Detection and Isolation

Model-based fault detection and isolation (FDI) relies on the use of a mathematical model to monitor the consistency between the expected and the measured behaviour of a dynamical system. The fundamental idea is to generate diagnostic signals, commonly referred to as *residuals*, which are sensitive to faults while remaining insensitive to disturbances, noise, and modelling uncertainties to the extent possible.

Consider a linear time-invariant system described by:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + E_f f(t) + Ed(t), \\ y(t) &= Cx(t) \end{aligned} \quad (12)$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ the control input, $y(t) \in \mathbb{R}^p$ the measured output, and $f(t) \in \mathbb{R}^q$ represents faults. The matrix E_f defines how faults affect the system dynamics, $d(t) \in \mathbb{R}^d$ is an unknown input.

A residual generator is typically constructed using an observer:

$$\begin{aligned} \dot{\hat{x}}(t) &= A\hat{x}(t) + Bu(t) + L(y(t) - \hat{y}(t)), \\ \hat{y}(t) &= C\hat{x}(t) \end{aligned} \quad (13)$$

where $\hat{x}(t)$ is the estimated state and L is the observer gain matrix. The residual is defined as:

$$r(t) = y(t) - \hat{y}(t) \quad (14)$$

Under nominal (fault-free) conditions, the residual converges to zero asymptotically, assuming appropriate observer design. In the presence of faults, the residual deviates from zero, providing a mechanism for fault detection. The sensitivity of the residual to faults depends on the system structure and the design of the observer.

For fault isolation, a structured residual approach is often employed. This consists in designing a set of residuals $\{r_1(t), r_2(t), \dots, r_k(t)\}$ such that each residual is sensitive to a specific subset of faults and decoupled from others. The resulting residual pattern allows identification of the fault location based on a predefined fault signature matrix.

The effectiveness of model-based FDI depends on several factors, including model accuracy, disturbance rejection, and robustness to noise. Trade-offs typically arise between sensitivity to faults and robustness, requiring careful design of the residual generator, often through techniques such as observer-based design, parity space methods, or optimization-based approaches.

In summary, model-based FDI provides a systematic framework for detecting and isolating faults by exploiting analytical redundancy in system models, enabling early fault diagnosis and improved system reliability.

2.3 Structured residual generation

In this section, the results from Alcorta Garcia and Frank (1999) are adapted. Let us consider an observer-based residual generator (12).

Lemma. Consider the system (12). A robust observer-based residual generator robust to the unknown inputs $d(t)$ is given by

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + L(y(t) - C\hat{x}(t)) \quad (15)$$

$$z_i(t) = T_i\hat{x}(t) \quad (16)$$

$$r_i(t) = Q_i z_i(t) + P_i y(t) \quad (17)$$

with

$$Q_i T_i + P_i C = 0 \quad (18)$$

$$T_i E_d = 0 \quad (19)$$

$$T_i E_f \neq 0 \quad (20)$$

when the eigenvectors of $(A - LC)$ can be selected as the columns of $\{E_f, E_d\}$. $\square\square$

2.4 Problem formulation

Consider a continuous-time linear time-invariant system affected by faults:

$$\dot{x}(t) = Ax(t) + Bu(t) + E_f f(t) + E_d d(t),$$

$$y(t) = Cx(t)$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^m$ is the known input, $y(t) \in \mathbb{R}^p$ is the measured output, and $f(t) \in \mathbb{R}^q$ represents unknown faults. The matrices A , B , C , E_f and E_d are assumed to be known and of appropriate dimensions.

In practice, measurements are not available continuously but only at discrete-time instants $t_k = kT_s$, where T_s is the sampling period. Thus, the available data are given by:

$$y_k := y(t_k), \quad u_k := u(t_k) \quad (21)$$

A common approach consists in deriving a discrete-time equivalent model of the system. However, when multiple faults are present, the discretization process may introduce coupling effects that obscure the individual influence of each fault on the measurements. This can degrade the fault isolation capability, even if detectability is preserved.

The problem addressed in this work is therefore the design of a residual generator that:

- exploits sampled input-output data $\{u(t_k), y(t_k)\}$,
- preserves the fault sensitivity structure of the original continuous-time model,
- enables both fault detection and isolation without explicitly relying on a fully discretized fault model.

More specifically, the objective is to construct a set of residual signals $\{r_i(t_k)\}$ such that:

$$r_i(t_k) \approx 0 \quad \text{in the absence of faults,}$$

$$r_i(t_k) \neq 0 \quad \text{when fault } f_i(t) \text{ is present}$$

while maintaining robustness with respect to measurement noise and modelling uncertainties.

Existing approaches to fault detection and isolation using sampled data can be broadly classified into discrete-time and continuous-time based methods. In the first category, the system is discretized using techniques such as zero-order hold, and standard discrete-time observer or parity space methods are applied. While these methods are

well established, it has been reported that discretization may introduce fault coupling, particularly in multi-fault scenarios, which complicates isolation.

Alternatively, several works propose to retain a continuous-time model structure while processing sampled data. Observer-based approaches that estimate continuous-time states directly from sampled measurements have shown improved fault isolation properties. In particular, structured residual design methods have demonstrated the ability to decouple fault effects by exploiting the original continuous-time fault distribution matrix E_f .

Recent contributions also investigate subspace-based and integral methods, where continuous-time models are identified or utilized directly from sampled data without explicit discretization. These approaches can mitigate numerical issues and preserve structural properties relevant for FDI.

Overall, the literature indicates a trade-off between implementation simplicity and fault isolation performance. Discrete-time methods are straightforward but may suffer from structural limitations, whereas continuous-time formulations using sampled data offer improved isolation capabilities at the cost of increased computational and design complexity.

3. MAIN RESULTS

3.1 Structured fault detection and isolation

Several approaches have been proposed to generate structured residuals; see, for instance, Blanke et al. (2016); Chen and Patton (1999); Ding (2021). Many of these approaches are, to some extent, equivalent. In this work, the formulation in Ding (2021) is adopted due to its favorable numerical properties, its suitability for optimization-based design, and its connection with parity space methods.

Consider the following observer-based residual generator Ding (2021):

$$\dot{z}(t) = Gz(t) + Hu(t) + Ly(t)$$

$$r(t) = -wz(t) + vy(t) + qu(t)$$

subject to:

$$TA - GT = LC$$

$$vC - wT = 0$$

$$H = TB - LD$$

where G is Hurwitz. Defining $z(t) \triangleq T\hat{x}(t)$, the dynamics of the residual generator can be written as:

$$\dot{z}(t) = T\dot{\hat{x}}(t) = GT\hat{x}(t) + TBu(t) - LDu(t) + Ly(t) \quad (22)$$

Using (22), i.e., $GT = TA - LC$, it follows that:

$$T\dot{\hat{x}}(t) = TA\hat{x}(t) - LC\hat{x}(t) + TBu(t) + Ly(t)$$

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) + T^{-1}L(y(t) - C\hat{x}(t)) \quad (23)$$

which corresponds to (15), thus establishing the equivalence.

It is important to note that the performance of the residual generator strongly depends on the selection of the observer gain. In particular, through the appropriate assignment of the eigenstructure of the matrix $A - LC$, it is possible to influence both robustness properties and fault isolation capabilities. However, while the eigenvalues can be arbitrarily assigned under observability conditions, the associated eigenvectors are subject to structural constraints. This limitation plays a central role in the design of structured residuals.

3.2 Proposed approach

Previous approaches to observer-based fault detection and isolation in sampled data systems follow the idea of use a discrete time representation of the system, in some cases approximated in other exact. This ways has some advantages, but early approaches have the inconvenient that the fault distribution matrix, approximated as well as exact becomes coupled with respect to the continuous time faults.

The proposed approach use an idea used in the parameter identification for continuous systems from sampled data Garnier and Wang (2008), in which based on an output error equation a state variable filter is used to get filtering signals to be used instead of derivatives of the input-output signals.

The starting point is the residual equation (17):

$$r_i(t) = Q_i T_i \hat{x}(t) + P_i y(t) \quad (24)$$

this equation is actually valid for all $t \geq 0$, in particular for the time $t_k = kT_s$, i.e.

$$r_i(t_k) = Q_i T_i \hat{x}(t_k) + P_i y(t_k) \quad (25)$$

this observation has an important consequence: in order to get the residual signal (25) only the value of the signals for the $\hat{x}$ and y at the time instants $t_k = kT_s$.

On the other size, in order to get the residual signal (25) only the samples of the input and outputs: $(u(t_k), y(t_k))$ available.

The differential equation (15) requires the signals $u(t)$ and $y(t)$, i.e. the continuous time signals. Instead, and because only the samples are available, the discrete time version of (15) is proposed to be used:

$$\hat{x}(t_{k+1}) = \mathcal{A}x(t_k) + \mathcal{B}u(t_k) + \mathcal{K}y(t_k) \quad (26)$$

where

$$\mathcal{A} = e^{(A-LC)T_s} \quad \mathcal{B} = \int_0^{T_s} e^{(A-LC)\tau} B d\tau \quad \mathcal{K} = \int_0^{T_s} e^{(A-LC)\tau} L d\tau$$

Note that the solution of the difference equation (26) is the required variable for the residual as defined by (25). The decoupling, however, can be done on the basis of the continuous time model of the system. A schema of the proposed approach can be seen in figure 1.

4. APPLICATION EXAMPLE

In order to show the above results lets us consider the Take-Home-Laboratory-FDI (THL-FDI) facility, Alcorta-Garcia et al. (2026). The THL-FDI is a portable laboratory

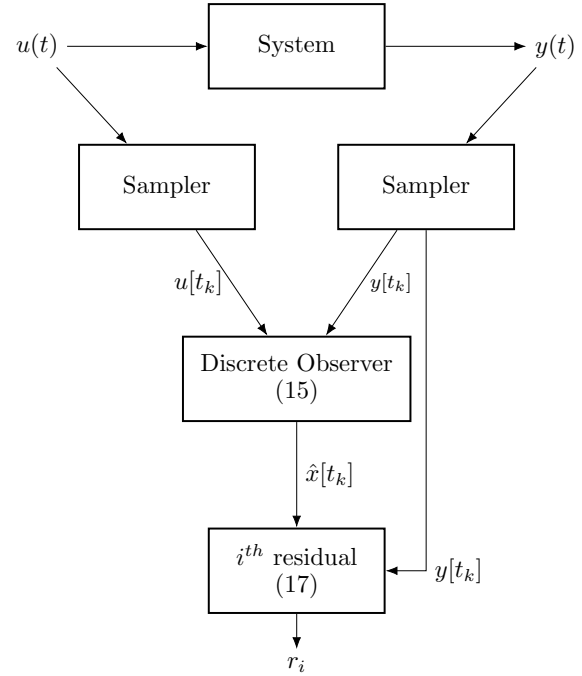


Fig. 1. Structured residual generation scheme.

which allows us to obtain input-output data for different fault situations in order to get data for testing of fault detection and diagnosis approaches. The implementation of the THL-FDI discussed in Alcorta-García et al. (2025) can be seen in figure 2.

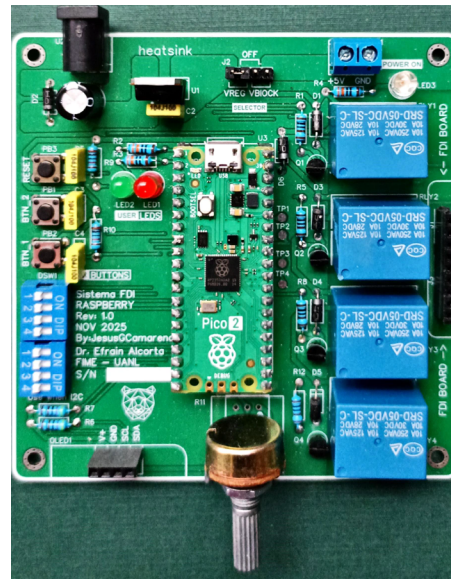


Fig. 2. Photo of the THL-FDI.

4.1 Nominal system model

The "plant" consists in an electronic circuit formed by two resistors and two capacitors connected as shown in the figure 3.

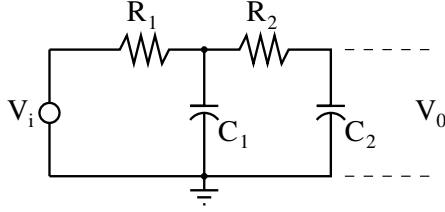


Fig. 3. Schematic of the nominal system

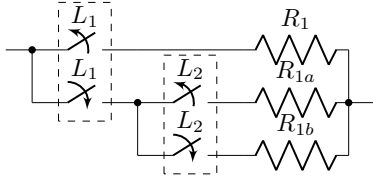


Fig. 4. Fault implementation: changes in R_1 .

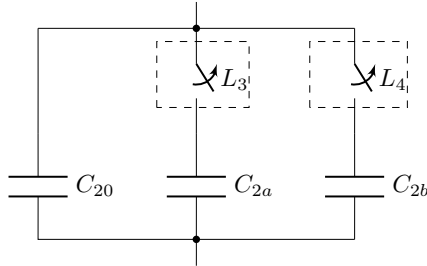


Fig. 5. Fault implementation: changes in C_2 .

where $R_1 = R_2 = 1M\Omega$, $C_1 = C_2 = 1\mu F$. The state space description of the nominal system is given by:

$$\dot{x}(t) = \begin{bmatrix} -\frac{1}{R_1 C_1} - \frac{1}{R_2 C_1} & \frac{1}{R_2 C_1} \\ \frac{1}{R_2 C_1} & -\frac{1}{R_2 C_2} \end{bmatrix} x(t) + \begin{bmatrix} \frac{1}{R_1 C_1} \\ 0 \end{bmatrix} V_i(t)$$

4.2 Fault representation

A key aspect in fault diagnosis is the way in which faults are modelled Ding (2021). The THL-FDI consider changes in the value of two components: changes in the value of R_1 as well as in C_2 . See the figure 4 and 5.

As can be seen in the nominal system equations the parameters appear in the denominator. Following Alcorta Garcia (1999), the effect of the faults are represented (in an approximated way) as:

$$\frac{1}{R_1} \approx \frac{1}{R_{10}} - \frac{1}{R_{10}^2} \Delta R_1 \quad \frac{1}{C_2} \approx \frac{1}{C_{20}} - \frac{1}{C_{20}^2} \Delta C_2 \quad (27)$$

4.3 Residual design

The nominal model results:

$$\dot{x}(t) = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix} x(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} V_i \quad (28)$$

$$y(t) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x(t) \quad (29)$$

and consider for the output the measure of the two states (to make the eigenvector selection possible). The design of the general observer requires the selection of the observer poles as well as the corresponding eigenvectors. The required Eigenvectors are $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ and the corresponding eigenvalues are: $\{-15, -15\}$

The observer gain is obtained by $L = WZ^\dagger$, where the columns of Z are given by $z_i = C\nu_i$, with ν_i is a eigenvector, and each column of W is given by $w_i = (A - \lambda_i I)\nu_i$, i.e. for the system:

$$W = \begin{pmatrix} 13 & 1 \\ 1 & 14 \end{pmatrix}; \quad Z = Z^\dagger = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad L = WZ^\dagger = \begin{pmatrix} 13 & 1 \\ 1 & 14 \end{pmatrix}$$

The general observer becomes

$$\dot{\hat{x}}(t) = \begin{bmatrix} -15 & 0 \\ 0 & -15 \end{bmatrix} \hat{x}(t) + \begin{bmatrix} 1 \\ 0 \end{bmatrix} V_i + \begin{bmatrix} 13 & 1 \\ 1 & 14 \end{bmatrix} y(t) \quad (30)$$

For the residual robust to the capacitor faults:

$$T_1 = [1 \ 0]; \quad P_1 = [-L_1 \ 0]; \quad \text{for a given } L_1$$

For the residual robust to the resistor faults:

$$T_2 = [0 \ 1]; \quad P_2 = [0 \ -L_2]; \quad \text{for a given } L_2$$

4.4 Observer discretization

Discretizing equation (30) results in

$$\mathcal{A} = \begin{bmatrix} 0.7408 & 0 \\ 0 & 0.7408 \end{bmatrix} \quad \mathcal{B} = \begin{bmatrix} 0.0207 \\ 0 \end{bmatrix} \quad \mathcal{K} = \begin{bmatrix} 0.2333 & 0.0173 \\ 0.0190 & 0.2471 \end{bmatrix}$$

4.5 Results

Simulation results are shown in figure 6, for the fault free case. See figure 7 for a fault in R_1 as well as figure 8 for a fault in C_2 .

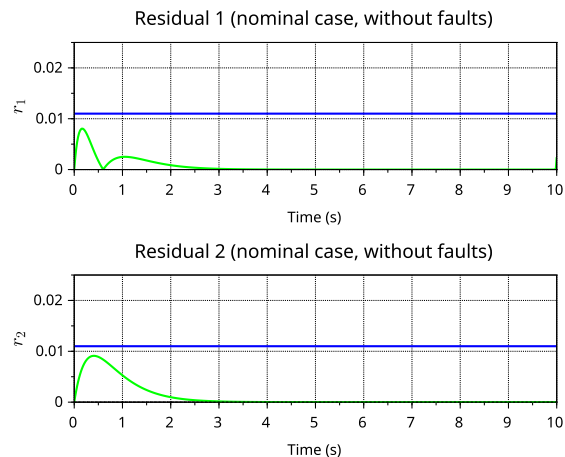


Fig. 6. Residual behaviour in the fault free case

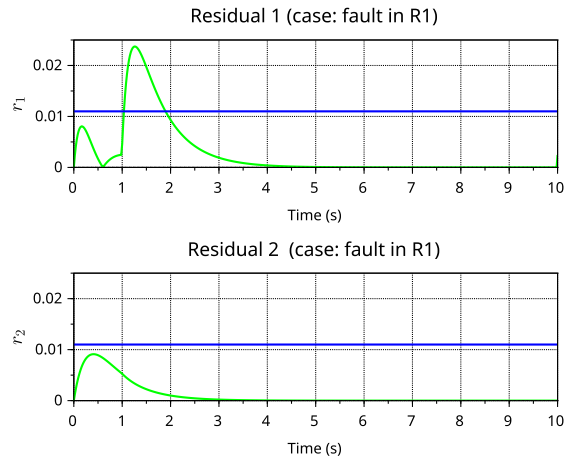


Fig. 7. Residual behaviour in the case of fault in R_1

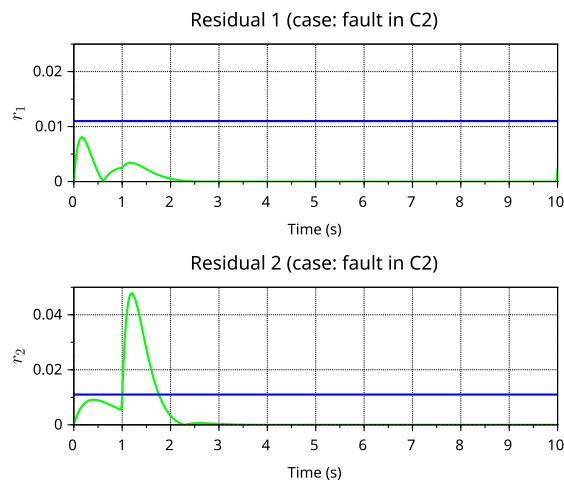


Fig. 8. Residual behaviour in the case of fault in C_2

A threshold value of 0.011 has been included in order to avoid false alarms. As can be seen in Figures 6–8, the residuals respond according to the expected design.

5. CONCLUSION

This work presents an approach for implementing observer-based residuals for fault isolation using sampled data. Although discretization introduces an approximation, the method remains simple and, when combined with a structural approach to residual generation, results in a practical solution. The proposed algorithm has been validated using a THL-FDI facility. The results indicate that the approach is suitable for practical applications.

Future work should include comparisons with other approaches reported in the literature.

REFERENCES

Alcorta-García, E., Platas-Garza, M.A., Díaz-Romero, D.A., Zambrano-Serrano, E., and Garza-Camarena,

- J.D. (2026). *Complex Systems and Their Applications: Sixth International Conference (EDIESCA2025)*, chapter Portable Laboratory for Self-Learning FaultDiagnosis, 1–13. Springer.
- Alcorta-García, E., Vázquez-Gonzalez, E.E., García-Lazaro, A., Platas-Garza, M.A., Díaz-Romero, D.A., and G.Romero-Galván (2025). Portable laboratory for fault diagnosis education: Component faults module. *Tecnología y ciencia aplicada*, 8(1), 488–493.
- Alcorta Garcia, E. (1999). *Modelgestützte Residuengenerierung für die Diagnose von additiven und multiplicativen Fehler in dynamischen Systemen (in german)*. Ph.D. thesis, Gerhard-Mercator University of Duisburg (Duisburg-Essen Universität), Germany.
- Alcorta Garcia, E. and Frank, P.M. (1999). A novel design of structured observer-based residuals for fdi. In *Proceedings of the American Control Conference*, 1341–1345. San Diego, Cal. USA.
- Aldrini, J., Chihi, I., and Sidhom, L. (2024). Fault diagnosis and self-healing for smart manufacturing: a review. *Journal of Intelligent Manufacturing* (, 35, 2441–2473.
- Blanke, M., Kinnaert, M., Lunze, J., and Staroswiecki, M. (2016). *Diagnosis and Fault-Tolerant Control*. Springer.
- Chen, J. and Patton, R.J. (1999). *Robust model based fault diagnosis for dynamic systems*. Kluwer Academic Publishers Group.
- Ding, S.X. (2021). *Advanced methods for fault diagnosis and fault-tolerant control*. Springer.
- Garnier, H. and Wang, L. (eds.) (2008). *Identification of Continuous-time Models from Sampled Data*. Advances in Industrial Control. Springer.
- Girard, P., Blanton, S., and Wang, L.C. (2023). *Machine Learning Support for Fault Diagnosis of System-on-Chip*. Springer.
- Izadi, I., Zhao, Q., and Chen, T. (2005). An optimal scheme for fast rate fault detection based on multirate sampled data. *Journal of Process Control*, 15, 307–319.
- Izadi, I., Zhao, Q., and Chen, T. (2006). An h_∞ approach to fast rate fault detection for multirate sampled-data systems. *Journal of Process Control*, 16, 651–658.
- Izadi, I., Zhao, Q., and Chen, T. (2007). Analysis of performance criteria in sampled-data fault detection. *Systems & Control letters*, 56, 320–325.
- Kong, X., Luo, J., and Feng (2024). *Process Monitoring and FaultDiagnosis Basedon Multivariable Statistical Analysis*. Springer.
- Niemann, H.H. and Poulsen, N.K. (2015). Active fault diagnosis in sampled-data systems. In *IFAC-PapersOnLine* 48-21, 883–888.
- Viswanadham, N. and Minto, K.D. (1990). Fault diagnosis in multirate sampled data systems. In *29th IEEE Conference on Decision and Control*, 3666–3671. Honolulu, Hawaii.
- Zhang, P., Ding, S.X., Wang, G.Z., and Zhou, D.H. (2001). An FDI approach for sampled-data systems. In *Proceedings of the American Control Conference 2001*, 2702–2707. VA, USA.
- Zhang, P., Ding, S.X., Wang, G.Z., and Zhou, D.H. (2003). A frequency domain approach to fault detection in sampled-data systems. *Automatica*, 39, 1303–1307.

Design of a Buck-Boost Converter Based on the Concept of Reduced Redundant Power Processing

Juan L. Alemán-Martínez* Santiago A. Gallardo-García*
Alberto Maldonado-Aguilar* Edgar S. Cano-Enríquez*
Emilio Hernández-Zúñiga* Iván A. Reyes-Portillo*,**

* *Departamento de Electrónica, Centro de Formación Profesional (CFP)/ Bachillerato Tecnológico Salesiano Carlos Gómez, 78240, San Luis Potosí, S.L.P., México.*

** *Academia de Sistemas y Tecnologías Industriales/Universidad Politécnica de San Luis Potosí, 78363, San Luis Potosí, S.L.P., México, (e-mail: ia.reyes@ieee.org).*

Abstract: This article presents the design and analysis of a high-conversion-ratio buck-boost converter with reduced redundant power processing, aimed at improving efficiency in applications that require wide voltage conversion ranges. The proposed topology minimizes the fraction of processed power by partially transferring energy directly from the input to the output, thereby reducing conduction losses and improving the overall performance of the system. A steady-state analysis is performed to determine the voltage conversion ratio and identify the operating conditions. Additionally, the influence of passive components on the system's dynamic response is evaluated. Simulation results validate the theoretical analysis, demonstrating high voltage gain, greater efficiency, and lower stress on the semiconductor devices compared to conventional buck-boost converters. The proposed converter is suitable for applications such as renewable energy systems and DC microgrids, where high efficiency and wide conversion ranges are required.

Keywords: Power electronics, Buck-Boost, DC-DC converter, R^2P^2 .

1. INTRODUCTION

Recent advances in semiconductor technologies, particularly silicon carbide (SiC) and gallium nitride (GaN) devices, have enabled the development of dc/dc converters that operate at higher switching frequencies, featuring reduced size and enhanced energy efficiency (Roccaforte et al., 2022). In this context, modern power conversion systems are required to simultaneously achieve high efficiency, high power density, wide voltage conversion range, and robust dynamic performance under disturbances and load variations (Singh et al., 2022). However, fulfilling all these requirements within a single topology remains a significant challenge due to inherent device limitations, as well as thermal, electromagnetic, and control constraints, which inevitably lead to trade-offs among key design parameters. Cascaded converter configurations have been widely adopted to achieve high voltage conversion ratios. These structures, formed by interconnecting basic converter stages, allow the electrical stress to be distributed among semiconductor devices (Zhang et al., 2022). Nevertheless, they suffer from a major drawback associated

with redundant power processing, since the input energy is processed through multiple stages, thereby increasing overall losses and reducing system efficiency (Reddy et al., 2024). In some cases, voltage multiplier cells based on diodes and capacitors are incorporated to further increase the conversion gain; however, this approach results in a higher component count, increased system volume, and reduced power density (Andres et al., 2023).

To overcome these limitations, converter architectures based on the concept of partial power processing (PPP) have recently emerged. In such structures, only a fraction of the total power is processed by active devices, while the remaining power is directly transferred from the source to the load through passive or coupled paths (Heo and Kim, 2026). This principle is implemented in topologies known as reduced redundant power processing (R^2P^2), which significantly reduces conduction and switching losses, as well as the electrical stress on semiconductor devices. As a result, higher efficiency can be achieved without compromising converter functionality. Moreover, this approach enhances energy utilization and is particularly attractive for applications requiring high conversion ratios (Reyes-Portillo et al., 2023b, 2025).

In this paper, a buck-boost converter based on the R^2P^2 concept is proposed. A steady-state analysis is carried out to derive the voltage conversion ratio and to establish

* The authors thank the Centro de Formación Profesional (CFP) / Bachillerato Tecnológico Salesiano Carlos Gómez for providing the resources used in this work, and the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) for its contribution to the development of this research..

the operating conditions. Finally, simulation results are presented to validate the performance of the proposed converter and to demonstrate its advantages in terms of efficiency and overall behavior.

2. PROPOSED BUCK-BOOST CONVERTER

The topology proposed in this paper is based on the concept of R^2P^2 , which has emerged as an effective strategy for improving energy efficiency by minimizing unnecessary power processing stages (Santosh Kumar Reddy and Obulesu, 2023). Unlike conventional converters, where the entire input power is processed by semiconductor devices, the R^2P^2 approach optimizes the energy transfer paths, significantly reducing conduction and switching losses, as well as the thermal stress on components. The fundamental principle of R^2P^2 lies in allowing a portion of the input power (P_{in}) to be transferred directly to the load without being fully processed by the active conversion stage. Under this scheme, only the differential power is regulated by the converter, leading to substantial improvements in overall efficiency. The topology can be interpreted as a multiport system characterized by the interaction of three main energy nodes: the input source, the load, and a storage element (buffer). The latter acts as a coupling port that enables alternative power flow paths (Reyes-Portillo et al., 2023a). Due to this simultaneous power distribution, the converter is classified as a Type I-IIA energy transfer structure. Fig. 1 shows the corresponding generalized structure alongside the proposed configuration.

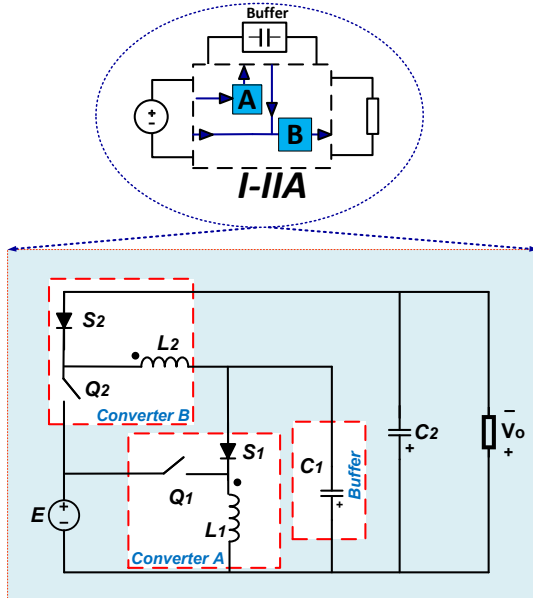


Fig. 1. General structure of the I-IIA configuration with R^2P^2 , on which the proposed converter is based..

The system comprises an input voltage source E , two active switches (Q_1 and Q_2), two switching diodes (S_1 and S_2), two inductors (L_1 , L_2), and two capacitors (C_1 , C_2). It is worth noting that capacitor C_1 serves as the buffer between converter A and converter B.

2.1 Steady-State Analysis

The converter exhibits two conduction modes: Continuous Conduction Mode (CCM), in which the inductor currents never reach zero, and Discontinuous Conduction Mode (DCM), where the inductor currents fall to zero during an instant within the switching period. This paper addresses only the CCM operation. In CCM, the converter operates in two states. During the ON state, switches Q conduct while diodes S remain off. During the OFF state, the diodes conduct and switches Q are turned off. Through a steady-state analysis, it is possible to determine the equilibrium points of the semiquadratic buck-boost converter. This analysis is valid only for CCM operation and under the assumption of ideal components, neglecting parasitic elements such as resistances, capacitances, and inductances. Fig. 2 shows the steady-state waveforms over one switching period.

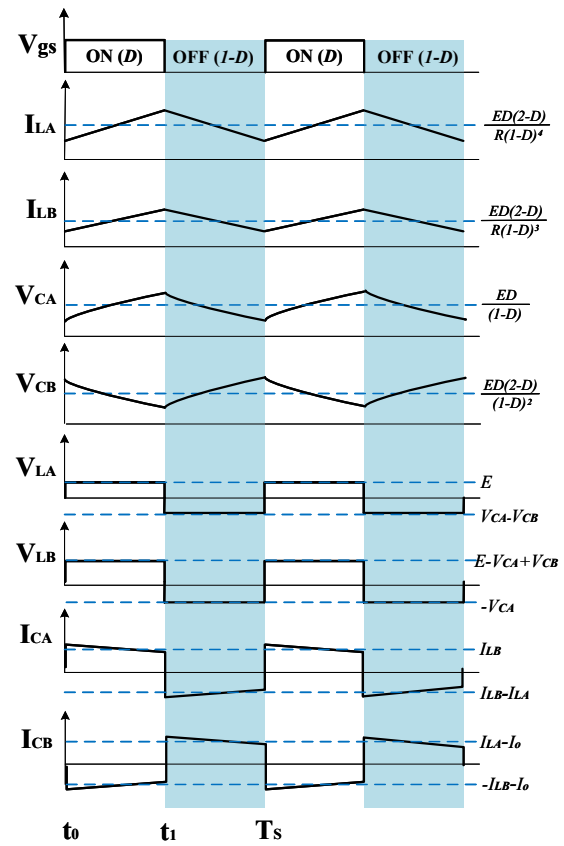


Fig. 2. Steady-state waveforms over one switching period.

During the ON state, the reactive elements L_1 , L_2 , and C_1 are charged, while C_2 is discharged. In the OFF state, L_1 , L_2 , and C_1 deliver their stored energy, and C_2 is charged. Expressions (1-4) represent the steady-state nominal voltages and currents of the capacitors and inductors in terms of the input voltage (E), duty ratio (D), and load resistance (R).

$$V_{C1} = \frac{ED}{(1-D)} \quad (1)$$

$$V_{C2} = \frac{ED(2-D)}{(1-D)^2} \quad (2)$$

$$I_{L1} = \frac{ED(2-D)}{R(1-D)^4} \quad (3)$$

$$I_{L2} = \frac{ED(2-D)}{R(1-D)^3} \quad (4)$$

The following expressions establish the sizing of the capacitors required for the converter to operate in CCM. The percentage of the voltage ripple (ΔV) must be considered according to the load being fed. These Expressions (5-6) are based on the steady-state analysis of the converter.

$$C_1 = \frac{ED^2(2-D)}{\Delta V_{C1} f_s (1-D)^3 R} \quad (5)$$

$$C_2 = \frac{ED^2(2-D)}{\Delta V_{C2} f_s (1-D)^2 R} \quad (6)$$

where f_s is the switching frequency. The following expressions are proposed to determine the value of the inductors:

$$L_1 = \frac{ED}{\Delta I_{L1} f_s} \quad (7)$$

$$L_2 = \frac{ED}{(1-D)\Delta I_{L2} f_s} \quad (8)$$

The voltage gain or transformation ratio (M) is given by:

$$M = \frac{V_o}{E} = \frac{(2-D)D}{(1-D)^2} \quad (9)$$

The proposed converter can step-up or step-down the input voltage. The converter operates as a mode Buck for duty ratios in the range $0.01 < D < 0.29$, as shows in Fig. 3. For duty ratios in the range $0.3 < D < 0.99$, the converter operates as a step-up, as shown in Fig. 4. To compute the duty ratio as a function of the desired voltage gain of the proposed converter, Expression (10) is proposed:

$$D = 1 - \frac{1}{\sqrt{1+M}} \quad (10)$$

3. DESIGN OF THE PROPOSED CONVERTER

Based on the previous design equations, the converter components are proposed. Table 1 lists the design parameters used to perform the comparative analysis with the switched model provided by PowerSim software. This converter is designed for a 200 W, 24 V photovoltaic panel to supply a single-phase inverter that requires an output voltage of 170 V from the converter.

Figs. 5 and 6 show the simulated waveforms of the converter operating in step-up mode with the parameters

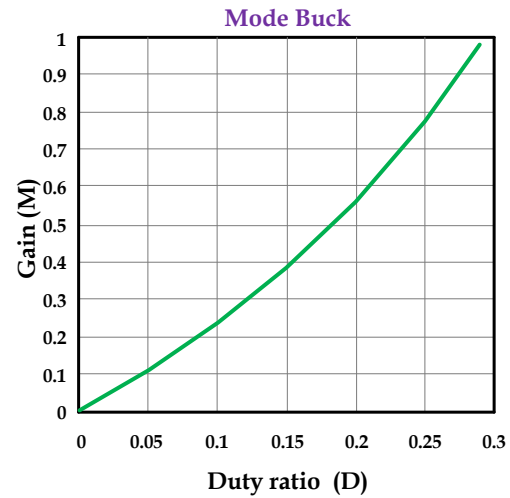


Fig. 3. Voltage gain of the converter in Buck mode.

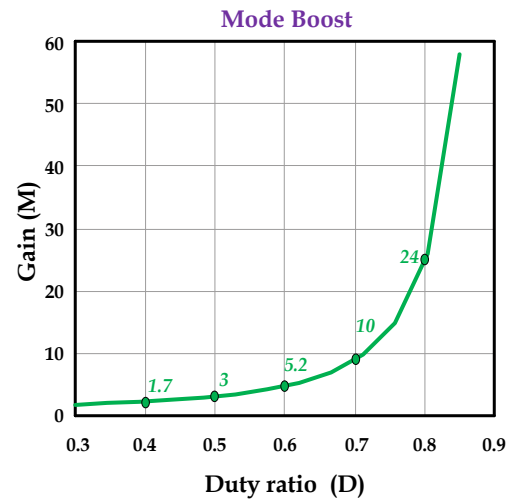


Fig. 4. Voltage gain of the converter in Boost mode.

Table 1. Design parameters for 200 W.

Symbol	Parameter	Value
E	Input voltage	24 V
I_o	Output current	1.176 A
P_o	Output power	200 W
V_o	Output voltage	170 V
f_s	Switching frequency	50 kHz
L_1, L_2	Inductors	82 μ H and 680 μ H
C_1, C_2	Capacitors	47 μ F and 4.7 μ F
D	Duty ratio	0.65

listed in Table 1. The input voltage remains constant at 24 V, while the output voltage reaches an average value close to 170 V, exhibiting low ripple. The PWM control signal operates with a duty ratio of 0.65, which is consistent with the required voltage gain.

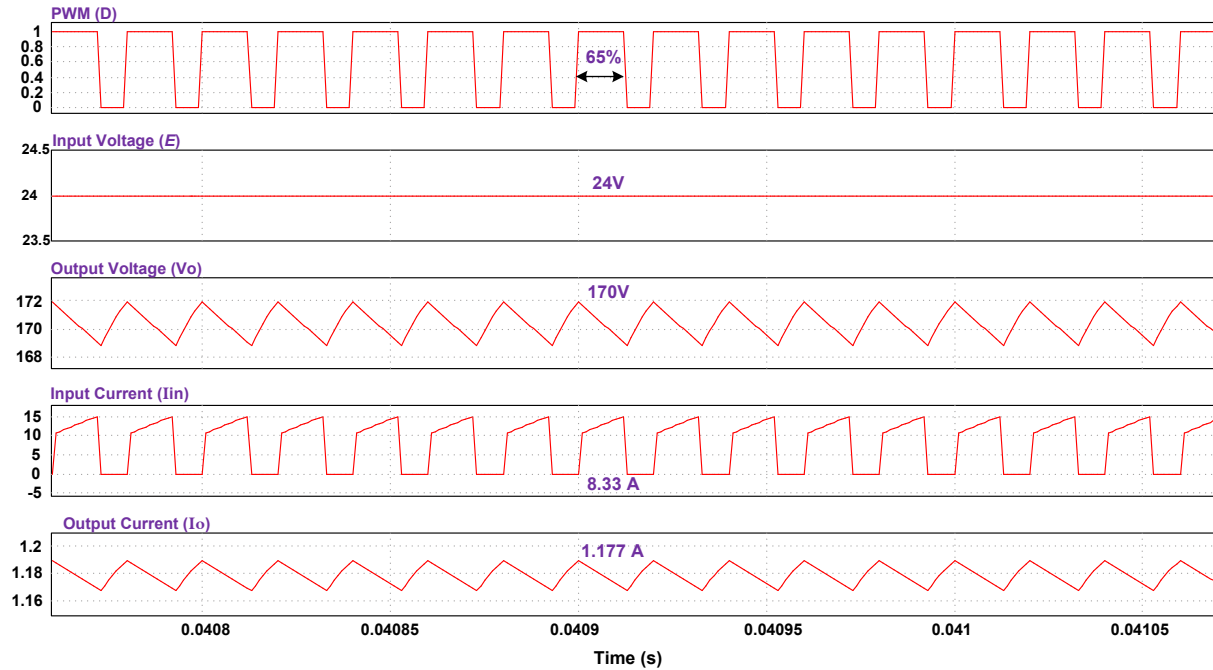


Fig. 5. Voltage and current transformation ratio waveforms obtained in PSIM.

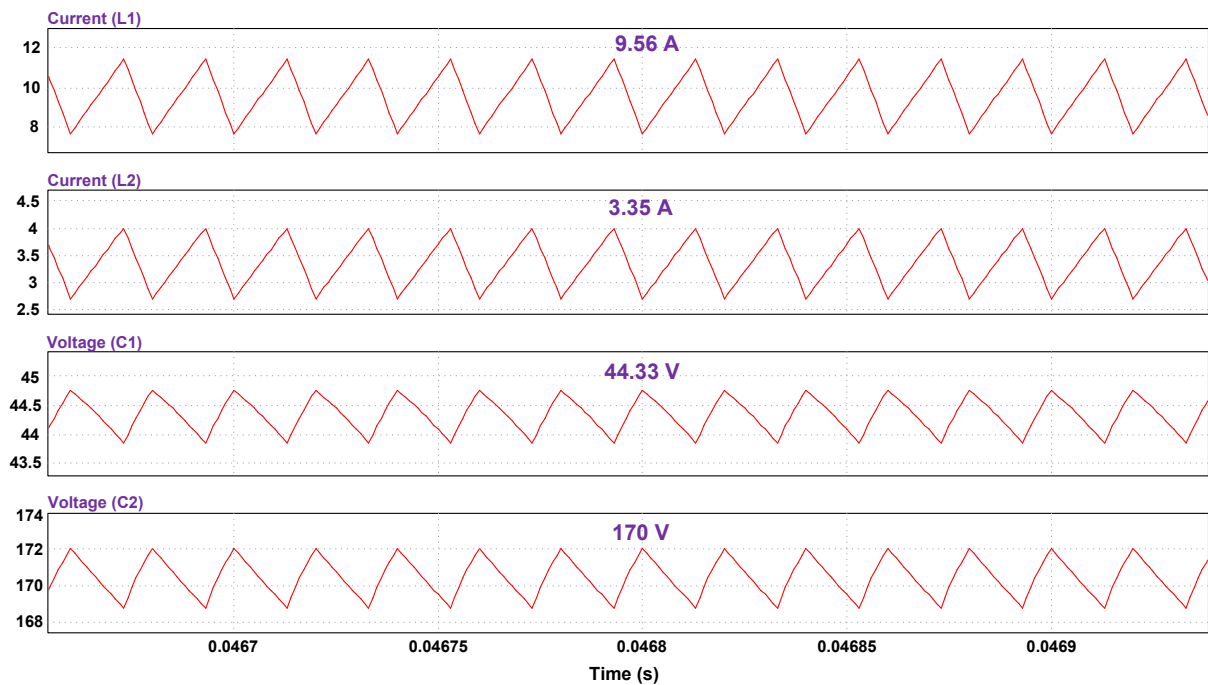


Fig. 6. Voltage and current waveforms of inductors and capacitors obtained in PSIM.

Table 2. Comparison of buck-boost converters proposed in the literature.

Converter	C	L	Q	S	(M)	Current Q_1	Current Q_2	Efficiency(%)
Proposed	2	2	2	2	$\frac{D(2-D)}{(1-D)^2}$	$\frac{ED^2(2-D)}{R(1-D)^3}$	$\frac{ED^2(2-D)}{R(1-D)^4}$	95.3
Buck-Boost (Kahani et al., 2022)	1	1	1	1	$\frac{D}{1-D}$	$\frac{ED}{R(1-D)^2}$	X	90
Buck-Boost R^2P^2 (Villanueva-Loredo et al., 2023)	2	2	2	2	$\frac{D}{1-D}$	$\frac{ED^3}{(1-D)^2 R}$	$\frac{ED^2}{(1-D)R}$	90.5
Quadratic Buck-Boost (Gholizadeh et al., 2023)	3	3	2	2	$\frac{D^2}{(1-D)^2}$	$\frac{ED^4}{R(1-D)^4}$	$\frac{ED^3}{R(1-D)^3}$	96.96
Quadratic Boost (Solis-Rodriguez et al., 2023)	2	2	2	2	$\frac{1}{(1-D)^2}$	$\frac{ED^2}{R(1-D)^4}$	$\frac{ED}{R(1-D)^3}$	—
New Buck-Boost (Abdel-Rahim et al., 2022)	1	2	3	2	$\frac{2D}{1-D}$	$\frac{4ED}{R(1-D)}$	$\frac{2ED}{R(1-D)}$	95.4

The input current exhibits a pulsating behavior with an average value of approximately 8.33 A. The output current is 1.177 A with a ripple of 1%. Fig. 6 shows the simulated waveforms of internal variables of the converter operating in step-up mode. Specifically, the current through inductor L_2 and the voltages across capacitors C_1 and C_2 are presented. The inductor current L_2 exhibits a characteristic triangular waveform typical of continuous conduction mode (CCM), with an average value close to 3.35 A, indicating proper sizing of the inductive element and stable operation. Furthermore, the voltage across capacitor C_1 remains around 44.33 V, exhibiting low ripple, which evidences its correct selection for stabilizing the intermediate node. Similarly, the voltage across capacitor C_2 reaches an average value close to 170 V, corresponding to the converter output voltage, with low ripple. Collectively, these waveforms validate the expected behavior of the converter and confirm the consistency between the theoretical design and the simulation results.

Table 2 shows a comparison between several buck-boost converters described in the literature and the proposed converter. This comparison is based on the number of capacitors (C), the number of inductors (L), the gain in terms of duty cycle (M), the number of diodes (S), the number of switches (Q), and the converter's efficiency.

4. RESULTS

The theoretical results are validated through an experimental prototype using the parameters listed in Table 1. Fig. 7 shows the prototype of the Buck-Boost converter with R^2P^2 .

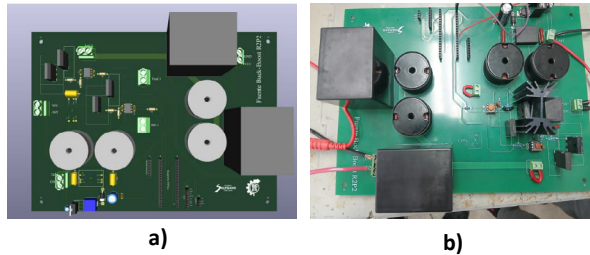


Fig. 7. Experimental prototype of the Buck-Boost converter: (a) PCB design and (b) experimental implementation.

The pulse-width modulation (PWM) control signal is presented in Fig. 8. A switching frequency of 50.12 kHz and a duty ratio of $D = 65.27\%$ are observed.

The obtained output voltage is shown in Fig. 9, where an average value of approximately 170 V is observed.

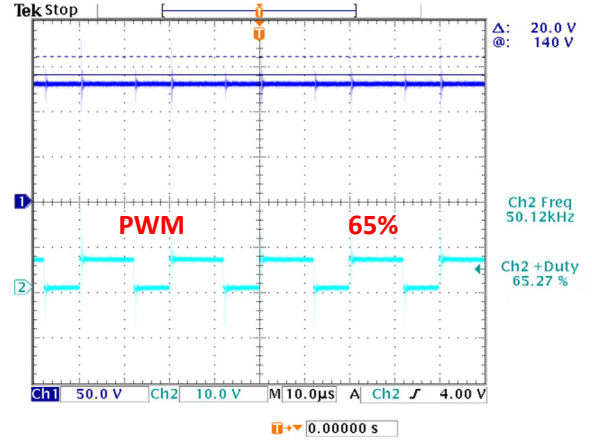


Fig. 8. PWM control signal applied to the converter switches.

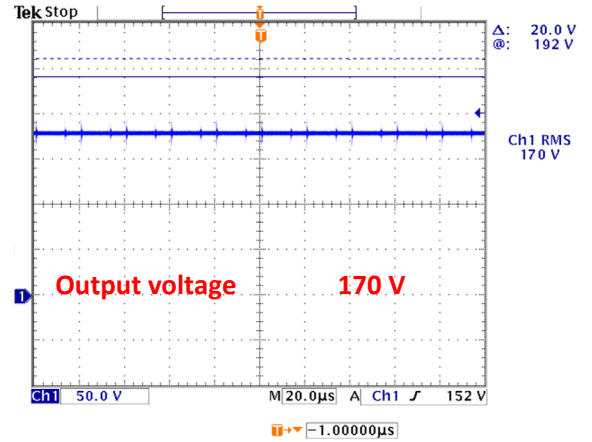


Fig. 9. Experimental output voltage in step-up mode.

Fig. 10 shows an RMS output current of 1.19 A. Fig. 11 presents the input current of 8.52 A. This current value confirms continuous conduction mode (CCM) operation. Additionally, a pulsating current that is unsuitable for some applications without an LC filter is observed.

5. CONCLUSION

This work presented the design, implementation, and experimental validation of a DC-DC Buck-Boost converter based on the reduced redundant power processing (R^2P^2) concept, aimed at applications requiring wide voltage

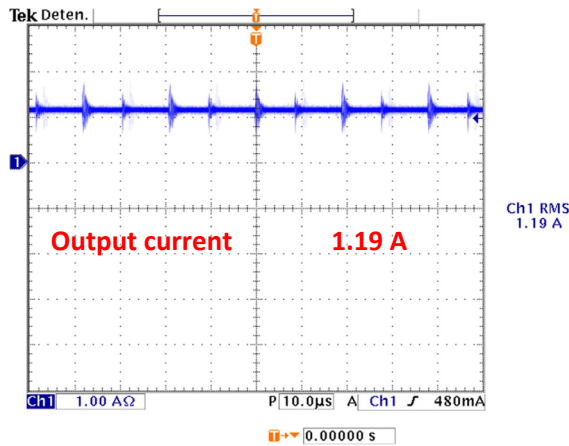


Fig. 10. Experimental output current of the converter.

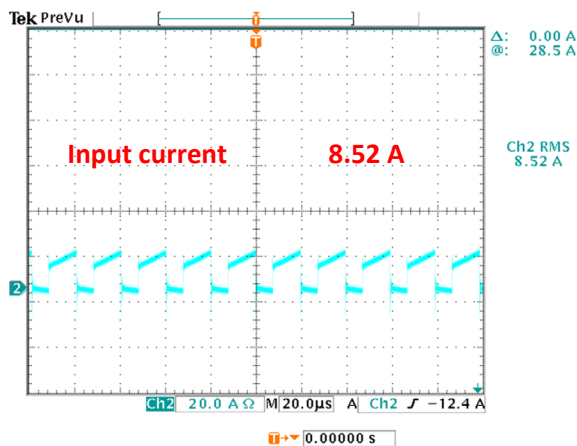


Fig. 11. Input current waveform.

conversion ranges and high efficiency. The steady-state analysis enabled precise determination of the operating conditions and design equations, which were corroborated through simulation and experimental results. The developed prototype demonstrated high agreement between theoretical and experimental outcomes, achieving an output voltage close to 170 V from a 24 V input, with stable operation in CCM. As a notable result, the converter achieved a peak efficiency of 95.3%, evidencing the effectiveness of the proposed approach compared to conventional topologies. This performance, together with the reduced electrical stress on the semiconductors, positions the proposed topology as a highly competitive alternative for applications in photovoltaic systems and DC microgrids. Finally, future work includes the implementation of closed-loop control strategies and the evaluation of dynamic behavior under load variations, aiming to further optimize system performance under real operating conditions.

ACKNOWLEDGEMENTS

The authors thank the Centro de Formación Profesional (CFP) / Bachillerato Tecnológico Salesiano Carlos Gómez

for providing the resources used in this work, and the Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) for its contribution to the development of this research.

REFERENCES

- Abdel-Rahim, O., Chub, A., Blinov, A., Vinnikov, D., and Peftitsis, D. (2022). An efficient non-inverting buck-boost converter with improved step up/down ability. *Energies*, 15(13), 4550. doi: <https://doi.org/10.3390/en15134550>.
- Andres, B., Romitti, L., Dupont, F.H., Roggia, L., and Schuch, L. (2023). A high step-up isolated dc-dc converter based on voltage multiplier cell. *International Journal of Circuit Theory and Applications*, 51(2), 557–578. doi: <https://doi.org/10.1002/cta.3444>.
- Gholizadeh, H., Gorji, S.A., and Sera, D. (2023). A quadratic buck-boost converter with continuous input and output currents. *IEEE Access*, 11, 22376–22393. doi: <https://doi.org/10.1109/ACCESS.2023.3253102>.
- Heo, J.I. and Kim, J.W. (2026). Llc converter using phase-shifted partial power processing circuit for on-board electric vehicle charger. *Journal of Electrical Engineering & Technology*, 1–17. doi: <https://doi.org/10.1007/s42835-026-02636-9>.
- Kahani, R., Jamil, M., and Iqbal, M.T. (2022). Direct model reference adaptive control of a boost converter for voltage regulation in microgrids. *Energies*, 15(14), 5080. doi: <https://doi.org/10.3390/en15145080>.
- Reddy, B.N., Shetty, K.B., Jalli, R., Akhila, K., Prudhviraj, K.S., and Jagan, V. (2024). Reduced redundant power processing in dc-dc converters: A comprehensive review. In *2024 3rd International conference on Power Electronics and IoT Applications in Renewable Energy and its Control (PARC)*, 361–365. IEEE. doi: <https://doi.org/10.1109/PARC59193.2024.10486338>.
- Reyes-Portillo, I.A., Méndez-Elizondo, S.R., Morales-Saldaña, J.A., Rivera-Romero, C.A., and Castro-López, D.L. (2025). Design of a transformerless dc regulator with reduced redundant power processing for a residential dc microgrid. *IEEE Latin America Transactions*, 23(5), 427–436. doi: <https://doi.org/10.1109/TLA.2025.10974364>.
- Reyes-Portillo, I.A., Morales-Saldaña, J., Romero-Rivera, C., and Palacios-Hernández, E. (2023a). Design and modeling of a high current ratio converter for pv applications. *IEEE Latin America Transactions*, 21(10), 1144–1155. doi: <https://doi.org/10.1109/TLA.2023.10255452>.
- Reyes-Portillo, I.A., Saldaña, J.A.M., Netzahuatl-Huerta, E.M., and Loera-Palomo, R. (2023b). Design and analysis of quadratic buck converter based on the reduced redundant power processing. *IEEE Journal of Emerging and Selected Topics in Power Electronics*, 12(1), 803–814. doi: <https://doi.org/10.1109/JESTPE.2023.3322707>.
- Roccaforte, F., Giannazzo, F., and Greco, G. (2022). Ion implantation doping in silicon carbide and gallium nitride electronic devices. In *Micro*, volume 2, 23–53. MDPI. doi: <https://doi.org/10.3390/micro2010002>.
- Santosh Kumar Reddy, P. and Obulesu, Y.P. (2023). A non-cascading dc/dc quadratic boost converter with high voltage gain for pv applications. *International Journal of Electronics*, 110(10), 1756–1777. doi: <https://doi.org/10.1080/00207217.2022.2118849>.



- Singh, S., Chaudhary, T., and Khanna, G. (2022). Recent advancements in wide band semiconductors (sic and gan) technology for future devices. *Silicon*, 14(11), 5793–5800. doi:<https://doi.org/10.1007/s12633-021-01362-3>.
- Solis-Rodriguez, J., Rosas-Caro, J.C., Alejo-Reyes, A., and Valdez-Resendiz, J.E. (2023). Optimal selection of capacitors for a low energy storage quadratic boost converter (les-qbc). *Energies*, 16(6), 2510. doi: <https://doi.org/10.3390/en16062510>.
- Villanueva-Loredo, J.A., Ortiz-Lopez, M.G., Leyva-Ramos, J., and Diaz-Saldierna, L.H. (2023). Switching regulator based on a non-inverting step-down/up dc-dc converter for lithium-ion battery applications. *Micromachines*, 14(6), 1144. doi: <https://doi.org/10.3390/mi14061144>.
- Zhang, X., Wang, T., Bao, H., Hu, Y., and Bao, B. (2022). Stability effect of load converter on source converter in a cascaded buck converter. *IEEE transactions on power electronics*, 38(1), 604–618. doi: 10.1109/TPEL.2022.3199234.

Use of Comparative LPWAN Studies as a Didactic Resource in Engineering Education

Wendy Janett Guzmán González*. Lucrecia Guadalupe Valenzuela Segura*. Cynthia Liliana Guzmán González*. Elizabeth Payan Nevárez*. Brenda Eunice Guzmán González**. Juan Alberto Luna Sandoval***.

* *Tecnológico Nacional de México / Instituto Tecnológico de Chihuahua*

Av. Tecnológico No. 2909, Chihuahua, Chih, México. C.P. 31310.

(Tel. +52 (614) 2-01-2000; e-mail: wendy.gg@chihuahua.tecnm.mx).

** *Colegio de Estudios Científicos y Tecnológicos del Estado de Durango plantel 09.*

Avenida del Tecnológico s/n, Lerdo Durango, México. C.P. 35150

(Tel. +52 (871) 9-43-1523; e-mail: brenda.guzman.cecyted@cecyte.edu.mx).

*** *Teléfonos de México SAB de CV.*

Calle garita y molino del rey s/n, Hidalgo del parral chihuahua. C.P. 33800

(Tel. +52 (627) 1-11-8584; e-mail: juan.luna.830101@gmail.com).

Seoul, Korea, (author@snu.ac.kr)

Abstract: Engineering education in the context of the Internet of Things (IoT) requires not only theoretical knowledge but also the ability to select appropriate technologies based on performance criteria. In this context, LPWAN networks, such as LoRaWAN and NB-IoT, represent a relevant case due to their differences in energy efficiency, coverage, and reliability. This study presents and preliminarily evaluates a didactic strategy based on comparative LPWAN studies for engineering education. Through the analysis of scientific literature, key performance metrics such as energy consumption, battery lifetime, latency, and coverage were identified. Based on these elements, a case-based learning strategy was designed in which students analyze IoT scenarios and select technologies using scientific evidence. A pilot evaluation involving engineering students was conducted to identify perceptions regarding the usefulness of the proposed strategy for technological decision-making and critical analysis. The pilot evaluation indicated a favorable perception toward the use of comparative studies as a pedagogical resource in engineering education.

Keywords: IoT, LPWAN, engineering education, LoRaWAN, NB-IoT

1. INTRODUCCIÓN

El crecimiento acelerado del Internet de las Cosas (IoT) ha impulsado el desarrollo y adopción de tecnologías de comunicación capaces de operar con bajo consumo energético y amplia cobertura. En este contexto, las redes de área amplia de bajo consumo LPWAN (*Low Power Wide Area Network*) han adquirido un papel fundamental, destacando tecnologías como LoRaWAN y NB-IoT (*Narrowband Internet of Things*), ampliamente utilizadas en aplicaciones industriales, ciudades inteligentes y monitoreo ambiental (Masek et al., 2021), (Soy, 2023), (Quispe-Vilca et al., 2024). Estas tecnologías presentan diferencias relevantes en términos de eficiencia energética, cobertura, latencia y confiabilidad, lo que implica que su selección depende del contexto de aplicación y de los requerimientos específicos del sistema (Kim and Jo, 2022), (Michelinakis et al., 2020).

Diversos estudios han analizado el desempeño de estas tecnologías desde perspectivas comparativas, evidenciando que LoRaWAN (*Long Range Wide Area Network*) suele ser más eficiente en escenarios de baja frecuencia de transmisión, mientras que NB-IoT ofrece ventajas en cobertura y confiabilidad, particularmente en entornos complejos (Piyare

et al., 2018), (Kim and Jo, 2022), (Urabe et al., 2023). Asimismo, investigaciones recientes han resaltado la importancia de métricas como el consumo energético, la duración de batería y la calidad del enlace para la evaluación del desempeño en redes LPWAN (Ugwuanyi et al., 2021), (Rahman and Saifullah, 2023), (Michelinakis et al., 2020). Sin embargo, la literatura presenta una alta heterogeneidad en los enfoques metodológicos y en las métricas utilizadas, lo que dificulta la comparación directa entre tecnologías y la interpretación de resultados (Quispe-Vilca et al., 2024).

En el ámbito educativo, la enseñanza de tecnologías emergentes relacionadas con el Internet de las Cosas suele centrarse principalmente en aspectos conceptuales y descriptivos, limitando la participación activa de los estudiantes en procesos de análisis, comparación y toma de decisiones tecnológicas (Michael, 2024). Aunque tecnologías como LoRaWAN y NB-IoT son ampliamente utilizadas en aplicaciones reales, su abordaje en programas de ingeniería frecuentemente se reduce a la revisión teórica de sus características generales, sin incorporar actividades que permitan interpretar evidencia científica o evaluar el desempeño de dichas tecnologías en distintos contextos de aplicación.



Esta situación representa una limitación importante en la formación de ingenieros, ya que el entorno profesional actual demanda competencias relacionadas con el pensamiento crítico, el análisis de información técnica y la capacidad de seleccionar soluciones tecnológicas fundamentadas en criterios de desempeño. En particular, el análisis de métricas como consumo energético, cobertura, latencia y confiabilidad resulta esencial para comprender las ventajas y limitaciones de las tecnologías LPWAN en escenarios reales.

Asimismo, diversos estudios han señalado la importancia de incorporar estrategias de aprendizaje activo y análisis basado en evidencia científica en la educación en ingeniería (Hmelo, 2004), con el propósito de fortalecer la vinculación entre teoría y práctica y favorecer el desarrollo de competencias para la resolución de problemas. En este contexto, los estudios comparativos reportados en la literatura científica representan una oportunidad para transformar información técnica en recursos didácticos orientados al análisis crítico y a la toma de decisiones tecnológicas.

Con base en lo anterior, el objetivo general de esta investigación es proponer una estrategia didáctica basada en el análisis de estudios comparativos de tecnologías LPWAN para fortalecer competencias de análisis, interpretación de evidencia científica y toma de decisiones tecnológicas en estudiantes de ingeniería.

De manera específica, se busca:

- Identificar métricas de desempeño utilizadas en estudios comparativos de tecnologías LPWAN, tales como consumo energético, cobertura, latencia y confiabilidad.
- Analizar el potencial de los estudios comparativos de LoRaWAN y NB-IoT como recurso didáctico en la enseñanza de tecnologías IoT.
- Diseñar una propuesta didáctica basada en análisis de casos que permita a los estudiantes evaluar y seleccionar tecnologías LPWAN en función de criterios técnicos y contextuales.

1. CONTEXTO TECNOLÓGICO DE LPWAN

El desarrollo del Internet de las Cosas IoT (*Internet of Things*) ha impulsado la necesidad de tecnologías de comunicación que permitan la conexión de un gran número de dispositivos con bajo consumo energético y cobertura extendida. En este contexto, LPWAN han emergido como una solución clave para aplicaciones donde la eficiencia energética y la autonomía de los dispositivos son factores críticos (Masek et al., 2021), (Soy, 2023).

Las tecnologías LPWAN se caracterizan por ofrecer comunicación de largo alcance, bajo consumo de energía y capacidad para soportar dispositivos con limitaciones de recursos. Estas características las hacen especialmente adecuadas para aplicaciones como monitoreo ambiental, ciudades inteligentes, agricultura de precisión y sistemas

industriales (Soussi et al., 2024), (Stanco et al., 2024). Entre las tecnologías más representativas se encuentran LoRaWAN y NB-IoT, las cuales han sido ampliamente estudiadas en la literatura científica debido a su relevancia en el ecosistema IoT (Piyare et al., 2018), (Kim and Jo, 2022).

LoRaWAN es una tecnología basada en comunicación de espectro no licenciado, que permite la transmisión de datos a largas distancias con bajo consumo energético. Su funcionamiento se apoya en el uso de factores de propagación que permiten adaptar la tasa de transmisión y el alcance de la comunicación, favoreciendo la eficiencia energética en escenarios de baja frecuencia de envío de datos (Andres-Maldonado et al., 2017), (Loh et al., 2022). Esta característica la hace particularmente adecuada para aplicaciones donde los dispositivos transmiten información de manera esporádica, permitiendo extender la duración de la batería (Anbazhagan and Mugelan, 2024).

Por otro lado, NB-IoT es una tecnología basada en redes celulares que opera en espectro licenciado, lo que le permite ofrecer mayor confiabilidad, calidad de servicio y cobertura, especialmente en entornos urbanos densos o con condiciones de propagación adversas (Ballerini and Magno, n.d.), (Ugwuanyi et al., 2021). Sin embargo, estas ventajas suelen implicar un mayor consumo energético en comparación con tecnologías como LoRaWAN, particularmente en escenarios con transmisiones frecuentes o condiciones de señal desfavorables (Rahman and Saifullah, 2023), (Michelinakis et al., 2020).

Diversos estudios comparativos han analizado el desempeño de ambas tecnologías considerando métricas como consumo energético, cobertura, latencia y capacidad de transmisión. Estos estudios evidencian que no existe una tecnología universalmente superior, ya que el desempeño depende en gran medida del contexto de aplicación y de los requerimientos del sistema (Kim and Jo, 2022), (Urabe et al., 2023). Por ejemplo, LoRaWAN presenta ventajas en términos de eficiencia energética en aplicaciones de bajo *duty cycle*, mientras que NB-IoT ofrece mejor desempeño en términos de confiabilidad y cobertura en aplicaciones críticas (Piyare et al., 2018), (Michelinakis et al., 2020).

Asimismo, la literatura ha identificado una amplia diversidad de métricas utilizadas para evaluar el desempeño energético, incluyendo consumo total, duración de batería y energía por transmisión. Esta heterogeneidad dificulta la comparación directa entre estudios y resalta la necesidad de interpretar los resultados en función de las condiciones específicas de cada evaluación (Quispe-Vilca et al., 2024).

2. METODOLOGÍA

La presente investigación se desarrolló con el propósito de analizar el potencial de los estudios comparativos de tecnologías LPWAN como recurso didáctico en la formación de ingenieros. Para ello, se realizó un análisis documental de literatura científica relacionada con tecnologías como



LoRaWAN y NB-IoT, considerando tanto sus características técnicas como sus posibilidades de aplicación en contextos educativos asociados al Internet de las Cosas.

A partir de la información analizada, se diseñó una propuesta didáctica orientada al fortalecimiento de competencias relacionadas con el análisis crítico, la interpretación de evidencia científica y la toma de decisiones tecnológicas. La metodología empleada se estructuró en las etapas descritas a continuación.

A. Enfoque de la investigación

El presente trabajo se desarrolló bajo un enfoque descriptivo y documental, orientado al análisis de estudios comparativos relacionados con tecnologías LPWAN y su integración como recurso didáctico en la formación de ingenieros. La investigación se centró en la revisión documental de literatura científica especializada sobre tecnologías como LoRaWAN y NB-IoT, considerando tanto sus características técnicas como su potencial aplicación en contextos educativos asociados al Internet de las Cosas (IoT).

B. Proceso de búsqueda y selección de literatura

La búsqueda de literatura científica se realizó mediante bases de datos académicas especializadas en ingeniería, telecomunicaciones y tecnologías IoT. Para ello, se consideraron artículos publicados principalmente entre los años 2017 y 2024, debido al crecimiento reciente de investigaciones relacionadas con tecnologías LPWAN. La selección de los estudios se realizó considerando los siguientes criterios:

- Estudios comparativos entre tecnologías LPWAN, particularmente LoRaWAN y NB-IoT.
- Investigaciones enfocadas en métricas de desempeño como consumo energético, cobertura, latencia, confiabilidad y duración de batería.
- Artículos relacionados con aplicaciones IoT en escenarios reales, simulados o experimentales.
- Trabajos con información relevante para el análisis técnico y educativo de tecnologías LPWAN.

Por otra parte, se excluyeron investigaciones que se enfocaban únicamente en aspectos de implementación sin análisis comparativo o sin evaluación de desempeño.

C. Análisis de la información

Una vez seleccionada la literatura, se realizó un análisis de contenido orientado a identificar las principales métricas utilizadas para evaluar el desempeño de las tecnologías LPWAN. Entre las variables analizadas se consideraron:

- Consumo energético.
- Cobertura de comunicación.
- Latencia.
- Duración de batería.
- Confiabilidad del enlace.

Adicionalmente, se revisaron los enfoques metodológicos empleados en los estudios analizados, incluyendo simulación, modelado analítico y validación experimental, con el propósito de identificar diferencias en los criterios de evaluación y en la interpretación de resultados reportados en la literatura científica.

D. Diseño de la propuesta didáctica

Con base en el análisis documental realizado, se diseñó una propuesta didáctica orientada al fortalecimiento de competencias relacionadas con el análisis crítico, la interpretación de evidencia científica y la toma de decisiones tecnológicas en estudiantes de ingeniería.

La propuesta se estructuró mediante actividades basadas en casos, en las cuales los estudiantes analizan escenarios IoT y seleccionan tecnologías LPWAN utilizando información proveniente de estudios comparativos reportados en la literatura científica.

El diseño de la estrategia consideró elementos como:

- Definición de objetivos de aprendizaje.
- Identificación de métricas técnicas relevantes.
- Comparación de tecnologías LPWAN.
- Análisis de escenarios de aplicación.
- Justificación de decisiones tecnológicas.

De esta manera, la propuesta busca fortalecer la vinculación entre teoría y práctica mediante el uso de evidencia científica como recurso didáctico en la educación en ingeniería.

3. ESTUDIOS COMPARATIVOS COMO RECURSO DIDÁCTICO

La formación de ingenieros en el contexto actual requiere el desarrollo de competencias que trascienden la adquisición de conocimientos teóricos, incorporando habilidades como el análisis crítico, la interpretación de información técnica y la toma de decisiones fundamentadas. En el ámbito del Internet de las Cosas, estas competencias resultan especialmente relevantes debido a la diversidad de tecnologías disponibles y a la necesidad de seleccionar soluciones adecuadas en función de múltiples criterios de desempeño.

Los estudios comparativos reportados en la literatura científica constituyen una herramienta valiosa para abordar esta problemática. Estos estudios analizan el desempeño de distintas tecnologías bajo diversas condiciones, considerando métricas como consumo energético, cobertura, latencia y confiabilidad (Kim and Jo, 2022), (Urabe et al., 2023), (Michelinakis et al., 2020). A través de estos análisis, es posible identificar patrones, ventajas y limitaciones que permiten comprender el comportamiento de las tecnologías en distintos escenarios.

Desde una perspectiva educativa, el uso de este tipo de estudios permite transformar la literatura científica en un recurso didáctico, facilitando la construcción de actividades centradas en el análisis de casos. En este enfoque, los estudiantes no solo reciben información, sino que interactúan



con datos reales, interpretan resultados y desarrollan argumentos para justificar la selección de una tecnología determinada. Este proceso contribuye al fortalecimiento del pensamiento crítico y a la aplicación contextual del conocimiento.

Asimismo, la diversidad de enfoques metodológicos presentes en la literatura —incluyendo estudios basados en simulación, modelado analítico y validación experimental— permite a los estudiantes comprender las implicaciones de cada tipo de análisis y sus limitaciones (Quispe-Vilca et al., 2024). Esta comprensión resulta fundamental para interpretar adecuadamente los resultados y evitar conclusiones generalizadas que no consideren el contexto de evaluación.

El uso de estudios comparativos como recurso didáctico también favorece la vinculación entre la teoría y la práctica, al permitir que los estudiantes enfrenten situaciones similares a las que encontrarán en el ámbito profesional. En particular, el análisis de métricas energéticas y de desempeño en tecnologías LPWAN proporciona un marco adecuado para desarrollar actividades orientadas a la toma de decisiones tecnológicas, integrando conocimientos técnicos con habilidades analíticas.

En este contexto, resulta pertinente diseñar estrategias didácticas que incorporen de manera estructurada el análisis de estudios comparativos en el aula. Estas estrategias pueden basarse en enfoques como el aprendizaje basado en problemas o el aprendizaje basado en casos, donde los estudiantes analizan escenarios específicos y seleccionan soluciones tecnológicas a partir de evidencia científica. De esta manera, se promueve una formación más integral, orientada no solo al conocimiento, sino también a la aplicación y la toma de decisiones en contextos reales de ingeniería.

4. PROPUESTA DIDÁCTICA

A. Objetivos de aprendizaje

La propuesta didáctica tiene como objetivo principal fortalecer en los estudiantes de ingeniería la capacidad de analizar, comparar y seleccionar tecnologías de comunicación en el contexto de IoT. En particular, se busca que los estudiantes desarrollen las siguientes competencias:

- Interpretar métricas de desempeño en tecnologías LPWAN, tales como consumo energético, cobertura, latencia y confiabilidad.
- Analizar estudios comparativos basados en evidencia científica.
- Evaluar el comportamiento de tecnologías como LoRaWAN y NB-IoT en distintos escenarios de aplicación.
- Tomar decisiones tecnológicas fundamentadas en criterios técnicos y contextuales.

Estas competencias son fundamentales en la formación de ingenieros, ya que permiten integrar conocimientos teóricos con habilidades de análisis y resolución de problemas en contextos reales.

B. Diseño de la actividad basada en casos

La estrategia didáctica propuesta se basa en el enfoque de aprendizaje basado en casos, en el cual los estudiantes enfrentan una situación real o simulada que requiere la selección de una tecnología LPWAN adecuada.

La actividad se estructura en las siguientes etapas:

1. **Planteamiento del problema:** Se presenta a los estudiantes un escenario de aplicación IoT. Por ejemplo, un sistema de monitoreo ambiental en una zona rural que requiere dispositivos de bajo consumo energético y cobertura extendida.
2. **Identificación de variables clave:** Los estudiantes identifican los criterios relevantes para la selección tecnológica, tales como frecuencia de transmisión, consumo energético, cobertura requerida y condiciones del entorno.
3. **Análisis de evidencia científica:** Se proporciona a los estudiantes un conjunto de estudios comparativos extraídos de la literatura, en los cuales se analizan tecnologías como LoRaWAN y NB-IoT bajo diferentes condiciones (Kim and Jo, 2022), (Urabe et al., 2023), (Michelinakis et al., 2020). Los estudiantes deben interpretar los resultados y extraer conclusiones relevantes.
4. **Comparación de tecnologías:** A partir de la información analizada, los estudiantes comparan las tecnologías considerando sus ventajas y limitaciones en función del escenario planteado.
5. **Toma de decisión y justificación:** Finalmente, los estudiantes seleccionan la tecnología más adecuada y justifican su decisión con base en las métricas y resultados obtenidos de la literatura.

Este proceso promueve un aprendizaje activo, donde el estudiante participa en la construcción del conocimiento mediante el análisis y la toma de decisiones.

La Figura 1 muestra el flujo general de la estrategia didáctica propuesta, integrando las etapas de análisis de escenarios IoT, revisión de literatura científica, comparación de tecnologías



LPWAN, análisis de métricas y toma de decisiones tecnológicas fundamentadas.

C.



Integración de estudios comparativos de LPWAN

Los estudios comparativos de tecnologías LPWAN constituyen el núcleo de la estrategia didáctica. Estos estudios permiten analizar el desempeño de LoRaWAN y NB-IoT en términos de métricas clave como consumo energético, duración de batería, cobertura y latencia (Piyare et al., 2018), (Kim and Jo, 2022), (Michelinakis et al., 2020).

En particular, la literatura reporta que LoRaWAN presenta ventajas en términos de eficiencia energética en escenarios de baja frecuencia de transmisión, lo que favorece una mayor duración de la batería (Andres-Maldonado et al., 2017), (Loh et al., 2022). Por otro lado, NB-IoT ofrece mayor confiabilidad y cobertura, especialmente en entornos urbanos o con condiciones de propagación adversas, aunque con un mayor consumo energético en ciertos casos (Ballerini and Magno, n.d.), (Ugwuanyi et al., 2021), (Rahman and Saifullah, 2023). Asimismo, la diversidad de enfoques metodológicos presentes en los estudios —como simulación, modelado analítico y validación experimental— permite a los estudiantes comprender cómo el tipo de evaluación influye en los resultados reportados (Quispe-Vilca et al., 2024). Este análisis contribuye a desarrollar una visión crítica sobre la interpretación de la evidencia científica.

De esta manera, los estudios comparativos no solo aportan información técnica, sino que también funcionan como herramientas para el desarrollo de habilidades analíticas en los estudiantes.

D. Herramienta didáctica para la toma de decisiones

Como parte de la propuesta, se plantea el uso de una herramienta didáctica que sintetiza las principales

características de las tecnologías LPWAN, facilitando su análisis comparativo. La Tabla I presenta un ejemplo de esta herramienta.

Tabla I Comparación de tecnologías LPWAN para uso didáctico

Tecnología	Consumo energético	Cobertura	Latencia	Aplicación recomendada
LoRaWAN	Bajo	Media	Alta	Monitoreo ambiental, sensores remotos
NB-IoT	Medio/Alto	Alta	Baja	Aplicaciones críticas, entornos urbanos

Esta tabla permite a los estudiantes visualizar de manera clara las diferencias entre tecnologías y utilizarlas como base para la toma de decisiones. Además, puede ser adaptada a distintos escenarios de aprendizaje, incorporando nuevas variables o criterios según el contexto del curso.

E. Evaluación piloto de la propuesta

Con el propósito de realizar una validación inicial de la propuesta didáctica, se desarrolló una evaluación piloto exploratoria con estudiantes de ingeniería en un contexto académico relacionado con tecnologías IoT y redes de comunicación.

La evaluación piloto se realizó con un grupo de 37 estudiantes de ingeniería pertenecientes a una asignatura relacionada con tecnologías IoT y redes de comunicación. La actividad se desarrolló en un contexto académico orientado al análisis de tecnologías LPWAN y consistió en la evaluación de un escenario de aplicación IoT en el cual los estudiantes debían comparar tecnologías como LoRaWAN y NB-IoT considerando métricas de desempeño como consumo energético, cobertura, latencia y confiabilidad. Posteriormente, se aplicó un instrumento de percepción basado en escala tipo Likert, orientado a identificar la opinión de los participantes respecto a la utilidad de la estrategia didáctica para fortalecer competencias relacionadas con el análisis crítico, la interpretación de evidencia científica y la toma de decisiones tecnológicas.

De manera general, los resultados obtenidos permitieron identificar una percepción favorable hacia el uso de estudios comparativos como recurso didáctico, particularmente en aspectos relacionados con la comprensión de tecnologías LPWAN y el análisis de escenarios de aplicación. Asimismo, los estudiantes señalaron que el uso de evidencia científica facilitó la comparación entre tecnologías y la justificación de decisiones técnicas.

Aunque la evaluación realizada corresponde a una validación exploratoria inicial y no pretende establecer resultados concluyentes, los hallazgos obtenidos sugieren que la propuesta presenta potencial para fortalecer procesos de aprendizaje activo en la formación de ingenieros.

F. Evaluación del aprendizaje

Para evaluar el desempeño de los estudiantes en la actividad propuesta, se sugiere el uso de una rúbrica basada en los siguientes criterios:



- **Análisis de información:** capacidad para interpretar correctamente las métricas y resultados presentados en los estudios.
- **Comparación de tecnologías:** identificación de ventajas y limitaciones en función del contexto.
- **Justificación de la decisión:** coherencia y fundamentación de la tecnología seleccionada.
- **Aplicación contextual:** adecuación de la solución propuesta al escenario planteado.

Este enfoque de evaluación permite valorar no solo el conocimiento adquirido, sino también las habilidades analíticas y de toma de decisiones, alineadas con las competencias requeridas en la formación de ingenieros.

G. Implementación en el contexto educativo

La propuesta didáctica puede implementarse en asignaturas relacionadas con redes de comunicación, Internet de las Cosas (IoT), telecomunicaciones o sistemas embebidos dentro de programas de ingeniería. Su aplicación es flexible, pudiendo desarrollarse en modalidades presencial, virtual o híbrida.

En entornos virtuales, esta estrategia puede integrarse mediante plataformas de gestión del aprendizaje, donde se facilite el acceso a artículos científicos, actividades de análisis y espacios de discusión. Asimismo, el uso de herramientas digitales como simuladores o actividades interactivas permite complementar el análisis teórico con la exploración de escenarios prácticos.

La implementación puede realizarse como una actividad de clase, práctica evaluable o proyecto integrador, dependiendo de los objetivos del curso. Esto permite adaptar la estrategia a distintos niveles de profundidad y complejidad.

H. Rol del docente en la estrategia didáctica

En esta propuesta, el docente asume el papel de facilitador del aprendizaje, guiando a los estudiantes en la interpretación de la información y en el análisis crítico de los estudios comparativos. Su función principal es orientar el proceso de aprendizaje, plantear preguntas detonadoras y promover la discusión fundamentada.

Además, el docente es responsable de seleccionar y adaptar la literatura científica que será utilizada en la actividad, asegurando que sea pertinente y accesible para los estudiantes. También debe brindar acompañamiento durante el proceso de análisis, ayudando a los estudiantes a identificar las variables clave y a interpretar correctamente los resultados.

Este rol favorece un ambiente de aprendizaje activo, donde el estudiante participa de manera significativa en la construcción del conocimiento.

I. Ventajas de la propuesta didáctica

La estrategia propuesta presenta diversas ventajas en la formación de ingenieros:

- Fomenta el aprendizaje activo y centrado en el estudiante.
- Desarrolla habilidades de pensamiento crítico mediante el análisis de evidencia científica.
- Integra conocimientos teóricos con su aplicación en escenarios reales.

- Promueve la toma de decisiones tecnológicas fundamentadas.
- Facilita la comprensión de tecnologías complejas a través del análisis comparativo.

Estas ventajas contribuyen a una formación más integral, alineada con las demandas actuales del entorno tecnológico.

J. Limitaciones de la propuesta

A pesar de sus beneficios, la propuesta presenta algunas limitaciones que deben considerarse en su implementación. En primer lugar, la lectura e interpretación de literatura científica puede representar un desafío para estudiantes con poca experiencia en análisis técnico.

En segundo lugar, la diversidad de métricas y enfoques metodológicos en los estudios comparativos puede generar dificultades para establecer comparaciones directas entre tecnologías, como se ha señalado en la literatura (Quispe-Vilca et al., 2024). Esto requiere una adecuada orientación por parte del docente.

Finalmente, la implementación de esta estrategia puede demandar mayor tiempo en comparación con enfoques tradicionales, lo que implica ajustes en la planificación del curso.

5. DISCUSIÓN

Los resultados de la propuesta didáctica planteada evidencian la pertinencia de integrar estudios comparativos de tecnologías LPWAN como recurso en la formación de ingenieros. A diferencia de los enfoques tradicionales centrados en la transmisión de contenidos, esta estrategia promueve un aprendizaje activo basado en el análisis de evidencia científica, lo cual resulta fundamental en contextos tecnológicos dinámicos como el Internet de las Cosas.

Los resultados preliminares obtenidos durante la evaluación piloto permitieron identificar una percepción favorable por parte de los estudiantes hacia el uso de estudios comparativos como recurso didáctico, particularmente en aspectos relacionados con el análisis de tecnologías LPWAN y la toma de decisiones tecnológicas fundamentadas.

El uso de estudios comparativos permite a los estudiantes comprender que el desempeño de tecnologías como LoRaWAN y NB-IoT no es absoluto, sino dependiente del contexto de aplicación y de variables específicas como el consumo energético, la cobertura y la latencia (Kim and Jo, 2022), (Urabe et al., 2023). Esta comprensión favorece una visión más crítica y evita la adopción de soluciones tecnológicas de manera generalizada sin considerar sus limitaciones.

Asimismo, la incorporación de métricas de desempeño reportadas en la literatura contribuye a que los estudiantes desarrollen habilidades para interpretar datos técnicos y tomar decisiones fundamentadas. En este sentido, se observa que el análisis de variables como la duración de la batería y la calidad del enlace permite establecer criterios más claros para la selección de tecnologías LPWAN en diferentes escenarios (Ugwuanyi et al., 2021), (Rahman and Saifullah, 2023), (Michelinakis et al., 2020).

No obstante, la implementación de esta estrategia también pone en evidencia ciertos desafíos. La heterogeneidad en los



enfoques metodológicos y en las métricas utilizadas en los estudios comparativos puede dificultar la interpretación de resultados, especialmente para estudiantes que no están familiarizados con la lectura de literatura científica (Quispe-Vilca et al., 2024). Esto resalta la importancia del acompañamiento docente para guiar el proceso de análisis y asegurar una adecuada comprensión de los contenidos.

Por otra parte, el enfoque propuesto se alinea con tendencias actuales en educación en ingeniería que promueven el aprendizaje basado en problemas y el uso de evidencia para la toma de decisiones. La integración de literatura científica en el aula no solo fortalece el conocimiento técnico, sino que también acerca a los estudiantes a prácticas propias de la investigación, favoreciendo una formación más integral.

Finalmente, es importante señalar que, aunque la propuesta se centra en tecnologías LPWAN, el enfoque didáctico puede extenderse a otras áreas de la ingeniería donde existan múltiples alternativas tecnológicas y la necesidad de seleccionar soluciones con base en criterios de desempeño. De esta manera, la estrategia propuesta no solo es pertinente en el contexto del IoT, sino que también tiene potencial de aplicación en diversos campos de la ingeniería.

6. CONCLUSIONES

El presente trabajo propone el uso de estudios comparativos de tecnologías LPWAN como recurso didáctico en la formación de ingenieros, destacando su potencial para fortalecer competencias clave como el análisis crítico, la interpretación de evidencia científica y la toma de decisiones tecnológicas fundamentadas.

A partir del análisis de la literatura, se identificó que tecnologías como LoRaWAN y NB-IoT presentan características diferenciadas en términos de consumo energético, cobertura y confiabilidad, lo que evidencia la necesidad de evaluar su desempeño en función del contexto de aplicación (Kim and Jo, 2022), (Urabe et al., 2023), (Michelinakis et al., 2020). Esta diversidad de resultados representa una oportunidad para su aprovechamiento en entornos educativos, permitiendo a los estudiantes comprender que no existe una solución única, sino que la selección tecnológica debe basarse en criterios específicos.

La estrategia didáctica propuesta, basada en el análisis de casos y el uso de evidencia científica, contribuye a una formación más activa y contextualizada. Mediante esta aproximación, los estudiantes no solo adquieren conocimientos técnicos, sino que también desarrollan habilidades analíticas y de argumentación, fundamentales en el ejercicio profesional de la ingeniería.

Asimismo, la incorporación de herramientas como tablas comparativas y rúbricas de evaluación facilita la estructuración del proceso de aprendizaje y permite una evaluación integral de las competencias desarrolladas. Este enfoque favorece la integración entre teoría y práctica, fortaleciendo la capacidad de los estudiantes para enfrentar problemas reales.

Asimismo, la evaluación piloto exploratoria realizada con estudiantes de ingeniería permitió identificar una percepción positiva respecto a la utilidad de la estrategia didáctica para fortalecer competencias relacionadas con el pensamiento

crítico, la interpretación de evidencia científica y la selección de tecnologías LPWAN.

Como parte de futuras líneas de investigación, la estrategia didáctica propuesta podría fortalecerse mediante la incorporación de herramientas digitales interactivas, simuladores de redes LPWAN y plataformas de análisis de datos IoT que permitan a los estudiantes experimentar con escenarios más dinámicos y cercanos a contextos reales. Asimismo, futuras versiones podrían integrar actividades colaborativas, análisis de datos en tiempo real y evaluaciones prácticas orientadas al desarrollo de competencias en resolución de problemas tecnológicos complejos.

Finalmente, se concluye que el uso de estudios comparativos como recurso didáctico representa una estrategia viable y pertinente en la educación en ingeniería, con potencial de aplicación en diversas áreas tecnológicas. Como trabajo futuro, se propone la implementación y evaluación de esta estrategia en contextos educativos reales, con el fin de medir su impacto en el desarrollo de competencias en los estudiantes.

REFERENCIAS

- Anbazzhagan, S., Mugelan, R.K., 2024. Energy efficiency optimization of NB-IoT using integrated Proxy & ERAI technique. *Results in Engineering* 23, 102419. <https://doi.org/10.1016/j.rineng.2024.102419>
- Andres-Maldonado, P., Ameigeiras, P., Prados-Garzon, J., Navarro-Ortiz, J., Lopez-Soler, J.M., 2017. Narrowband IoT Data Transmission Procedures for Massive Machine-Type Communications. *IEEE Network* 31, 8–15. <https://doi.org/10.1109/MNET.2017.1700081>
- Ballerini, M., Magno, M., n.d. NB-IoT vs. LoRaWAN: An Experimental Evaluation for Industrial Applications.
- Hmelo, S., 2004. Problem-Based Learning: What and How Do Students Learn? *Educational Psychology Review* 16, 235–266. <https://doi.org/https://doi.org/10.1023/B:EDPR.0000034022.16470.f3>
- Kim, M.-G., Jo, H.-S., 2022. Performance Analysis of NB-IoT Uplink in Low Earth Orbit Non-Terrestrial Networks. *Sensors* 22, 7097. <https://doi.org/10.3390/s22187097>
- Loh, F., Mehling, N., Hoßfeld, T., 2022. Towards LoRaWAN without Data Loss: Studying the Performance of Different Channel Access Approaches. *Sensors* 22, 691. <https://doi.org/10.3390/s22020691>
- Masek, P., Stusek, M., Svertoka, E., Pospisil, J., Burget, R., Lohan, E.S., Marghescu, I., Hosek, J., Ometov, A., 2021. Measurements of LoRaWAN Technology in Urban Scenarios: A Data Descriptor. *Data* 6, 62. <https://doi.org/10.3390/data6060062>
- Michael, P., 2024. Does Active Learning Work? A Review of the Research. *Journal of Engineering Education* 223–231.
- Michelinakis, F., Al-selwi, A.S., Capuzzo, M., Zanella, A., Mahmood, K., Elmokashfi, A., 2020. Dissecting Energy



- Consumption of NB-IoT Devices Empirically.
<https://doi.org/10.48550/arXiv.2004.07127>
- Piyare, R., Murphy, A.L., Magno, M., Benini, L., 2018. On-Demand TDMA for Energy Efficient Data Collection with LoRa and Wake-up Receiver.
<https://doi.org/10.48550/arXiv.1809.04142>
- Quispe-Vilca, J.L., Moreno-Cardenas, E., Sacoto-Cabrera, E.J., Moreno-Cardenas, Y., 2024. Integrating IoT and Image Processing for Crop Monitoring: A LoRa-Based Solution for Citrus Pest Detection. *Electronics* 13, 4863.
<https://doi.org/10.3390/electronics13244863>
- Rahman, M., Saifullah, A., 2023. Boosting Reliability and Energy-Efficiency in Indoor LoRa, in: *Proceedings of the 8th ACM/IEEE Conference on Internet of Things Design and Implementation*. Presented at the IoTDI '23: International Conference on Internet-of-Things Design and Implementation, ACM, San Antonio TX USA, pp. 396–409. <https://doi.org/10.1145/3576842.3582327>
- Soussi, A., Zero, E., Sacile, R., Trincherro, D., Fossa, M., 2024. Smart Sensors and Smart Data for Precision Agriculture: A Review. *Sensors* 24, 2647.
<https://doi.org/10.3390/s24082647>
- Soy, H., 2023. Coverage Analysis of LoRa and NB-IoT Technologies on LPWAN-Based Agricultural Vehicle Tracking Application. *Sensors* 23, 8859.
<https://doi.org/10.3390/s23218859>
- Stanco, G., Navarro, A., Frattini, F., Ventre, G., Botta, A., 2024. A comprehensive survey on the security of low power wide area networks for the Internet of Things. *ICT Express* 10, 519–552.
<https://doi.org/10.1016/j.icte.2024.03.003>
- Ugwuanyi, S., Paul, G., Irvine, J., 2021. Survey of IoT for Developing Countries: Performance Analysis of LoRaWAN and Cellular NB-IoT Networks. *Electronics* 10, 2224. <https://doi.org/10.3390/electronics10182224>
- Urabe, I., Li, A., Fujisawa, M., Kim, S.-J., Hasegawa, M., 2023. Combinatorial MAB-Based Joint Channel and Spreading Factor Selection for LoRa Devices. *Sensors* 23, 6687. <https://doi.org/10.3390/s23156687>

Performance Evaluation of Fermat Spiral Distributions for Grating Lobe Suppression in Sparse UAV Networks

Kleiverg Eulalio Encino-Morales*, Miguel Ángel Sidón-Ayala *, Rolando Díaz-Castillo*

**Tecnológico Nacional de México/I.T. Ensenada, Boulevard Tecnológico No. 150, Ex-Ejido Chapultepec, 22780 Ensenada, Baja California, México (e-mail: kencino@ite.edu.mx, msidon@ite.edu.mx, rdiaz@ite.edu.mx)*

Abstract: Distributed antennas arrays in Unmanned Aerial Vehicles (UAV) swarms are a key technology for high-gain beamforming in next-generation networks. However, maintaining the Nyquist-Criterion spacing ($\lambda/2$) between aerial platforms is physically impractical due to rotor clearance and flight safety constraints. This results in ultra-sparse distributed arrays where conventional periodic geometries, such as rectangular grids, suffer from severe spatial aliasing and the emergence of high-magnitude grating lobes, which impair system performance. This paper presents a performance evaluation of Fermat spiral distributions as an aperiodic solution to mitigate these effects at 5.8 GHz. By integrating individual antenna element patterns characterized in CST Studio Suite with array factor synthesis in MATLAB, we analyze the suppression of grating lobes under metric-scale inter-element spacing. Evaluating the spiral array with this methodology, a Side Lobe Level (SLL) of -6 dB and a Half-Power Beamwidth (HPBW) of 2.4° are achieved for a geometry with 16 elements maintaining a minimum distance of 5λ . Furthermore, variations of 0.5 dB are obtained for greater distances between elements, showing good performance for array scaling. An SLL of -6 dB is also reached in scanning for a maximum value of $\theta_0 = 45^\circ$. No grating lobes were obtained in the cases evaluated. The results demonstrate that the Fermat Spiral effectively redistributes the interference energy into a controlled side-lobe floor, providing a robust radiation pattern for high-directivity links that is not achievable with periodic distributions.

Keywords: Antenna Arrays; UAV Swarms; Fermat Spiral; Beamforming; Grating Lobes; 6G Networks

1. INTRODUCTION

The rapid evolution of 6G wireless communications has positioned Non-Terrestrial Networks (NTN) as a fundamental pillar for achieving global connectivity. In this context, swarms of Unmanned Aerial Vehicles (UAVs) are being extensively researched to serve as aerial relay nodes, providing high-speed data links in remote areas or disaster zones where terrestrial infrastructure is absent (Yang et al., 2017). A novel and disruptive approach in this field is the concept of Aerial Distributed Antenna Arrays, where a fleet of drones is deployed in an open-air environment to synthesize a large-scale radiation aperture (Jia et al., 2019).

In this architecture, each individual drone acts as a single active element of a larger phased array, carrying its own antenna and RF chain. This allows the swarm to behave as a dynamic and reconfigurable antenna system capable of focusing energy toward specific targets through collaborative beamforming, like when it is desired to establish reliable long-range communications, such as UAV-to-Satellite or UAV-to-Ground Backhaul. The primary advantage of this deployment is its flexibility; unlike fixed ground-base stations, a swarm can adjust its geometry in real-time to optimize link quality or bypass obstacles (Ju & Son, 2019). However, the physical implementation of distributed antenna arrays in aerial platforms faces a significant constraint: the Nyquist-criterion spacing.

In conventional array theory, elements must be separated by no more than half a wavelength ($\lambda/2$) to avoid spatial aliasing, (Balanis, 2015). For a system operating at 5.8 GHz, this corresponds to a separation of approximately 2.5 cm. Maintaining such proximity between multiple autonomous UAVs is technically impractical due to aerodynamic turbulence, rotor clearance, and flight safety protocols, which typically demand metric-scale separations (Jia et al., 2019).

Consequently, when these drones are deployed to form a collective geometry in the air, the resulting array is inherently ultra-sparse (Ullah et al., 2017). In periodic configurations, such as the standard rectangular grid, this high degree of sparsity leads to the emergence of grating lobes, which are interference peaks with magnitudes nearly identical to the main beam (Malek et al., 2018). These lobes severely degrade the signal-to-interference ratio (SIR) and lead to inefficient energy distribution, rendering periodic distributions unsuitable for safe and effective UAV-based beamforming.

Despite the known benefits of aperiodic distributions in theory, there is a significant research gap regarding their characterization in the context of UAV networks. Most existing studies rely on optimization methods that are computationally expensive for real-time coordination (Zhang et al., 2023), or use ideal isotropic models that neglect the impact of antenna element patterns (Bui et al., 2018). Furthermore, the evidence of how these geometries scale as the array aperture expands in dynamic environments is scarce.

Therefore, this paper addresses these gaps by providing a performance evaluation of the Fermat Spiral as a low-complexity, deterministic alternative for grating lobe suppression.

The objective of this research is to evaluate and validate the effectiveness of the Fermat spiral geometry in suppressing grating lobes within ultra-sparse aerial arrays. To achieve this, a cross-platform numerical approach is employed, integrating CST Studio for the electromagnetic characterization of the individual antenna elements and MATLAB for the mathematical synthesis of the collective array factor and total radiation pattern. By strategically positioning each drone-element in a deterministic aperiodic distribution, it is possible to break the coherence in the interference patterns and suppress grating lobes across the angular domain. Through this methodology, it is possible to eliminate grating lobes and to obtain an SLL of approximately -6 dB in fixed beam and scanning for distances equal to and greater than 5λ . The results demonstrate that Fermat Spiral based aerial arrays provide a robust and high-directivity solution that overcomes the physical limitations of drone flight while maintaining superior radiation characteristics.

2. AERIAL ARRAY GEOMETRY

2.1 Theoretical Constraints of Aerial Arrays

The synthesis of a distributed antenna array from a swarm of UAVs introduces a fundamental conflict between electromagnetic theory and aeronautical safety. In conventional phased array design, the suppression of spatial aliasing is governed by the Nyquist-Shannon sampling theorem (Cardama et al., 2002), which dictates that the inter-element spacing (d) must satisfy:

$$d \leq \frac{\lambda}{1 + |\sin \theta_0|} \quad (1)$$

where λ is the operating wavelength and θ_0 is the maximum steering angle in elevation. Is important to note that the criteria of distance also depend of the maximum angle of direction of the main lobe, and is in the range $\theta_0 = [-\pi/2, \pi/2]$, with $\theta_0 = 0^\circ$ for a perpendicular radiation to the array. Therefore, if the array needs to point in different directions, the elements must be placed at smaller distances to avoid grating lobes. The maximum allowed distance occurs when there is no scanning. For a carrier frequency of 5.8 GHz, this theoretical limit is approximately 2.58 cm. However, in an aerial environment, maintaining such proximity is physically impossible. Factors such as aerodynamic coupling, GPS positioning error, and the need for rotor clearance impose a minimum safety distance (d_{min}).

When a periodic distribution, such as a uniform rectangular array (URA), is scaled to meet these safety requirements, the array becomes ultra-sparse. In the spatial frequency domain, this over-sampling leads to the emergence of grating lobes. These are constructive interference maxima that appear at angles, $\theta_{p,q}$ defined by:

$$\sin \theta_{p,q} = \sin \theta + p \frac{\lambda}{d_x} + q \frac{\lambda}{d_y} \quad (2)$$

Where p, q , are integers, and d_x, d_y are the distances between adjacent elements in the axis of the plane. In an ultra-sparse periodic regime, these lobes exhibit the same magnitude as the main beam, resulting in a degradation of the Signal-to-Interference Ratio (SIR) and a significant loss of energy efficiency, as the radiated power is divided among multiple redundant beams. Consequently, periodic geometries are inherently unsuitable for UAV-based collaborative beamforming, necessitating a transition toward deterministic aperiodic distributions. Among the various geometries reported in the literature, spiral antenna arrays have attracted increasing attention due to their ability to provide aperiodic element distributions and improved radiation characteristics compared to conventional linear and rectangular configurations.

2.2 Fermat Spiral Modeling and Aperiodicity

To mitigate the emergence of grating lobes in ultra-sparse conditions, this study adopts a deterministic aperiodic distribution based on the Fermat Spiral (a specific form of phyllotaxis), (Vogel, 1979). A critical advantage of this deterministic approach is its computational efficiency compared to stochastic methods. Traditional techniques for optimizing sparse arrays, such as Genetic Algorithms (GA) or Particle Swarm Optimization (PSO), require intensive iterative processes to determine optimal element positions that minimize sidelobe levels (SLL). As the number of elements (N) increases, the search space for these metaheuristic methods grows exponentially, leading to significant computational overhead that is often impractical for real-time UAV swarm coordination.

In contrast, the Fermat spiral formulation offers the following benefits. It offers low complexity, because the coordinates are generated through a direct closed-form expression, and the time required to calculate positions scales linearly and remains negligible even for large-scale swarms. The model allows for dynamic expansion; new UAVs can be integrated into the array by simply incrementing the index n , maintaining the overall aperiodic properties without the need to recalculate the positions of the existing element.

The spatial coordinates (x_n, y_n) for each UAV-element in the swarm are derived from the polar representation, (Gabielli & Hernandez-Figueroa, 2016):

$$\rho_n = \frac{d_{min}}{d_{14}} \sqrt{n} \quad (3)$$

$$\alpha_n = n\pi(3 - \sqrt{5}) \quad (4)$$

where ρ_n is the distance of the n th element to the center of the array, ϕ_n the angle of two adjacent elements (through the spiral), $\pi(3 - \sqrt{5})$ is the golden angle, $d_{14} = \sqrt{5 - 4\cos\alpha_3}$ and $n=1, 2, \dots, N$. These equations can be set for a minimal distance between elements d_{min} . Figure 1 exhibits the geometry when $d_{min} = 5\lambda$ for $N=16$.

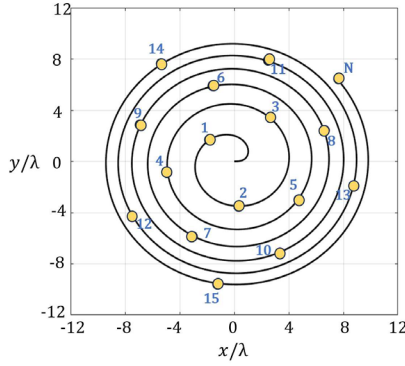


Figure 1. Geometry of Fermat spiral for UAV, $d_{min} = 5\lambda$ for $N=16$

The advantage of this geometry is its intrinsic aperiodicity. Unlike rectangular or circular grids, the Fermat spiral ensures that the inter-element distances are non-uniform and non-repetitive across the array aperture. In terms of the Array Factor, this spatial “jitter” prevents the coherent summation of phases in off-axis directions. Consequently, the power that would otherwise be concentrated into discrete grating lobes is instead redistributed into a lower, nearly uniform sidelobe floor. This property allows the swarm to maintain a highly directive main beam even when the average inter-element spacing is higher than $\lambda/2$. This characteristic is shown in next sections.

2.3. Array factor and total radiation pattern

The far-field radiation characteristics of the aerial swarm are determined by the interaction between the individual antenna response and the spatial distribution of the nodes. The total radiation pattern $E_{total}(\theta, \phi)$ on the distributed array is governed by the pattern principle (Kildal, 2015):

$$E_{total}(\theta, \phi) = EF(\theta, \phi) \cdot AF(\theta, \phi) \quad (5)$$

where $EF(\theta, \phi)$ represents the Element Factor (the radiation pattern of the single antenna mounted on each UAV) and $AF(\theta, \phi)$ is the Array Factor. The spherical coordinate system is used like in Figure 2, where there is a point of observation P in the far-field and a distance R from the origin (Martínez et al., 2010).

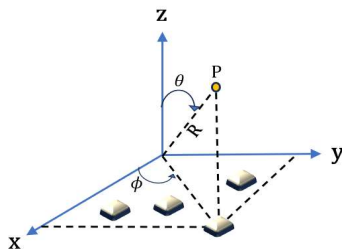


Figure 2. Spherical coordinate system for distributed array antenna

In this coordinate system, ϕ corresponds to the azimuth angle, and θ corresponds to elevation angle. For a swarm of N elements distributed in the xy -plane, the Array Factor is expressed as:

$$AF(\theta, \phi) = \sum_{n=1}^N A_n e^{j\varphi_n} \cdot e^{jk[x_n \sin\theta \cos\phi + y_n \sin\theta \sin\phi]} \quad (6)$$

Where $k=2\pi/\lambda$ is the wavenumber, (x_n, y_n) are the Cartesian coordinates of each drone-element derived from the equations 3 and 4 for the Fermat spiral, and A_n is the excitation amplitude with respective phase φ_n for the n th antenna element. The value φ_n is calculated according to the desired direction angle of the main beam. Figure 3 exhibits a proposal to physical implementation of each swarm element. To achieve a stable radiation pattern required for communication, an antenna is integrated into the upper chassis of the UAV. This configuration avoids signal obstruction from the drone body and minimizes interference with the rotors, providing a clear line-of-sight for the array's beamforming operations.



Figure 3. Illustration of the UAV node architecture for the proposed Fermat Spiral array

To achieve a collaborative beamforming toward a specific target at (θ_0, ϕ_0) , the phase excitation φ_n of each drone must be dynamically adjusted to compensate for the path differences:

$$\varphi_n = -k(x_n \sin\theta_0 \cos\phi_0 + y_n \sin\theta_0 \sin\phi_0) \quad (7)$$

The inclusion of the Element Factor is critical in this study, as the individual antenna's directivity acts as a spatial filter. Even if the Array Factor presents sidelobes at wide angles due to the ultra-sparse nature of the distribution, the EF (characterized via CST Studio Suite) provides an additional layer of suppression, effectively tapering the radiation outside the region of interest and optimizing the overall efficiency of the UAV-to-Satellite or UAV-to-Ground link.

3. PERFORMANCE METRICS FOR ARRAY EVALUATION

To quantitatively evaluate the radiation performance of the proposed Fermat Spiral UAV array, two fundamental metrics are considered: Sidelobe Level (SLL) and Half-Power Beamwidth (HPBW). These parameters are essential to characterize the beam-scanning capabilities and the interference suppression potential of the swarm in sparse conditions.

3.1 Sidelobe Level (SLL)

The SLL is defined as the ratio of the power of the secondary lobes to the peak power of the main beam. In sparse arrays, managing the SLL is critical to avoid potential interference with adjacent communications links (Cardama et al., 2002). For the purpose of this study, the SLL is defined in decibels (dB) as:

$$SLL(dB) = 10\log_{10}\left(\frac{|AF(\theta, \phi)|^2_{sidelobe}}{|AF(\theta, \phi)|^2_{max}}\right) \quad (8)$$

3.2 Half-Power Beamwidth (HPBW)

The HPBW measures the angular separation between the two points on the main lobe where the power density drops to 50% (-3 dB) of its maximum value (Balanis, 2015). This metrics define the resolution and spatial selectivity of the beam. A narrower HPBW indicates a more directive beam, which is desirable for long-range Non-Terrestrial Network links. The HPBW is fundamentally determined by the total array aperture and the effective distribution of the N elements. Figure 4 provides a graphical definition of these metrics.

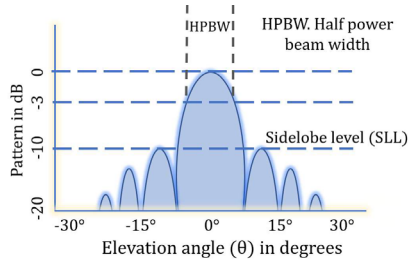


Figure 4. SLL and HPBW for antenna arrays

3.3 Radiation pattern characterization and SLL suppression

The primary objective of employing the Fermat spiral geometry is to transform the coherent interference (grating lobes) into a non-coherent spatial distribution of sidelobes. In ultra-sparse arrays, the total radiated power is conserved; however, the aperiodic nature of the Fermat distribution ensures that the energy not contained in the main beam is spread across the entire angular domain rather than being concentrated in spurious maxima. Mathematically, this result leads to a nearly uniform Sidelobe level (SLL) floor. While a periodic array with high scale would exhibit grating lobes with a relative magnitude of 0 dB, the Fermat-based aerial array is designed to suppress these peaks significantly. In the case of deterministic aperiodic arrays, the SLL improves as the number of elements N increases, following a relationship approximately proportional to $10\log_{10}(N)$. By strategically expanding the scaling factor d_{min}/d_{14} to meet flight safety requirements, the effective aperture size of the swarm increases, leading to an extremely narrow main beam. This high-resolution focus is what enables the high-gain links require for long-range communication in Non-Terrestrial Networks, effectively overcoming the sparsity constraints.

4. METHODOLOGY

The performance of the Fermat spiral array was evaluated through a co-simulation approach combining MATLAB and CST Studio Suite. The methodology followed a three-stage process. First, the spatial coordinates of the UAV-elements were generated in MATLAB using the Fermat spiral formulation (equations 3 and 4), allowing for the calculation of the theoretical Array Factor (equation 6) to provide a preliminary assessment of the SLL and HPBW. Subsequently, a single antenna element was designed and simulated in CST Studio Suite to obtain its complex radiation pattern, accounting for realistic electromagnetic properties and

physical dimensions. Finally, the individual element pattern was exported back to MATLAB, where the total radiation pattern of the swarm was computed by applying the Pattern Multiplication Principle (equation 5). This workflow ensures that the final results effectively combine the high-fidelity CST element data with the geometric properties of the Fermat distribution, providing a realistic characterization of the physical antenna array.

5. RESULTS AND DISCUSSION

5.1 Geometry Visualization and Grating Lobe Analysis

To evaluate the impact of geometry on spatial aliasing, a comparative analysis was performed between a Uniform Rectangular Array (URA) and the proposed Fermat Spiral distribution. In this scenario, both configurations were designed with a small swarm of $N=16$ elements and an inter-element spacing of 5λ (≈ 25.8 cm at 5.8 GHz). While this distance is below the typical aeronautical safety margin, it serves as a critical benchmark to observe the onset of grating lobes in sparse periodic structures, even for a reduced number of nodes. Figure 5 and Figure 6 illustrate the physical distribution of the drone nodes. The periodic URA (Figure 5) aligns the 16 drones in a rigid 4x4 grid, creating a highly structured and symmetric layout. In contrast, the Fermat Spiral geometry (Figure 6) for the same number of drones adopts a deterministic aperiodic layout derived from the Golden Angle, resulting in a more dispersed distribution of elements.

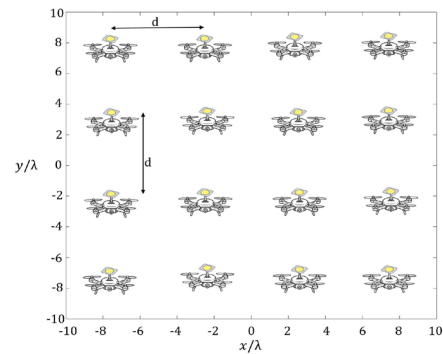


Figure 5. Spatial distribution of N=16 UAV nodes in a periodic rectangular array (URA) with 5λ inter-element spacing

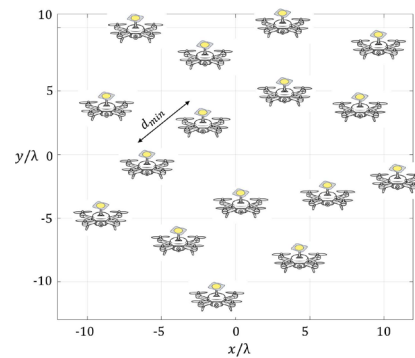


Figure 6. Spatial distribution of N=16 UAV nodes in aperiodic Fermat Spiral geometry with $d_{min}=5\lambda$

The array factor resulting from these two specific geometries are compared in Figure 7. It is important to emphasize that for both cases, no amplitude tapering or phase optimization methods were applied to the antenna elements. All nodes are excited with uniform amplitude and the phases are set solely to satisfy the constructive interference condition at the boresight ($0^\circ, 0^\circ$). This approach ensures that the observed differences in performance are strictly attributable to the spatial distribution of the nodes and not to external signal processing enhancements.

The periodic URA pattern exhibits multiple grating lobes at 0 dB, a direct consequence of the 5λ periodicity. Conversely, the Fermat Spiral achieves a Peak Sidelobe Level (PSL) of approximately -6 dB. This reduction is achieved purely through the deterministic aperiodicity of the Fermat geometry. Achieving a 6 dB suppression margin without any complex weight optimization demonstrates the inherent efficiency of the Fermat distribution for sparse aerial arrays.

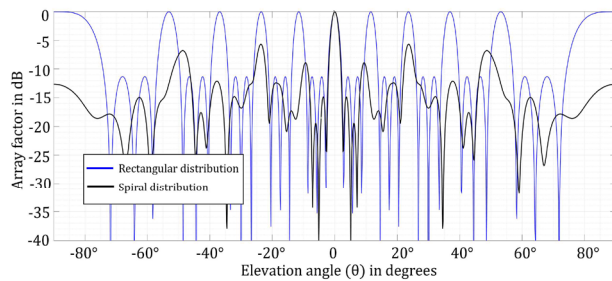


Figure 7. Normalized radiation patterns for $N=16$ with 5λ spacing

5.2 Beamwidth and Spatial Resolution Analysis

The secondary advantage of the Fermat-based distribution is its ability to synthesize a large virtual aperture, which directly dictates the Half-Power Beamwidth (HPBW) and the spatial resolution of the aerial array. The large physical footprint of the swarm, enforced by the 5λ minimum spacing, results in an extremely narrow beam, which is essential for high-precision UAV-to-Satellite links. Quantitatively, the periodic URA exhibited an HPBW of 2.6° . In contrast, the Fermat Spiral geometry achieved an even narrower beamwidth of 2.4° . This slight reduction in HPBW is directly attributed to the spatial distribution of the nodes.

To satisfy the minimum distance constraint ($d_{min} \approx 5\lambda$) while breaking the periodic grid, the Fermat Spiral occupies a slightly larger total aperture area compared to the URA. In antenna theory, a larger effective aperture leads to a higher spatial resolution and a more concentrated main lobe. Consequently, the Fermat-based virtual antenna not only suppresses interference but also provides a more directive beam, maximizing the link budget and efficiency for Non-Terrestrial Network (NTN) applications.

5.3 Beam steering and Scanning Stability

To evaluate the operational flexibility of the aerial array, the collaborative beamforming capability was tested by steering the main beam across an angular range of $\theta_0 = [0, 45^\circ]$. This simulation is critical for Non-Terrestrial Networks (NTN), where the swarm must maintain a continuous, high-gain link with a moving satellite.

The stability of the radiation pattern during beam steering is summarized in Table 1. As the main beam is steered from 0° to 45° , the Fermat Spiral geometry exhibits remarkable consistency. The Peak Sidelobe Level (PSL) remains effectively locked at approximately -6 dB throughout the entire scanning range. This demonstrates that the deterministic aperiodic distribution prevents the formation of grating lobes that would typically emerge in periodic arrays.

Table 1. Scanning performance analysis of the 16-element Fermat Spiral array, $d_{min} = 5\lambda$

Steering angle (θ_0)	Peak Sidelobe level (PSL)	Steering angle (θ_0)	Peak Sidelobe level (PSL)
0°	$\approx -6 \text{ dB}$	30°	$\approx -6 \text{ dB}$
10°	$\approx -6 \text{ dB}$	40°	$\approx -6 \text{ dB}$
20°	$\approx -6 \text{ dB}$	45°	$\approx -6 \text{ dB}$

Figure 8 illustrates the radiation patterns for the Fermat Spiral distribution during the scanning process. The periodic URA configuration is omitted from this comparison as it consistently exhibits grating lobes of 0 dB across the entire steering range, rendering it impractical for dynamic link tracking. In contrast, the Fermat Spiral results demonstrate remarkable advantage. As the beam is steered, the energy remains successfully redistributed into a controlled, non-coherent sidelobe floor. This constant 6 dB suppression margin ensures that the main beam remains the dominant and clearly distinguishable path for power delivery, proving that the Fermat based swarm can effectively track orbital targets without the risk of spatial aliasing or significant gain fluctuations.

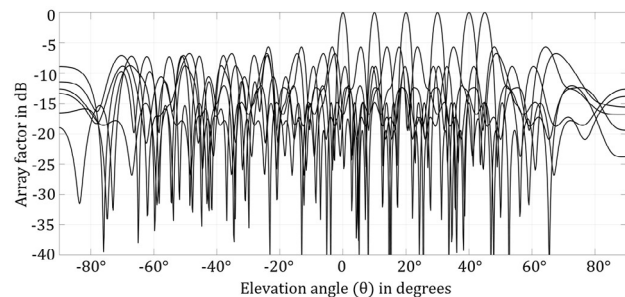


Figure 8. Beam steering performance of the $N=16$ Fermat Spiral array for θ_0 steered from 0° to 45°

5.4 Scalability of the Fermat Spiral

The data presented in Table 2 confirms the invariance of the sidelobe suppression with the distance between elements within the Fermat Spiral. Despite increasing the swarm footprint significantly, the PSL remains stable within the range of -5.5 to -6 dB. This stability is critical; it demonstrates that the deterministic aperiodic nature of the Fermat distribution effectively prevents the formation of dominant grating lobes even as the array aperture expands. Consequently, the proposed geometry is not merely restricted to the simulated 5λ configuration but is fully scalable for diverse UAV swarm operational requirements, ensuring consistent link performance across various physical deployment densities.

Table 2. Impact of inter-element spacing on SLL performance

Spacing (d/λ)	Physical distance (5.8 GHz)	Peak Sidelobe level (PSL)
5λ	$\approx 26 \text{ cm}$	$\approx -6 \text{ dB}$
10λ	$\approx 52 \text{ cm}$	$\approx -6 \text{ dB}$
20λ	$\approx 104 \text{ cm}$	$\approx -5.5 \text{ dB}$
30λ	$\approx 155 \text{ cm}$	$\approx -6 \text{ dB}$

5.5 Impact of the Antenna Element Factor (CST Characterization)

The ideal isotropic sources were replaced by a realistic antenna element. The element selected is a circular microstrip patch antenna characterized in CST Studio Suite, operating in a center frequency of 5.8 GHz, using an Arlon 25N substrate with the following characteristics: substrate thickness $h=1.54 \text{ mm}$, $\epsilon_{\text{r}}=3.38$, and $\text{tangential losses}=0.0025$. The design in CST is shown in Figure 9. The values of r and p' are 7.06 mm and 1.18 mm , respectively.

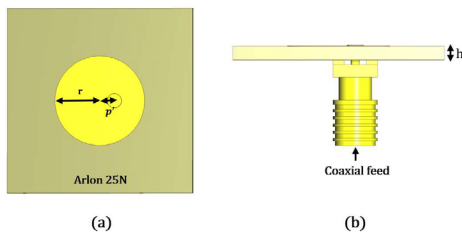


Figure 9. Design of circular patch antenna in CST for 5.8 GHz, (a) Top view, (b) Side view

The Figure 10 and Figure 11 exhibit the radiation pattern in 3D in 5.8 GHz and the reflection coefficient S_{11} for the antenna patch, respectively. It can be noted that the circular patch exhibits a narrow operational bandwidth. While this constraint is inherent to the chosen element, the Fermat Spiral geometry is element-agnostic. Future iterations of this design could integrate wideband elements, such as Vivaldi or log-periodic antennas, to support multi-band operations without requiring modifications to the underlying spatial distributions. These alternatives, while outside the scope of this study, provide a

clear path for expanding the frequency agility of the aerial array.

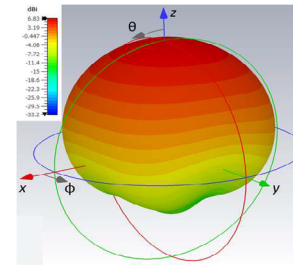


Figure 10. Radiation pattern of the antenna element in CST

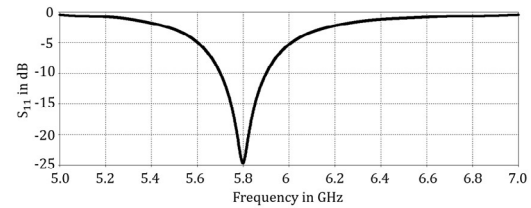


Figure 11. Reflection coefficient S_{11} of the antenna element in CST

The influence of the individual antenna element characteristics on the overall array performance is shown in Figure 12. This comparison highlights the transition from theoretical geometric analysis to a realistic system model. The array factor (black line) represents the pure effect of the 16-element Fermat distribution using ideal isotropic radiators. In this case, the side-lobe floor is dictated strictly by the spatial layout of the nodes. The total radiation pattern (red line) is obtained by integrating the simulated element factor of the circular patch antenna. As observed, the total pattern has slightly lower levels than the array factor ($\approx -6.8 \text{ dB}$). It is critical to not that achieving low SLL in designs with this distance between elements is inherently challenging.

Furthermore, while the current design utilizes a standard patch, the overall radiation pattern could be further optimized by integrating more sophisticated antenna elements. These elements, when combined with the inherent spatial suppression of the Fermat geometry, represent a viable path for advanced array synthesis in future works.

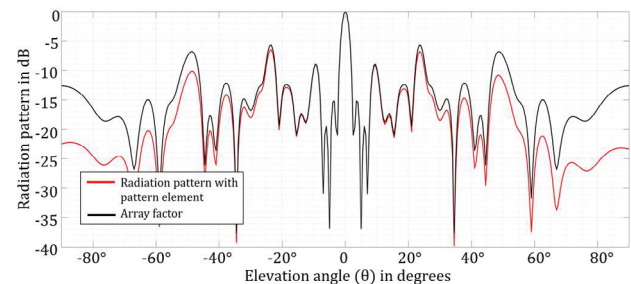


Figure 12. Comparative radiation patterns for the 16-element Fermat Spiral Swarm.

As a final result, to evaluate the performance benefits of the proposed Fermat Spiral distribution, a comparative analysis was performed against conventional uniform rectangular arrays. As summarized in Table 3, the Fermat distribution is a superior candidate for dynamic, sparse swarm operations.

Table 3. Comparison of Array Geometries for Sparse UAV Swarms

Parameter	Uniform Rectangular Array (URA)	Fermat Spiral Distribution
Geometry	Periodic (grid)	Deterministic aperiodic
Grating Lobe Suppression	Low (if spacing $> \lambda/2$)	High (suppressed by design)
Spatial Aliasing	High risk in sparse configurations	Negligible due to non-uniformity
Beam Steering Complexity	Moderate	Moderate (scale-invariant properties)
Scalability	Limited by grating lobes constraints	High (scale-invariant properties)
Optimal Applications	Fixed infrastructure	Dynamic sparse aerial networks

5.6 Discussion

The results demonstrate that the Fermat Spiral distribution provides a robust, scale-invariant solution for large-aperture UAV swarms in non-terrestrial networks. Unlike periodic architectures, which are susceptible to grating lobe formation when inter-element spacing exceeds $\lambda/2$, the deterministic aperiodicity of the Fermat geometry effectively suppresses spatial aliasing. This ensures beamforming stability as the swarm aperture scales from $d_{min} = 5\lambda$ to 30λ , confirming the suitability of this layout for dynamic operational environments. By treating the UAV swarm as a reconfigurable virtual antenna, this approach shifts the paradigm from static ground-based infrastructure to dynamic, intelligent aerial formations.

6. CONCLUSIONS

This study has successfully evaluated the performance of Fermat spiral distributions as a deterministic solution for grating lobe suppression in ultra-sparse UAV-based antenna arrays. The results demonstrate that the intrinsic aperiodicity of this geometry effectively breaks the phase coherence that leads to spatial aliasing in periodic structures. Through a co-simulation workflow integrating MATLAB and CST Studio Suite, it was possible to validate that the proposed distribution maintains a robust radiation pattern under realistic flight safety constraints, achieving a Sidelobe Level (SLL) of approximately -6 dB and a Half-Power Beamwidth (HPBW) of 2.4° for a 16-element swarm with a minimum inter-element spacing of 5λ . These values were achieved solely through geometric distribution, without applying any additional sidelobe reduction techniques such as amplitude tapering. Maintaining uniform excitation across all elements, the complexity of the feeding system and the power management of the UAVs are significantly reduced, which is a major advantage for practical deployments.

The Fermat spiral provides a scale-invariant architecture, where the sidelobe suppression levels remain stable even as the swarm footprint expands. Furthermore, beamforming capabilities were confirmed during scanning, showing that the main lobe remains the dominant energy path without the emergence of grating lobes. This research establishes a good foundation for the deployment of reconfigurable virtual antenna arrays in Non-Terrestrial Networks, offering a computationally efficient alternative to stochastic optimization methods.

REFERENCES

- Balanis, C. A. (2015). *Antenna Theory. Analysis and Design* (4th ed.). Wiley & Sons. https://books.google.com.mx/books/about/Antenna_Theory.html?id=v1PSZ48DnuEC&redir_esc=y
- Bui, P. L. T., Rocca, P., & Haupt, R. L. (2018). Aperiodic Planar Array Synthesis Using Pseudo-Random Sequences. *2018 International Applied Computational Electromagnetics Society Symposium - China (ACES)*, 1–2. <https://doi.org/10.23919/ACCESS.2018.8669272>
- Cardama, A., Jofre, L., Rius, J. M., Romeu, J., Blanch, S., & Ferrando, M. (2002). *Antenas* (2nd ed., Vol. 1). Ediciones UPC. https://books.google.com.mx/books/about/Antenas.html?id=AlQ8p2C4sCoC&redir_esc=y
- Gabrielli, L. H., & Hernandez-Figueroa, H. E. (2016). Aperiodic Antenna Array for Secondary Lobe Suppression. *IEEE Photonics Technology Letters*, 28(2), 209–212. <https://doi.org/10.1109/LPT.2015.2492419>
- Jia, Z., Wang, L., Yu, J., & Ai, X. (2019). Distributed adaptive neural networks leader-following formation control for quadrotors with directed switching topologies. *ISA Transactions*, 93, 93–107. <https://doi.org/10.1016/j.isatra.2019.02.030>
- Ju, C., & Son, H. Il. (2019). A distributed swarm control for an agricultural multiple unmanned aerial vehicle system. *Proceedings of the Institution of Mechanical Engineers, Part I: Journal of Systems and Control Engineering*, 233(10), 1298–1308. <https://doi.org/10.1177/0959651819828460>
- Kildal, P.-S. (2015). *Foundations of Antenna Engineering: a Unified Approach for Line-of-Sight and Multipath*. Kildal Antenna AB. https://books.google.com.mx/books/about/Foundations_of_Antenna_Engineering.html?id=3pm_jwEACAAJ&redir_esc=y
- Malek, N. A., Khalifa, O. O., Abidin, Z. Z., Mohamad, S. Y., & Rahman, N. A. A. (2018). Beam Steering using the Active Element Pattern of Antenna Array. *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, 16(4), 1542. <https://doi.org/10.12928/telkomnika.v16i4.9040>
- Martinez, O., Martín, C. J., Godoy, G., & Ullate, L. G. (2010). 2D array based on Fermat spiral. *Physics Procedia*, 3(1), 399–406. <https://doi.org/10.1016/j.phpro.2010.01.052>
- Ullah, N., Huiling, Z., Rahim, T., ur Rahman, S., & MuhammadKamal, M. (2017). Reduced side lobe level of sparse linear antenna array by optimized spacing and excitation amplitude using particle swarm optimization. *2017 7th IEEE International Symposium on Microwave, Antenna, Propagation, and EMC Technologies (MAPE)*, 96–99. <https://doi.org/10.1109/MAPE.2017.8250806>
- Vogel, H. (1979). A better way to construct the sunflower head. *Mathematical Biosciences*, 44(3–4), 179–189. [https://doi.org/10.1016/0025-5564\(79\)90080-4](https://doi.org/10.1016/0025-5564(79)90080-4)
- Yang, H., Lee, Y., Jeon, S.-Y., & Lee, D. (2017). Multi-rotor drone tutorial: systems, mechanics, control and state estimation. *Intelligent Service Robotics*, 10(2), 79–93. <https://doi.org/10.1007/s11370-017-0224-y>
- Zhang, J., Zhang, H., Chen, X., Wu, F., Liu, Y., & Zhang, W. (2023). A Scalable Arrangement Method for Aperiodic Array Antennas to Reduce Peak Sidelobe Level. *IEEE Access*, 11, 59899–59908. <https://doi.org/10.1109/ACCESS.2023.3282167>

Spatiotemporal Modeling of Cloudiness Using ConvLSTM Networks: Applications to Photovoltaic Energy

V. Rivera Altamirano*, C.D. Garcia-Beltran *, M. Borunda*

**Tecnológico Nacional de México/Centro Nacional de Investigación y Desarrollo Tecnológico, Interior
Internado Palmira s/n, Col. Palmira, C.P. 62490, Cuernavaca, Morelos, México. (e-mail:
m24ce074@cenidet.tecnm.mx, monica.bp@cenidet.tecnm.mx, carlos.gb@cenidet.tecnm.mx)*

Abstract: This work presents a spatiotemporal model based on Convolutional Long Short-Term Memory (ConvLSTM) networks for cloudiness forecasting using satellite imagery. The proposed approach leverages the ability of ConvLSTM to capture both spatial and temporal dependencies in image sequences, which is essential for modeling complex atmospheric phenomena. Using a dataset of hourly satellite images, the architecture employs stacked ConvLSTM layers followed by a convolutional decoder to predict future cloud patterns from historical observations. Model performance was evaluated using Mean Squared Error (MSE), Mean Absolute Error (MAE), and Structural Similarity Index (SSIM), demonstrating high agreement between predicted and real images. These results indicate that the model effectively captures spatiotemporal patterns of cloud dynamics. This type of prediction is exceptionally relevant for renewable energy applications, as it can help anticipate solar irradiance variability and improve the management of photovoltaic systems.

Keywords: ConvLSTM, Cloudiness prediction, Satellite imagery, Deep learning, Spatiotemporal modeling, renewable energy.

1. INTRODUCCIÓN

Los sistemas fotovoltaicos representan una de las principales fuentes de energía renovable debido a la capacidad que tiene de generar electricidad de manera limpia y sostenible. Sin embargo, su desempeño depende directamente de la irradiación solar, la cual presenta una alta variabilidad asociada a factores ambientales como la temperatura, la radiación atmosférica y, principalmente, la presencia de nubosidad. Esta última genera fluctuaciones críticas en la potencia de salida. (Inman et al., 2013).

Una variabilidad alta de irradiación solar introduce incertidumbre en la producción de energía solar, lo que representa un gran desafío para la planificación, operación y control de los sistemas fotovoltaicos, especialmente en escenarios de alta penetración de energías renovables en la red eléctrica (Antonanzas et al., 2016). Por ello, la predicción de la nubosidad se ha convertido en una herramienta fundamental para anticipar cambios en la irradiación solar y mejorar la gestión energética.

Si bien ya existen diversos métodos para la predicción de nubosidad, que incluyen modelos estadísticos, probabilísticos y enfoques físicos tradicionales. Estos métodos presentan limitaciones al no capturar de manera adecuada la dinámica no lineal espacio-temporal de formaciones nubosas, lo que reduce su precisión en escenarios complejos.

En los últimos años se han propuesto enfoques basados en aprendizaje profundo, los cuales han demostrado una mayor capacidad para modelar la evolución espacio-temporal de

variables atmosféricas, mediante el uso de redes neuronales convolucionales y recurrentes (Huang et al., 2023).

El aprendizaje profundo ha emergido como una alternativa para el modelado de fenómenos atmosféricos, permitiendo el análisis de grandes secuencias de datos y la extracción de patrones complejos. En particular, la integración de imágenes de cielo y datos satelitales han demostrado ser una buena estrategia para la predicción de variables meteorológicas, al capturar la evolución temporal y también la estructura espacial de la nubosidad (Ayzel et al., 2020).

Dentro de las técnicas de Deep learning, las redes del tipo Convolutional Long Short-Term Memory (ConvLSTM) han destacado por su eficacia en tareas de predicción basadas en secuencias de imágenes, al integrar operaciones de convolución para el procesamiento espacial y estructuras recurrentes para modelar dependencias temporales. Esta combinación permite capturar de manera efectiva la dinámica de la nubosidad en el tiempo (Shi et al., 2015).

En este trabajo se propone la implementación de una arquitectura ConvLSTM para la predicción de nubosidad a corto plazo, a partir de secuencias de imágenes satelitales. El objetivo es modelar la evolución de la nubosidad en el espacio y el tiempo, con el fin de generar predicciones que puedan ser utilizadas en aplicaciones relacionadas con sistemas energéticos, con el fin de mejorar la gestión y estimación de la irradiación solar en sistemas fotovoltaicos.

2. IMÁGENES SATELITALES

Las imágenes empleadas en este trabajo fueron obtenidas a partir de datos satelitales de precipitación y nubosidad provenientes del sistema Global Satellite Mapping of Precipitation (GSMaP), desarrollado por la Agencia Japonesa de Exploración Aeroespacial (JAXA). El sistema GSMaP forma parte de la misión internacional Global Precipitation Measurement (GPM), la cual incluye información de diferentes satélites de observación terrestre (JAXA, 2024).

El sistema GSMaP ofrece información con una resolución espacial de aproximadamente 0.1° por píxel, lo que equivale a cerca de 11 kilómetros, y una resolución temporal de una hora, esta información funciona para realizar análisis de fenómenos atmosféricos.

Las imágenes utilizadas en este trabajo permiten representar la distribución espacial de la nubosidad y precipitación, donde los diferentes niveles de intensidad están asociados a la presencia y densidad de las nubes. Esto permite identificar patrones relacionados con su formación, desplazamiento y disipación.

El uso de datos satelitales como GSMaP resulta relevante en aplicaciones de predicción, ya que proporciona cobertura global, continua y homogénea permitiendo capturar la dinámica de la nubosidad de manera más completa en comparación con enfoques basados únicamente en mediciones puntuales a nivel local.

La adquisición de las imágenes se realizó mediante procesos automatizados de descarga, permitiendo la construcción de secuencias temporales de datos. En la Fig. 1 se muestra un ejemplo del proceso de obtención de las imágenes, y en la Fig. 2 algunas muestras representativas del conjunto de datos utilizado.

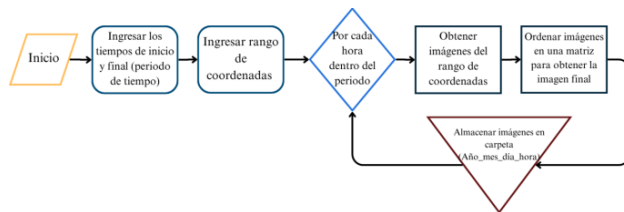


Figura 1. Proceso de adquisición y procesamiento de imágenes satelitales utilizadas en el estudio.

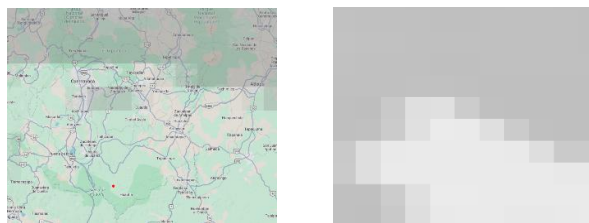


Figura 2. Ejemplo de la región de estudio (izquierda) y muestra de imagen satelital procesada (derecha).

3. CONJUNTO DE DATOS

El dataset (conjunto de datos), está compuesto por imágenes de nubosidad con una resolución temporal de una hora, correspondientes al periodo comprendido entre enero y diciembre 2024, lo que representa 8,760 imágenes. Estas imágenes posteriormente se utilizaron para construir un conjunto de datos orientado al entrenamiento de los modelos de predicción basados en secuencias temporales, este enfoque es utilizado en el modelado de fenómenos atmosféricos, donde se emplean secuencias de imágenes para capturar la evolución espacio-temporal de variables meteorológicas (An et al., 2025).

La región de estudio se localiza en Cuernavaca, Morelos, Mexico con las siguientes coordenadas geográficas 18.9483611°N y $99.21523056^\circ\text{O}$. Está región fue seleccionada debido a la disponibilidad de datos satelitales.

Para generar las secuencias temporales, se utilizó una estructura de ventana deslizante, común en problemas de predicción con series de tiempo e información visual (An et al., 2025). En este esquema, un grupo de imágenes consecutivas se utiliza como entrada del modelo, mientras que la imagen siguiente corresponde a la salida deseada.

Las imágenes fueron procesadas con el objetivo de facilitar su uso de entrada en el modelo. En primer lugar, se realizó un redimensionamiento desde la resolución original de 764×764 píxeles a 224×224 píxeles. el uso de resoluciones reducidas es una práctica común de modelos de aprendizaje profundo a imágenes satelitales, debido al alto costo computacional asociado al procesamiento de secuencias espacio-temporales de alta dimensión.

Aunque esta reducción implica pérdida de detalle espacial fino, las principales estructuras de nubosidad y patrones globales permanecen visibles, permitiendo conservar información relevante asociada a la distribución espacial de la nubosidad y su evolución temporal.

Posteriormente, se hizo un proceso de normalización, con el objetivo de estandarizar los valores de entrada y facilitar el proceso de entrenamiento del modelo.

En este trabajo se utilizó una configuración de entrada de seis imágenes consecutivas ($x_{t-5}, x_{t-4}, x_{t-3}, x_{t-2}, x_{t-1}, x_t$) con el objetivo predecir la imagen correspondiente al siguiente instante temporal $\hat{x}_{t+1}$. De esta manera, el dataset final puede representarse como un conjunto de tensores de dimensión $(N \cdot T \cdot H \cdot W)$, donde N corresponde al número de muestras, T al número de pasos temporales de entrada, y H y W representan las dimensiones espaciales de cada imagen. Esta estructura permite preservar tanto la información temporal como espacial de la nubosidad, facilitando su procesamiento mediante redes neuronales del tipo ConvLSTM.

Posteriormente, el conjunto de datos fue dividido en subconjuntos de entrenamiento, validación y prueba, respetando el orden temporal de las secuencias. El 75% de las muestras iniciales se utilizó para el entrenamiento del modelo, el 15 % siguiente al proceso de validación, y el 10% restante correspondiente, se destinó al proceso de prueba. Con este

enfoque se evita la mezcla de información futura durante el entrenamiento, lo cual es relevante en problemas de predicción, donde el objetivo es estimar estados futuros a partir de información pasada.

Por último, durante el proceso de entrenamiento, las muestras del conjunto de entrenamiento fueron organizadas por lotes (batches) y presentadas al modelo de forma aleatorio. Esto se hace para mejorar la capacidad de generalización del modelo, al mismo tiempo se mantiene la estructura temporal definida durante la construcción del dataset.

4. MODELO CONVLSTM

El modelo propuesto se basa en una red neuronal del tipo ConvLSTM, propuesta originalmente por (Shi et al., 2015). Este tipo de red aprende a generar la imagen correspondiente al siguiente instante temporal, mediante operaciones de convolución y las dependencias temporales a través de estructuras recurrentes.

Mientras que las redes LSTM tradicionales, donde las operaciones internas se realizan mediante productos matriciales, en las redes ConvLSTM se integran capas convolucionales, que permite preservar la estructura espacial de los datos de entrada (Huang et al., 2023b). Esta característica resulta adecuada en aplicaciones donde la información está representada en forma de imágenes, como en el caso de la nubosidad.

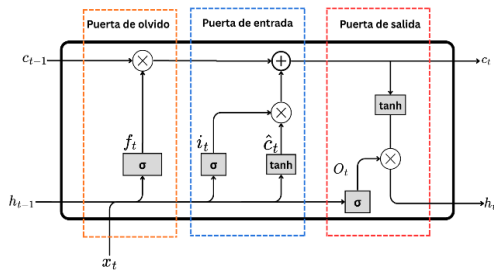


Figura 3. Estructura interna de una celda ConvLSTM.

En la Fig. 3 se muestra la arquitectura interna de una celda ConvLSTM, en la cual se puede observar las principales compuertas que regulan el flujo de información a lo largo del tiempo. Esta arquitectura integra la gestión de dependencias temporales a largo plazo, de las redes LSTM para modelar dependencias temporales con operaciones de convolución que actúan directamente sobre los tensores de entrada. Esto asegura la conservación de la jerarquía espacial mientras la red procesa la evolución cronológica de la serie

La arquitectura de la celda está compuesta por tres compuertas principales: la compuerta de olvido, la compuerta de entrada y la compuerta de salida. La compuerta de olvido controla que información del estado anterior se descarta, la compuerta de entrada gestiona la incorporación de nuevos datos al estado de la celda, mientras que la compuerta de salida filtra la información procesada para generar la predicción correspondiente al instante en curso.

El estado de la celda se actualiza mediante la combinación de la información previa y la nueva información procesada, lo que permite mantener memoria a lo largo de la secuencia. En el caso de la arquitectura ConvLSTM, este proceso se realiza mediante operadores convolucionales, lo que le permite al modelo capturar simultáneamente la jerarquía de patrones espaciales y las dependencias temporales en los datos de entrada.

El funcionamiento interno de la celda ConvLSTM queda definida por el sistema de ecuaciones (1)–(6), las cuales modelan la evolución del estado interno a lo largo del tiempo. En cada paso temporal t , la celda recibe como entrada la información actual X_t , junto con el estado oculto H_{t-1} y el estado de memoria C_{t-1} , asegurando una transición coherente de la información a través de la secuencia.

$$i_t = \sigma(W_{xi} * X_t + W_{hi} * H_{t-1} + W_{ci} \odot C_{t-1} + b_i) \quad (1)$$

$$f_t = \sigma(W_{xf} * X_t + W_{hf} * H_{t-1} + W_{cf} \odot C_{t-1} + b_f) \quad (2)$$

$$\tilde{C}_t = \tanh(W_{xc} * X_t + W_{hc} * H_{t-1} + b_c) \quad (3)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \quad (4)$$

$$o_t = \sigma(W_{xo} * X_t + W_{ho} * H_{t-1} + W_{co} \odot C_t + b_o) \quad (5)$$

$$H_t = o_t \odot \tanh(C_t) \quad (6)$$

En estas ecuaciones, los términos W representan los núcleos o filtros convolucionales que la red optimiza durante el proceso de entrenamiento. Dichos filtros permiten la extracción características relevantes tanto de la entrada actual X_t como del estado oculto previo H_{t-1} , capturando patrones espaciales en las imágenes.

Los términos W_{x*} están asociados a la información de entrada, mientras que W_{h*} corresponden a la información derivada del estado oculto anterior. Por otro lado, los términos W_{c*} regulan la interacción con el estado de memoria de la celda, permitiendo ajustar la influencia de la información previamente almacenada.

El operador $*$ representa la convolución entre los filtros y los tensores de entrada, este mecanismo permite conservar la estructura espacial de la imagen durante el procesamiento. En cambio, el operador $\odot$ representa el producto de Hadamard (elemento a elemento), mediante el cual las compuertas ejecutan una modulación puntual de la información que se conserva, se incorpora o se transmite en cada paso temporal.

Las funciones de activación también cumplen un papel importante en la dinámica de la celda. La función sigmoide σ restringe sus valores al intervalo $[0, 1]$, por lo que actúa como un mecanismo de compuerta al determinar qué proporción de la información debe pasar. Por otro lado, la función tangente hiperbólica $\tanh$ genera valores en el intervalo $[-1, 1]$. Esta combinación asegura que el modelo pueda capturar tanto la



presencia como la usencia de patrones atmosféricos, manteniendo el equilibrio numérico durante el entrenamiento.

A partir de estas operaciones, la celda calcula las compuertas de entrada, olvido y salida, las cuales regulan el flujo de información. La compuerta de olvido determina qué información previa se conserva, la compuerta de entrada selecciona qué nueva información se incorpora al estado de memoria, y la compuerta de salida controla la generación del estado oculto. De esta forma, los filtros convolucionales permiten aprender representaciones espaciales relevantes, mientras que la estructura recurrente modela la evolución temporal de la información, lo cual resulta fundamental para capturar patrones dinámicos en secuencias de imágenes de nubosidad (Shi et al., 2015).

4.1 Arquitectura del modelo

Se implementó una arquitectura ConvLSTM compuesta por dos capas recurrentes convolucionales apiladas, seguidas de un bloque decodificador encargado de generar la imagen de salida como se muestra en la Fig.4.

Como parte de la aportación de este trabajo, se definió una configuración específica de la arquitectura orientada al equilibrio entre capacidad de representación y eficiencia computacional, considerando las características del dataset y la naturaleza del problema.

La primera capa ConvLSTM recibe como entrada imágenes de un solo canal y genera una representación interna con 32 mapas de características, utilizando un kernel convolucional de tamaño 3×3 . Esta capa se encarga de extraer características espaciales relevantes de cada imagen, al mismo tiempo que incorpora información temporal proveniente de estados anteriores.

La segunda capa ConvLSTM toma como entrada la salida de la primera capa y produce una representación de menor dimensión con 16 mapas de características, permitiendo una compresión progresiva de la información espacio-temporal. Esta reducción en la dimensionalidad contribuye a obtener una representación más abstracta y compacta de la dinámica de la nubosidad, facilitando la eliminación del ruido e información redundante presente en las secuencias de imágenes originales. Al proyectar los datos hacia un espacio de menor dimensión, el modelo se ve obligado a conservar únicamente las características más importantes, lo cual es determinante para la fidelidad de la predicción temporal.

La red opera de forma iterativa, actualizando los estados internos en cada paso temporal. Al terminar el ciclo de entrada, el último estado oculto de la segunda capa ConvLSTM se extrae como la representación final; este actúa como un vector de contexto (embeddings) que resume la información acumulada y las dependencias de largo plazo de la secuencia completa.

La etapa de decodificación procesa el vector de contexto mediante capas convolucionales, cuya función es proyectar los mapas de características hacia el dominio de la imagen de salida. El decodificador integra una capa convolucional de 16 filtros con un kernel de 3×3 , apoyada por una función ReLU

para introducir no linealidad en el proceso de reconstrucción. Finalmente, una capa de convolución de 1×1 reduciendo la dimensionalidad a un solo canal, generando la predicción de la final de la cobertura nubosa.

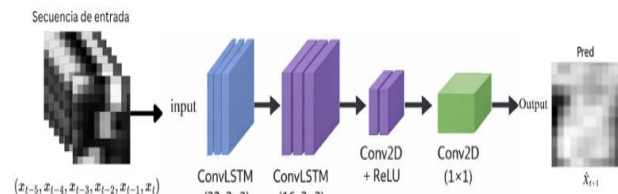


Figura 4. Arquitectura del Modelo ConvLSTM propuesto para la predicción de imágenes de nubosidad.

Para la generación de la salida final (output), se emplea una función de activación sigmoide para normalizar los valores de salida en el rango $[0, 1]$. Esta operación permite interpretar la respuesta del modelo como un mapa de intensidad de nubosidad, facilitando la comparación directa con las observaciones satelitales originales.

En conjunto la integración de capas ConvLSTM con el bloque decodificador permite extraer y proyectar eficazmente las dependencias de la serie temporal, demostrando como una herramienta eficaz para la predicción de fenómenos dinámicos y modelación de la evolución atmosférica.

En la fig. 4 se muestra la arquitectura del modelo propuesto. En ella se detalla la configuración de hiperparámetros, incluyendo el número de filtros y el tamaño de los kernels de convolución en cada etapa. El diagrama muestra la progresión desde el volumen de datos de entrada hasta la generación de imagen de salida, destacando la función del decodificador en la obtención de la predicción final.

4.2 Entrenamiento

El modelo fue implementado en PyTorch y entrenado utilizando aceleración por GPU cuando estuvo disponible. Para la optimización de los parámetros se empleó el algoritmo Adam con una tasa de aprendizaje de 1×10^{-3} . Como función de pérdida se utilizó el error cuadrático medio (Mean Squared Error, MSE), con el objetivo de minimizar la diferencia entre la imagen predicha y la imagen real.

El proceso de entrenamiento se llevó a cabo durante 30 épocas, utilizando lotes de datos obtenidos mediante el conjunto train_loader. Además, se empleó entrenamiento de precisión mixta automática (automatic mixed precision) mediante autocast y GradScaler, con el fin de optimizar el uso de memoria y acelerar el entrenamiento en GPU.

Durante cada iteración, las secuencias de entrada X y las imágenes objetivo Y fueron transferidas al dispositivo de cómputo correspondiente. Posteriormente, se generó la predicción $\hat{X}_{t+1}$ a través del modelo y se calculó la función de pérdida definida como:

$$\mathcal{L} = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W (X_{t+1}(i,j) - (\hat{X}_{t+1}(i,j)))^2 \quad (7)$$

donde $X_{t+1}(i,j)$ representa el valor real del píxel en la posición (i,j) de la imagen objetivo, y $\hat{X}_{t+1}(i,j)$ la predicción generada por el modelo (Goodfellow, 2016). Esta formulación permite evaluar la diferencia entre la imagen predicha y la real de manera detallada a nivel espacial. Esta función de pérdida es adecuada para tareas de predicción de imágenes, ya que penaliza directamente las diferencias de intensidad entre píxeles.

5. RESULTADOS

Se analizaron los resultados obtenidos durante el proceso de entrenamiento, así como la capacidad del modelo para predecir imágenes de nubosidad.

5.1 Función de pérdida

En la Fig. 5 se muestra la evolución de la función de pérdida (MSE) durante el entrenamiento del modelo. Se observa una disminución progresiva del error a lo largo de las épocas, lo que indica que el modelo logra aprender patrones relevantes en la secuencia de datos y converge de manera adecuada.

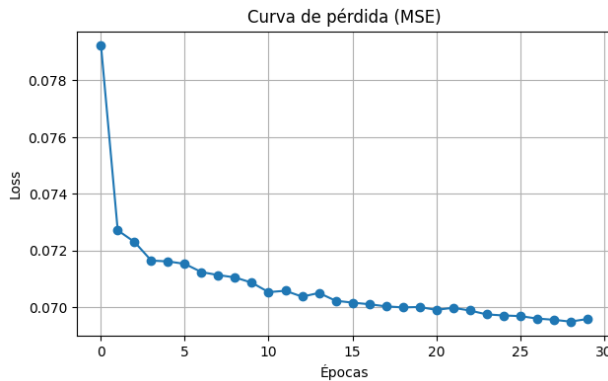


Figura 5. Gráfica de la función de pérdida Modelo ConvLSTM.

Como se observa en la Fig. 5, la función de pérdida presenta una disminución significativa en las primeras épocas, pasando de un valor cercano a 0.079 a aproximadamente 0.072. Este comportamiento indica que el modelo aprende rápidamente patrones relevantes en las etapas iniciales del entrenamiento.

Posteriormente, la curva muestra una disminución más gradual, estabilizándose alrededor de un valor cercano a 0.070. Esta tendencia sugiere que el modelo alcanza un proceso de convergencia, en el cual los ajustes en los parámetros se vuelven más pequeños conforme se aproxima a una solución óptima.

Además, la curva no presenta oscilaciones abruptas ni incrementos significativos en la pérdida, lo que indica estabilidad en el proceso de entrenamiento y una adecuada

elección de los hiperparámetros, como la tasa de aprendizaje y el optimizador.

En conjunto, estos resultados muestran que el modelo ConvLSTM logra aprender representaciones espacio-temporales de la nubosidad, reduciendo progresivamente el error de predicción durante el entrenamiento.

5.2 Desempeño del modelo

Para evaluar el desempeño del modelo, se calcularon métricas de error entre las imágenes reales y las imágenes predichas. En particular, se emplearon el error cuadrático medio (MSE), previamente definido como función de pérdida durante el entrenamiento, así como el Mean Absolute Error (MAE) y el índice de Structural Similarity Index Measure (SSIM), ampliamente utilizados en tareas de procesamiento de imágenes (Wang et al., 2004).

Es importante señalar que, aunque el Mean Squared Error MSE, fue utilizado como función de pérdida durante el entrenamiento del modelo, en esta sección se emplea como métrica de evaluación con el fin de cuantificar el desempeño final sobre los datos analizados.

El Error Absoluto medio (MAE) se define como (Goodfellow, 2016):

$$MAE = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W |X_{t+1}(i,j) - \hat{X}_{t+1}(i,j)| \quad (8)$$

Por su parte, el índice de SSIM evalúa la similitud entre dos imágenes considerando aspectos de luminancia, contraste y estructura, y se define como (Wang et al., 2004):

$$SSIM(X, \hat{X}) = \frac{(2\mu_X\mu_{\hat{X}} + C_1)(2\sigma_{X\hat{X}} + C_2)}{(\mu_X^2 + \mu_{\hat{X}}^2 + C_1)(\sigma_X^2 + \sigma_{\hat{X}}^2 + C_2)} \quad (9)$$

donde μ_X y $\mu_{\hat{X}}$ representan la media de las imágenes real y predicha, respectivamente; σ_X^2 y $\sigma_{\hat{X}}^2$ corresponden a sus varianzas, las cuales miden el contraste; y $\sigma_{X\hat{X}}$ representa la covarianza entre ambas imágenes, asociada a la similitud estructural (Wang et al., 2004).

Las constantes C_1 y C_2 son términos de estabilización que evitan divisiones por valores cercanos a cero. En general, el índice SSIM toma valores en el rango $[-1,1]$, donde valores cercanos a 1 indican una alta similitud entre las imágenes (Wang et al., 2004).

5.3 Métricas

Como modelo de referencia se utilizó un enfoque de persistencia, donde la predicción del siguiente instante temporal $t+1$ corresponde directamente a la última imagen observada X_t . Este enfoque es utilizado como punto de referencia en problemas de predicción temporal, ya que asume continuidad entre estados consecutivos. La Tabla 1 muestra la comparación entre el modelo de persistencia y la arquitectura CovLSTM propuesta.

Tabla 1. Comparación de las métricas de evaluación del modelo ConvLSTM y el modelo de persistencia.

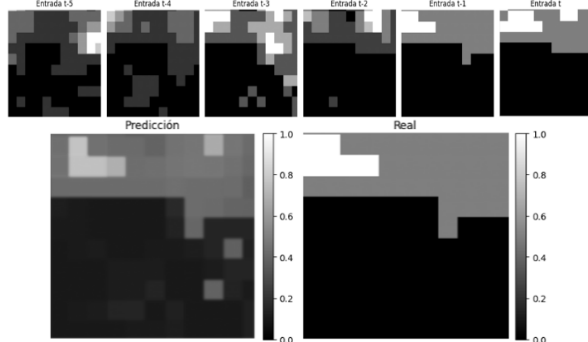
Modelo	MSE	MAE	SSIM
Persistencia	0.0773	0.1368	0.6758
ConvLSTM	0.0030	0.0259	0.9366

Los resultados indican que el modelo ConvLSTM supera al punto de referencia en todas las métricas. La reducción del error y el aumento del SSIM (0.9366) validan que el modelo propuesto aprendió los patrones complejos de la nubosidad, superando la tendencia de copiar la imagen previa para realizar la predicción.

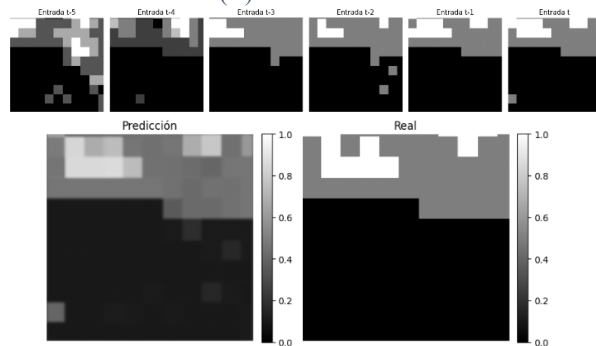
En la Fig. 6 se presentan diferentes escenarios representativos de predicción nubosidad utilizando distintas secuencias temporales de entrada. En cada caso se muestra la secuencia por el modelo ConvLSTM, así como la comparación entre la imagen predicha y la imagen real correspondiente al siguiente instante temporal.

Figura 6. Escenarios representativos de predicción de nubosidad.

(A). Escenario 1



(B). Escenario 2



En el primer escenario (A) se observa una distribución parcial de nubosidad concentrada principalmente en la región superior de la imagen, mientras que el segundo escenario (B) presenta

una cobertura nubosa más extensa y una variación más marcada.

En ambos casos, se observa que el modelo es capaz de reproducir la distribución espacial general de la nubosidad y conservar la coherencia con las principales regiones de intensidad presentes en las imágenes objetivo. Además, las predicciones mantienen consistencia espacial respecto a las imágenes reales, lo que indica que el modelo ha aprendido patrones espacio-temporales relevantes asociados a la evolución de la nubosidad.

Sim embargo, se aprecian variaciones ligeras en la precisión de los bordes y en los niveles de intensidad. Este resultado es esperado al emplear funciones de pérdida basadas en MSE, las cuales tienden a generar imágenes ligeramente difuminadas.

6. CONCLUSIONES

En este trabajo se desarrolló un modelo basado en redes neuronales ConvLSTM para la predicción de imágenes de nubosidad a partir de secuencias temporales de datos satelitales. Este enfoque responde a la necesidad reducir la incertidumbre en la generación de energía fotovoltaica, cuya eficiencia se ve directamente afectada por la variabilidad de la irradiación solar causada principalmente por la presencia de nubosidad.

La arquitectura propuesta permitió modelar de manera conjunta las dependencias espaciales y temporales de las formaciones nubosas, los resultados obtenidos muestran que el modelo es capaz de capturar la evolución de la nubosidad, reproduciendo la estructura global de las imágenes y la localización de las principales regiones de mayor intensidad. Este comportamiento se refleja tanto en el análisis visual como en las métricas cuantitativas empleadas.

En términos cuantitativos, el modelo alcanzó valores bajos de error y un elevado índice de similitud estructural (SSIM), indicando una buena correspondencia entre las imágenes reales. También, la curva de aprendizaje mostró una convergencia estable, esto confirma que la estructura del modelo es capaz de representar correctamente el movimiento y cambio de nubosidad.

Los resultados obtenidos muestran que las redes de aprendizaje profundo representan una alternativa efectiva frente a métodos tradicionales, al permitir capturar patrones no lineales asociados a la evolución atmosférica. Para el sector fotovoltaico, esta capacidad de predicción a corto plazo resulta relevante para anticipar variaciones en la irradiación solar y contribuir a una mejor gestión de la energía generada.

Entre las principales limitaciones del modelo se encuentra el horizonte de predicción de un solo paso temporal ($t + 1$), equivalente a una hora. Aunque este horizonte resulta útil en aplicaciones de nowcasting atmosférico y predicción fotovoltaica de corto plazo, aún no permite evaluar el desempeño del modelo en horizontes temporales mayores.

Como trabajo futuro, se plantea la exploración de arquitecturas más avanzadas, incluyendo modelos basados en Transformers, así como la integración de variables físicas adicionales, tales



como presión, humedad y velocidad del viento, con el objetivo de mejorar la precisión del modelo. Asimismo, se espera extender este enfoque hacia predicciones de múltiples pasos con horizontes de al menos cuatro horas, fortaleciendo su aplicación en sistemas fotovoltaicos y gestión energética.

Adicionalmente, se considera relevante poner a prueba el desempeño del modelo bajo diferentes condiciones climáticas y escenarios temporales, tales como distintas estaciones del año o periodos con alta presencia de nubosidad.

REFERENCIAS

- An, S., Oh, T.-J., Sohn, E., & Kim, D. (2025). Deep learning for precipitation nowcasting: A survey from the perspective of time series forecasting. *Expert Systems with Applications*, 268, 126301. <https://doi.org/10.1016/j.eswa.2024.126301>
- Antonanzas, J., Osorio, N., Escobar, R., Urraca, R., Martínez-de-Pison, F. J., & Antonanzas-Torres, F. (2016). Review of photovoltaic power forecasting. *Solar Energy*, 136, 78–111. <https://doi.org/10.1016/j.solener.2016.06.069>
- Ayzel, G., Scheffer, T., & Heistermann, M. (2020). RainNet v1.0: A convolutional neural network for radar-based precipitation nowcasting. *Geoscientific Model Development*, 13(6), 2631–2644. <https://doi.org/10.5194/gmd-13-2631-2020>
- Goodfellow, I., B. Y., & C. A. (2016). *Deep Learning*. MIT Press. MIT Press.
- Huang, X., Liu, J., Xu, S., Li, C., Li, Q., & Tai, Y. (2023). A 3D ConvLSTM-CNN network based on multi-channel color extraction for ultra-short-term solar irradiance forecasting. *Energy*, 272, 127140. <https://doi.org/10.1016/j.energy.2023.127140>
- Inman, R. H., Pedro, H. T. C., & Coimbra, C. F. M. (2013). Solar forecasting methods for renewable energy integration. In *Progress in Energy and Combustion Science* (Vol. 39, Number 6, pp. 535–576). <https://doi.org/10.1016/j.pecs.2013.06.002>
- Japan Aerospace Exploration Agency (JAXA). (2024). *Guidelines for using the Global Satellite Mapping of Precipitation (GSMaP)*. <https://sharaku.eorc.jaxa.jp/GSMaP/>
- Shi, X., Chen, Z., Wang, H., Yeung, D.-Y., Wong, W., & Woo, W. (2015). *Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting*. <http://arxiv.org/abs/1506.04214>
- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. In *IEEE TRANSACTIONS ON IMAGE PROCESSING* (Vol. 13, Number 4). <http://www.cns.nyu.edu/~lcv/ssim/>.

Energy Management Strategies in Water Pumping Systems Supplied by Renewable Energy: A Review

Marifer Jiménez Luna*, Jesús Darío Mina Antonio*

*Tecnológico Nacional de México / Centro Nacional de Investigación y Desarrollo Tecnológico, Cuernavaca, Morelos, C.P. 62490 México; (e-mail: m24ce121@cenidet.tecnm.mx, jesus.ma@cenidet.tecnm.mx)

Abstract: This paper reviews the main energy management strategies (EMS) used in renewable energy-based water pumping systems. It examines different system configurations, including photovoltaic, hybrid, and grid-connected setups, and categorizes EMS approaches into rule-based methods, artificial intelligence techniques, optimization methods, and advanced control strategies. Key variables, operational constraints, and decision-making processes are also discussed. In addition, the performance of these strategies is assessed from a techno-economic perspective, highlighting their advantages, limitations, and current trends. The findings emphasize the importance of selecting an EMS that balances efficiency, cost, and complexity to enhance system reliability and sustainability.

Keywords: Energy management systems (EMS), control strategies, water pumping systems, renewable energy sources, hybrid systems, optimization

1. INTRODUCCIÓN

En los últimos años, el uso de sistemas de bombeo de agua impulsados por energías renovables ha cobrado una relevancia creciente ante la necesidad de implementar soluciones sostenibles y eficientes para el abastecimiento de agua en sectores como el agrícola, doméstico e industrial (Sallem et al., 2009). Este tipo de sistemas permite reducir la dependencia de fuentes energéticas convencionales, lo que se traduce en menores costos operativos y una reducción del impacto ambiental asociado al uso de combustibles fósiles. Dentro de las fuentes renovables más utilizadas, los sistemas fotovoltaicos destacan por su amplia disponibilidad y facilidad de integración (Amrouche et al., 2024). No obstante, tanto la energía solar como otras fuentes renovables, como la eólica, presentan un carácter intermitente que introduce desafíos importantes para garantizar una operación continua y estable.

En este contexto, los sistemas de gestión de energía (EMS, por sus siglas en inglés) cumplen un papel fundamental al coordinar la generación, el almacenamiento y el consumo de energía (Rezk et al., 2020). Un EMS define cómo se distribuye la energía entre los distintos componentes del sistema, con el objetivo de asegurar un funcionamiento eficiente, confiable y seguro, además de proteger elementos críticos como los sistemas de almacenamiento (Ben Safia et al., 2023). En la literatura se han propuesto diversas estrategias de EMS, entre las que se incluyen métodos basados en reglas, técnicas de inteligencia artificial, enfoques de optimización y estrategias de control avanzado (Khiareddine et al., 2014). Estas metodologías presentan diferencias significativas en cuanto a su complejidad, capacidad de adaptación y desempeño frente a condiciones operativas variables (Mahinda et al., 2023).

El presente artículo tiene como objetivo realizar una revisión de las principales estrategias de gestión de energía aplicadas a sistemas de bombeo de agua con fuentes renovables,

analizando su clasificación, principios de operación y desempeño técnico-económico.

2. SISTEMAS DE ALIMENTACIÓN PARA BOMBEO DE AGUA

Los sistemas de bombeo de agua basados en energías renovables pueden clasificarse en función de la configuración de sus fuentes de energía y su forma de integración con el sistema eléctrico. Esta clasificación permite identificar las características operativas de cada sistema, así como los requerimientos asociados a la gestión de energía.

2.1 Sistemas fotovoltaicos

Los sistemas fotovoltaicos (PVS, por sus siglas en inglés) utilizan la energía solar como única fuente de alimentación para el bombeo de agua, aprovechando la conversión de energía solar en electricidad mediante el efecto fotovoltaico (Joshua et al., 2024). En la Figura 1 se muestra cómo están compuestos estos sistemas: paneles solares, un sistema de conversión de energía (inversor o controlador) y la bomba de agua.

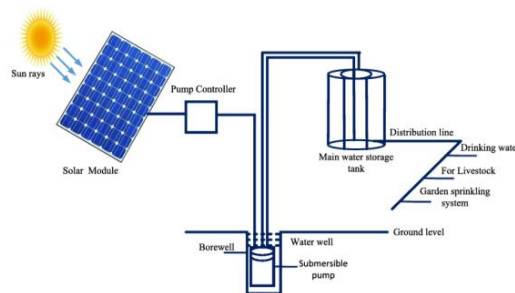


Figure 1. Diagrama de Sistema fotovoltaico (Poompavai and Kowsalya, 2019).

En la literatura, estos sistemas son ampliamente utilizados en aplicaciones donde se busca una solución autónoma con bajo mantenimiento, aunque presentan limitaciones en términos de continuidad del suministro energético, ya que la generación depende directamente de la disponibilidad de radiación solar, lo que introduce variabilidad en la operación del sistema.

2.2 Sistemas híbridos

Los sistemas híbridos combinan múltiples fuentes de energía, tales como sistemas fotovoltaicos, aerogeneradores, baterías y, en algunos casos, generadores diésel (Gueye et al., 2024). Esta configuración permite mejorar la confiabilidad del sistema y reducir la dependencia de una única fuente energética. Una de las principales ventajas de estos sistemas es su capacidad para garantizar un suministro energético más estable, lo que los hace adecuados para aplicaciones con mayor demanda o requerimientos de operación continua.

En la literatura se han propuesto configuraciones como en la Figura 2 que muestra una combinación PV-batería; y otros autores proponen combinaciones como: PV-eólico-batería (Amrouche et al., 2024) y PV-diésel-batería (Rezk et al., 2020) para bombeo de agua. No obstante, su implementación implica una mayor complejidad en el diseño y dimensionamiento del sistema, así como un incremento en los costos de inversión inicial.

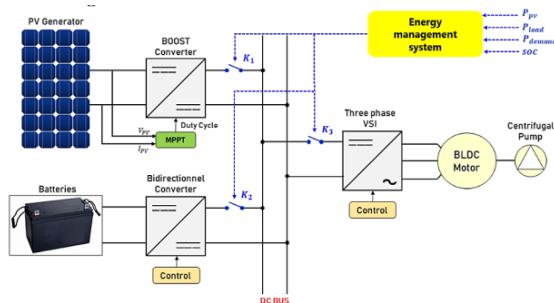


Figure 2. Diagrama de Sistema Híbrido (Benzaouia et al., 2021).

2.3 Sistemas conectados a la red

Los sistemas conectados a la red integran fuentes renovables con el sistema eléctrico convencional como se muestra en la Figura 3, permitiendo intercambiar energía con la red según las condiciones de generación y demanda (Mahinda et al., 2023). En este tipo de sistemas, la red actúa como respaldo en condiciones de déficit energético, lo que mejora significativamente la confiabilidad y facilita la operación del sistema.

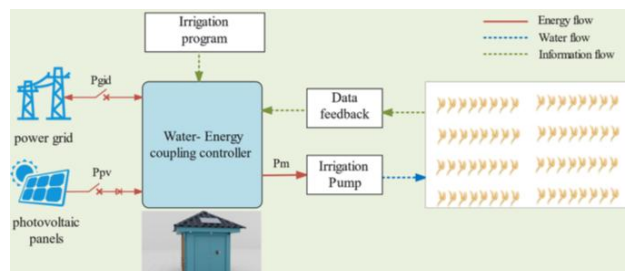


Figure 3. Diagrama de Sistema Conectado a la Red (Xu and Chen, 2024).

La elección de la configuración del sistema depende de factores como la disponibilidad de recursos energéticos, los requerimientos de la carga y las condiciones de operación, aunque su implementación también depende de la infraestructura eléctrica disponible y puede implicar costos asociados por consumo de red.

3. ESTRATEGIAS EMS

Las estrategias de gestión de energía constituyen el núcleo operativo de los sistemas de bombeo de agua basados en energías renovables, ya que deciden cómo se distribuye la energía entre las fuentes disponibles, el almacenamiento y el sistema de bombeo. Estas estrategias permiten mejorar la eficiencia del sistema, garantizar la continuidad del suministro y optimizar el uso de los recursos energéticos (Poompavai and Kowsalya, 2019).

De acuerdo a la revisión de la literatura las estrategias EMS se pueden clasificar en cuatro grupos: métodos basados en reglas, técnicas de inteligencia artificial, enfoques de optimización y estrategias de control avanzado. Cada una tiene sus propias características en cuanto a complejidad, rendimiento y casos de aplicación. A continuación, se describe las variables principales, las restricciones operativas y las decisiones que generalmente se toman en cuenta en cualquier estrategia de EMS; posteriormente se discute brevemente cada uno de estos grupos de estrategias EMS.

3.1 Variables, Restricciones y Toma de Decisiones en esquemas EMS

En la literatura, se observa que la mayoría de los esquemas de EMS se fundamentan en tres elementos clave: la evaluación de las variables del sistema (fuentes de alimentación, almacenamiento de energía, bomba de agua), la consideración de restricciones operativas y la ejecución de decisiones que definen en cada instante, el flujo de energía o la estrategia de control más adecuada.

3.1.1 Variables principales

Las variables utilizadas en los EMS reflejan el estado energético del sistema y permiten definir las condiciones de operación. Entre las más empleadas se encuentran:

- Estado de carga de la batería (SOC): Es una variable crítica en la mayoría de los esquemas de EMS, ya que indica la cantidad de energía disponible en el almacenamiento o la cantidad de energía que puede recibir para almacenamiento.
- Potencia generada (P_{Gen}): Incluye la energía de fuentes renovables como sistemas fotovoltaicos (P_{PV}) y eólicos (P_{WT}), la cual depende de condiciones ambientales variables.
- Demanda del sistema (P_L): Representa la potencia requerida por el sistema de bombeo, que puede variar según las condiciones de operación.
- Potencia de red (P_{Grid}): Energía suministrada por la red eléctrica para satisfacer la demanda del sistema de bombeo cuando la energía de las fuentes renovables o el



sistema de almacenamiento no es suficiente. O energía inyectada a la red cuando se tiene excedente de generación.

- Variables adicionales: En algunos estudios se consideran condiciones climáticas, temperatura, costos energéticos o nivel de agua, especialmente en sistemas más avanzados.

3.1.2 Restricciones del sistema

Las restricciones definen los límites operativos bajo los cuales el sistema debe funcionar, garantizando la seguridad y prolongando la vida útil de los componentes. Entre las principales restricciones se encuentran:

- Límites del SOC (SOC_{min} y SOC_{max}): Evitan la sobrecarga y la descarga profunda de la batería, protegiendo así su vida útil.
- Balance de potencia: La energía generada, almacenada y consumida debe mantenerse en equilibrio en todo momento.
- Capacidad de los componentes: Incluye límites en la potencia de generadores, convertidores y bombas.
- Restricciones operativas: Condiciones mínimas necesarias para el funcionamiento del sistema, como niveles de energía o demanda.

3.1.3 Toma de decisiones en el EMS

La toma de decisiones en un EMS consiste en definir cómo se distribuye la energía disponible en función del estado energético del sistema. Las acciones más comunes incluyen:

- Carga de la batería cuando hay exceso de generación.
- Descarga de la batería en condiciones de déficit energético en las fuentes renovables.
- Suministro directo a la carga desde las fuentes renovables, cuando la generación es suficiente.
- Desconexión o limitación de la carga en condiciones críticas.
- Activación de fuentes auxiliares (red eléctrica o generadores e.g. diésel).

Este tipo de acciones son las que determina el EMS, de manera general, independientemente del tipo de estrategia que se use (basado en reglas, algoritmos o modelos matemáticos). Desde luego, el EMS también debe determinar la duración de estas acciones y la cantidad de energía que se maneje para cada acción.

3.2 Métodos basados en reglas

Los métodos basados en reglas son ampliamente utilizados debido a su simplicidad de implementación y bajo requerimiento computacional. Estos enfoques se basan en condiciones lógicas predefinidas, generalmente relacionadas con el SOC, la potencia generada y la demanda del sistema de bombeo (Zaouche et al., 2016).

En este tipo de estrategias, la toma de decisiones se realiza mediante diagramas de flujo o estructuras condicionales que definen modos de operación en el sistema de bombeo. Un ejemplo de este tipo de métodos basado en reglas (diagramas de flujo) se muestra en la Figura 4.

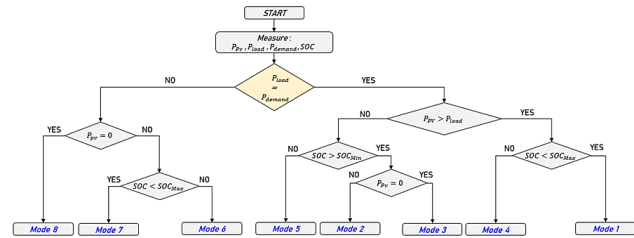


Figure 4. Diagrama de flujo del EMS (Benzaouia et al., 2021).

En (Sallem et al., 2009) se presenta un sistema de bombeo aislado de la red, alimentado mediante un sistema híbrido PVS-Batería. El sistema emplea 3 relevadores para gestionar el flujo de energía entre subsistemas y establecer 4 modos de operación: $Bat \rightarrow Bomba$, $PVS + Bat \rightarrow Bomba$, $PVS \rightarrow Bomba + Bat$ y $PVS \rightarrow Bat$. La estrategia de EMS busca maximizar el tiempo de operación de la bomba, manteniendo el equilibrio entre la potencia generada y la potencia demandada, mediante la evaluación de las variables como el SOC, P_{PV} (potencia fotovoltaica) y P_p (potencia de bomba); para decidir qué relevadores activar o qué modo seleccionar.

En (Khiareddine et al., 2014) se presenta un sistema de bombeo para irrigación agrícola; el cual está acondicionado con 3 interruptores para controlar el flujo de energía y definir los modos de operación: $PVS \rightarrow Bomba$, $Bat \rightarrow Bomba$, $PVS \rightarrow Bat$. La estrategia de EMS busca mantener un nivel de referencia del tanque de agua, aprovechar la energía solar y reducir el uso de las baterías (para prolongar su vida útil). Además, se asegura el equilibrio entre generación y demanda; y se evalúan los valores de las variables de SOC, irradiancia solar (G) y el nivel del tanque (H), para decidir el estado de los interruptores.

El trabajo de (Benzaouia et al., 2021) presenta un sistema de bombeo aislado de la red para irrigación agrícola, similar al propuesto por (Sallem et al., 2009), alimentado mediante un sistema PVS-Batería. El sistema incorpora 3 interruptores para gestionar los 8 modos de operación: $PVS \rightarrow Bomba + Bat$, $PVS \rightarrow Bomba$, $Bat \rightarrow Bomba$, $PVS + Bat = Bomba$, ($P_{PV} = 0$, $P_{Bat} = 0$, $P_p \neq 0$), $PVS \rightarrow Bat$, Desconexión de Bat ($P_{PV} \neq 0$, $SOC = 1$, $P_p = 0$), Desconexión de Sistema ($P_{PV} = 0$, $P_p = 0$). La estrategia de EMS busca aprovechar la energía solar y usar las baterías solo como respaldo; mediante la evaluación del SOC (restringida entre 0.2-0.8), P_{PV} y P_p , para determinar el modo de operación.

En (Cacciato et al., 2016) se aborda un sistema de bombeo alimentado mediante PVS, aislado de la red, para abastecer una comunidad rural. El trabajo evalúa la capacidad anual de bombeo al incorporar un sistema de almacenamiento basado en baterías (BEES, por sus siglas en inglés) y una estrategia EMS, en comparación con una configuración sin ello. Además, se busca maximizar la eficiencia, limitando el uso del BESS y definiendo un rango de operación de la potencia de bomba P_p .

en valores medios y altos (mayor eficiencia del motor). En este contexto, el EMS plantea 4 modos de operación: $PVS \rightarrow Bat$, $PVS \rightarrow Bomba$, $PVS \rightarrow Bomba + Bat$, $PVS + Bat \rightarrow Bomba$; cuya activación depende del P_{PV} disponible, las cotas de potencia de P_p y del SOC .

El sistema de bombeo y esquema de EMS presentado en (Zaouche et al., 2016) son similares a los de (Sallem et al., 2009) y (Benzaouia et al., 2021). El sistema PVS-Batería-Bomba incorpora 3 interruptores para establecer 5 modos de operación: $PVS \rightarrow Bomba + Bat$, $PVS + Bat \rightarrow Bomba$, $Bat \rightarrow Bomba$, Desconexión de Bat ($P_{PV} \neq 0$, $SOC = 1$, $P_p = 0$), Desconexión de Bomba ($P_{PV} = 0$, $P_{Bat} = 0$). La estrategia de EMS busca maximizar el uso de la generación P_{PV} y mantener la operación continua de la bomba; considerando los valores de P_{PV} , SOC , P_p y el balance $P_{PV} + P_{Bat} = P_p$, para decidir entre los modos.

El trabajo de (Ben Safia et al., 2023) se enfoca en un sistema híbrido de bombeo en zona rural sin acceso a la red eléctrica, alimentado por PVS-Batería-Turbina Eólica (WT por sus siglas en inglés). El sistema cuenta con 5 modos de operación: $PVS + WT \rightarrow Bomba$ ($P_{Bat} = 0$), $PVS + WT \rightarrow Bomba + Bat$, $PVS + WT + Bat \rightarrow Bomba$, Disipación Resistiva $PVS + WT \rightarrow Bomba$ ($SOC = 1$), Desconexión ($P_{PV} = 0$, $P_{WT} = 0$, $P_{bat} = 0$). El EMS busca maximizar la transferencia de potencia; y usa variables como la irradiación solar, velocidad del viento, P_p , P_{PV} , P_{WT} y SOC , además del balance de potencia para decidir los modos de operación.

El trabajo de (Gueye et al., 2024) propone un sistema de bombeo híbrido para irrigación agrícola, alimentado con PVS-Generador Diésel (DG por sus siglas en inglés). La estrategia EMS se basa en la conmutación automatizada de las fuentes energéticas mediante un algoritmo por condiciones, el cual prioriza el uso de la generación renovable y reduce el consumo de diésel. El sistema considera 3 modos de operación: $DG \rightarrow Bomba$ ($G < 300 \frac{W}{m^2}$), $PVS \rightarrow Bomba$ ($G > 300 \frac{W}{m^2}$) y Desconexión ($H = H_{ref}$). Para la toma de decisiones se evalúan variables como la irradiancia solar, P_{PV} , P_{DG} , H y el nivel de combustible diésel.

Como se puede ver, los métodos basados en reglas son sencillos y flexibles; dependiendo de la cantidad y tipo de fuentes de alimentación se establecen múltiples modos de operación para definir el flujo de energía. Sin embargo, requieren la medición de muchas variables y suelen no incluir funciones de costo ni resolver problemas de optimización.

3.3 Métodos basados en inteligencia artificial

Las técnicas de inteligencia artificial, en el caso de sistemas de bombeo han sido ampliamente exploradas para mejorar el desempeño de los sistemas EMS, especialmente en entornos con alta incertidumbre. Estos métodos permiten modelar comportamientos no lineales y tomar decisiones más flexibles que los enfoques basados en reglas.

El trabajo de (Moyo, 2024) propone un sistema de bombeo para diversos usos, alimentado por un sistema híbrido PVS-Batería-DG. El EMS está basado en reglas difusas y considera 3 modos de operación: $PVS \rightarrow Bomba$, $PVS + Bat \rightarrow$

$Bomba$, $DG + PVS + Bat \rightarrow Bomba$ (uso/arranque de DG). El EMS se implementa en MATLAB/Simulink mediante reglas IF-THEN que relacionan variables de entrada (P_{PV} , P_{DG} , P_{Bat}) con funciones de pertenencia (bajo, medio, alto) y variables de salida, como el estado y velocidad del DG, así como el estado de carga/descarga de la batería. En la Figura 5, mediante Rule Viewer, se muestran las reglas activas del EMS. El objetivo es maximizar el uso de la generación renovable y minimizar el tiempo de operación del DG.



Figure 5. Visor de reglas basado en Lógica Difusa en MATLAB/Simulink (Moyo, 2024).

En (Khiareddine et al., 2015) se aborda un sistema de bombeo para una estación agrícola experimental, alimentado por PVS-Batería. El sistema híbrido cuenta con 4 modos de operación: $PVS \rightarrow Bomba$, $Bat \rightarrow Bomba$, $PVS \rightarrow Bat$ ($H = H_{ref}$, $SOC < SOC_{max}$), y Desconexión por: ($P_{PV} = 0$, $P_{Bat} = 0$), ($P_{PV} \neq 0$, $SOC < SOC_{max}$, $H = H_{ref}$), o ($P_{PV} \neq 0$, $SOC = 1$, $H = H_{ref}$). Para el EMS se comparan tres estrategias: control difuso (variables G y H), control neuro-difuso (reglas difusas con aprendizaje neuronal) y control algorítmico (relación matemática entre G y H). En las tres estrategias se utilizan como entradas: H , SOC y G ; y salidas: la velocidad de la bomba y el ciclo de trabajo, priorizando el uso de energía renovable y la protección de la batería.

Por su parte, (Ouachani et al., 2017) presenta un sistema de bombeo en áreas rurales sin acceso a la red eléctrica, alimentado por PVS-WT-Batería. El trabajo evalúa distintas estrategias EMS para garantizar el suministro continuo de agua, proteger las baterías y maximizar el uso de energías renovables. La primera estrategia, denominada Umbral de las Renovables (TOR, por sus siglas en inglés), busca satisfacer la demanda, mantener el SOC entre 0.2–0.8 y usar el excedente de energía para carga de baterías. La segunda, basada en lógica difusa Takagi-Sugeno, emplea variables como P_{PV} , P_{WT} , SOC y P_p , para decidir entre alimentar la bomba, cargar/descargar la batería o disipar excedentes. La tercera utiliza algoritmos genéticos para optimizar los modos de operación mediante procesos de selección, cruce y mutación. Los tres métodos permiten operar el sistema bajo variaciones de irradiancia y velocidad del viento.

En general, estos métodos son muy parecidos a los basados en reglas; sin embargo, estos últimos operan bajo un esquema simple de activación/desactivación, mientras que los métodos basados en inteligencia artificial, generalmente apoyados en lógica difusa, utilizan funciones de pertenencia o reglas difusas para definir las acciones de salida. Desde luego, este proceso requiere un diseño e implementación más complejos del EMS.

3.4 Métodos de optimización

Los métodos de optimización buscan determinar la mejor estrategia de operación del sistema de bombeo, minimizando o maximizando una función objetivo, como el costo total, el consumo de combustible o la degradación de la batería (Xu and Chen, 2024). Entre las técnicas más utilizadas se encuentran la programación matemática, los algoritmos evolutivos y la optimización no lineal.

Trabajos recientes como (Joshua et al., 2024), se centran en sistemas de bombeo para riego, ganadería y suministro en áreas remotas, alimentados por PVS. El EMS busca optimizar el suministro energético mediante una estrategia que combina el algoritmo de Optimización de Escarabajo de Namibia (NBO, por sus siglas en inglés), para generar parámetros de control, y Redes Residuales Profundas con Picos (SDRN, por sus siglas en inglés), para aprender de los datos generados por NBO y predecir valores óptimos de operación. En la búsqueda de solución se garantiza estabilidad y la maximización del uso de la energía solar bajo variaciones de irradiación solar (ej. $200W/m^2$ a $1000W/m^2$).

En (Mahinda et al., 2023), el sistema es un suministro de agua para riego agrícola y uso doméstico, alimentado por PVS-Red eléctrica. El EMS resuelve un problema de optimización con programación lineal entera no binaria “intlinprog” (MATLAB, solver MILP para optimización con restricciones y variables enteras). La función objetivo minimiza el costo total de operación, considerando el consumo de energía de la red y costos de mantenimiento. Las variables consideradas son $G, P_L, \$/kWh$; las restricciones son capacidades máximas/mínimas de los tanques, el balance de potencia y periodos de riego preestablecidos. Las variables de decisión son el apagado/encendido del sistema de bombeo y el uso o no de la red.

$$F = \min \sum_{i=1}^{SN} \{ \sigma \cdot P_{gid}(n_i) \cdot C_i^{sell} + (1 - \sigma) \cdot P_{gid}(n_i) \cdot C_i^{buy} \} \cdot \Delta T_i$$

Figure 6. Función objetivo de minimización del costo operativo diario del modelo de bombeo fotovoltaico (Xu and Chen, 2024).

Por otro lado, en (Xu and Chen, 2024) se presenta un sistema de bombeo para irrigación de alta precisión en agricultura, alimentado por PVS-Red eléctrica. El enfoque principal consiste en ajustar la potencia de la bomba para sincronizar el consumo eléctrico con la generación fotovoltaica, minimizando los costos operativos diarios sin requerir almacenamiento adicional. En la Figura 6 se muestra la función objetivo, donde el costo se calcula en función de la energía de red (P_{gid}) comprada o vendida (C_i^{buy}, C_i^{sell}) durante un periodo de tiempo. Además, se consideran restricciones asociadas a la capacidad de generación P_{PV} , y al balance entre generación y demanda.

En estos enfoques de EMS, basados en optimización, se requiere la definición explícita de una función objetivo, junto con restricciones y variables de decisión. En general, el problema de optimización suele ser *off line*, es decir, la solución es una programación *a priori* sobre la operación del sistema, a partir de datos de pronóstico como e.g. de irradiancia, P_{PV} , P_P , etc. En este sentido, el problema de optimización no solo suele ser complejo de resolverse, razón por la cual se hace uso de métodos metaheurísticos o bio-inspirados, sino que, además, requieren de mucha información y de mayor capacidad computacional, lo que puede limitar su implementación en aplicaciones prácticas.

3.5 Estrategias de control avanzado

Las estrategias de control avanzado, como el control predictivo y el control basado en pasividad, han surgido recientemente como alternativas para mejorar la eficiencia y estabilidad de los sistemas EMS. Estos métodos consideran la dinámica del sistema y permiten anticipar su comportamiento futuro, lo que resulta en una gestión de energía más eficiente. Sin embargo, su implementación es compleja y aún se encuentra limitada principalmente a estudios teóricos o simulaciones.

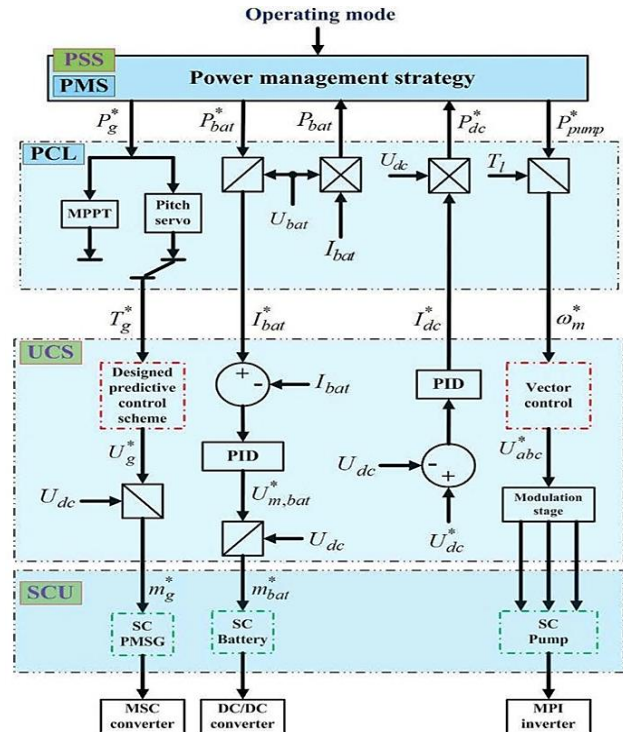


Figure 7. Función objetivo de minimización del costo operativo diario del modelo de bombeo (Mossa et al., 2022).

El trabajo de (Mossa et al., 2022), presenta un sistema de bombeo en zonas remotas alimentado por WT-Batería, orientado a maximizar la eficiencia mediante un esquema jerárquico y control predictivo. Como se muestra en la Figura 7, el EMS se compone de: una estrategia de reparto de potencia (PSS) como supervisor general; un sistema de gestión de potencia (PMS) que controla el flujo energético y la carga/descarga de la batería; un lazo de control de potencia (PCL) que regula la estabilidad del sistema; una etapa de



control de unidades (UCS) que coordina los subsistemas; y una unidad de gestión de conmutación (SMU) encargada de las señales de conmutación. Para la toma de decisiones se consideran variables como velocidad del viento, SOC , H , P_{WT} , P_p , límites de SOC , balance de potencia y voltaje en el bus DC.

En (Shchur et al., 2021) se estudia un sistema de bombeo con bomba centrífuga accionada por un motor DC sin escobillas (BLDC, por sus siglas en inglés), alimentado por un sistema PVS-Batería. Su enfoque se centra en integrar las funciones de los subsistemas mediante un control basado en pasividad (PBC, por sus siglas en inglés) permitiendo una mejor coordinación y estabilidad en el sistema bajo las condiciones variables de G , P_p , SOC , voltajes del bus interno y externo. A partir de estas variables se realiza la toma de decisiones según los modos de operación y las acciones de control (ciclos de trabajo) de los tres convertidores DC-DC.

Por su parte, (Kumar et al., 2024) aborda un sistema de bombeo con motores BLDC, para aplicaciones agrícolas, alimentado por PVS-Batería. El trabajo busca maximizar el uso de energía solar mediante un algoritmo MPPT tipo P&O, y controladores Proporcional-Integral (PI) para regular la carga/descarga de las baterías y el voltaje del bus DC. Con ello se garantiza la máxima extracción de P_{PV} y la estabilidad del bus para la operación continua del motor BLDC, además de controlar el flujo energético entre paneles y baterías. Las variables consideradas para la toma de decisiones son G , SOC , P_p , temperatura ambiente, límite de potencia de salida.

En este enfoque de EMS se emplean estructuras de control jerárquico, con distintos niveles y funciones específicas de control avanzado. Sin embargo, el diseño de estos controladores requiere un conocimiento detallado de la dinámica de los subsistemas, cuyos modelos pueden ser altamente complejos según el nivel jerárquico considerado.

En general, la elección de la estrategia EMS depende del equilibrio entre la complejidad y el desempeño que se busque. La Tabla 1 muestra que las estrategias EMS presentan diferencias importantes en términos de complejidad, eficiencia energética y aplicabilidad. Los métodos basados en reglas continúan siendo los más utilizados debido a su simplicidad de implementación y bajo costo computacional, especialmente en sistemas aislados y aplicaciones rurales. Sin embargo, su capacidad de adaptación frente a variaciones de irradiancia, demanda o condiciones climáticas es limitada, ya que las decisiones se basan en reglas predefinidas.

Por otro lado, las estrategias basadas en inteligencia artificial y optimización ofrecen un mejor aprovechamiento de la energía renovable y una gestión más eficiente del sistema, al considerar múltiples variables y criterios de decisión. No obstante, estos enfoques requieren modelos más complejos, mayor capacidad computacional y procesos de ajuste más elaborados, lo que puede dificultar su implementación práctica en sistemas de baja escala.

Finalmente, las estrategias de control avanzado presentan el mayor desempeño dinámico y estabilidad operativa, debido a su capacidad de anticipar el comportamiento futuro del sistema. A pesar de ello, la mayoría de los trabajos revisados

se encuentran todavía en etapas de simulación, por lo que su validación experimental y aplicación en escenarios reales continúa siendo un desafío importante.

Tabla 1. Comparación de estrategias de gestión de energía (EMS)

Estrategia EMS	Principio de operación	Variables principales	Ventajas Desventajas	Aplicación típica
Basado en reglas	Condiciones lógicas (if-else)	SOC , P_{gen} , P_L	- Simple, bajo costo - No óptimo, poca adaptabilidad	Sistemas básicos
Inteligencia artificial	Lógica difusa, redes neuronales	SOC , clima, generación	- Adaptativo, flexible - Diseño complejo	Sistemas híbridos
Optimización	Función objetivo	Costos, energía, demanda	- Óptimo, eficiente - Alta complejidad	Sistemas con red
Control avanzado	Modelos dinámicos	Variables del sistema	- Alta precisión - Difícil implementación	Investigación

5. EVALUACIÓN TECNOECONÓMICA

La evaluación tecnoeconómica de los sistemas de gestión de energía (EMS) es fundamental para determinar su viabilidad en aplicaciones de bombeo de agua con energías renovables, al relacionar aspectos técnicos, como eficiencia y confiabilidad, con indicadores económicos de operación, inversión y mantenimiento (Poompavai and Kowsalya, 2019).

5.1 Indicadores tecnoeconómicos

Para evaluar el desempeño de las estrategias de gestión de energía (EMS) en sistemas de bombeo alimentados con energías renovables, es necesario considerar indicadores tanto técnicos como económicos que permiten cuantificar tanto la eficiencia operativa como la viabilidad económica del sistema.

A partir de los indicadores mostrados en la Tabla 2, se observa que las estrategias EMS tiene un impacto directo tanto en el desempeño energético como en la viabilidad económica del sistema. Variables como la eficiencia energética, el consumo eléctrico y el estado de la carga de la batería influyen significativamente en los costos operativos, la continuidad del suministro y la vida útil de los componentes.

Tabla 2. Indicadores tecnoeconómicos en sistemas EMS

Tipo de indicador	Parámetro	Descripción	Impacto del EMS
Técnico	Eficiencia	Relación energía útil/generada	Mejora con optimización
	Confiabilidad	Satisfacción de la demanda	Aumenta con mejor control
	Uso de batería	Ciclos y profundidad de descarga	Reduce degradación
	Energía renovable	Aprovechamiento de generación	Minimiza pérdidas
Económico	Costo total	Inversión + operación	Disminuye con buena gestión
	LCOE	Costo por energía suministrada	Optimización lo reduce
	Costos operativos	Uso de respaldo	Disminuye con EMS eficiente
	Vida útil	Duración de componentes	Se extiende con control adecuado

Los trabajos revisados muestran que las estrategias más avanzadas permiten mejorar el aprovechamiento de la energía renovable y reducir el uso de fuentes auxiliares; sin embargo,

esto generalmente implica un incremento en la complejidad computacional y en los requerimientos de implementación. Por ello, la selección de una estrategia EMS debe considerar un equilibrio entre desempeño técnico, costo y condiciones operativas de cada aplicación.

5.2. Comparación de estrategias EMS desde el enfoque tecnoeconómico

Las estrategias EMS, desde el punto de vista tecnoeconómico, presentan ventajas y limitaciones que influyen en el desempeño del sistema. Los métodos basados en reglas son de bajo costo y fácil implementación, aunque no optimizan los recursos energéticos, aumentando los costos operativos a largo plazo. La inteligencia artificial mejora la eficiencia y adaptabilidad, pero requiere mayor complejidad de diseño. Los métodos de optimización minimizan costos y maximizan el aprovechamiento energético, especialmente en sistemas híbridos o conectados a red, aunque su complejidad computacional limita su aplicación en sistemas de baja escala. Finalmente, el control avanzado ofrece alta precisión, pero su implementación práctica aún es limitada y mayormente simulada.

La Tabla 3 presenta una comparación tecnoeconómica general de las principales estrategias EMS aplicadas en sistemas de bombeo de agua con energías renovables, considerando los aspectos más representativos de cada enfoque en términos de costo de implementación, complejidad computacional, eficiencia energética e impacto económico. Asimismo, esta comparación engloba las características técnicas más relevantes en la literatura para cada estrategia y su aplicación en diferentes configuraciones de sistemas aislados, híbridos y conectados a la red.

Tabla 3. Comparación tecnoeconómica de estrategias EMS

Estrategia EMS	Costo de implementación	Complejidad	Eficiencia energética	Impacto económico
Basado en reglas	Bajo	Baja	Media	Medio
Inteligencia artificial	Medio	Media-Alta	Alta	Alto
Optimización	Alto	Alta	Muy alta	Muy alto
Control avanzado	Alto	Muy alta	Muy alta	Alto

En general, la evaluación tecnoeconómica demuestra que no existe una única estrategia óptima para todos los escenarios, sino que la selección del EMS depende de un equilibrio entre costo, complejidad y desempeño.

6. RETOS Y TENDENCIAS

A pesar de los avances en estrategias de gestión de energía (EMS) para sistemas de bombeo con energías renovables, aún existen limitaciones técnicas y operativas que dificultan su aplicación a gran escala. Uno de los principales retos es la variabilidad de las fuentes solar y eólica (Amrouche et al., 2024), lo que genera fluctuaciones de potencia que afectan la estabilidad energética e hidráulica, especialmente en sistemas híbridos o aislados.

Otro desafío importante es la dependencia de modelos simplificados para representar baterías, bombas y convertidores electrónicos. Muchos trabajos utilizan escenarios idealizados sin considerar degradación, envejecimiento, variabilidad hidráulica, transitorios eléctricos ni incertidumbre meteorológica, limitando la validez de los resultados en condiciones reales. Asimismo, estrategias avanzadas como MPC, optimización multiobjetivo (Mahinda et al., 2023), algoritmos evolutivos o inteligencia artificial presentan altos requerimientos computacionales, dependencia de datos y dificultades de implementación en plataformas embebidas de bajo costo, especialmente en zonas rurales o sistemas autónomos.

En particular, la inteligencia artificial aún enfrenta limitaciones asociadas a disponibilidad de datos, interpretabilidad y capacidad computacional (Moyo, 2024). Además, existen pocos estudios enfocados en EMS resilientes ante fallas, eventos climáticos extremos o incertidumbre operativa prolongada. Como tendencia reciente, destacan las arquitecturas IoT para monitoreo y supervisión en tiempo real mediante adquisición continua de datos operativos y climáticos.

Desde el enfoque tecnoeconómico, la degradación de las baterías continúa siendo un problema importante, debido a los ciclos de carga y descarga reducen su vida útil e incrementan los costos de reemplazo. Además, la mayoría de las estrategias EMS se enfocan principalmente en variables energéticas y económicas, dejando en segundo plano parámetros hidráulicos como caudal, presión y disponibilidad de agua.

A pesar de estos avances, aún existen vacíos de investigación importantes. La validación experimental de EMS avanzados continúa siendo limitada, ya que muchos estudios permanecen en simulación. También son escasos los trabajos que integran degradación de baterías, dinámica hidráulica y variabilidad climática dentro de una misma arquitectura de control, así como marcos estandarizados para comparar estrategias EMS mediante indicadores energéticos, hidráulicos, económicos y de confiabilidad.

Como líneas futuras de investigación, se recomienda desarrollar EMS híbridos con capacidades predictivas y adaptativas, incorporar herramientas digitales para supervisión y optimización en tiempo real, emplear modelos electrotérmicos avanzados para estimar degradación del almacenamiento, e implementar estrategias inteligentes en plataformas embebidas de bajo costo orientadas a aplicaciones reales de bombeo de agua con energías renovables.

7. CONCLUSIONES

En este artículo se realizó una revisión de las principales estrategias de gestión de energía (EMS) aplicadas a sistemas de bombeo de agua basados en energías renovables, abordando su clasificación, variables de operación, modos de control y desempeño tecnoeconómico.

Se identificó que los sistemas híbridos son los más robustos por integrar múltiples fuentes de energía. Los métodos basados en reglas son ampliamente usados por su simplicidad, mientras que enfoques avanzados (inteligencia artificial, optimización,



control avanzado) ofrecen mejor desempeño a costa de mayor complejidad. Variables como el SOC, la potencia generada y la demanda son fundamentales en la toma de decisiones.

Desde el punto de vista tecnoeconómico, se evidenció que la selección de la estrategia EMS tiene un impacto directo en la eficiencia del sistema, los costos operativos y la vida útil de los componentes, especialmente del sistema de almacenamiento. En este sentido, las estrategias más avanzadas tienden a ofrecer mejores beneficios a largo plazo. Finalmente, se identificaron retos asociados a la variabilidad renovable y la complejidad computacional. Como tendencias futuras destacan la inteligencia artificial, optimización multiobjetivo, IoT y estrategias híbridas. En este contexto, los modos de operación representan una solución equilibrada entre simplicidad y desempeño.

REFERENCIAS

- Amrouche, K., Abdelli, R. & Benhacine, N.-E. 2024. Study and sizing of a multi-source pumping system. *Journal of Renewable Energies*.
- Ben Safia, Z., Allouche, M. & Chaabane, M. 2023. Renewable energy management of an hybrid water pumping system (photovoltaic/wind/battery) based on Takagi–Sugeno fuzzy model. *Optimal Control Applications and Methods*, 44, 373-390.
- Benzaouia, M., Hajji, B., Rabhi, A., Mellit, A., Benslimane, A. & Dubois, A. M. Energy Management Strategy for an Optimum Control of a Standalone Photovoltaic-Batteries Water Pumping System for Agriculture Applications. In: Hajji, B., Mellit, A., Marco Tina, G., Rabhi, A., Launay, J. & Naimi, S. E., eds. *Proceedings of the 2nd International Conference on Electronic Engineering and Renewable Energy Systems*, 2021// 2021 Singapore. Springer Singapore, 855-868.
- Cacciato, M., Finocchiaro, L., Nobile, G., Scarcella, G. & Scelba, G. Assessment of energy management strategies for battery assisted solar pumping systems. 2016 International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), 22-24 June 2016 2016. 575-581.
- Gueye, I., Kébé, A., Dia, O. & Sakho, A. 2024. Intelligent Management of the Pumping System for Irrigation of MBAWAANDE Lands in Podor within the NANN-K Project. *The Eurasia Proceedings of Science Technology Engineering and Mathematics*, 29, 219-225.
- Joshua, K. P., Rangasamy, L. V., Reddy, C. V. K. & Veeruchinnan, R. 2024. Energy management of solar photovoltaic fed water pumping system based BLDC motor drive using NBO–SDRN approach. *Electrical Engineering*, 106, 3045-3059.
- Khiareddine, A., Ben Salah, C. & Mimouni, M. F. 2015. Power management of a photovoltaic/battery pumping system in agricultural experiment station. *Solar Energy*, 112, 319-338.
- Khiareddine, A., Salah, C. B. & Mimouni, M. F. Strategy of energy control in PVP/battery water pumping system. 2014 First International Conference on Green Energy ICGE 2014, 25-27 March 2014 2014. 49-54.
- Kumar, V., Patel, R., Sahu, L. & Kishor, Y. 2024. Solar photovoltaic-integrated energy storage system with a power electronic interface for operating a brushless DC drive-coupled agricultural load. *Energy Harvesting and Systems*, 11.
- Mahinda, M. W., Wanjiru, E. M. & Njiri, J. G. 2023. Optimal control of a grid-connected photovoltaic agricultural water pumping system. *Journal of Engineering and Applied Science*, 70, 85.
- Mossa, M. A., Gam, O., Bianchi, N. & Quynh, N. V. 2022. Enhanced Control and Power Management for a Renewable Energy-Based Water Pumping System. *IEEE Access*, 10, 36028-36056.
- Moyo, R. 2024. Fuzzy logic-based energy management controller for a hybrid solar water pumping system. *International Journal of Sustainable Energy*, 15, 1-11.
- Ouachani, I., Rabhi, A., Yahyaoui, I., Tidhaf, B. & Tadeo, T. F. 2017. Renewable Energy Management Algorithm for a Water Pumping System. *Energy Procedia*, 111, 1030-1039.
- Poompavai, T. & Kowsalya, M. 2019. Control and energy management strategies applied for solar photovoltaic and wind energy fed water pumping system: A review. *Renewable and Sustainable Energy Reviews*, 107, 108-122.
- Rezk, H., Al-Dhaifallah, M., Hassan, Y. B. & Ziedan, H. A. 2020. Optimization and Energy Management of Hybrid Photovoltaic-Diesel-Battery System to Pump and Desalinate Water at Isolated Regions. *IEEE Access*, 8, 102512-102529.
- Sallem, S., Chaabene, M. & Kamoun, M. B. A. 2009. Optimum energy management of a photovoltaic water pumping system. *Energy Conversion and Management*, 50, 2728-2731.
- Shchur, I., Lis, M. & Biletskyi, Y. 2021. Passivity-Based Control of Water Pumping System Using BLDC Motor Drive Fed by Solar PV Array with Battery Storage System. *Energies* [Online], 14.
- Xu, Z. & Chen, X. 2024. Optimization of the Coupling between Water and Energy Consumption in a Smart Integrated Photovoltaic Pumping Station System. *Water* [Online], 16.
- Zaouche, F., Mokrani, Z. & Rekioua, D. Control and energy management of photovoltaic pumping system with battery storage. 2016 International Renewable and Sustainable Energy Conference (IRSEC), 14-17 Nov. 2016 2016. 917-922.

Contribution to the Design Methodology of DC-DC LLC Resonant Converter for Electric Vehicles

Dara Naomi Ventura Guzmán*. Carlos Daniel Garcia-Beltran*. Mario Ponce Silva*

*Tecnológico Nacional de México/CENIDET, Interior Internado Palmira S/N, Col. Palmira, C.P. 62490 Cuernavaca, Morelos, México.

(e-mail: m24ce080@cenidet.tecnm.mx, carlos.gb@cenidet.tecnm.mx, mario.ps@cenidet.tecnm.mx)

Abstract: This paper presents a design methodology for a transformer-based DC-DC LLC resonant converter for electric vehicle applications. The proposed approach is based on phasor analysis, first-harmonic approximation, rectifier efficiency evaluation, and resonant tank design constraints, allowing the systematic calculation of the main converter parameters. The converter is designed for an input voltage of 400 V, an output voltage target of 12 V, and an operating frequency of 500 kHz. The methodology includes determination of the quality factor, equivalent reflected resistance, transformer turns ratio, magnetizing inductance, resonant inductor, resonant capacitor and output filter. The complete converter is validated through SPICE simulations, where resonant behavior and soft switching operation are observed. Simulation results show an output power of 309.147 W, an efficiency of 96.52%, and an output voltage ripple close to 3.3%, confirming the usefulness of the proposed methodology for the design of high-frequency and high-efficiency LLC resonant converters for power conversion applications.

Keywords: Resonance, DC-DC converter, LLC topology, soft switching, design methodology

1. INTRODUCCIÓN

El acelerado crecimiento de la economía mundial ha impulsado un incremento significativo en el consumo de recursos naturales. El sector del transporte, como uno de los principales responsables del cambio climático global y de las emisiones de CO₂, ha motivado la búsqueda de alternativas más sostenibles y eficientes para la sustitución de motores de combustión interna alimentados por combustibles fósiles (Knez et al., 2019). En (Hoque et al., 2017) se destaca que el desarrollo de las tecnologías de los vehículos eléctricos (VE), sus aplicaciones, la gestión de la energía y los sistemas de almacenamiento representan áreas clave para la industria automotriz en sus aspectos medioambientales y económicos globales.

Los convertidores CD-CD resonantes han emergido como una solución fundamental en electrónica de potencia, particularmente en aplicaciones que demandan alta eficiencia y operación a frecuencias elevadas. En particular, los convertidores LLC con transformador se destacan por su capacidad para lograr conmutación suave, minimizar pérdidas, y filtrar armónicos de manera natural, gracias al tanque resonante compuesto por inductancias y capacitores (Rocha et al., 2025). Esta topología, ampliamente estudiada, permite no solo adaptarse a cargas variables, sino también garantizar aislamiento galvánico mediante el transformador, lo que lo hace ideal para sistemas fotovoltaicos, cargadores rápidos y fuentes conmutadas de alta densidad de potencia (Alatai et al., 2021).

Sin embargo, el diseño de este tipo de convertidores no es directo, ya que su comportamiento depende de la interacción entre la etapa de conmutación, el tanque resonante, el

transformador y el rectificador. En muchos casos, los enfoques reportados en la literatura se apoyan en simplificaciones que dificultan relacionar de manera clara las especificaciones del sistema con los parámetros reales de diseño, especialmente cuando se consideran pérdidas en diodos y transistores MOSFET, variaciones de carga y restricciones de operación en resonancia. Esto limita tanto la precisión del análisis como la obtención de convertidores con desempeño robusto y alta eficiencia.

Bajo este panorama, el presente trabajo propone una metodología de diseño para un convertidor resonante CD-CD LLC con transformador que funciona como un reductor a alta frecuencia. La propuesta parte del análisis por fasores y de la aproximación al primer armónico, e incorpora de manera explícita el estudio del factor de calidad, la eficiencia del rectificador, las restricciones de ganancia y la reflexión de impedancias hacia el devanado primario. A partir de ello se obtienen los parámetros principales del convertidor, tales como la inductancia resonante, la capacitancia resonante, la inductancia magnetizante, la relación de vueltas del transformador y el filtro de salida. Las especificaciones del convertidor se plantean con orientación a aplicaciones de vehículos eléctricos, considerando un nivel de entrada asociado a un bus de corriente directa y una salida adecuada para alimentar subsistemas auxiliares y electrónica de bajo voltaje.

La importancia de este trabajo radica en el enfoque metodológico para el diseño de convertidores resonantes LLC, en donde se abordan conceptos prácticos mediante simulaciones que permiten validar los conceptos del funcionamiento con conmutación a corriente cero y voltaje cero, la eficiencia del rectificador y la resonancia frente a la no

linealidad de componentes que introducen pérdidas como son los MOSFETs y los diodos en sus procesos de conmutación, en adición al estudio y análisis de todas las etapas del convertidor.

En la Figura 1 se resume de forma general las etapas funcionales del convertidor resonante analizado en este trabajo. Se identifican la etapa de conmutación, encargada de generar la excitación de alta frecuencia; el tanque resonante LLC, que transfiere la energía bajo condiciones cercanas a la resonancia; el transformador, que proporciona aislamiento galvánico, y adaptación del nivel de voltaje; y la etapa de rectificación y filtrado, mediante la cual se obtiene el voltaje de salida en corriente directa (Outeiro et al., 2016).

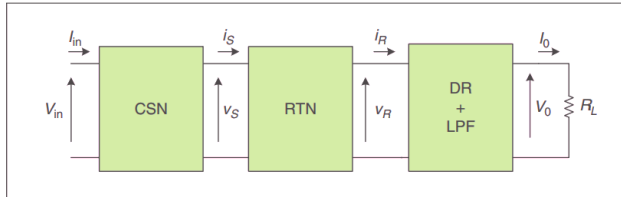


Figura 1. Etapas del convertidor resonante.

2. DESARROLLO

2.1 Estructura del circuito

Los convertidores resonantes pueden ser clasificados como convertidores serie, paralelo o serie-paralelo según la conexión de la red resonante con respecto a la carga, siendo este último el que permite lograr un mayor equilibrio en la regulación del voltaje y la corriente (Salem et al., 2018). Con base en esta característica, la estructura seleccionada se presenta en la Figura 2, que contempla una etapa de rectificación de puente completo y un filtro pasa bajas en paralelo a la carga.

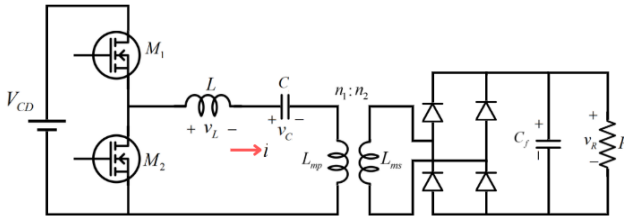


Figura 2. Topología del convertidor resonante LLC con transformador.

El tanque, por su parte, se compone de elementos inductivos y capacitivos que realizan la transferencia de energía para permitir que el sistema funcione en regímenes cercanos a la resonancia. Además, dentro de esta configuración se puede utilizar el amplificador clase D que contiene dos dispositivos de conmutación gracias al capacitor en serie (C) del tanque resonante que filtra las componentes de corriente directa.

2.2 Factor de calidad (Q)

Existe una relación directa entre el valor del factor de calidad (Q) y la cantidad de armónicos presentes en el circuito, siendo

que, a mayor Q, los armónicos se reducen sustancialmente permitiendo el paso únicamente a la componente fundamental como se presenta en la Figura 3, y en donde se concluye que un valor de Q=5 es el valor óptimo para el diseño siguiente (Guerrero, 2025).

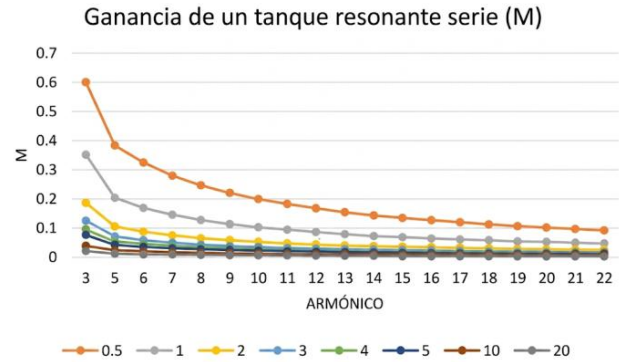


Figura 3. Atenuación del contenido armónico a distintos valores de factor de calidad (Guerrero, 2025).

2.3 Eficiencia del rectificador

Como parte inicial del análisis del circuito, se estudian los efectos del rectificador de puente completo. De forma inicial se consideran las pérdidas generadas en cada diodo como:

$$P_D = \frac{1}{T} \int_0^T v_D i_D dt \quad (1)$$

Al aproximar las pérdidas del diodo en su estado activo al comportamiento de una fuente de voltaje (V_F) y asignar un valor de $Q > 4$, la corriente en el diodo se comporta como:

$$i_D = \begin{cases} I \sin(\omega t) & 0 < \omega t < \pi \\ 0 & \pi < \omega t < 2\pi \end{cases} \quad (2)$$

Añadiendo el término de la eficiencia del rectificador (η_r) a la definición de las pérdidas en cada diodo, se establece la igualdad como:

$$P_D = \frac{V_R V_F}{2\eta_r R} \quad (3)$$

Con esto, la eficiencia del rectificador se define como:

$$\eta_r = \frac{P_R}{P_o} = 1 - \frac{2V_F}{V_R} \quad (4)$$

Para evitar que el nivel de la eficiencia sea negativo se añade la condición donde:

$$\begin{aligned} \frac{2V_F}{V_R} &\leq 1 \\ V_R &\geq 2V_F \end{aligned} \quad (5)$$

Que implica que la potencia que entra en el rectificador debe ser mayor que las pérdidas, ya que, en caso contrario, las fuentes de voltaje de los diodos entregarían energía en lugar de

recibirla, condición que físicamente resulta imposible y lleva a definirse la potencia de forma final como:

$$P_o \geq 4P_D \quad (6)$$

2.4 Análisis del circuito

A partir del análisis fasorial desarrollado en (Guerrero, 2025) para la reducción de la topología del convertidor resonante, se considera a la salida del amplificador clase D una señal cuadrada, generada por la activación y desactivación de los MOSFET. En dicha señal, se contempla una aproximación a la fundamental de la fuente de voltaje de entrada que permite reestructurar el circuito como se observa en la Figura 4.

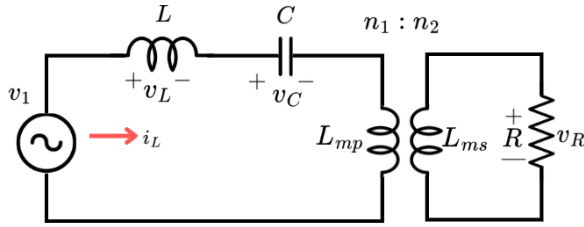


Figura 4. Reducción del convertidor resonante LLC.

Considerando un factor de acoplamiento $k = 1$, se realiza la reflexión al devanado primario, retomando la igualdad de potencias establecida como:

$$P = \frac{V_2^2}{R} = \frac{V_1^2}{R_{eq1}} \quad (7)$$

Bajo el análisis de fasores, R_{eq1} se determina como:

$$R_{eq1} = R \left(\frac{V_1}{V_2} \right)^2 = R \left(\frac{n_1}{n_2} \right)^2 \quad (8)$$

Lo que permite reestructurar el circuito como se muestra en la Figura 5.

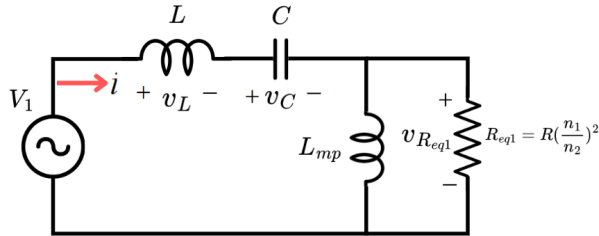


Figura 5. Segunda reducción del convertidor resonante LLC.

Al encontrarse la carga equivalente en paralelo al inductor L_{mp} , el tanque resonante se puede comportar como una fuente de voltaje o una fuente de corriente, dependiendo del valor de las impedancias X_{Lmp} y R_{eq1} .

Al cumplirse $R_{eq1} > X_{Lmp}$ predomina la reactancia X_{Lmp} y el comportamiento se da como una fuente de voltaje, mientras

que, en el caso de $R_{eq1} < X_{Lmp}$, el comportamiento se presenta como una fuente de corriente.

Bajo estas consideraciones se utiliza un rectificador sin inductancia que permite tomar el primer comportamiento de fuente de voltaje, y que de forma final genera un comportamiento de fuente de corriente en el tanque resonante, respetando así las leyes eléctricas que establecen como incorrecta la conexión de dos fuentes de voltaje en paralelo. Esto se presenta en el circuito final dado en la Figura 6, en donde debe de cumplirse que $R_{eq1} \ll X_{Lmp}$.

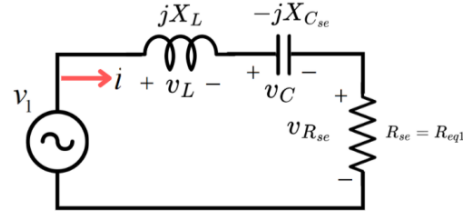


Figura 6. Reducción por impedancias considerando una fuente de corriente en el devanado secundario.

Con esto, la etapa de rectificación del circuito original obtiene su resistencia equivalente como:

$$R_{eq1} = \frac{8R\eta_{rec}}{\pi^2} \quad (9)$$

2.5 Restricciones de diseño

Bajo el análisis del tanque resonante, se considera la expresión de la reactancia de magnetización del devanado primario del transformador como:

$$X_{Lmp} = \frac{R_{eq1}}{\sqrt{M^2 - 1}} \quad (10)$$

Al asegurar que $R_{eq1} \ll X_{Lmp}$, la ganancia del convertidor se establece como:

$$\begin{aligned} \sqrt{M^2 - 1} &\ll 1 \\ M &\ll \sqrt{2} \end{aligned} \quad (11)$$

Por otro lado, para que X_{Lmp} sea real, y de acuerdo con la relación dada en (11), la ganancia debe de cumplir con ser mayor a la unidad, característica que permite anular términos negativos en la raíz cuadrada. Bajo las consideraciones previas, el valor de la ganancia del convertidor debe comprenderse dentro del rango:

$$1 < M \leq \sqrt{2} \quad (12)$$

La segunda restricción de diseño tiene se enfoca en la condición $R_{eq1} < X_{Lmp}$, en donde al añadir el término auxiliar a , se puede conocer la relación de disminución para la resistencia equivalente, expresada como:

$$R_{eq1} = \frac{X_{Lmp}}{a} \quad (13)$$

A partir de esta expresión, se observa que a mayor valor de X_{Lmp} , menor valor tomará R_{eq1} . Al realizar la sustitución de (13) en (10) y despejar el valor de la ganancia del convertidor, se obtiene que:

$$M = \frac{\sqrt{a^2 + 1}}{a} \quad (14)$$

Por lo que ante un valor de $a = 10$, la ganancia obtenida es $M = 1.0049$, cumpliendo con la condición dada en (12).

2.6 Análisis de ganancias

La ganancia total del convertidor CD-CD con tanque resonante LLC con transformador se define como:

$$M_{rect} = \frac{V_R}{V_{CD}} \quad (15)$$

Por otro lado, la ganancia del tanque resonante LLC con transformador sin considerar el rectificador se define como:

$$M_t = \frac{V_{Rect1}}{V_{CD}} = \frac{n_2}{n_1} \frac{2M}{\pi} \quad (16)$$

donde, V_{Rect1} es la componente fundamental del voltaje rectificado que se establece como:

$$V_{Rect1} = \frac{4V_R}{\pi} \quad (17)$$

Sustituyendo (17) en (16) se obtiene la expresión de la ganancia que permite añadir el término de la relación de vueltas del transformador, como:

$$M_{rect} = \frac{n_2}{n_1} \frac{M}{2} \quad (18)$$

2.7 Diseño del convertidor resonante CD-CD LLC

El diseño del convertidor resonante CD-CD LLC se desarrolla mediante la metodología secuencial que permite obtener los parámetros principales del circuito a partir de las especificaciones eléctricas de entrada y salida. El procedimiento propuesto considera el análisis de la componente fundamental de la señal de excitación, la selección del factor de calidad, la estimación de la eficiencia del rectificador, la reflexión de impedancias hacia el devanado primario y el dimensionamiento del tanque resonante y del filtro de salida.

La metodología de diseño propuesta se resume en los siguientes pasos principales presentes en la Figura 7, y que se utiliza a continuación.

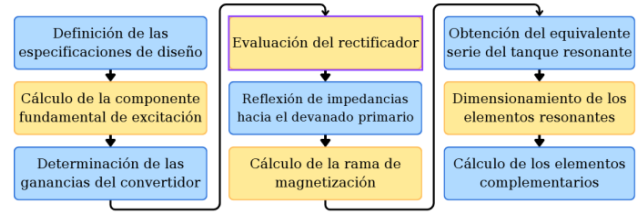


Figura 7. Metodología de diseño del convertidor resonante.

El circuito a diseñar parte de las especificaciones dadas en la Tabla 1, que sirven para completar el cálculo de los elementos que componen al circuito original de la Figura 2. Dicho sistema se llevará a la simulación en SPICE, en donde se propone un inversor clase D.

Tabla 1. Especificaciones del convertidor resonante LLC

Variable	Símbolo	Valor
Condición de resonancia		$\omega = \omega_0$
Potencia	P	300 W
Voltaje de entrada	V_{CD}	400 V
Frecuencia	f	500 kHz
Voltaje de salida	V_R	12 V
Factor de calidad	Q	5
Relación de impedancias	a	10
Caída de voltaje en diodos	V_F	0.7 V
Porcentaje de rizo	$\%v_r$	5%

A continuación, y tomando como base las fórmulas obtenidas del análisis del circuito, se comienza a obtener los valores correspondientes a cada parámetro del sistema que permiten llegar al dimensionamiento de los componentes electrónicos correspondientes al convertidor resonante.

Voltaje máximo de la componente fundamental:

$$V_1 = \frac{2V_{CD}}{\pi} = 254.6473 \text{ V} \quad (19)$$

Ganancia total del circuito:

$$M_{Rect} = \frac{V_R}{V_{CD}} = 0.03 \quad (20)$$

Ganancia del tanque resonante:

$$M = \frac{\sqrt{a^2 + 1}}{a} = 1.0049 \quad (21)$$

Factor de calidad mínimo (se selecciona $Q = 5$).

$$Q_{min} = \sqrt{M^2 - 1} = 0.1 \quad (22)$$

Resistencia de carga:



$$R = \frac{V_R^2}{P} = 0.48 \, \Omega \quad (23)$$

Relación de vueltas:

$$\frac{n_2}{n_1} = \frac{2M_{\text{Rect}}}{M} = 0.0597 \quad (24)$$

Eficiencia del rectificador:

$$\eta_r = \frac{V_R}{2V_F + V_R} = 0.8955 \quad (25)$$

Resistencia equivalente del rectificador:

$$R_{eq} = \frac{8\eta_r}{\pi^2} R = 0.348422 \, \Omega \quad (26)$$

Resistencia equivalente del rectificador reflejada hacia el primario:

$$R_{eq1} = \left(\frac{n_1}{n_2} \right)^2 R_{eq} = 97.7518 \, \Omega \quad (27)$$

Reactancia de magnetización del devanado primario:

$$X_{Lmp} = \frac{R_{eq1}}{\sqrt{M^2 - 1}} = 97.518 \quad (28)$$

Inductancia magnetizante del primario:

$$L_{mp} = \frac{X_{Lmp}}{\omega} = 311.153 \, \mu H \quad (29)$$

Reactancia inductiva serie:

$$X_{Lse} = \frac{jX_{Lmp} R_{eq1}^2}{R_{eq1}^2 + X_{Lmp}^2} = 9.67839 \quad (30)$$

Inductancia serie:

$$L_{se} = \frac{X_{Lse}}{\omega} = 3.08072 \, \mu H \quad (31)$$

Resistencia serie:

$$R_{se} = \frac{X_{Lmp}^2 R_{eq1}}{R_{eq1}^2 + X_{Lmp}^2} = 96.7839 \, \Omega \quad (32)$$

Reactancia del inductor resonante:

$$X_{Lr} = X_{cs} = Q R_{se} = 483.9198 \quad (33)$$

Reactancias del inductor:

$$X_L = X_{Lr} - X_{Lse} = 474.2414 \quad (34)$$

Valor del inductor:

$$L = \frac{X_L}{\omega} = 150.955 \, \mu H \quad (35)$$

Valor del capacitor:

$$C = \frac{1}{X_{cs} \omega} = 657.773 \, pF \quad (36)$$

Inductor magnetizante del secundario:

$$L_{ms} = L_{mp} \left(\frac{n_2}{n_1} \right)^2 = 1.10906 \, \mu H \quad (37)$$

Corriente del tanque resonante:

$$I_p = \frac{V_1}{R_{se}} = 2.63109 \, A \quad (38)$$

Capacitor de filtrado:

$$C_f \geq \frac{10.52}{\%V_R f R} = 8.76667 \, \mu F \quad (39)$$

Resistencia de encendido del MOSFET:

$$R_{DS(on)} \leq \frac{0.04P}{I_p^2} = 1.7334 \, \Omega \quad (40)$$

Corriente promedio en los diodos:

$$\bar{I}_D = \frac{V_R}{2R} = 12.5 \, A \quad (41)$$

La selección de los dispositivos de conmutación, diodos y transistores MOSFET, se seleccionan bajo los parámetros obtenidos, en donde se destaca la corriente que circula por estos componentes en el circuito, y para los MOSFET el valor de la resistencia de encendido.

Con los modelos seleccionados y los parámetros calculados, se obtiene el convertidor resonante LLC con transformador, como se muestra en la Figura 8.

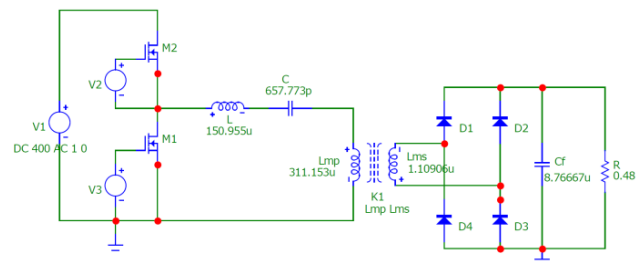


Figura 8. Convertidor resonante CD-CD LLC.

Con el fin de garantizar la conmutación a la frecuencia de resonancia de 500 kHz, se configuran las señales de V_2 y V_3 de acuerdo con los parámetros mostrados en la Tabla 2.

Tabla 2 Configuración de la fuente V_2 y V_3 del convertidor resonante.

Variable	V2	V3
V1	0 V	0 V
V2	12 V	12 V
TD	0 s	1 μ s
TR	1 ns	1 ns
TF	1 ns	1 ns
PW	0.97 μ s	0.97 μ s
PER	2 μ s	2 μ s

3. RESULTADOS

Para comenzar con la validación de resultados se analizan las señales provenientes de los dispositivos de conmutación en el medio puente. Dadas las características de la Tabla 2, se verifica que no existe un traslape en la activación de los transistores MOSFET, representado en la Figura 8. Esta característica es importante dado que, al variar el ancho del pulso, la frecuencia de resonancia en la que se diseñó el convertidor se ve afectada.

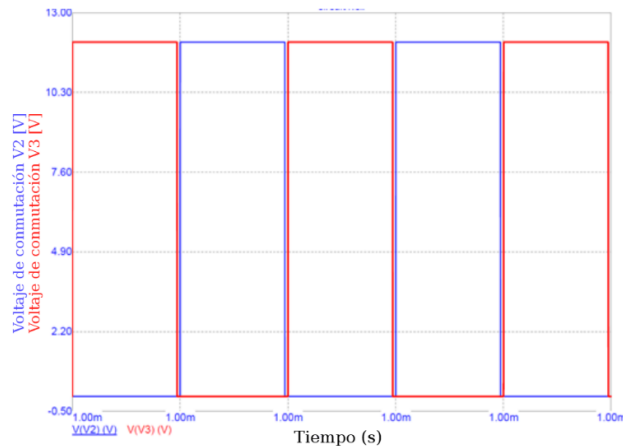


Figura 9. Voltaje de conmutación de los MOSFETs.

El fenómeno de conmutación a corriente cero y voltaje cero, se destaca principalmente en el comportamiento de la energía en el tanque resonante, presentado en la Figura 10 con las señales de corriente en el inductor y voltaje en el nodo a la entrada del tanque resonante.

Se observa un desfase por efecto de las capacitancias parásitas de los diodos, que, aunque generan una pequeña cantidad de pérdidas, también aseguran que el convertidor no entre en cortocircuito. Para poder visualizar correctamente la transición de las señales, se realiza un escalamiento x100 en la señal roja correspondiente a la corriente en el inductor del tanque resonante. Dicha señal toma valor de 2.5482 A sin el escalamiento, lo que permite corroborar la similitud a los valores obtenidos teóricamente en donde se consideraba una corriente pico en el tanque resonante de 2.63 A.

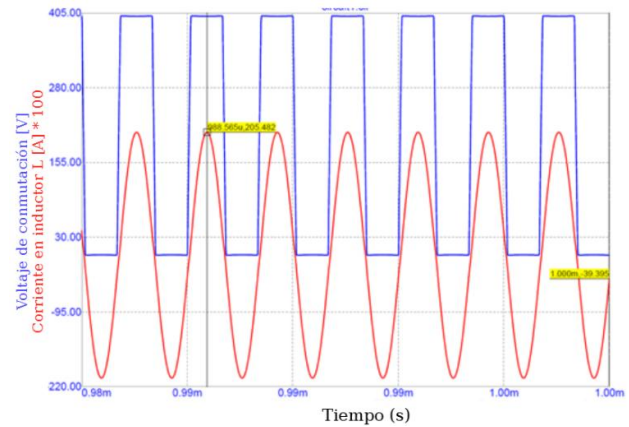


Figura 10. Conmutación en el convertidor resonante.

Con un error menor al 4% en la corriente generada en el tanque resonante se corroboran conceptos abordados en el análisis del circuito. La selección de un factor de calidad ($Q=5$) logra incorporar casi en su totalidad la energía de la señal de entrada del convertidor, siendo que se realiza una aproximación a la fundamental, sin tomar en cuenta el resto de los armónicos, concepto que facilita el análisis y genera muy ligeras variaciones en la implementación.

Con el fin de comparar como afecta esta consideración al convertidor, se realiza la simulación de ambas entradas como se muestra en la Figura 11. La parte del circuito superior considera como alimentación una fuente de voltaje senoidal de amplitud de la fundamental (V_1), mientras que el circuito inferior comprende la entrada con la fuente de voltaje V_{CD} alimentando a los MOSFET en conexión de medio puente.

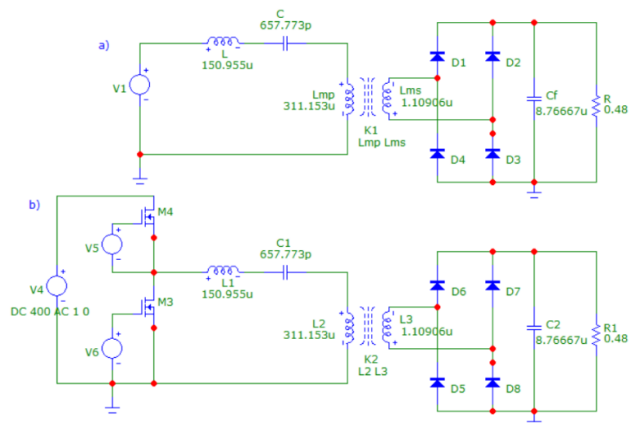


Figura 11. Convertidor resonante considerando a) Voltaje del armónico fundamental y b) Voltaje de entrada con etapa de conmutación.

Al simular ambas condiciones se logra ver que el voltaje del capacitor, que permite observar la energía que oscila en el tanque resonante, presenta diferencias de amplitud de alrededor del 20%, como se observa en la Figura 12.

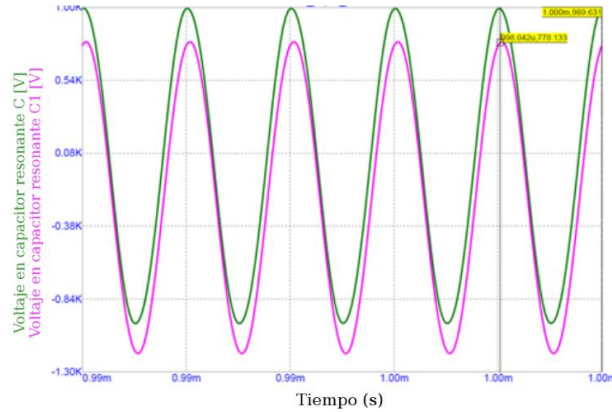


Figura 12. Voltaje en el capacitor del tanque resonante ante distintas fuentes de entrada.

Aunque la magnitud del error suena considerable, se destaca que la importancia radica en el comportamiento de las señales, siendo que ambas presentan dinámica senoidal similar, tanto en fase como en amplitud.

Siguiendo con el análisis, se observan las señales presentes en la carga del convertidor. Para el caso del voltaje se espera un valor de 12 V y un rizo del 5%, mientras que la simulación, presente en la Figura 13, arroja un valor final de 10.40 V y un rizo de voltaje de aproximadamente 3.3%, como se muestra en la Figura 14 al realizar un acercamiento a la señal obtenida.

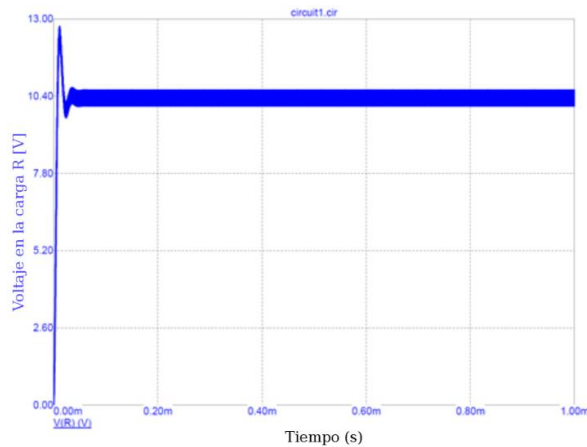


Figura 13. Voltaje en la carga del convertidor resonante.

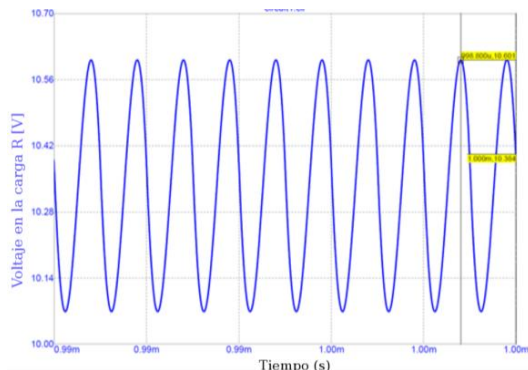


Figura 14. Rizo de voltaje en la carga.

Por su parte, la corriente obtiene un valor pico de 21.634 A, que difiere por 3.3 A del valor esperado, tal como se observa en la Figura 15.

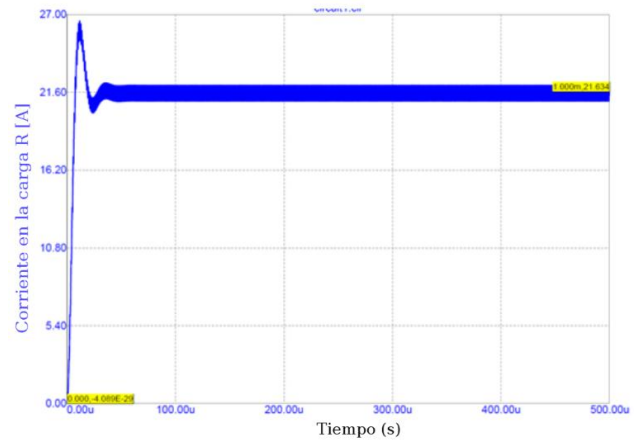


Figura 15. Corriente en la carga del convertidor resonante.

Las pérdidas observadas en la carga se dan principalmente por los dispositivos de conmutación (diodos y MOSFET), siendo estos los principales puntos de mejora al hablar del diseño de convertidores. En la Figura 16 se observan las pérdidas generadas en los MOSFET que impactan en la variación de los parámetros obtenidos, valor que se observa en la potencia que cae en dichos dispositivos.

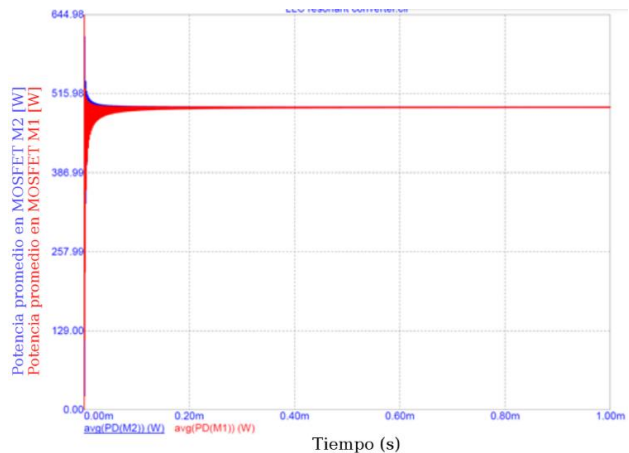


Figura 16. Pérdidas en los MOSFET del convertidor resonante.

La principal desviación se presenta en el voltaje de salida, lo cual puede atribuirse a las idealizaciones consideradas durante el análisis teórico. En el modelo de diseño se emplea la aproximación al primer armónico y se consideran parámetros equivalentes del rectificador y del tanque resonante; sin embargo, en la simulación se incorporan efectos no ideales como la caída directa de voltaje en los diodos, la resistencia en conducción de los MOSFET, las capacitancias parásitas y las pérdidas durante la conmutación. Estos elementos modifican la transferencia de energía hacia la carga y reducen el nivel de voltaje obtenido.



Además, el convertidor se encuentra operando en lazo abierto, por lo que no cuenta con una estrategia de regulación que compense las variaciones causadas por las pérdidas y por los parámetros no ideales de los dispositivos. A pesar de ello, la potencia de salida se mantiene cercana al valor nominal de diseño y el rizo de voltaje obtenido es menor al límite establecido, lo que confirma que la metodología propuesta permite observar un dimensionamiento inicial adecuado.

4. CONCLUSIONES

En este trabajo se presentó una metodología de diseño para un convertidor resonante LLC con transformador, orientada a una aplicación de conversión CD-CD de alta frecuencia. A partir del análisis fasorial, la aproximación al primer armónico, el estudio de la eficiencia del rectificador y las restricciones impuestas por el tanque resonante, fue posible establecer una secuencia de diseño para determinar los parámetros principales del convertidor y llevarlos posteriormente a simulación.

Los resultados obtenidos en SPICE mostraron un comportamiento cercano al previsto teóricamente. Se obtuvo un nivel de eficiencia del sistema de 96.52% a una frecuencia de operación de 500 kHz, lo que confirma que la topología LLC constituye una alternativa adecuada para aplicaciones que demandan operación eficiente a frecuencias elevadas. De igual forma, el voltaje de salida presentó un rizo aproximado de 3.3%, inferior al valor de diseño considerado, lo cual evidencia la efectividad del filtrado y la adecuada transferencia de energía en el convertidor.

Asimismo, la simulación permitió identificar el cumplimiento de conmutación suave y corroborar la importancia del factor de calidad seleccionado y del dimensionamiento del tanque resonante en el desempeño global del sistema. Aunque se observó un ligero desfase asociado a capacitancias parásitas en los diodos y pérdidas en los dispositivos de conmutación, dicho efecto no comprometió la operación resonante del convertidor. Por el contrario, estos resultados permiten reconocer con mayor claridad los elementos susceptibles de mejora en futuras implementaciones físicas.

En conjunto, el trabajo demuestra que la metodología propuesta permite relacionar de manera clara las especificaciones eléctricas con el dimensionamiento del convertidor y con su respuesta en simulación. Por ello, constituye una base útil para el diseño de convertidores resonantes LLC con transformador, y abre la posibilidad de continuar hacia etapas posteriores de optimización experimental, selección más precisa de semiconductores y desarrollo de estrategias de control para mejorar la regulación del voltaje de salida.

REFERENCIAS

Alatai, S., Salem, M., Ishak, D., Das, H. S., Nazari, M. A., Bughneda, A., & Kamarol, M. (2021). A review on state-of-the-art power converters: Bidirectional, resonant,

multilevel converters and their derivatives. In *Applied Sciences (Switzerland)* (Vol. 11, Number 21). MDPI. <https://doi.org/10.3390/app112110172>

Guerrero, L. (2025). Convertidor resonante bidireccional multipuerto para sistemas fotovoltaicos [Tesis de maestría, Centro Nacional de Investigación y Desarrollo Tecnológico]. Repositorio institucional del TecNM. México.

Hoque, M. M., Hannan, M. A., Mohamed, A., & Ayob, A. (2017). Battery charge equalization controller in electric vehicle applications: A review. In *Renewable and Sustainable Energy Reviews* (Vol. 75, pp. 1363–1385). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2016.11.126>

Knez, M., Zevnik, G. K., & Obrecht, M. (2019). A review of available chargers for electric vehicles: United States of America, European Union, and Asia. *Renewable and Sustainable Energy Reviews*, 109, 284–293. <https://doi.org/10.1016/j.rser.2019.04.013>

Outeiro, M. T., Buja, G., & Czarkowski, D. (2016). Resonant Power Converters: An Overview with Multiple Elements in the Resonant Tank Network. *IEEE Industrial Electronics Magazine*, 10(2), 21–45. <https://doi.org/10.1109/MIE.2016.2549981>

Rocha, J., Amin, S., Coelho, S., Rego, G., Afonso, J. L., & Monteiro, V. (2025). Design and Implementation of a DC–DC Resonant LLC Converter for Electric Vehicle Fast Chargers. *Energies*, 18(5). <https://doi.org/10.3390/en18051099>

Salem, M., Jusoh, A., Idris, N. R. N., Das, H. S., & Alhamrouni, I. (2018). Resonant power converters with respect to passive storage (LC) elements and control techniques – An overview. In *Renewable and Sustainable Energy Reviews* (Vol. 91, pp. 504–520). Elsevier Ltd. <https://doi.org/10.1016/j.rser.2018.04.020>

Yao, X., Han, X., Liao, Y., & Wang, J. (2022). PWM Strategy for Start-Up Process of LLC Resonant Converter Based on Event Trigger Control Method. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 69(3), 1357–1361. <https://doi.org/10.1109/TCSII.2021.3121075>

Zhang, Y., Liang, W., Wang, X., & Lifan, L. I. (2024). Passivity-based control strategy for resonant converter based on Euler–Lagrange model. *Archives of Electrical Engineering*, 73(1), 105–119. <https://doi.org/10.24425/aee.2024.148860>

A Single Source Nine Levels Inverter for Electromobility Applications

Salgado-Bajonero, Oswaldo*, Aguilar-Castillo, Carlos**, Arau-Roffiel, Jaime E.***, Vargas-Méndez, Rodolfo A.****

* Centro Nacional de Investigación y Desarrollo Tecnológico, 62493.

* m23ce113@cenidet.tecnm.mx, ** carlos.ac@cenidet.tecnm.mx

** Centro Nacional de Investigación y Desarrollo Tecnológico, 62943

*** jaim.e.ar@cenidet.tecnm.mx, **** rodolfo.vm@cenidet.tecnm.mx

Abstract: It is well known that the use of internal combustion vehicles is associated with several environmental problems; therefore, extensive research has been conducted to replace conventional combustion vehicles with electric vehicles. In this context, this paper introduces a new nine-level multilevel inverter topology with a single power source. The inverter is intended for electromobility applications, where a single source (batteries) is available. The design incorporates a unique auxiliary source which is added or subtracted from the main source to get the necessary levels, up to nine levels. The proposed inverter uses a fundamental modulation technique whose switching angles are calculated with the triangular number series, resulting in a very simple method but with excellent results about output voltage THD. The simulation results show the effectiveness of the proposed topology, achieving a THD of 8.65% without extra filtering. In addition, the topology is with less power switching, diodes and capacitors than previous proposals. In addition, the proposed topology requires fewer power switches, diodes, and capacitors than conventional multilevel inverter structures.

Keywords: Power electronics, Electromobility, Inverters, Multilevel inverters.

1. INTRODUCCIÓN

No cabe la menor duda que el desarrollo de la máquina de combustión interna fue crucial para impulsar el mundo de la movilidad. Sin embargo, actualmente se está cuestionando su vigencia dados los problemas causados colateralmente por los residuos generados por la quema de combustibles fósiles, por lo que es necesario considerar otras alternativas que permitan sustituir los autos basados en combustión interna por movilidad eléctrica o vehículos eléctricos (VE). De acuerdo con Blázquez y Martín (2010), los VE tienen menor costo de energía por kilómetro cuando se comparan con automóviles convencionales; el ahorro en costo por kilómetros de distintas tecnologías es: gasolina – 0% (tecnología base de comparación), etanol – 1%, vehículos híbridos – 20%, diésel – 23%, biodiésel – 27%, gas licuado del petróleo – 33%, gas natural – 38% y eléctricos – 75%. Es claro que la tecnología que mayor ahorro presenta es la de los VE.

La fuente de energía de los VE actualmente está basada en baterías. Este elemento es fundamental para el desarrollo y uso masivo de la electromovilidad, ya que de él depende la autonomía, el tiempo entre recargas, el volumen y el peso. Aunque existen otras tecnologías como la de supercapacitores híbridos e hidrógeno —celdas de combustible o motores de combustión de hidrógeno— (Tagore et al., 2022), las baterías de ion-litio son las que actualmente tienen mayor aceptación.

La Figura 1 muestra una configuración típica del tren motriz de un VE, en donde se observa el inversor (impulsor del motor)

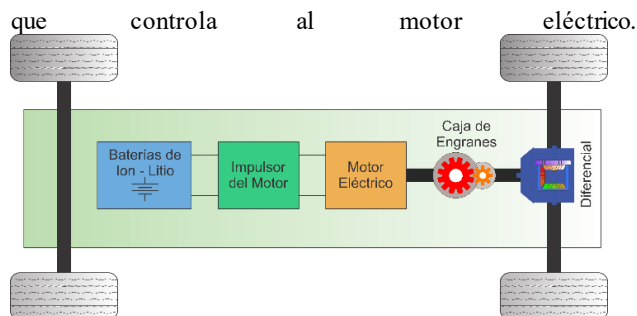


Fig. 1. Configuración de un VE basado en baterías.

La Electrónica de Potencia es fundamental en el desarrollo de los vehículos eléctricos (VE), ya que influye directamente en el diseño de los sistemas de conversión de energía y en el desempeño del tren motriz. Los avances en dispositivos semiconductores basados en carburo de silicio (SiC) y nitruro de galio (GaN) han permitido mejorar el rendimiento de los convertidores electrónicos, especialmente en aplicaciones de control de motores. Asimismo, nuevas configuraciones de inversores impactan en parámetros clave como la vida útil de la batería, la reducción de armónicos, la eficiencia del sistema y la disminución de efectos mecánicos adversos (Bucher et al., 2016; Katebi, He y Weise, 2018). Diversas topologías de inversores han sido propuestas para sistemas de electromovilidad. Por ejemplo, Bucher et al. (2016) presentan un inversor trifásico de dos niveles basado en MOSFETs de SiC que permite operar a mayores frecuencias de conmutación que los dispositivos IGBT tradicionales. Por otro lado, Katebi,

He y Weise (2018) proponen un inversor de tres niveles con capacidad de tolerancia a fallas. En aplicaciones con baterías de bajo voltaje, Merkert, Müller y Mertens (2016) plantean el uso de un convertidor elevador CD-CD combinado con un inversor de dos niveles, aunque la inclusión de esta etapa adicional puede reducir la eficiencia global. Alternativamente, Alfariy y Bhattacharya (2017) proponen un inversor con red quasi Z, que mejora el desempeño de conmutación pero introduce elementos adicionales que también generan pérdidas. Aunque las configuraciones tradicionales aprovechan la naturaleza inductiva del motor para filtrar el voltaje de salida, la presencia de armónicos no se elimina completamente. Estos armónicos pueden generar pérdidas adicionales, calentamiento, degradación del aislamiento y vibraciones mecánicas. Por ello, es fundamental desarrollar topologías de inversores que produzcan señales con baja distorsión armónica total (THD), con el fin de mejorar la eficiencia y confiabilidad del sistema de accionamiento eléctrico..

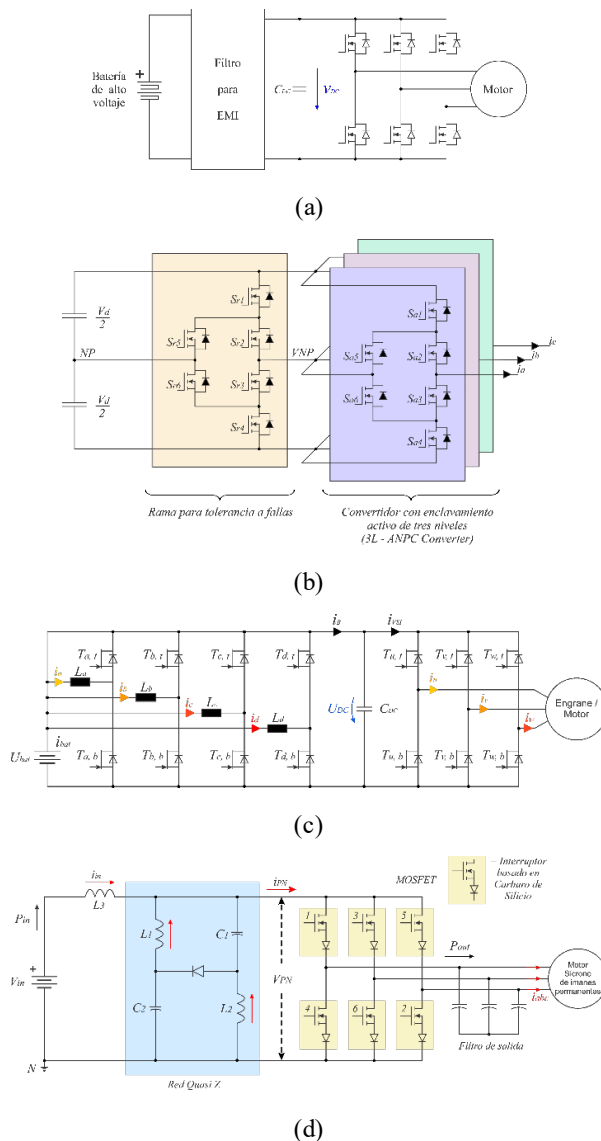


Figura 2. Configuraciones típicas de impulsor del motor de un VE. (a) Inversor típico de dos niveles con MOSFETs de SiC [3], (b) Inversor de tres niveles con enclavamiento activo tolerante a fallas [4], (c) Inversor de dos niveles con convertidor elevador [5], (d) Inversor alimentado en corriente con red quasi Z [6].

Una estructura multinivel genera una forma de onda de voltaje que busca parecerse a una sinusoidal, y se acerca más en la medida que se incrementan la cantidad de niveles. Cabe señalar que la cantidad de niveles no necesariamente está asociada a la frecuencia de conmutación, es decir, no se requiere frecuencia de conmutación alta para generar formas de onda con baja THD. Esencialmente existen cuatro configuraciones base de convertidores multi fases: capacitores flotados, con diodos de enclavamiento, en cascada y capacitores conmutados. En este trabajo se propone una nueva topología de inversor multinivel que opera con una única fuente de alimentación y reducida cantidad de componentes activos. La estructura propuesta emplea una estrategia de conmutación tipo modulación tipo step-voltage para el control de los interruptores. Esencialmente, la estructura combina (sumando o restando) el voltaje de la fuente principal con un voltaje auxiliar y de esta manera simple se generan los nueve niveles que se pretenden. La validación mediante simulación en PSpice® respalda la efectividad de esta propuesta, demostrando su viabilidad para aplicaciones en energías renovables, transporte eléctrico y sistemas industriales. En conjunto, esta investigación ofrece una solución más compacta y eficiente en términos de conversión de potencia, representando un avance significativo en la optimización de los inversores multinivel.

2. METODOLOGÍA

La Figura 3 muestra la idea conceptual de la topología de inversor multinivel que se propone, en donde la integración estratégica de una única fuente de alimentación auxiliar permite la generación de un voltaje de salida con hasta nueve niveles. Para ello, una etapa auxiliar genera una tensión auxiliar de $\pm 3 V_{bat}$ y mediante interruptores (tipo IGBT o MOSFET con tecnología Si o SiC) se suma o resta para generar los niveles del voltaje de salida. A diferencia de otras arquitecturas previamente reportadas en la literatura en las cuales se requieren múltiples fuentes de voltaje (Cheng et al., 2006) o configuraciones de interruptores más complejas (Singh, Boora y Kumar, 2022), esta propuesta simplifica la estructura sin comprometer la eficiencia o la complejidad de la estrategia de conmutación.

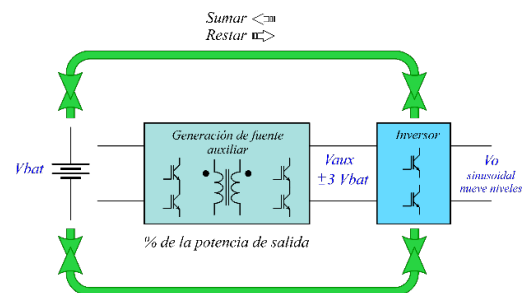


Figura 3. Configuración conceptual del inversor multinivel propuesto

La Figura 4 muestra la estructura de electrónica de potencia para la implementación del inversor propuesto. La frecuencia de conmutación es de baja frecuencia (algunas conmutaciones en un ciclo de red), con lo que las pérdidas de conmutación se mantienen bajas. Cabe señalar, que la fuente auxiliar podría ser del tipo reductora y emplearse en situaciones en donde se requiere reducción del voltaje de salida, como podría ser en configuraciones de paneles fotovoltaicos interconectados con la red. O bien, en donde el voltaje de la batería del vehículo sea alto y el motor sea de menor voltaje a la de la batería.

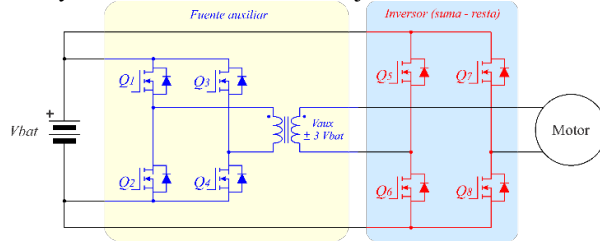


Figura 4. Inversor multinivel propuesto de nueve niveles con fuente única.

La Figura 5 muestra la forma de onda del voltaje de salida generada mediante el inversor multinivel y el voltaje que debe generar la etapa auxiliar.

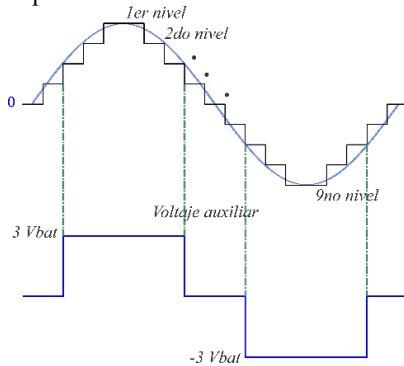


Figura 5. Voltaje generado por el inversor multinivel de nueve niveles propuesto y fuente auxiliar a generar.

La Tabla 1 muestra la acción que debe realizar el inversor propuesto para generar los niveles del voltaje de salida. Como se observa, a partir del voltaje auxiliar único es posible generar hasta nueve niveles en la tensión de salida. Se ha demostrado que con nueve niveles, optimizando los ángulos de conmutación, es posible obtener voltaje de salida con THD suficientemente bajo que incluso cumple con la norma IEEE 519 sin requerir filtro adicional.

Tabla 1. Acción de la etapa del inversor de acuerdo con el nivel del voltaje de salida.

Vaux	Acción del Inversor	Nivel	Vo
3 Vbat	Sumar	Vbat + 3 Vbat	4 Vbat
3 Vbat	Inactividad	3 Vbat	3 Vbat
3 Vbat	Restar	3 Vbat - Vbat	2 Vbat
0	Inactividad	Vbat	Vbat
0	Inactividad	0	0

0	Inactividad	- Vbat	6	- Vbat
- 3 Vbat	Sumar	- 3 Vbat + Vbat	7	- 2 Vbat
- 3 Vbat	Inactividad	- 3 Vbat	8	- 3 Vbat
- 3 Vbat	Restar	- 3 Vbat - Vbat	9	- 4 Vbat

El encendido de los interruptores es de acuerdo con el nivel de voltaje que se requiere generar y también de esto depende el voltaje que debe generar la fuente auxiliar.

2.3 Ecuaciones

La Figura 6 muestra la forma de onda de corriente de los interruptores Q1 - Q4. En esta Figura se considera que la carga tiene factor de potencia unitario. Donde T representa el periodo de la señal de salida, Zequiv corresponde a la impedancia equivalente de carga reflejada en el primario del transformador y Δtn representa el intervalo de conducción asociado a cada nivel de voltaje. El análisis de corriente considera operación balanceada en estado estable y una carga con factor de potencia unitario. En el caso de la corriente del interruptor Q1 (ver Figura 6), considerando la simetría de la forma de onda, se tiene:

$$I_{Q1} = \frac{2}{T} \int_{t_2}^{\frac{T}{4}} i_{Q1}(t) dt \quad (1)$$

$$I_{Q1} = \frac{2}{T} \left[\int_{t_2}^{t_3} i_{Q1_{3niv}}(t) dt + \int_{t_3}^{t_4} i_{Q1_{2niv}}(t) dt + \int_{t_4}^{\frac{T}{4}} i_{Q1_{1niv}}(t) dt \right] \quad (2)$$

La corriente en cada sección o nivel es la corriente de la carga reflejada al primario del transformador. Recuerdese que se está asumiendo que el transformador eleva el voltaje de la batería tres veces, es decir:

$$I_{Q1} = \frac{6}{T} \left[\frac{2V_{bat}}{Z_{equiv}} t \Big|_{t_2}^{t_3} + \frac{3V_{bat}}{Z_{equiv}} t \Big|_{t_3}^{t_4} + \frac{4V_{bat}}{Z_{equiv}} t \Big|_{t_4}^{\frac{T}{4}} \right] \quad (3)$$

$$I_{Q1} = \frac{6V_{bat}}{T Z_{equiv}} \left[2(t_3 - t_2) + 3(t_4 - t_3) + 4\left(\frac{T}{4} - t_4\right) \right] \quad (4)$$

$$I_{Q1} = \frac{6V_{bat}}{T Z_{equiv}} [2\Delta t_2 + 3\Delta t_3 + 4\Delta t_4] \quad (5)$$

Como se observa en la Figura 6, la corriente en el interruptor Q3 es similar a la corriente en Q1, pero ésta se presenta durante la parte negativa del ciclo de línea. Por lo tanto, la corriente en Q3 será también:



$$I_{Q3} = \frac{6V_{bat}}{T Z_{equiv}} [2\Delta t_2 + 3\Delta t_3 + 4\Delta t_4] \quad (6)$$

Siguiendo un procedimiento similar, se evalúa la corriente en Q2.

$$I_{Q2} = \frac{2}{T} \left[- \int_{t_1}^{t_2} i_{Q2_{4niv}}(t) dt + \int_{t_1}^{t_2} i_{Q2_{6niv}}(t) dt + \int_{t_2}^{t_3} i_{Q2_{7niv}}(t) dt + \int_{t_3}^{t_4} i_{Q2_{8niv}}(t) dt + \int_{t_4}^{\frac{T}{4}} i_{Q2_{9niv}}(t) dt \right] \quad (7)$$

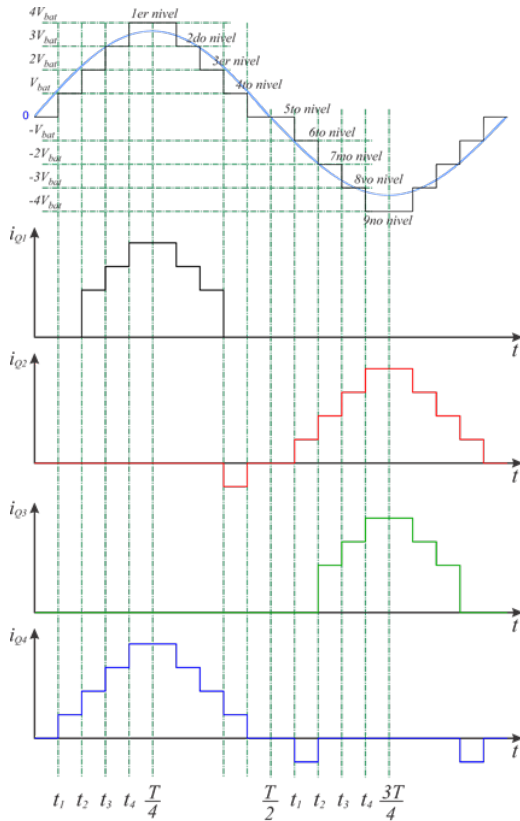


Figura 6. Formas de onda teóricas de la corriente en los interruptores de la fuente auxiliar.

$$I_{Q2} = \frac{6}{T} \left[- \frac{V_{bat}}{Z_{equiv}} t \Big|_{t_1}^{t_2} + \frac{V_{bat}}{Z_{equiv}} t \Big|_{t_1}^{t_2} + \frac{2V_{bat}}{Z_{equiv}} t \Big|_{t_2}^{t_3} + \frac{3V_{bat}}{Z_{equiv}} t \Big|_{t_3}^{t_4} + \frac{4V_{bat}}{Z_{equiv}} t \Big|_{t_4}^{\frac{T}{4}} \right] \quad (8)$$

$$I_{Q2} = \frac{6}{T} \left[\frac{2V_{bat}}{Z_{equiv}} t \Big|_{t_2}^{t_3} + \frac{3V_{bat}}{Z_{equiv}} t \Big|_{t_3}^{t_4} + \frac{4V_{bat}}{Z_{equiv}} t \Big|_{t_4}^{\frac{T}{4}} \right] \quad (9)$$

$$I_{Q2} = \frac{6V_{bat}}{T Z_{equiv}} \left[2(t_3 - t_2) + 3(t_4 - t_3) + 4\left(\frac{T}{4} - t_4\right) \right] \quad (10)$$

$$I_{Q2} = \frac{6V_{bat}}{T Z_{equiv}} [2\Delta t_2 + 3\Delta t_{32} + 4\Delta t_4] \quad (11)$$

De igual manera, como se observa en la Figura 8, la corriente en el interruptor Q2 es similar a la corriente del interruptor Q4, pero uno en el semiciclo positivo, y el otro en el semiciclo negativo. Por lo tanto se tiene:

$$I_{Q4} = \frac{6V_{bat}}{T Z_{equiv}} [2\Delta t_2 + 3\Delta t_{32} + 4\Delta t_4] \quad (12)$$

Como se muestra en las ecuaciones (5), (6), (11), y (12), todos los interruptores conducen corriente promedio iguales. Esto es adecuado ya que puede utilizar el mismo modelo de interruptor, teniendo las mismas pérdidas, y por lo tanto se ecualiza la disipación de temperatura.

En el caso de los interruptores del inversor que suma y resta (ver Figura 7), la corriente del interruptor Q5 es:

$$I_{Q5} = \frac{2}{T} \left[\int_{t_1}^{t_2} i_{Q5_{4niv}}(t) dt + \int_{t_4}^{\frac{T}{4}} i_{Q5_{1niv}}(t) dt - \int_{t_2}^{t_3} i_{Q5_{7niv}}(t) dt \right] \quad (13)$$

$$I_{Q5} = \frac{2}{T} \left[\frac{V_{bat}}{Z_{equiv}} t \Big|_{t_1}^{t_2} + \frac{4V_{bat}}{Z_{equiv}} t \Big|_{t_4}^{\frac{T}{4}} - \frac{2V_{bat}}{Z_{equiv}} t \Big|_{t_2}^{t_3} \right] \quad (14)$$

$$I_{Q5} = \frac{2V_{bat}}{T Z_{equiv}} \left[(t_2 - t_1) + 4\left(\frac{T}{4} - t_4\right) - (t_3 - t_2) \right] \quad (15)$$

$$I_{Q5} = \frac{2V_{bat}}{T Z_{equiv}} (\Delta t_1 - 2\Delta t_2 + 4\Delta t_4) \quad (16)$$

La corriente de Q7 es similar a la de Q5, es decir:

$$I_{Q7} = \frac{2 V_{bat}}{T Z_{equiv}} (\Delta t_1 - 2\Delta t_2 + 4\Delta t_4) \quad (17)$$

Mediante un procedimiento similar, la corriente en Q6 es:

$$I_{Q6} = \frac{2 V_{bat}}{T Z_{equiv}} (-2\Delta t_2 - 3\Delta t_3 + \Delta t_1 + 3\Delta t_3 + 4\Delta t_4) \quad (18)$$

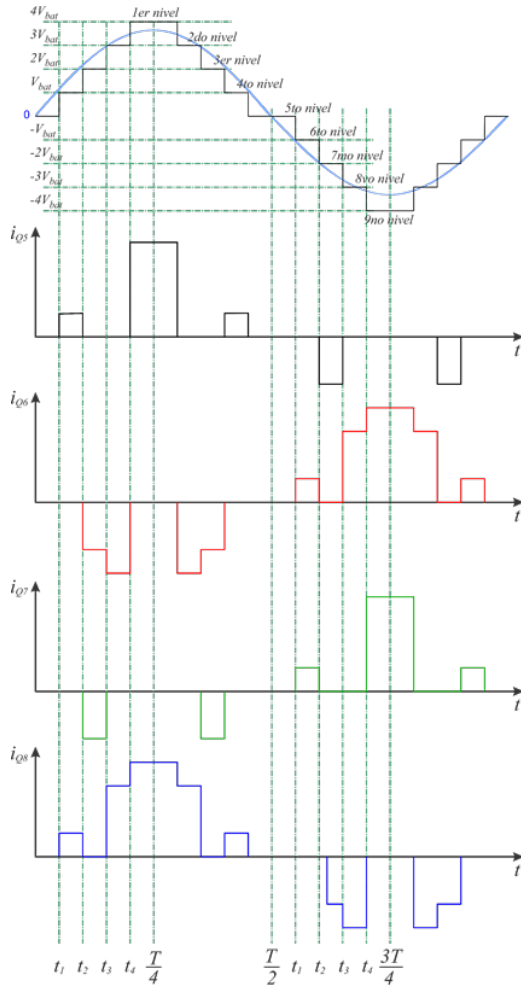


Figura 7. Formas de onda teóricas de la corriente en los interruptores del inversor suma – resta.

$$I_{Q6} = \frac{2 V_{bat}}{T Z_{equiv}} (\Delta t_1 - 2\Delta t_2 + 4\Delta t_4) \quad (19)$$

Y la corriente en Q8 será:

$$I_{Q8} = \frac{2 V_{bat}}{T Z_{equiv}} (\Delta t_1 - 2\Delta t_2 + 4\Delta t_4) \quad (20)$$

Todos los interruptores del inversor que suman – resta conducen la misma cantidad de corriente promedio.

3. RESULTADOS Y/O DISCUSIÓN

Se realizó la simulación de la estructura propuesta utilizando el software OrCAD PSpice®. Para ello se considerarán las especificaciones que se muestran en la siguiente tabla.

Tabla 3. Especificaciones de ejercicio de simulación del inversor de 9 niveles.

Especificaciones		
Potencia de salida máxima P_o	1000	watts (1 fase)
Frecuencia de voltaje de salida	60	Hz
Voltaje de salida	440	Vrms
Número de niveles	9	
V_{bat}	144	V

Los ángulos de conmutación empleados son los que muestra la Tabla 2. La Figura 8 muestra la forma de onda del voltaje de salida del inversor obtenida mediante simulación; en dicha figura se notan los niveles y la forma de onda cercana a una sinusoidal. La Figura 9 muestra el análisis de Fourier del voltaje de salida. El contenido armónico total es de 8.65%, muy por debajo del contenido armónico de una señal de 2 o 3 niveles, que puede ser entre 60 y 140%, dependiendo del índice de modulación. De acuerdo con el análisis de Fourier, la fundamental tiene un valor de aproximadamente 422 volts rms, 4% menor a lo especificado; esto se debe a las desviaciones en la precisión para implementar en la simulación los ángulos de conmutación. Existen algunas propuestas para calcular los ángulos de conmutación que utilizan técnicas de optimización, tales como redes neuronales artificiales (Ghadi et al., 2023); con estas técnicas se puede obtener una THD ligeramente menor, pero la complejidad de la implementación es alta.

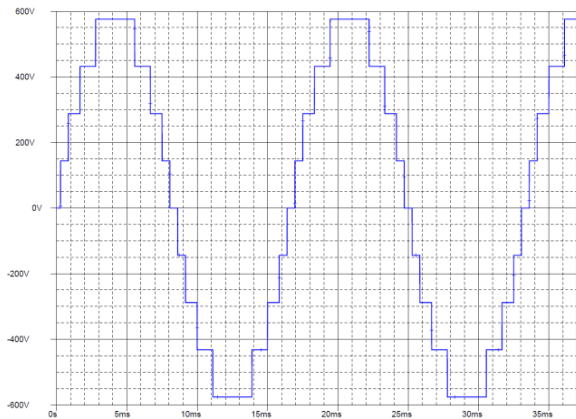


Figura 8. Forma de onda del voltaje de salida del inversor propuesto de nueve niveles obtenido mediante simulación.

La evaluación de la THD se realizó considerando condiciones de estado estable y carga resistiva nominal. De acuerdo con las

recomendaciones de la norma IEEE 519, la distorsión armónica debe mantenerse dentro de límites aceptables en el punto de acoplamiento común; por ello, el valor obtenido de 8.65% representa una reducción significativa respecto a inversores convencionales de dos niveles sin requerir filtros adicionales. Se espera que, bajo condiciones de carga inductiva como las presentes en motores eléctricos, la distorsión de corriente se atenúe adicionalmente debido al efecto natural de filtrado de la inductancia de la máquina. Los componentes armónicos dominantes observados en el espectro corresponden principalmente a armónicos de bajo orden asociados a la forma escalonada del voltaje generado.

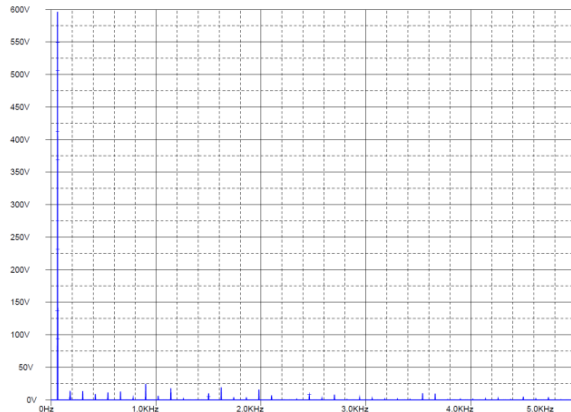


Figura 9. Análisis de Fourier de la forma de onda del voltaje de salida mediante la Transformada Rápida de Fourier.

La Figura 10 muestra las formas de onda obtenidas mediante simulación de la corriente de los interruptores del convertidor que genera la fuente auxiliar. Como se observa, las formas de onda coinciden con las estimadas en el análisis del convertidor. La Figura 11 muestra las formas de onda obtenidas en la simulación de corriente en los interruptores del inversor que suma – resta para generar los multiniveles. De igual manera, coinciden con las que se muestran en el análisis del convertidor.

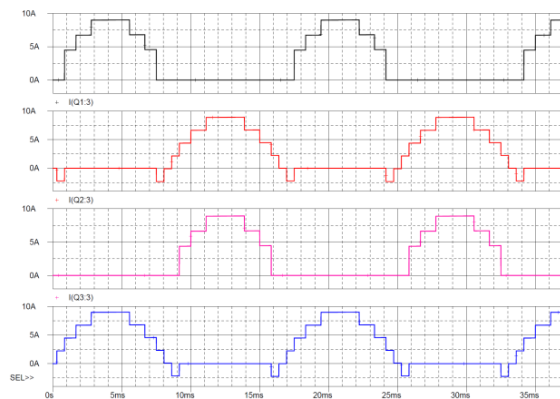


Figura 10. Formas de onda de las corrientes en los interruptores del convertidor fuente auxiliar obtenidas mediante simulación.

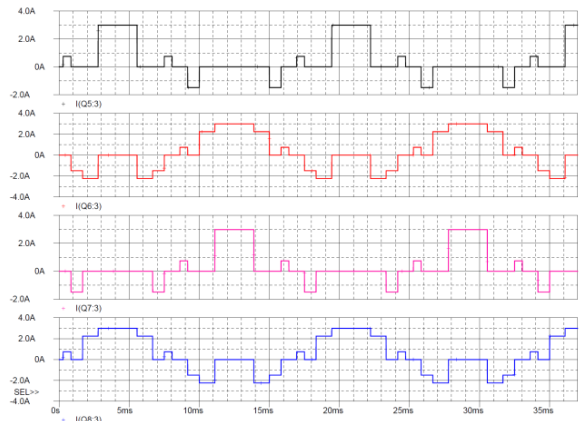


Figura 11. Formas de onda de la corriente en los interruptores del inversor que “suma – resta” obtenidas mediante simulación.

En cuanto a la cantidad de componentes empleados para la implementación de la topología propuesta, la Tabla 4 muestra un comparativo con topologías clásicas. En general se disminuye la cantidad de componentes empleados para la implementación.

Tabla 4. Comparativo de componentes empleados en las estructuras de potencia clásicas e inversor propuesto.

	Topología			
	Diodos de enclavamiento	Capacitores flotados	Convertidores en cascada	Inversor propuesto
Fuentes	1	1	≥ 4	1
Interruptores	16	16	16	8
Diodos	56	0	0	0
Capacitores	10	10	0	0

Además de la reducción en la cantidad de dispositivos activos, la topología propuesta presenta diversas ventajas prácticas respecto a estructuras multinivel convencionales. Debido a que el inversor opera con baja frecuencia de conmutación, las pérdidas por conmutación se reducen significativamente, mejorando potencialmente la eficiencia global del convertidor. Asimismo, la menor cantidad de dispositivos semiconductores disminuye los requerimientos de circuitos de disparo y simplifica la estrategia de control. Otra ventaja importante es la distribución balanceada de corriente entre los interruptores, lo cual contribuye a una distribución térmica más uniforme y facilita la gestión térmica del sistema. Sin embargo, la topología requiere una etapa auxiliar de generación de voltaje, lo cual introduce consideraciones adicionales relacionadas con el dimensionamiento del transformador y la regulación de voltaje.

4. CONCLUSIONES

La electromovilidad es una tecnología alternativa para disminuir la contaminación por gases que contribuyen al cambio climático. Cada país, al ritmo adecuado con cada

economía, tiene que transitar hacia ese tipo de tecnologías y no cabe duda de que existe la capacidad técnica para desarrollar proyectos de infraestructura para contribuir en la innovación y mejoras en el sistema de tracción de los vehículos eléctricos e híbridos. La investigación de inversores que mejoran el performance es un claro ejemplo de temáticas en el que se puede contribuir.

Este artículo presenta una estructura de electrónica de potencia para el inversor con la capacidad de generar formas de onda de voltaje con mucho menor THD al que exhiben los inversores tradicionales de dos o tres niveles. La estructura propuesta contiene menor cantidad de interruptores y menor cantidad de diodos o capacitores que se requieren con estructuras con capacidades similares. Particularmente se analiza y simula un inversor de nueve niveles con una única fuente de alimentación (banco de baterías de un vehículo eléctrico). Para ello se genera un voltaje bipolar adicional con la que se puede sumar o restar el voltaje de baterías para obtener hasta nueve niveles.

Desde la perspectiva de eficiencia, la topología propuesta presenta ventajas derivadas de la operación a baja frecuencia de conmutación y de la reducción en la cantidad de dispositivos semiconductores respecto a convertidores multinivel convencionales. Estas características contribuyen potencialmente a menores pérdidas de conmutación y a una gestión térmica más sencilla. Además, la distribución uniforme de corriente entre los interruptores reduce esfuerzos térmicos localizados y puede incrementar la confiabilidad del convertidor en aplicaciones prácticas.

El artículo presenta resultados de simulación en el software OrCAD PsPice® para comprobar la operación del inversor previamente a la implementación. Los resultados obtenidos muestran excelente funcionamiento respecto a lo esperado en forma de onda de voltaje multinivel, THD y esfuerzos de corriente. La THD obtenida mediante simulación fue de 8.65%, lo cual representa una reducción significativa respecto a inversores convencionales de dos niveles y se aproxima a los límites recomendados por normas como la IEEE 519 sin requerir filtrado adicional. Y aun cuando no existen normas específicas en aplicaciones de electromovilidad, se tiene conocimiento de las ventajas que esto conlleva en el funcionamiento del motor y tren motriz en general.

REFERENCIAS

- Aguayo, J., De León, S., Calleja, H. y Claudio, A. (2020) 'Switching angles calculation in multilevel inverters using triangular number sequence - A THD minimization approach', *European Journal of Electrical Engineering*, 22(1), pp. 49–55.
- Alfaris, F.E. y Bhattacharya, S. (2017) 'A current-fed quasi Z-source inverter with SiC power modules for EV/HEV applications', en *IEEE Energy Conversion Congress and Exposition (ECCE)*.
- Attique, Q.M., Wang, K., Zheng, Z., Zhu, H., Li, Y. y Rodríguez, J. (2022) 'A generalized simplified virtual vector PWM to balance the capacitor voltages of multilevel diode-clamped converters', *IEEE Transactions on Power Electronics*, 37, pp. 9377–9391.
- Blázquez, J. y Martín, J.M. (2010) 'Eficiencia energética en la automoción: el vehículo eléctrico, un reto del presente', *Economía Industrial*, 377, pp. 76–85.
- Bucher, A., Schmidt, R., Werner, R., Leipenat, M., Hasenohr, C., Werner, T., Schmitz, S. y Heitmann, A. (2016) 'Design of a full SiC voltage source inverter for electric vehicle applications', en *European Conference on Power Electronics and Applications (EPE)*.
- Cheng, Y., Qian, C., Crow, M.L., Pekarek, S. y Atcitty, S. (2006) 'A comparison of diode-clamped and cascaded multilevel converters for a STATCOM with energy storage', *IEEE Transactions on Industrial Electronics*, 53, pp. 1512–1521.
- Ferreira, A., Rodríguez, C. y Gomis, O. (2018) 'Modulation techniques applied to medium voltage modular multilevel converters for renewable energy integration: A review', *Electric Power Systems Research*, 155, pp. 21–39.
- Ghadi, Y.Y., Iqbal, M.S., Adnan, M., Amjad, K., Ahmad, I. y Farooq, U. (2023) 'An improved artificial neural network-based approach for total harmonic distortion reduction in cascaded H-bridge multilevel inverters', *IEEE Access*, 11, pp. 127348–127363.
- Kampitsis, G., Batzelis, E.I., Kolokasis, A., Matioli, E. y Pal, B.C. (2022) 'A generalized phase-shift PWM extension for improved natural and active balancing of flying capacitor multilevel inverters', *IEEE Open Journal of Power Electronics*, 3, pp. 621–634.
- Katebi, R., He, J. y Weise, N. (2018) 'An advanced three-level active neutral-point-clamped converter with improved fault-tolerant capabilities', *IEEE Transactions on Power Electronics*, 33, pp. 6897–6909.
- Merkert, A., Müller, J. y Mertens, A. (2016) 'Component design and implementation of a 60 kW full SiC traction inverter with boost converter', en *IEEE Energy Conversion Congress and Exposition (ECCE)*.
- Muthyala, U.R. y Dhabale, A. (2020) 'Numerical series based switching angle calculation for multilevel inverters', en *IEEE International Conference on Smart Technologies for Power, Energy and Control (STPEC)*.
- Singh, P., Boora, K. y Kumar, V. (2022) 'Optimization of cascaded H-bridge multilevel inverter using particle swarm optimization technique for reduction of THD', en *International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT)*.
- Sivapriya, A., Kalaiarasi, N., Verma, R., Chokkalingam, B. y Munda, J.L. (2023) 'Fault diagnosis of cascaded multilevel inverter using multiscale kernel convolutional neural network', *IEEE Access*, 11, pp. 79513–79530.
- Tagore, R., Kotapati, A., Kandipati, R. y Sekhar, C. (2022) 'A review on energy efficient technologies for electric vehicle applications', *Journal of Energy Storage*, 50, pp. 1–27.

Effect of the Modeling Approach on Porous Hip Implants: Structured vs. Homogenized *

Magali E. García-Gutiérrez * Arturo Gómez Ortega *
Astrid Lorena Giraldo Betancur ** Enrique Cuan Urquiza ***
Jesús González Hernández ** José María Rodríguez Lelis ****

* Centro de Ingeniería y Desarrollo Industrial (CIDESI), Querétaro,
Av. Pie de la Cuesta 702, 76125 Santiago de Querétaro, México
(inv.asoc57mgarcia@cidesi.edu.mx, arturo.gomez@cidesi.edu.mx).

** Centro de Investigación y de Estudios Avanzados del IPN
(Cinvestav), Unidad Querétaro, Libramiento Norponiente 2000,
Fraccionamiento Real de Juriquilla, 76230, Santiago de Querétaro,
México (agiraldo@cinvestav.mx, jesusgonzalez@cinvestav.mx).

*** Institute of Advanced Materials for Sustainable Manufacturing,
Tecnológico de Monterrey, Monterrey, Mexico (ecuanurqui@tec.mx)

**** Tecnológico Nacional de México/Centro Nacional de Investigación
y Desarrollo Tecnológico, Interior Internado Palmira S/N, Col.
Palmira, Cuernavaca 62490, Mexico, (jose.rl@cenidet.tecnm.mx)

Abstract: *Stress shielding* remains a major challenge in porous hip implant design. This study compares structured and homogenized finite element models of femoral stems with TPMS-based porous architectures (*diamond* and *primitive*) at 50 % relative density and 5 mm unit cell size under osteoporotic conditions. Mechanical responses, including displacement, von Mises stress, strain energy, and strain energy density (SED), were evaluated using median and 95th percentile statistics. Homogenized models eliminated mechanical differences between topologies and underestimated median SED by 23 % relative to structured models. These results suggest that homogenized approaches may affect predictions of mechanical stimuli and bone resorption in implant evaluation.

Keywords: Numerical homogenization, TPMS structures, hip implants, osteoporotic bone, strain energy density, finite element analysis

1. INTRODUCCIÓN

Las prótesis de cadera son uno de los implantes ortopédicos más utilizados a nivel mundial, en gran medida debido al alto éxito inicial que presentan. Este procedimiento es particularmente frecuente en poblaciones con deterioro de la calidad ósea, como pacientes con osteoporosis, condición ampliamente asociada a fracturas de cadera y degeneración articular severa. Sin embargo, la evidencia indica que la vida útil del implante puede ser inferior a la expectativa de vida del paciente, especialmente en poblaciones jóvenes y activas Abdel et al. (2000); Nugent et al. (2021); Smith et al. (2023). Se ha estimado un riesgo de revisión de ~ 30 % en individuos menores de 50 años Abdel et al. (2000), mientras que incluso en prótesis modernas la incidencia de revisión a 10 años puede variar entre 2,8 % y 9,6 % dependiendo del diseño y la combinación protésica. Como consecuencia, el diseño de vástagos de cadera ha evolucionado significativamente en las últimas décadas con el objetivo de mejorar la integración biológica y prolongar la vida útil del sistema hueso-implante.

Uno de los principales desafíos en el diseño de vástagos femorales es el fenómeno de *stress shielding*, que se origina cuando el implante, al ser más rígido que el hueso circundante, asume la mayor parte de la carga mecánica, reduciendo el estímulo mecánico en el tejido óseo y provocando resorción ósea progresiva Niinomi et al. (2016). Este fenómeno resulta particularmente crítico en hueso osteoporótico, donde la reducción de la densidad y de la rigidez ósea incrementa el desajuste mecánico entre el implante y el tejido circundante, intensificando la pérdida ósea y comprometiendo la estabilidad a largo plazo del implante Abdel et al. (2000); Nugent et al. (2021); Smith et al. (2023).

Para mitigar el fenómeno de *stress shielding*, las generaciones más recientes de implantes han incorporado estructuras reticulares o porosas en el vástago femoral. Estas estructuras permiten reducir el desajuste de rigidez y aumentar la transferencia de carga hacia el hueso Wang et al. (2018); Arabnejad Khanoki and Pasini (2012); Jetté et al. (2018). En este contexto, las superficies mínimas triplemente periódicas (TPMS, por sus siglas en inglés) han emergido como una familia de geometrías con propiedades mecánicas ajustables, continuidad superficial y características favorables para la osteointegración. Entre las

* Este trabajo fue realizado con el apoyo de la Secretaría de Ciencia, Humanidades, Tecnología e Innovación (SECIHTI) a través de una Estancia Posdoctoral por México.

topologías TPMS más estudiadas destacan las configuraciones *diamond* y *primitive*, las cuales presentan distintas distribuciones de material y, por tanto, comportamientos mecánicos diferenciados bajo carga.

El análisis por elemento finito (FEA) es la herramienta computacional predominante para evaluar la respuesta mecánica de implantes estructurados. Sin embargo, el modelado explícito de la geometría reticular implica un costo computacional elevado, especialmente cuando se trabaja con celdas unitarias de tamaño reducido en dominios de escala macroscópica. Como alternativa, la homogeneización numérica permite representar la microestructura como un medio continuo equivalente con propiedades anisotrópicas efectivas, reduciendo significativamente el costo computacional. Esta aproximación es ampliamente utilizada en la literatura para el diseño y optimización de implantes con estructuras reticulares Wang et al. (2018); Alkhatib et al. (2019); Müller et al. (2024).

No obstante, la validez del enfoque homogeneizado para estimar parámetros biomecánicos locales, como la densidad de energía de deformación (SED, por sus siglas en inglés), ampliamente utilizada como indicador del estímulo mecánico asociado a la remodelación ósea, no ha sido evaluada de forma sistemática frente al modelado estructurado equivalente, particularmente bajo condiciones representativas de hueso osteoporótico. Si bien existen modelos por elemento finito que incorporan propiedades dependientes de la densidad ósea para representar el comportamiento del hueso, los estudios que analizan de manera comparativa el efecto del enfoque de modelado, homogeneizado frente a estructurado, en la estimación de parámetros mecánicos en condiciones osteoporóticas son limitados. La mayoría de los trabajos emplea modelos homogeneizados como herramienta principal de análisis, reservando los modelos estructurados para validaciones parciales o dominios reducidos Alkhatib et al. (2019); Müller et al. (2024), lo que deja abierta la interrogante sobre el impacto del enfoque de modelado en la estimación del estímulo mecánico y, en consecuencia, en la interpretación del *stress shielding* en hueso con propiedades degradadas.

El presente trabajo tiene como objetivo comparar la respuesta mecánica de un vástago femoral mediante dos enfoques de modelado. En el primero, el vástago se representa como un modelo estructurado, en el cual se incorpora explícitamente un núcleo poroso basado en estructuras TPMS. En el segundo, dicha estructura interna se reemplaza por un material continuo equivalente, dando lugar a un modelo homogeneizado. A partir de esta comparación, se evalúa el efecto del enfoque de modelado sobre parámetros mecánicos clave, como desplazamientos, esfuerzos equivalentes de von Mises (vMS, por sus siglas en inglés), energía de deformación (SE, por sus siglas en inglés) y SED, bajo condiciones representativas de hueso osteoporótico, escenario clínico predominante en pacientes sometidos a reemplazo de cadera y aún limitado en estudios comparativos entre ambos enfoques.

2. METODOLOGÍA

El presente estudio compara dos enfoques de modelado por elemento finito para evaluar la respuesta mecánica de un vástago de prótesis de cadera total con estructuras internas

de tipo TPMS. A continuación se describe el proceso de construcción geométrica, la generación de los modelos de elemento finito y las condiciones de análisis empleadas.

2.1 Modelo geométrico del fémur

El modelo tridimensional del fémur se obtuvo a partir de un proceso de segmentación de una tomografía computarizada de nivel clínico utilizando el software 3D Slicer (Fedorov et al. (2012)). El modelo considera la diferenciación entre hueso cortical y trabecular, con un espesor característico asociado al tamaño de la celda unitaria de las estructuras analizadas (5 mm).

2.2 Modelos de implantes estructurados

Los implantes estructurados considerados en este estudio corresponden a vástagos femorales que incorporan un núcleo poroso basado en estructuras TPMS, el cual define su estructura interna. En este contexto, las topologías *diamond* y *primitive* fueron seleccionadas debido a que se encuentran entre las geometrías TPMS más reportadas en aplicaciones de implantes ortopédicos, con propiedades mecánicas bien caracterizadas y comportamientos diferenciados en la distribución de material y transferencia de carga Wang et al. (2018); Jetté et al. (2018). Estas topologías TPMS se muestran en la Figura 1, junto con su implementación en el núcleo del implante.

Para generar el núcleo poroso, primero se construye un volumen de estructura TPMS de forma rectangular mediante la repetición de la celda unitaria. Este volumen se obtiene en MATLAB (The MathWorks Inc. (2023)) a partir de las ecuaciones implícitas de las TPMS, mostradas a continuación:

Schwarz Primitive

$$f(x, y, z) = \cos(px) + \cos(py) + \cos(pz) = c \quad (1)$$

Schwarz Diamond

$$\begin{aligned} &\sin(px) \sin(py) \sin(pz) + \sin(px) \cos(py) \cos(pz) \\ &+ \cos(px) \sin(py) \cos(pz) + \cos(px) \cos(py) \sin(pz) = c \end{aligned} \quad (2)$$

donde $p = \frac{2\pi}{l}$ define la periodicidad espacial, l es la longitud de la celda unitaria y c es el *isovalor* que controla la densidad relativa. Estos volúmenes se generan haciendo uso de las funciones *isosurface* e *isocaps*, y presentan una densidad relativa del 50 % y una longitud de celda unitaria de 5 mm. Posteriormente, tanto los volúmenes con estructuras TPMS como el núcleo sólido del implante se convierten a una representación voxelizada. Mediante una operación de intersección entre ambos volúmenes, se obtiene el núcleo del implante con estructura TPMS.

Después, el núcleo poroso se une al resto del implante, previamente voxelizado, mediante una operación de unión booleana, obteniendo así el implante completo en formato voxel. Dicho implante voxelizado se reconvierte a formato de malla y se procesa en Fusion 360 (Autodesk Inc. (2026)) para reducir el número de caras y corregir defectos geométricos. El diseño final del vástago femoral con núcleo TPMS integrado se presenta en la Figura 1. Finalmente, el modelo se importa en ANSYS Workbench (ANSYS Inc. (2020)), donde se genera el sólido para su análisis numérico.

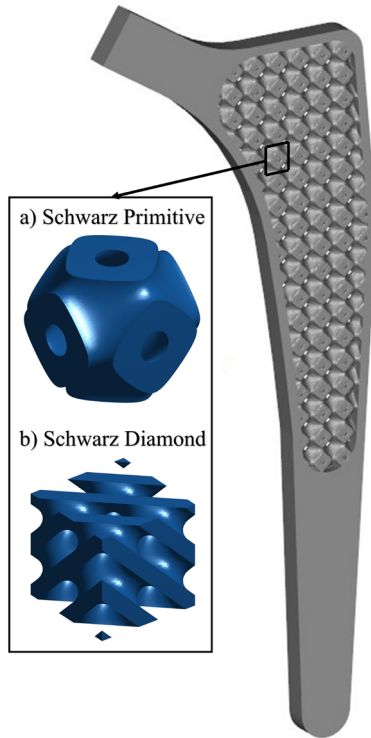


Figura 1. Estructuras TPMS empleadas en el diseño del núcleo poroso del implante: a) Schwarz Primitive y b) Schwarz Diamond, junto con su implementación en el núcleo del vástago femoral.

2.3 Modelos de implantes homogeneizados

Para los modelos homogeneizados, se utilizaron las celdas unitarias *Schwarz Primitive* y *Schwarz Diamond* como volúmenes elementales representativos (VER), mostradas en la Figura 1. A los VER se les asignaron propiedades elásticas isotrópicas de Ti6Al4V y se discretizaron mediante una malla tetraédrica con un tamaño de elemento de 0.1 mm.

El proceso de homogenización numérica se realizó en el software nTopology (nTopology Inc. (2023)), donde se aplicaron seis condiciones de frontera periódicas sobre los VER, correspondientes a tres estados de tensión uniaxial y tres de cortante uniaxial. Este proceso permitió obtener los tensores elásticos equivalentes de cada topología, los cuales presentan una simetría ortotrópica.

Finalmente, los tensores de rigidez obtenidos se asignaron a los núcleos sólidos de los implantes en ANSYS Workbench como propiedades de material equivalentes, dando lugar a los modelos homogeneizados correspondientes.

2.4 Propiedades de los materiales

Los materiales fueron modelados como lineales elásticos e isotrópicos, con propiedades mecánicas resumidas en la Tabla 1. Esta simplificación se adoptó con el objetivo de evaluar la transferencia de carga entre configuraciones estructurales, sin considerar efectos asociados a comportamiento no lineal o falla del material.

Tabla 1. Propiedades mecánicas de los materiales (Morgan et al. (2003); Rho et al. (1998); Zebaze et al. (2010)).

Material	E (GPa)	ν
Hueso trabecular	0.52	0.29
Hueso cortical	1.7	0.30
Ti6Al4V	110	0.30

Las propiedades mecánicas asignadas al tejido óseo corresponden a valores aparentes o efectivos, los cuales incorporan de manera implícita los efectos de la microestructura, tales como la porosidad y la conectividad interna del hueso. En particular, para el hueso cortical, el valor de módulo de Young corresponde a un caso crítico dentro de los rangos reportados, el cual refleja la reducción de rigidez asociada al incremento de la porosidad intracortical y permite incorporar de manera parcial el efecto del adelgazamiento cortical, el cual no fue modelado explícitamente en la geometría. De este modo, se mantiene la consistencia geométrica del modelo, al mismo tiempo que se captura de forma efectiva la degradación mecánica global asociada a condiciones osteoporóticas Morgan et al. (2003); Rho et al. (1998); Zebaze et al. (2010).

2.5 Discretización y mallado

El mallado se realizó mediante elementos tetraédricos cuadráticos para los modelos estructurados, con un tamaño de elemento de 1 mm en el fémur y 0.4 mm en el implante. Para los modelos homogeneizados, el implante se discretizó mediante elementos hexaédricos cuadráticos, manteniendo el mismo criterio de refinamiento en el fémur. Se realizó un análisis de convergencia de malla para verificar la independencia de los resultados respecto al tamaño de elemento seleccionado.

2.6 Condiciones de frontera y carga

Las condiciones de frontera empleadas en los modelos de elemento finito se muestran en la Figura 2, correspondiente a un caso de modelo homogeneizado. Estas consistieron en la aplicación de una carga sobre el vástago a la altura de la cabeza femoral, equivalente a tres veces el peso corporal (323.58 N en el eje x y -2302.4 N en el eje y), representando la carga máxima durante la fase de apoyo del ciclo de marcha Bergmann et al. (2001). Adicionalmente, se consideró la fuerza del músculo abductor aplicada sobre el hueso cortical a la altura del trocánter mayor, con componentes de -329.6 N en el eje x y 619.8 N en el eje y Kuiper and Huiskes (1992). El extremo distal del fémur se restringió completamente.

La interacción entre el implante y el hueso se modeló mediante una condición de contacto tipo *bonded*. Esta condición supone una interfaz completamente unida, representando una condición de osteointegración total, lo cual es consistente con los supuestos adoptados en estudios previos Alkhatib et al. (2019); Müller et al. (2024).

3. RESULTADOS

3.1 Respuesta mecánica en modelos estructurados

La Tabla 2 presenta los valores medianos y la Tabla 3 los valores del percentil 95 de los vMS y SE obtenidos

Tabla 2. Desplazamientos máximos, valores medianos de vMS y SE

Estructura	Despl. [mm]	vMS fémur [MPa]	vMS vástago [MPa]	SE trab. [mJ]	SE cort. [mJ]	SE fémur [mJ]	SE vástago [mJ]
Diamond	205.81	22.55	215.35	$4,91 \times 10^{-3}$	0.016	$8,53 \times 10^{-3}$	$2,75 \times 10^{-4}$
Primitive	206.36	23.23	200.41	$4,89 \times 10^{-3}$	0.017	$8,49 \times 10^{-3}$	$1,36 \times 10^{-3}$
Diamond homog.	205.81	19.26	128.78	$4,76 \times 10^{-3}$	0.016	$8,41 \times 10^{-3}$	$1,44 \times 10^{-2}$
Primitive homog.	205.82	19.39	125.98	$4,76 \times 10^{-3}$	0.016	$8,47 \times 10^{-3}$	$1,41 \times 10^{-2}$

Tabla 3. Valores del percentil 95 de vMS y SE

Estructura	vMS fémur [MPa]	vMS vástago [MPa]	SE trab. [mJ]	SE cort. [mJ]	SE fémur [mJ]	SE vástago [mJ]
Diamond	231.11	856.62	0.181	2.289	1.850	$2,45 \times 10^{-2}$
Primitive	231.11	856.62	0.181	2.289	1.850	$2,45 \times 10^{-2}$
Diamond homog.	229.75	918.42	0.182	2.283	1.846	$2,94 \times 10^{-1}$
Primitive homog.	229.75	921.74	0.182	2.283	1.846	$3,01 \times 10^{-1}$



Figura 2. Modelo de elemento finito del sistema implante-fémur de un modelo homogeneizado y condiciones de frontera aplicadas.

para los modelos estructurados *diamond* y *primitive* y sus correspondientes modelos homogeneizados. Los resultados se presentan de forma separada para el vástago y el fémur, considerando este último como la contribución conjunta del hueso cortical y trabecular.

En los modelos estructurados, los desplazamientos máximos fueron ligeramente mayores para la configuración

primitive en comparación con *diamond* ($\approx 0,3\%$). En cuanto a los vMS, el fémur presentó valores mayores para la configuración *primitive* ($\approx 3\%$), mientras que en el vástago los valores fueron superiores para la configuración *diamond* ($\approx 7,5\%$).

En cuanto a la SE, en el hueso trabecular ambas configuraciones mostraron valores similares (diferencias $< 1\%$). En el hueso cortical se observaron valores ligeramente mayores para la configuración *primitive* ($\approx 2 - 3\%$). A nivel del fémur, los valores fueron comparables entre ambas configuraciones ($\approx 1\%$). En contraste, en el vástago se observaron diferencias más marcadas, con valores superiores para la configuración *primitive* (aproximadamente cinco veces mayores), lo que sugiere una mayor acumulación de energía en el implante asociada a su arquitectura interna.

En términos del percentil 95, los vMS en el fémur y en el vástago fueron similares entre ambas configuraciones (diferencias $< 3\%$). La SED en el hueso trabecular, cortical y a nivel del fémur presentó valores comparables (diferencias $< 2\%$), mientras que en el vástago la configuración *primitive* presentó valores superiores (aproximadamente cinco veces mayores).

3.2 Respuesta mecánica en modelos homogeneizados

A diferencia de los modelos estructurados, los modelos homogeneizados mostraron un comportamiento prácticamente idéntico entre las configuraciones *diamond* y *primitive* en todos los parámetros evaluados. Los desplazamientos máximos fueron prácticamente idénticos entre las configuraciones *diamond* y *primitive* (diferencias $< 0,01\%$). Los vMS en el fémur y en el vástago también fueron muy similares, con diferencias menores al $1 - 2\%$.

En cuanto a la energía de deformación, en el hueso trabecular y cortical los valores fueron prácticamente iguales entre ambas configuraciones (diferencias $< 1\%$). A nivel del fémur, las diferencias se mantuvieron igualmente bajas ($\approx 1\%$). En el vástago se observaron variaciones ligeramente mayores, aunque aún reducidas ($\approx 2\%$).

En términos del percentil 95, los vMS en el fémur y en el vástago fueron prácticamente idénticos (diferencias $< 1\%$). La SE mostró valores coincidentes en el hueso trabecular, cortical y en el fémur total, mientras que en el vástago se observaron diferencias menores al 3% . Estos resultados indican que el enfoque homogeneizado no es

sensible a la arquitectura interna del implante, eliminando la diferenciación entre topologías observada en los modelos estructurados.

3.3 Comparación entre enfoques de modelado

Al comparar ambos enfoques, se observa que el efecto del modelado depende tanto de la variable analizada como de la región del sistema implante-fémur considerada. Los desplazamientos máximos no mostraron variaciones relevantes entre ambos enfoques, con diferencias menores al 0,3% para ambas configuraciones, lo que indica que la rigidez global del sistema es capturada de manera equivalente por ambos modelos.

En cuanto a los vMS, en el fémur se observaron valores mayores en los modelos estructurados ($\approx 15 - 20\%$), mientras que en el vástago las diferencias fueron más marcadas, alcanzando incrementos del orden de $\approx 60 - 70\%$ para la configuración *diamond* y $\approx 55 - 60\%$ para *primitive*.

Respecto a la SE, en el hueso trabecular y cortical los modelos estructurados presentaron valores ligeramente superiores ($\approx 2 - 5\%$), con un comportamiento similar a nivel del fémur. En contraste, en el vástago se identificaron diferencias significativamente mayores, con valores superiores en los modelos homogeneizados, del orden de ≈ 10 a 50 veces respecto a los modelos estructurados, lo que evidencia que el efecto del enfoque de modelado es considerablemente más relevante en el implante que en el hueso circundante.

En términos del percentil 95, los vMS mostraron variaciones menores al 5% tanto en el fémur como en el vástago. La SE fue comparable en el hueso trabecular, cortical y en el fémur total (diferencias $< 2 - 3\%$), mientras que en el vástago los modelos homogeneizados presentaron valores superiores (aproximadamente cinco veces mayores).

3.4 SED como indicador de estímulo mecánico

La SED constituye el parámetro más relevante para evaluar el estímulo mecánico en el tejido óseo, dado su uso generalizado en modelos de remodelación ósea. Los valores medianos de SED obtenidos para ambas configuraciones (*diamond* y *primitive*) fueron iguales en los modelos estructurados (0.26 mJ/mm^3), mientras que en los modelos homogeneizados se registraron valores inferiores (0.20 mJ/mm^3), lo que representa una reducción aproximada del 23%. Esta diferencia sugiere que el enfoque homogeneizado subestima el nivel general de estímulo mecánico distribuido en el hueso.

En términos del percentil 95, ambos enfoques arrojaron resultados prácticamente equivalentes, con diferencias menores al 0,3%. Los modelos estructurados presentaron valores de 19.07 mJ/mm^3 y 19.09 mJ/mm^3 para las configuraciones *diamond* y *primitive*, respectivamente, mientras que los modelos homogeneizados registraron un valor de 19.03 mJ/mm^3 en ambos casos. Esto indica que, si bien la homogeneización reproduce adecuadamente los valores máximos de SED, reduce el nivel general de estímulo mecánico distribuido en el tejido óseo.

3.5 Distribución de vMS

La Figura 3 ilustra la distribución espacial de los vMS en el ensamble implante-fémur para las cuatro configuraciones analizadas. En los modelos estructurados, cada topología exhibe un campo de esfuerzos característico con patrones diferenciados en la localización de las concentraciones de esfuerzo a lo largo del vástago y en la interfaz con el fémur, con valores máximos de 2151 MPa y 2089.6 MPa para las configuraciones *primitive* y *diamond*, respectivamente.

En los modelos homogeneizados, en cambio, la distribución de esfuerzos es visualmente similar entre ambas topologías, lo que es consistente con los resultados numéricos reportados en las secciones anteriores y confirma que la homogeneización elimina la influencia de la arquitectura interna sobre la distribución de cargas.

4. DISCUSIÓN

Los resultados obtenidos en los modelos estructurados evidencian que la arquitectura interna del implante influye en su respuesta mecánica global, lo cual es consistente con lo reportado en la literatura para estructuras reticulares de tipo TPMS Wang et al. (2018); Jetté et al. (2018). Aunque los desplazamientos máximos y la energía de deformación en el fémur presentan diferencias reducidas entre las configuraciones evaluadas, se observan variaciones en la distribución de esfuerzos y en la energía almacenada en el vástago. En particular, la configuración *primitive* muestra una mayor energía de deformación en el implante, mientras que la configuración *diamond* presenta mayores vMS en el vástago. Estas diferencias confirman que las topologías analizadas no son mecánicamente equivalentes cuando se considera su geometría estructurada, destacando el papel de la arquitectura interna en la distribución de cargas dentro del sistema implante-fémur.

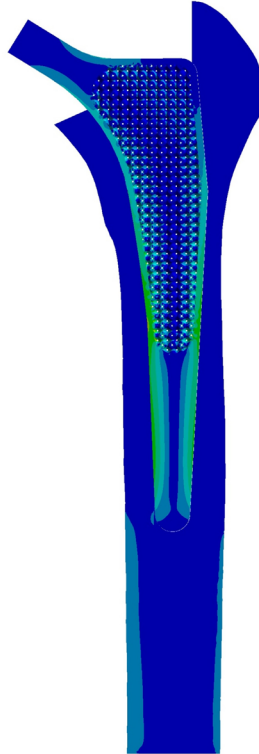
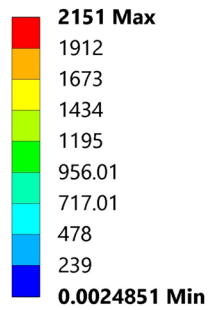
En contraste, esta diferenciación entre configuraciones desaparece al emplear el enfoque de modelado homogeneizado. En este caso, los resultados muestran un comportamiento prácticamente idéntico entre las topologías *diamond* y *primitive* en términos de desplazamientos, vMS y energía de deformación. Las diferencias se reducen a valores marginales, lo que indica que el modelo homogeneizado no captura la influencia de la arquitectura interna del implante, al representar el material como un medio continuo equivalente que elimina el efecto geométrico del reticulado.

Al comparar los modelos estructurados con los homogeneizados, se observa que las diferencias dependen de la variable analizada y de la región del sistema implante-fémur. Los desplazamientos máximos presentan variaciones despreciables entre ambos enfoques, lo que indica que la rigidez global del sistema es capturada de manera similar. En cambio, los vMS muestran discrepancias más relevantes, particularmente en el vástago, donde los modelos estructurados presentan valores considerablemente mayores ($\approx 60 - 70\%$ para *diamond* y $\approx 55 - 60\%$ para *primitive*), mientras que en el fémur las diferencias son más moderadas ($\approx 15 - 20\%$).

En cuanto a la SE, las diferencias entre enfoques son reducidas en el tejido óseo, tanto en la región cortical

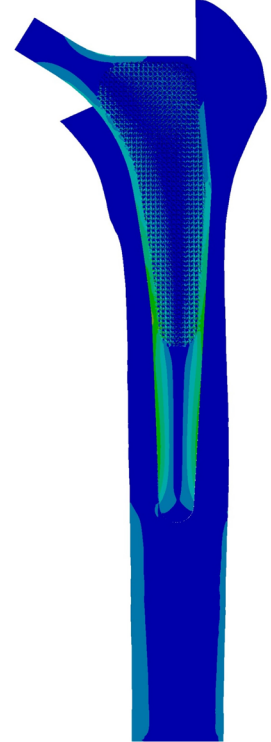
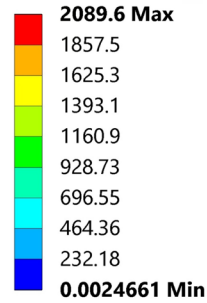
B: PRIMITIVE RD50 3MM

Equivalent Stress 3
Type: Equivalent (von-Mises) Stress
Unit: MPa
Time: 1 s
09/04/2026 06:04 a. m.



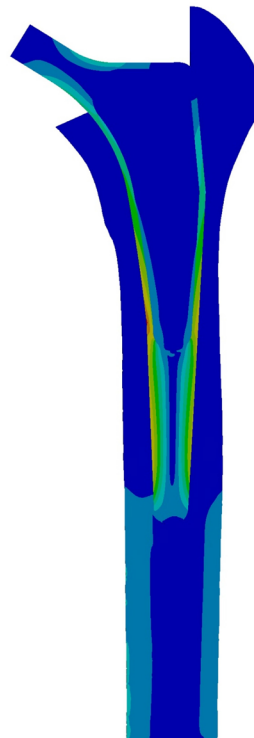
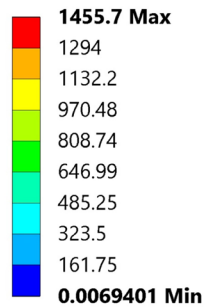
A: DIAMOND RD50 3MM

Equivalent Stress
Type: Equivalent (von-Mises) Stress
Unit: MPa
Time: 1 s
09/04/2026 06:03 a. m.



C: PRIMITIVE RD50 3MM

Equivalent Stress
Type: Equivalent (von-Mises) Stress
Unit: MPa
Time: 1 s
09/04/2026 06:01 a. m.



E: DIAMOND RD50 3MM

Equivalent Stress
Type: Equivalent (von-Mises) Stress
Unit: MPa
Time: 1 s
09/04/2026 06:01 a. m.

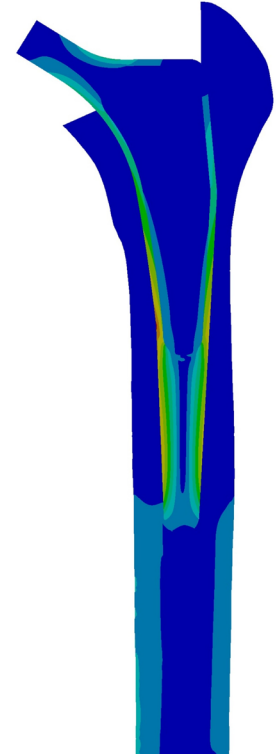
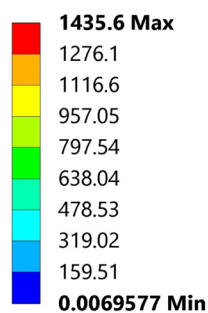


Figura 3. Distribución de vMS en el ensamble implante-fémur para las configuraciones *primitive* y *diamond*. La fila superior corresponde a los modelos estructurados y la fila inferior a los modelos homogeneizados.

como trabecular ($\approx 2 - 5\%$). Sin embargo, en el vástago se observan variaciones significativamente mayores, siendo los modelos homogeneizados los que presentan valores superiores, del orden de ≈ 10 a 50 veces respecto a los modelos estructurados. Este comportamiento indica que el efecto del enfoque de modelado es más relevante en el vástago que en el hueso circundante.

Cabe destacar que la distribución entre hueso cortical y trabecular no mostró una discrepancia dominante entre enfoques, lo que sugiere que la homogeneización afecta de manera global la respuesta del hueso, mientras que en el implante las diferencias se amplifican de forma más notable.

El análisis de la SED permite establecer una relación más directa con el estímulo mecánico en el tejido óseo. A nivel de valores medianos, los modelos estructurados presentan valores de SED superiores en comparación con los modelos homogeneizados (0.26 vs. 0.20 mJ/mm³, reducción del 23%), mientras que en el percentil 95 ambos enfoques arrojan resultados prácticamente equivalentes (diferencias $< 0,3\%$). Esto indica que la homogeneización reproduce adecuadamente los valores máximos de SED, pero subestima el nivel general de estímulo mecánico distribuido en el hueso.

Dado que la SED es un parámetro ampliamente utilizado en modelos de remodelación ósea, estas diferencias sugieren que el uso de modelos homogeneizados puede afectar la estimación del estímulo mecánico asociado a procesos de *stress shielding*. En particular, la reducción en los valores medianos de SED observada en los modelos homogeneizados podría conducir a una subestimación del estímulo mecánico en el hueso y, en consecuencia, a una sobreestimación del riesgo de resorción ósea, con implicaciones directas en la evaluación clínica del diseño del implante, especialmente bajo condiciones representativas de hueso osteoporótico.

Este comportamiento puede observarse de forma cualitativa en la Figura 3, donde los modelos estructurados presentan patrones de distribución de esfuerzos diferenciados entre las configuraciones *primitive* y *diamond*. En particular, se aprecia que cada topología genera un campo de esfuerzos característico, evidenciando que la arquitectura interna del implante modifica la forma en que se transmiten las cargas dentro del sistema implante-fémur.

En la literatura reciente, el diseño de implantes de cadera con estructuras reticulares se ha abordado principalmente mediante enfoques de optimización basados en la distribución de densidad y el uso de parámetros biomecánicos como la SED para evaluar el comportamiento del hueso y el fenómeno de *stress shielding* Wang et al. (2018); Vafaefar et al. (2025). En estos trabajos, la respuesta mecánica del sistema implante-fémur se evalúa comúnmente mediante modelos homogeneizados, en los cuales la microestructura se representa como un medio equivalente continuo Wang et al. (2018); Alkhatib et al. (2019); Müller et al. (2024).

Aunque algunos estudios han incorporado validaciones a nivel microestructural o modelos más detallados, éstos suelen emplearse como herramientas de verificación o en dominios limitados, mientras que el análisis global continúa realizándose bajo aproximaciones homogeneizadas

Alkhatib et al. (2019); Müller et al. (2024). Asimismo, los modelos que consideran la arquitectura interna de forma directa suelen restringirse a configuraciones simplificadas, densidades relativas bajas o tamaños de celda específicos, lo que limita su aplicación en condiciones más representativas de diseño Mehboob et al. (2020); Jetté et al. (2018).

En este contexto, el presente trabajo contribuye al estado del arte al evaluar de forma directa y sistemática el efecto del enfoque de modelado sobre parámetros biomecánicos clave en un sistema implante-fémur con estructuras TPMS de densidad relativa del 50 %, condición poco explorada en la literatura existente.

Sin embargo, aunque la literatura ha demostrado que las estructuras reticulares pueden mejorar la transferencia de carga y el estímulo mecánico en el hueso, en dichos trabajos la representación homogeneizada del material se emplea generalmente bajo condiciones ideales o simplificadas Vafaefar et al. (2025); Mehboob et al. (2020). Los resultados del presente estudio sugieren que, bajo condiciones representativas de hueso osteoporótico, este enfoque puede modificar la estimación del estímulo mecánico, particularmente cuando se emplea la SED como indicador de remodelación ósea. En este sentido, la ausencia de una representación explícita de las propiedades degradadas del hueso podría conducir a interpretaciones poco representativas del comportamiento mecánico del sistema implante-fémur en condiciones clínicas reales.

5. CONCLUSIONES

Los resultados demuestran que la arquitectura interna del implante influye de forma significativa en su respuesta mecánica, evidenciando diferencias entre las configuraciones *primitive* y *diamond* en los modelos estructurados, las cuales desaparecen al emplear el enfoque homogeneizado, confirmando que este último no captura la influencia geométrica del reticulado.

La homogeneización no modifica de forma significativa la rigidez global del sistema, pero sí altera la distribución de esfuerzos y la energía de deformación, con diferencias más marcadas en el vástago que en el hueso. En particular, los modelos homogeneizados presentan menores valores de SED, lo que implica una subestimación del estímulo mecánico en el tejido óseo y, en consecuencia, una posible sobreestimación del riesgo de resorción ósea, especialmente bajo condiciones representativas de hueso osteoporótico.

En este sentido, el enfoque homogeneizado puede ser adecuado para estimar la rigidez global del sistema, pero su uso en el análisis de la transferencia de carga y del estímulo mecánico en implantes con estructuras porosas debe considerarse con cautela, particularmente cuando la SED se emplea como indicador de remodelación ósea y bajo condiciones donde las propiedades mecánicas del hueso se encuentran degradadas.

Estos resultados tienen implicaciones directas para el diseño y evaluación computacional de implantes ortopédicos estructurados en contextos clínicos reales, donde la calidad del hueso juega un papel determinante en la respuesta mecánica del sistema. Como trabajo futuro, se propone cuantificar el impacto de estas diferencias en modelos de remodelación ósea adaptativa y extender el análisis a con-

diciones de carga más representativas del entorno fisiológico, incluyendo actividades de mayor demanda mecánica.

REFERENCIAS

- Abdel, M.P., Von Roth, P., and Harmsen, W.S. (2000). What is the lifetime risk of revision for patients undergoing total hip arthroplasty? *Bone and Joint Journal*, 1436–1440.
- Alkhatib, S.E., Mehboob, H., and Tarlochan, F. (2019). Finite element analysis of porous titanium alloy hip stem to evaluate the biomechanical performance during walking and stair climbing. *Journal of Bionic Engineering*, 16, 1103–1115.
- ANSYS Inc. (2020). ANSYS Workbench. URL <https://www.ansys.com>. Sustituir versión y año por los utilizados en el estudio.
- Arabnejad Khanoki, S. and Pasini, D. (2012). Multiscale design and multiobjective optimization of orthopedic hip implants with functionally graded cellular material. *Journal of Biomechanical Engineering*, 134(3), 1–10.
- Autodesk Inc. (2026). Autodesk Fusion 360. URL <https://www.autodesk.com/products/fusion-360>. Sustituir año por el utilizado en el estudio.
- Bergmann, G., Deuretzbacher, G., Heller, M., Graichen, F., Rohlmann, A., Strauss, J., and Duda, G.N. (2001). Hip contact forces and gait patterns from routine activities. *Journal of Biomechanics*, 34(7), 859–871. doi:10.1016/S0021-9290(01)00040-9. Mediciones in vivo de fuerzas de contacto en la cadera durante actividades diarias.
- Fedorov, A., Beichel, R., Kalpathy-Cramer, J., Finet, J., Fillion-Robin, J.C., Pujol, S., Bauer, C., Jennings, D., Fennessy, F.M., Sonka, M., Buatti, J., Aylward, S.R., Miller, J.V., Pieper, S., and Kikinis, R. (2012). 3D Slicer as an image computing platform for the quantitative imaging network. *Magnetic Resonance Imaging*, 30(9), 1323–1341.
- Jetté, B., Brailovski, V., Dumas, M., Simoneau, C., and Terriault, P. (2018). Femoral stem incorporating a diamond cubic lattice structure: Design, manufacture and testing. *Journal of the Mechanical Behavior of Biomedical Materials*, 77, 58–72.
- Kuiper, J.H. and Huiskes, R. (1992). Numerical optimization of hip-prosthetic stem material. In J. Middleton, G.N. Pande, and K.R. Williams (eds.), *Recent Advances in Computer Methods in Biomechanics and Biomedical Engineering*, 76–84. Books and Journals International, Swansea. Modelo FEM 2D del fémur con optimización del material del vástago y definición clásica de condiciones de carga.
- Mehboob, H., Tarlochan, F., Mehboob, A., Chang, S.H., Ramesh, S., Harun, W.S.W., and Kadirgama, K. (2020). A novel design, analysis and 3D printing of Ti-6Al-4V alloy bio-inspired porous femoral stem. *Journal of Materials Science: Materials in Medicine*, 31, 78.
- Morgan, E.F., Bayraktar, H.H., and Keaveny, T.M. (2003). Trabecular bone modulus-density relationships depend on anatomic site. *Journal of Biomechanics*, 36(7), 897–904. doi:10.1016/S0021-9290(03)00071-X. Relaciones módulo-densidad para hueso trabecular humano; base para estimar módulos aparentes dependientes de la densidad en modelos FEM.
- Müller, M. et al. (2024). Topology optimization of homogenized lattice structures for hip implants. *Scientific Reports*. Completar con volumen, número y páginas.
- Niinomi, M., Liu, Y., Nakai, M., Liu, H., and Li, H. (2016). Biomedical titanium alloys with young's moduli close to that of cortical bone. *Regenerative Biomaterials*, 173–185.
- nTopology Inc. (2023). nTopology. URL <https://ntopology.com>. Sustituir versión y año por los utilizados en el estudio.
- Nugent, M., Young, S.W., Frampton, C.M., and Hooper, G.J. (2021). The lifetime risk of revision following total hip arthroplasty. *Bone and Joint Journal*, 103-B(3), 479–485.
- Rho, J.Y., Kuhn-Spearing, L., and Zioupos, P. (1998). Mechanical properties and the hierarchical structure of bone. *Medical Engineering & Physics*, 20(2), 92–102. doi:10.1016/S1350-4533(98)00007-1. Revisión sobre propiedades mecánicas del hueso y su relación con la estructura jerárquica; incluye valores típicos de módulo de Young para hueso cortical y trabecular.
- Smith, W.C. et al. (2023). Hip, knee and shoulder arthroplasty: 2023 annual report. Adelaide, South Australia.
- The MathWorks Inc. (2023). MATLAB. Sustituir versión y año por los utilizados en el estudio.
- Vafaefar, M., Quinn, C., Moerman, K.M., and Vaughan, T.J. (2025). A computational optimisation study of hip implant using density mapping functionally graded biomimetic tpms-based lattice structures. *npj Metamaterials*, 1(3). doi:10.1038/s44455-025-00003-8.
- Wang, Y., Arabnejad, S., Tanzer, M., and Pasini, D. (2018). Hip implant design with three-dimensional porous architecture of optimized graded density. *Journal of Mechanical Design*, 140, 111406.
- Zebaze, R.M.D., Ghasem-Zadeh, A., Mbala, A., and See-man, E. (2010). Intracortical remodeling and porosity in the distal radius and post-mortem femurs of women. *Journal of Bone and Mineral Research*, 25(5), 1059–1068. doi:10.1002/jbmr.47. Estudio experimental que cuantifica el aumento de la porosidad intracortical en osteoporosis y su impacto en la resistencia mecánica del hueso.

Bioclimatic Evaluation of Thermal Comfort in Morelos Using the Givoni Chart: Comparison Between Historical Data and In-Situ Monitoring

Vázquez, Ch., H.*, Pachecano, L.*, Moya, S. L.*, Ascencio, J. A., Cruz-Sandoval, M.****

* *Tecnológico Nacional de México/Centro Nacional de Investigación y Desarrollo Tecnológico, Interior Internado Palmira S/N, Colonia Palmira, CP 62490, México.*

(e-mail: m24ce094@cenidet.tecnm.mx, luisana.cp@cenidet.tecnm.mx, sara.ma@cenidet.tecnm.mx).

** *Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Santa Fe, Av. Carlos Lazo 100, Santa Fe, La Loma, Álvaro Obregón, CP 01389, CDMX*

(e-mail: jascencio@tec.mx, cruzsandovalmarco@tec.mx)

Abstract: This study evaluates the use of bioclimatic charts as a preliminary tool for assessing the thermal behavior of sites intended for residential or commercial development in the state of Morelos, Mexico. The analysis is initially based on historical climate data obtained from official governmental sources, which are commonly used during early design stages to identify regional thermal conditions and potential comfort challenges throughout the year.

The Givoni bioclimatic chart is applied to characterize different climatic regions within the state, allowing the identification of periods associated with thermal comfort, natural ventilation potential, and overheating risk. While this approach provides a general understanding of the climatic context, it is limited by using average data, which can mask critical thermal extremes, such as unusually high temperatures during heatwaves or extreme cold spells that can significantly impact human comfort and energy consumption.

To address this limitation, in situ monitoring data collected over a full daily cycle in an urban environment are incorporated. Unlike historical averages, these measurements capture real-time variability, including maximum and minimum conditions, providing a more accurate representation of thermal behavior and the interaction between indoor and outdoor environments.

The comparison between both approaches shows that evaluations based solely on historical data may underestimate critical overheating periods. These findings highlight the importance of local monitoring for improving the reliability of bioclimatic assessments and for supporting the selection of effective passive design strategies in warm sub-humid climates.

Keywords: Bioclimatic analysis, thermal comfort, Givoni chart, natural ventilation, in-situ monitoring, historical climate data, urban heat island, and passive design strategies.

1. INTRODUCCIÓN

El incremento del consumo energético en edificaciones, particularmente en regiones de clima cálido, se encuentra asociado al uso intensivo de sistemas de climatización mecánica. Esta situación ha impulsado el desarrollo de estrategias pasivas orientadas a mejorar el confort térmico interior sin incrementar la demanda energética, en concordancia con recomendaciones internacionales para el diseño eficiente de edificios (ASHRAE, 2020). Entre estas estrategias, la ventilación natural representa una de las alternativas más utilizadas; sin embargo, su efectividad depende directamente de variables climáticas locales como la temperatura del aire, la humedad relativa y la disponibilidad de viento (Nicol & Humphreys, 2002).

En el estado de Morelos predominan condiciones de clima cálido subhúmedo, con temperaturas superiores a 30 °C durante los periodos más cálidos y velocidades de viento generalmente bajas (Gobierno del Estado de Morelos, 2018). Estas condiciones, combinadas con procesos de urbanización y densificación constructiva, favorecen la acumulación de calor y la presencia de episodios de sobrecalentamiento, particularmente en entornos urbanos.

Para el análisis climático aplicado al diseño pasivo, la carta bioclimática de Givoni ha sido ampliamente utilizada como herramienta de interpretación térmica en arquitectura y diseño ambiental (Givoni, 1992, 1998). Este enfoque relaciona temperatura y humedad relativa con diferentes estrategias de acondicionamiento pasivo, permitiendo identificar zonas de confort térmico, ventilación natural, enfriamiento evaporativo y sobrecalentamiento. Su aplicación resulta especialmente útil en etapas tempranas de diseño; sin embargo, normalmente se basa en datos climáticos históricos expresados mediante promedios mensuales o anuales, los cuales no representan necesariamente las condiciones térmicas extremas presentes en el entorno construido.

Diversos estudios sobre confort adaptativo han demostrado que la percepción térmica de los ocupantes depende de la variabilidad temporal de las condiciones ambientales y no únicamente de valores promedio (Nicol & Humphreys, 2002). De manera complementaria, guías internacionales de evaluación térmica establecen que la duración e intensidad de los eventos de sobrecalentamiento son factores críticos para evaluar el desempeño ambiental de las edificaciones (CIBSE, 2013).

En este contexto, el monitoreo in situ permite capturar la variabilidad térmica real del entorno construido, incluyendo

condiciones máximas y mínimas dentro del ciclo diario, proporcionando una representación más precisa del comportamiento térmico local y de la interacción entre el ambiente urbano y las edificaciones (De Buen, 2020).

La Figura 1 muestra la carta bioclimática base propuesta por Baruch Givoni (1992, 1998), fundamentada en la relación entre la temperatura de bulbo seco y la humedad relativa dentro del diagrama psicrométrico. A partir de estas variables se delimitan regiones asociadas a confort térmico y distintas estrategias pasivas de acondicionamiento ambiental, como ventilación natural, enfriamiento evaporativo, masa térmica y ventilación nocturna.

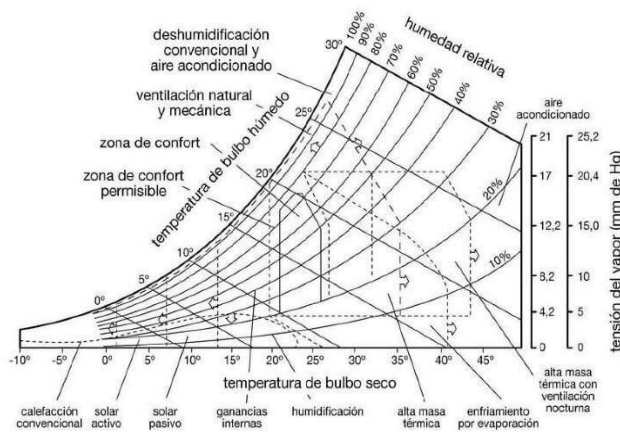


Figura 1: Carta bioclimática base propuesta por Givoni (1992, 1998), utilizada como referencia para la identificación de zonas de confort térmico y estrategias pasivas de acondicionamiento ambiental.).

A partir de lo anterior, el objetivo de esta investigación fue evaluar el comportamiento térmico de las condiciones climáticas del estado de Morelos mediante la carta bioclimática de Givoni, definiendo zonas bioclimáticas específicas para la región y contrastando los resultados obtenidos a partir de datos históricos con información derivada de un monitoreo térmico in situ realizado en una vivienda ubicada en un entorno urbano del sur del estado.

2. METODOLOGÍA

El estudio se estructuró en dos etapas principales. En una primera etapa se emplearon datos climáticos históricos de carácter gubernamental para caracterizar las condiciones térmicas regionales y su representación en la carta bioclimática de Givoni. En una segunda etapa, estos resultados fueron contrastados con datos obtenidos mediante un monitoreo in situ de alta resolución temporal, permitiendo identificar diferencias entre una base de datos históricos y un monitoreo específico y completo de un caso de estudio.

2.1 Área de estudio y caracterización climática

Como parte de la caracterización climática del estado de Morelos, distintas dependencias gubernamentales han definido zonas climáticas regionales (Figura 2), permitiendo identificar

la distribución espacial de los principales regímenes térmicos asociados a factores como altitud y topografía.

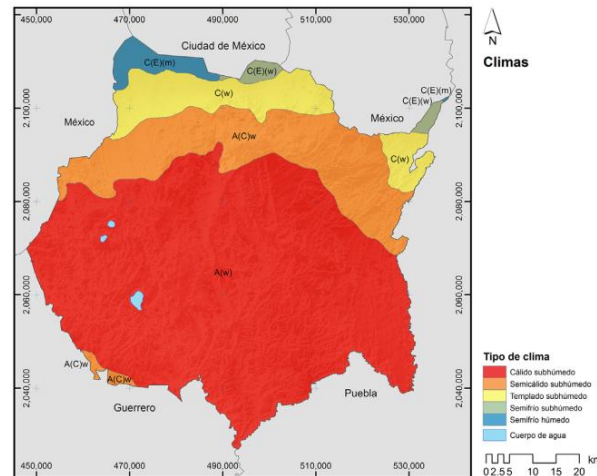


Figura 2: Tipo de clima en el estado de Morelos. Fuente: CINPRO (2017), con información de INEGI (2013). Base Histórica.

En Morelos predominan climas cálidos subhúmedos en regiones de menor altitud, mientras que en las zonas montañosas del norte se presentan climas templados subhúmedos y semifríos. Esta variabilidad climática influye directamente en el comportamiento térmico regional y en la efectividad de estrategias pasivas como la ventilación natural. De acuerdo con el Programa de Gestión para Mejorar la Calidad del Aire del Estado de Morelos (ProAire, 2018), en la entidad predominan tres regiones climáticas principales: cálida, semicálida y templada. La Tabla 1 presenta una síntesis de temperatura media, humedad relativa y rangos térmicos representativos para distintas localidades del estado.

Región	Localidades representativas	T media (°C)	HR media (%)	T máx (°C)	T mín (°C)
Cálida	Jojutla, Cuautla	24–26	45–60	34–38	14–18
Semicálida	Cuervavaca, Jonacatepec	20–23	50–65	30–34	10–14
Templada	Huitzilac, Tetela del Volcán	12–18	65–80	22–26	4–8

Tabla 1: Síntesis de datos climáticos. Fuente: Elaboración propia con base en INEGI (2022) y ProAire Morelos (2018).

Los datos presentados permitieron establecer una base representativa de las condiciones térmicas regionales utilizadas posteriormente en el análisis bioclimático mediante la carta de Givoni, facilitando la identificación de tendencias generales de confort térmico y riesgo de sobrecalentamiento. Respecto al comportamiento del viento, Uribe Indalecio (2018) desarrolló un análisis del recurso eólico estatal mediante un modelo numérico de mesoescala con resolución espacial de 3 km, utilizando condiciones de frontera derivadas del modelo GFS y extrapolación vertical mediante la ley de Hellmann ($\alpha = 0.13$).

Los resultados muestran que las mayores velocidades medias de viento se concentran en la región noroeste del estado, particularmente en municipios como Atlatláhuca, Totolapan

y Yecapixtla, con velocidades promedio cercanas a 2.7 m/s. Asimismo, se identificaron direcciones dominantes N–NW y SE, así como una limitada viabilidad eólica en otras regiones del estado debido a velocidades medias reducidas.

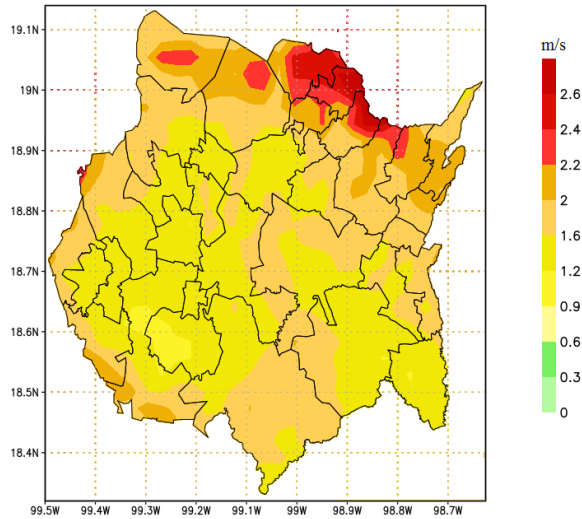


Figura 3 Velocidad del viento promedio anual en los municipios de Morelos, México. Uribe (2018).

Tetecala exhibe ráfagas muy altas (22 m/s) pero poco frecuentes, por lo que su velocidad media 1.15 m/s limita el recurso.

Existen gráficos (p. ej., Atlatlahucan: 62% de horas ≥ 2 m/s) y rosas que confirman direcciones dominantes N–NW/SE en zonas con mejor potencial. El trabajo concluye que la región NW es la única con viabilidad preliminar para sistemas de generación eólica a pequeña escala (≤ 20 kW) y recomienda campañas anemométricas in situ para validar las proyecciones antes de decisiones.

2.3 Monitoreo de datos experimentales para caso de estudio

El complejo habitacional cuenta aproximadamente con 2,250 viviendas de características constructivas similares, lo que permite considerar el caso de estudio como representativo de un modelo ampliamente replicado en la región. La metodología de monitoreo y adquisición de datos se basó en procedimientos previamente desarrollados para caracterización ambiental mediante sensores IoT en microescala urbana (Vázquez Chávez et al., 2025).

Las mediciones se realizaron a una altura aproximada de 1.8 m sobre el nivel del suelo utilizando sensores de temperatura y humedad relativa con precisión nominal de ± 0.5 °C y ± 3 % HR, respectivamente. Los registros fueron adquiridos cada 60 s durante 24 horas (Vázquez Chávez et al., 2025).

2.4 Caracterización de zonas bioclimáticas en la carta de Givoni para el estado de Morelos.

La definición de las zonas bioclimáticas se basó en los planteamientos originales de Givoni (1992, 1998) y su aplicación al diseño arquitectónico en climas cálidos

subhúmedos. A partir de este enfoque, se establecieron rangos de operación térmica adaptados a las condiciones climáticas predominantes del estado de Morelos.

➤ Zona de confort térmico

La zona de confort térmico se definió entre 22 y 26 °C, con una extensión adaptativa hasta aproximadamente 29 °C y humedades relativas entre 30 y 65 %. Este rango representa condiciones aceptables de confort para edificaciones ventiladas naturalmente (Givoni, 1998; Nicol & Humphreys, 2002; Fernández et al., 2010).

2. Zona de ventilación natural

Las temperaturas entre 26 y 32 °C y humedades relativas entre 20 y 35 % corresponden a condiciones donde el incremento de la velocidad del aire favorece la disipación de calor mediante mecanismos convectivos y evaporativos (Nicol & Humphreys, 2002; Sánchez Montes, 2021).

➤ Zona de masa térmica y ventilación nocturna

El intervalo de 20 a 26 °C y humedades relativas entre 20 y 55 % se asoció con estrategias de masa térmica y ventilación nocturna orientadas a disipar el calor acumulado en la envolvente durante periodos nocturnos (Givoni, 1992, 1998; Bastide et al., 2006).

➤ Zona de enfriamiento evaporativo

El rango de 27 a 36 °C y humedades relativas entre 35 y 75 % define condiciones favorables para estrategias de enfriamiento evaporativo pasivo, aprovechando la diferencia entre temperatura de bulbo seco y bulbo húmedo (Givoni, 1998; Cárdenas Rodríguez et al., 2023; García, 2024).

➤ Zona de sobrecalentamiento y enfriamiento asistido

Las condiciones superiores a 30 °C fueron clasificadas como fuera del rango de confort térmico, asociadas a riesgo de sobrecalentamiento y necesidad de estrategias adicionales de enfriamiento (CIBSE, 2013, 2017; World Health Organization, 2018).

❖ Zonas complementarias para condiciones de baja temperatura para la región norte del estado.

Con el objetivo de representar regiones de mayor altitud del estado, se incorporaron zonas complementarias asociadas a estrategias de calentamiento pasivo y protección térmica.

❖ Zona de ganancia solar pasiva

Se definió un intervalo entre 10 y 18 °C y humedades relativas entre 35 y 70 %, donde la radiación solar puede emplearse como fuente pasiva de calentamiento (Olgay, 1963; Givoni, 1998).

❖ Zona de masa térmica para calefacción

El rango de 14 a 20 °C y humedades relativas entre 30 y 60 % corresponde a condiciones donde la inercia térmica contribuye al almacenamiento y liberación gradual de calor (Givoni, 1998; De Buen, 2020).

❖ Zona de aislamiento térmico

La franja entre 9 y 16 °C y humedades relativas entre 25 y 55 % representa condiciones donde el aislamiento térmico resulta fundamental para reducir pérdidas energéticas de la envolvente (ASHRAE, 2020).

❖ Zona de protección frente a ventilación fría

El intervalo de 9 a 15 °C y humedades relativas entre 35 y 65 % corresponde a condiciones donde la infiltración y el movimiento del aire incrementan las pérdidas de calor corporal, requiriendo control de ventilación (ASHRAE, 2020). A partir de la definición de estas zonas bioclimáticas se estableció la base de clasificación utilizada para representar las condiciones climáticas del estado de Morelos dentro de la carta de Givoni, permitiendo identificar regiones asociadas a confort térmico, sobrecalentamiento y estrategias pasivas de acondicionamiento ambiental.

3. PROCESO DE CÓDIGO - R STUDIO

El procesamiento de la información y la generación de la carta bioclimática se realizaron mediante el entorno de programación RStudio, integrando datos climáticos históricos y registros obtenidos mediante monitoreo térmico in situ. Este enfoque permitió contrastar el comportamiento climático regional con condiciones reales del entorno construido.

En una primera etapa, se emplearon datos históricos de temperatura del aire y humedad relativa provenientes de fuentes oficiales para distintas estaciones meteorológicas representativas del estado de Morelos. Los registros fueron depurados, organizados y estandarizados para su procesamiento dentro del plano psicrométrico, siguiendo criterios utilizados en análisis climatológicos (ASHRAE, 2021).

Para la representación psicrométrica se calcularon la presión de saturación del vapor de agua y la humedad específica del aire, utilizada como variable principal debido a que representa de forma más precisa el contenido de humedad atmosférica en estudios de confort térmico (ASHRAE, 2021; Moran et al., 2018). A partir de estos cálculos se construyó la malla psicrométrica base, incluyendo curvas de humedad relativa constante y la curva de saturación, la cual define el límite termodinámico del sistema (Givoni, 1998).

Posteriormente, se incorporaron las zonas bioclimáticas de la carta de Givoni mediante polígonos definidos dentro del espacio psicrométrico, permitiendo identificar regiones asociadas a confort térmico, ventilación natural, enfriamiento evaporativo, masa térmica y sobrecalentamiento (Givoni, 1992, 1998; Fernández et al., 2010).

Los datos obtenidos mediante monitoreo y los registros climáticos históricos fueron sometidos a procesos de limpieza, normalización y verificación de consistencia temporal, incluyendo corrección de formatos, depuración de registros atípicos y estandarización de variables, con el objetivo de preservar la continuidad y coherencia de las series analizadas. Finalmente, tanto los datos históricos como los registros de monitoreo fueron integrados bajo un mismo formato gráfico dentro del espacio psicrométrico, permitiendo comparar el comportamiento promedio regional con las condiciones térmicas reales del entorno urbano e identificar periodos críticos asociados a condiciones de sobrecalentamiento (CIBSE, 2013; Nicol & Humphreys, 2002).

4. RESULTADOS

Con base en el procesamiento de datos y la implementación del modelo psicrométrico, se obtuvieron las representaciones de la carta bioclimática correspondientes tanto a los datos históricos como al monitoreo térmico in situ. Las Figuras 4 y 5 muestran ambas aproximaciones, permitiendo comparar la distribución de las condiciones térmicas dentro del espacio bioclimático.

La representación de los datos climáticos históricos mediante la carta bioclimática de Givoni permitió identificar el comportamiento térmico promedio de distintas regiones del estado de Morelos. En la Figura 4 se observa que la mayoría de las trayectorias mensuales se concentran entre 20 °C y 28 °C, ubicándose principalmente en regiones asociadas a confort térmico, ventilación natural y masa térmica con ventilación nocturna.

No obstante, estaciones como Cuautla y Jojutla presentan desplazamientos hacia regiones de mayor temperatura durante los meses más cálidos, alcanzando zonas cercanas al enfriamiento evaporativo y sobrecalentamiento. En contraste, estaciones de mayor altitud, como Huitzilac y Tetela del Volcán, muestran trayectorias asociadas a menores temperaturas y posibles requerimientos de ganancia solar pasiva y masa térmica para calefacción.

Por otro lado, la representación del monitoreo térmico in situ (Figura 5) evidencia una dinámica térmica distinta a la observada en los datos históricos. Durante las primeras horas del día, las condiciones se ubican próximas a las zonas de confort y masa térmica; sin embargo, conforme avanza el día, se observa un desplazamiento progresivo hacia ventilación natural y posteriormente hacia condiciones de sobrecalentamiento entre las 13:00 y 16:00 horas.

La representación horaria permitió identificar periodos específicos fuera del rango de confort y una recuperación térmica gradual durante la tarde y noche, comportamiento que no puede observarse mediante datos climáticos promedio.

En conjunto, los resultados muestran que las representaciones basadas únicamente en datos históricos tienden a concentrarse en regiones cercanas al confort térmico, mientras que el monitoreo in situ evidencia variaciones térmicas significativas dentro del ciclo diario, incluyendo periodos definidos de sobrecalentamiento.



Carta bioclimática de Givoni para estaciones representativas de Morelos

Base de datos: INEGI / SEMARNAT

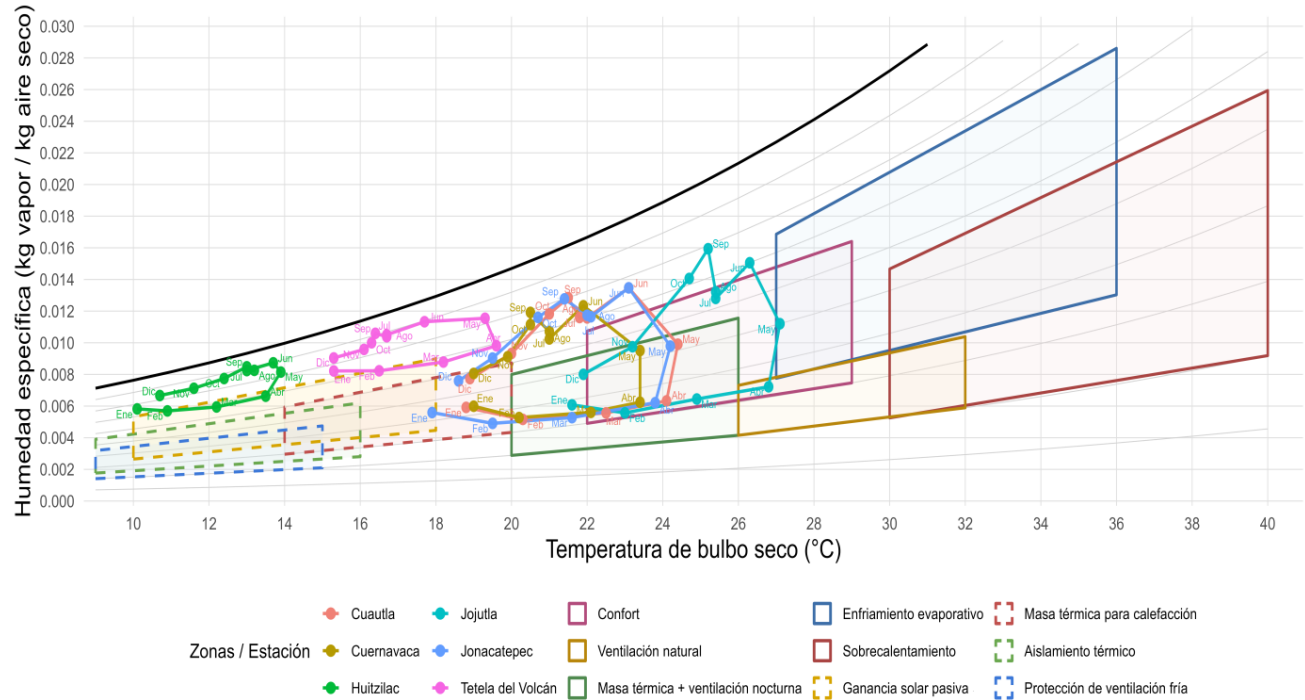


Figura 4 Carta bioclimática de Givoni adaptada a Morelos, México. Aplicación de datos promedio mensuales gubernamentales .

Carta bioclimática de Givoni para monitoreo térmico in situ

Cuarto crítico: interior segundo piso | 17 de mayo de 2025

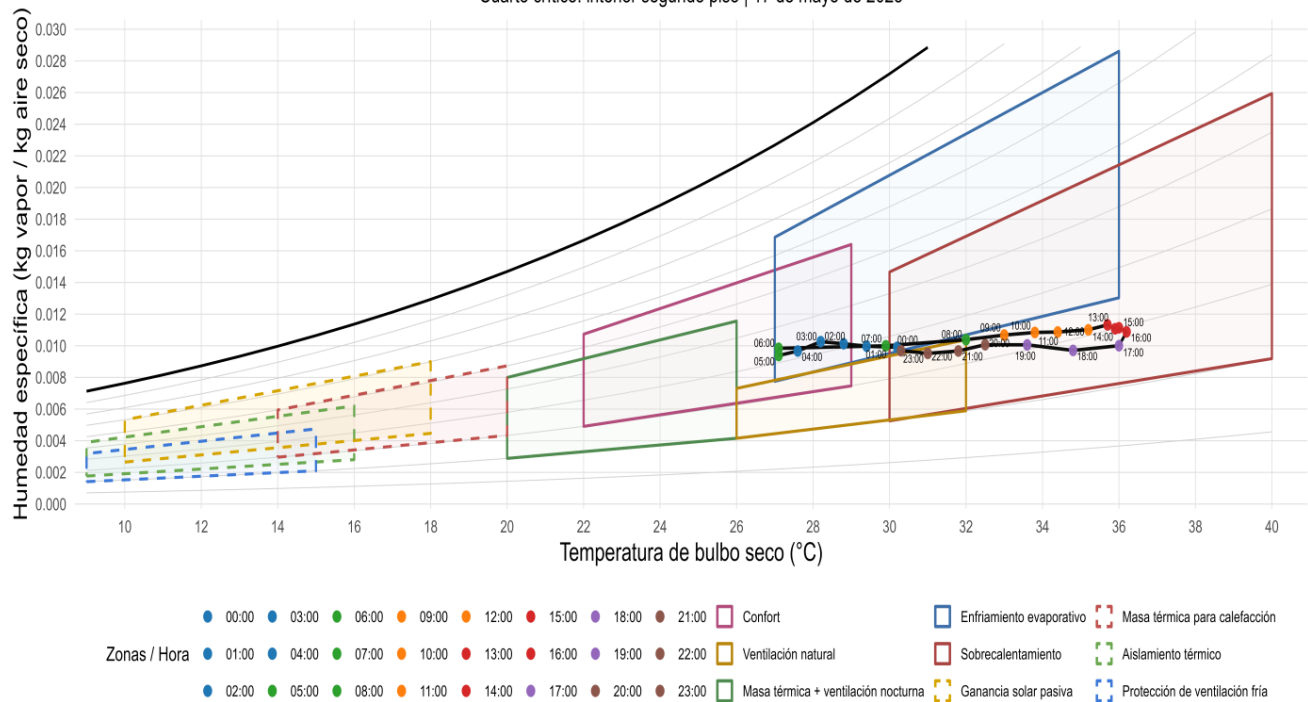


Figura 4 Carta bioclimática de Givoni adaptada a Morelos, México. Aplicación de datos obtenidos a partir de un monitoreo térmico *in situ*, correspondiente a un ciclo diario completo en el jardín de una vivienda de interés social ubicada en Emiliano Zapata, Morelos.



5. DISCUSIONES

Las diferencias observadas entre la caracterización basada en datos históricos y el monitoreo térmico in situ pueden explicarse a partir de la escala de análisis y de las condiciones específicas del entorno urbano donde se realizaron las mediciones. Mientras que los registros históricos representan condiciones promedio regionales, el monitoreo refleja el comportamiento térmico real de un entorno construido con influencia directa de superficies urbanas y limitada cobertura vegetal.

La presencia de materiales constructivos expuestos a radiación solar y la acumulación de calor en superficies urbanas favorecen fenómenos asociados a isla de calor, incrementando las temperaturas locales y modificando la dinámica térmica del aire. Esto permite explicar la presencia de valores máximos más elevados y una recuperación térmica más lenta en comparación con la representación derivada de datos promedio. Adicionalmente, la representación horaria del monitoreo permitió identificar periodos específicos de sobrecalentamiento dentro del ciclo diario, así como la evolución temporal de las condiciones térmicas, información que no puede obtenerse directamente mediante registros climáticos promedio.

En conjunto, los resultados evidencian que la interpretación de la carta bioclimática debe considerar las condiciones particulares del sitio de estudio, especialmente en entornos urbanos, donde el comportamiento térmico real puede diferir significativamente de las referencias climáticas regionales.

6. CONCLUSIONES

El presente estudio permitió contrastar el análisis bioclimático basado en datos climáticos históricos con información obtenida mediante monitoreo térmico in situ. Los resultados evidenciaron que, aunque los datos históricos permiten establecer una referencia general del comportamiento climático regional, su carácter promedio limita la identificación de condiciones críticas relacionadas con el confort térmico.

El monitoreo realizado en un entorno urbano permitió identificar la dinámica térmica real del ambiente, incluyendo variaciones máximas y mínimas dentro del ciclo diario, así como periodos específicos de sobrecalentamiento asociados a factores locales como acumulación de calor en superficies construidas y efectos de isla de calor urbana.

A partir de este contraste, se reconoce que la evaluación bioclimática del confort térmico requiere complementar los registros climáticos promedio con mediciones reales in situ, particularmente en entornos urbanos donde las condiciones térmicas presentan una variabilidad significativa.

Finalmente, la integración de herramientas bioclimáticas con monitoreo térmico permite una representación más precisa del comportamiento ambiental, facilitando la selección de estrategias pasivas más adecuadas para viviendas ubicadas en climas cálidos subhúmedos.

REFERENCIAS

- ASHRAE. (2020). *ASHRAE Handbook—HVAC Systems and Equipment*. American Society of Heating, Refrigerating and Air-Conditioning Engineers.
- ASHRAE. (2021). *ASHRAE Handbook—Fundamentals*. American Society of Heating, Refrigerating and Air-Conditioning Engineers.
- Bastide, A., Lauret, P., Garde, F., & Boyer, H. (2006). Building energy efficiency and thermal comfort in tropical climates. *Energy and Buildings*, 38(9), 1093–1103.
- Cárdenas-Rodríguez, Roberto & García, Luis & Esparza López, Carlos Javier & Peña, José & Arellano, José. (2023). *Sistemas de Enfriamiento Evaporativo para uso en elementos verticales: una revisión del estado del arte*. Ciencia Nicolaita.
- CIBSE. (2013). *Guide A: Environmental Design*. Chartered Institution of Building Services Engineers.
- CIBSE. (2017). *TM52: The Limits of Thermal Comfort: Avoiding Overheating in European Buildings*. Chartered Institution of Building Services Engineers.
- De Buen, O. (2020). Confort térmico y eficiencia energética en edificaciones. *Mundo HVAC&R*.
- Fernández Zayas, J. L., Martínez, F., & Hernández, A. (2010). Confort térmico y ahorro de energía en la vivienda económica en México. Universidad Nacional Autónoma de México.
- García, L. (2024). Evaluación del desempeño del enfriamiento evaporativo en climas cálidos. *Revista de Ingeniería Térmica*
- Givoni, B. (1992). Comfort, climate analysis and building design guidelines. *Energy and Buildings*, 18(1), 11–23.
- Givoni, B. (1998). *Climate considerations in building and urban design*. John Wiley & Sons.
- Gobierno del Estado de Morelos. (2018). *Programa de Gestión para Mejorar la Calidad del Aire del Estado de Morelos (ProAire)*.
- INEGI. (2013). *Conjunto de datos de uso de suelo y vegetación*. Instituto Nacional de Estadística y Geografía.
- INEGI. (2022). *Estadísticas climatológicas del estado de Morelos*. Instituto Nacional de Estadística y Geografía.
- Moran, M. J., Shapiro, H. N., Boettner, D. D., & Bailey, M. B. (2018). *Fundamentals of Engineering Thermodynamics* (9th ed.). Wiley.
- Nicol, J. F., & Humphreys, M. A. (2002). Adaptive thermal comfort and sustainable thermal standards for buildings. *Energy and Buildings*, 34(6), 563–572.
- Sánchez Montes, J. (2021). *Evaluación de estrategias pasivas de ventilación en vivienda en clima cálido subhúmedo*. Tesis de Maestría. (CENIDET)
- World Health Organization. (2018). *Housing and health guidelines*. WHO.
- CINPRO. (2017). *Mapa climático del estado de Morelos con base en datos de INEGI (2013)*.



Vázquez-Chávez, H., Pachecano, L., Moya, S. L., Ascencio, J. A., & Cruz-Sandoval, M. (2025). Micro-scale wind assessment with commercial UAVs and IoT sensors for comfort and environmental management. *Tecnología y Ciencia Aplicada*, 8(2).

ANEXO 1. PROMEDIOS GENERALES – ESTACIONES METEOROLÓGICAS EN MORELOS (BASE INEGI SEMARNAR)

Localidad	Altitud (m)	T media (°C)	HR media (%)	Rango térmico (°C)	Rango (HR (%))	Mes más cálido	T máx (°C)	HR mín. (%)	Mes más frío	T mín (°C)	HR máx. (%)	Precipitación (mm)	Tipo de clima
Cuernavaca	1,100 – 3,000	21	55 – 65	12 – 24	45 – 75	Mayo	30–32	40–50	Enero	10–12	70–80	800–1,600	Semicaldo / templado
Cuautla	1,200	24	50 – 60	20 – 24	40 – 70	Mayo	34–36	35–45	Enero	14–16	65–75	800–1,000	Cálido subhúmedo
Jojutla	900	25	45 – 55	24 – 26	35 – 65	Mayo	36–38	30–40	Enero	16–18	60–70	800–1,000	Cálido subhúmedo
Jonacatepec	1,500	23	50 – 60	18 – 24	40 – 70	Mayo	32–34	35–45	Enero	12–14	65–75	800–1,000	Cálido subhúmedo
Tetela del Volcán	2,200	16	65 – 75	12 – 20	55 – 85	Mayo	24–26	50–60	Enero	6–8	80–90	1,000–1,500	Templado subhúmedo
Huitzilac	2,500	14	70 – 80	10 – 18	60 – 90	Mayo	22–24	55–65	Enero	4–6	85–95	1,200–1,600	Templado / semifrio húmedo

ANEXO 2 : DATOS PROMEDIOS MENSUALES

Estación	Mes	Altitud_m	Tipo_clima	Temp_C	Precip_mm	HR_pct	Psat_kPa	W_kg_vapor_kg_aire_seco
Cuautla	Ene	1320	Cálido subhúmedo	18.8	8.9	44.0	2.170	0.0059
Cuautla	Feb	1320	Cálido subhúmedo	20.3	3.0	35.0	2.382	0.0052
Cuautla	Mar	1320	Cálido subhúmedo	22.5	4.3	33.0	2.725	0.0056
Cuautla	Abr	1320	Cálido subhúmedo	24.1	11.2	34.0	3.002	0.0063
Cuautla	May	1320	Cálido subhúmedo	24.4	62.8	52.0	3.056	0.0099
Cuautla	Jun	1320	Cálido subhúmedo	23.1	192.4	76.0	2.826	0.0135
Cuautla	Jul	1320	Cálido subhúmedo	22.0	171.4	70.0	2.644	0.0116
Cuautla	Ago	1320	Cálido subhúmedo	21.8	170.1	71.0	2.612	0.0116
Cuautla	Sep	1320	Cálido subhúmedo	21.5	185.7	80.0	2.564	0.0129
Cuautla	Oct	1320	Cálido subhúmedo	21.0	67.0	76.0	2.487	0.0118
Cuautla	Nov	1320	Cálido subhúmedo	20.0	14.2	64.0	2.338	0.0093
Cuautla	Dic	1320	Cálido subhúmedo	18.9	3.0	57.0	2.184	0.0077
Jojutla	Ene	915	Cálido subhúmedo	21.6	6.9	38.0	2.580	0.0061
Jojutla	Feb	915	Cálido subhúmedo	23.0	2.4	32.0	2.809	0.0056
Jojutla	Mar	915	Cálido subhúmedo	24.9	3.5	33.0	3.149	0.0064
Jojutla	Abr	915	Cálido subhúmedo	26.8	8.5	33.0	3.524	0.0072
Jojutla	May	915	Cálido subhúmedo	27.1	58.6	50.0	3.586	0.0112
Jojutla	Jun	915	Cálido subhúmedo	26.3	196.6	70.0	3.421	0.0151
Jojutla	Jul	915	Cálido subhúmedo	25.4	175.0	63.0	3.244	0.0128
Jojutla	Ago	915	Cálido subhúmedo	25.4	182.5	65.0	3.244	0.0132
Jojutla	Sep	915	Cálido subhúmedo	25.2	178.3	79.0	3.206	0.0159
Jojutla	Oct	915	Cálido subhúmedo	24.7	58.0	72.0	3.112	0.0141
Jojutla	Nov	915	Cálido subhúmedo	23.2	8.0	55.0	2.844	0.0098
Jojutla	Dic	915	Cálido subhúmedo	21.9	2.6	49.0	2.628	0.0080
Jonacatepec	Ene	1165	Cálido subhúmedo	17.9	8.7	44.0	2.051	0.0056
Jonacatepec	Feb	1165	Cálido subhúmedo	19.5	4.3	35.0	2.267	0.0049
Jonacatepec	Mar	1165	Cálido subhúmedo	21.6	7.4	33.0	2.580	0.0053
Jonacatepec	Abr	1165	Cálido subhúmedo	23.8	11.8	34.0	2.948	0.0062
Jonacatepec	May	1165	Cálido subhúmedo	24.2	66.1	52.0	3.020	0.0098
Jonacatepec	Jun	1165	Cálido subhúmedo	23.1	208.7	76.0	2.826	0.0135



Jonacatepec	Jul	1165	Cálido subhúmedo	22.1	165.2	70.0	2.660	0.0116
Jonacatepec	Ago	1165	Cálido subhúmedo	22.0	203.9	71.0	2.644	0.0117
Jonacatepec	Sep	1165	Cálido subhúmedo	21.4	197.3	80.0	2.549	0.0128
Jonacatepec	Oct	1165	Cálido subhúmedo	20.7	64.2	76.0	2.441	0.0116
Jonacatepec	Nov	1165	Cálido subhúmedo	19.5	16.0	64.0	2.267	0.0090
Jonacatepec	Dic	1165	Cálido subhúmedo	18.6	3.1	57.0	2.143	0.0076
Cuernavaca	Ene	1653	Semicálido subhúmedo / templado	19.0	10.4	44.0	2.197	0.0060
Cuernavaca	Feb	1653	Semicálido subhúmedo / templado	20.2	5.7	36.0	2.367	0.0053
Cuernavaca	Mar	1653	Semicálido subhúmedo / templado	22.1	6.8	34.0	2.660	0.0056
Cuernavaca	Abr	1653	Semicálido subhúmedo / templado	23.4	15.3	35.0	2.878	0.0062
Cuernavaca	May	1653	Semicálido subhúmedo / templado	23.4	63.1	53.0	2.878	0.0095
Cuernavaca	Jun	1653	Semicálido subhúmedo / templado	21.9	252.7	75.0	2.628	0.0123
Cuernavaca	Jul	1653	Semicálido subhúmedo / templado	21.0	242.9	66.0	2.487	0.0102
Cuernavaca	Ago	1653	Semicálido subhúmedo / templado	21.0	254.4	69.0	2.487	0.0107
Cuernavaca	Sep	1653	Semicálido subhúmedo / templado	20.5	255.9	79.0	2.412	0.0119
Cuernavaca	Oct	1653	Semicálido subhúmedo / templado	20.5	98.9	74.0	2.412	0.0112
Cuernavaca	Nov	1653	Semicálido subhúmedo / templado	19.9	18.6	63.0	2.324	0.0091
Cuernavaca	Dic	1653	Semicálido subhúmedo / templado	19.0	4.2	59.0	2.197	0.0081
Tetela del Volcán	Ene	2129	Templado subhúmedo	15.3	9.7	76.0	1.739	0.0082
Tetela del Volcán	Feb	2129	Templado subhúmedo	16.5	6.2	70.5	1.877	0.0082
Tetela del Volcán	Mar	2129	Templado subhúmedo	18.2	5.6	67.5	2.090	0.0088
Tetela del Volcán	Abr	2129	Templado subhúmedo	19.6	23.1	69.0	2.281	0.0098
Tetela del Volcán	May	2129	Templado subhúmedo	19.3	86.9	82.5	2.239	0.0115
Tetela del Volcán	Jun	2129	Templado subhúmedo	17.7	218.2	89.5	2.025	0.0113
Tetela del Volcán	Jul	2129	Templado subhúmedo	16.7	195.0	88.5	1.901	0.0105
Tetela del Volcán	Ago	2129	Templado subhúmedo	16.7	203.2	87.5	1.901	0.0104
Tetela del Volcán	Sep	2129	Templado subhúmedo	16.4	210.1	91.0	1.865	0.0106
Tetela del Volcán	Oct	2129	Templado subhúmedo	16.3	84.6	86.5	1.853	0.0100
Tetela del Volcán	Nov	2129	Templado subhúmedo	16.1	19.2	84.0	1.830	0.0096
Tetela del Volcán	Dic	2129	Templado subhúmedo	15.3	3.9	83.5	1.739	0.0090
Huitzilac	Ene	2612	Semifrio húmedo / subhúmedo	10.1	20.8	76.0	1.236	0.0058
Huitzilac	Feb	2612	Semifrio húmedo / subhúmedo	10.9	12.9	70.5	1.304	0.0057
Huitzilac	Mar	2612	Semifrio húmedo / subhúmedo	12.2	17.9	67.5	1.421	0.0059
Huitzilac	Abr	2612	Semifrio húmedo / subhúmedo	13.5	32.5	69.0	1.547	0.0066
Huitzilac	May	2612	Semifrio húmedo / subhúmedo	13.9	85.2	82.5	1.588	0.0081
Huitzilac	Jun	2612	Semifrio húmedo / subhúmedo	13.7	287.6	89.5	1.568	0.0087
Huitzilac	Jul	2612	Semifrio húmedo / subhúmedo	13.0	332.4	88.5	1.498	0.0082
Huitzilac	Ago	2612	Semifrio húmedo / subhúmedo	13.2	361.4	87.5	1.517	0.0083
Huitzilac	Sep	2612	Semifrio húmedo / subhúmedo	13.0	288.4	91.0	1.498	0.0085
Huitzilac	Oct	2612	Semifrio húmedo / subhúmedo	12.4	111.6	86.5	1.440	0.0077
Huitzilac	Nov	2612	Semifrio húmedo / subhúmedo	11.6	30.0	84.0	1.366	0.0071

Numerical Analysis of Mixing in Lug-Bolt Burners in Transition from Sub-critical to Trans-critical Pressure

M. De la Cruz-Ávila*, Fatima E. Carbajal-Arriaga**, Gustavo Arvizu-Corona***.

*Tecnológico Nacional de México-TES de Coacalco, Ingeniería Industrial CP 57700 México.
(e-mail: mauriciodlca@gmail.com).

**Tecnológico Nacional de México-TES de Coacalco, Ingeniería Química CP 57700 México.
(e-mail: 202220792@tesco.edu.mx).

***Tecnológico Nacional de México-TES de Coacalco, Ingeniería Química CP 57700 México.
(e-mail: gustav2247@gmail.com).

Abstract: Combustion during the transition from sub-critical to trans-critical pressure regimes remains underexplored in the open literature. This study aims to analyze the mixing process in central-peripheral (Lug-bolt) burners with inverse diffusion flames using numerical simulations. ANSYS is employed to evaluate flammability limits associated with mixture density, considering turbulence models only and excluding combustion models. The results identify stability and mixing zones within the domain, providing relevant insights for improving mixing performance and supporting more energy-efficient burner operation. Given the complexity of experimental analysis under these conditions, this numerical approach offers an efficient alternative for understanding the physical processes involved. The conclusions indicate that the turbulence model adequately captures the main flow dynamics, although further studies incorporating combustion models are required to support the design and optimization of high-efficiency combustion systems.

Keywords: Turbulent flow, Mixing species, Sub-critical pressure, Trans-critical pressure, Flammable limits.

1. INTRODUCCION

El proceso de mezcla de fluidos desempeña un papel fundamental en los fenómenos de combustión, particularmente en sistemas presurizados como los utilizados en generadores de vapor en fondo de pozo petrolero. En este contexto, la interacción entre metano y oxígeno en distintas condiciones de presiones es esencial para comprender la dinámica del mezclado, los límites de flamabilidad y la eficiencia del sistema. La presión modifica la densidad del fluido y sus propiedades de transporte, condicionando los mecanismos por difusión y la estabilidad de la combustión, especialmente en condiciones cercanas al régimen transcrítico, donde las propiedades termodinámicas presentan variaciones altamente no lineales (Banuti, 2015; Ma *et al.*, 2017).

Fossa (Fossa, 1995) y Cioncolini (Cioncolini, Thome and Lombardi, 2009) analizaron la mezcla de flujos en dos fases bajo condiciones de referencia o presiones moderadas, sin profundizar en la influencia de la presión sobre la estructura del flujo ni en los efectos de las variaciones de densidad. Estudios posteriores han mostrado que, en condiciones de alta presión, los gradientes de densidad modifican la dinámica del flujo y la tasa de mezcla, generando estructuras turbulentas distintas a las de regímenes subcríticos.

Existe además información relevante para combustibles en aire u oxígeno en condiciones atmosféricas. El modelo semiempírico de Jones (Coward and Jones, 1927) a sido

ampliamente utilizado para evaluar los límites de flamabilidad, y los datos de Lewis y von Elbe (Lewis and Von Elbe, 1937) indican que, a una atmósfera, los límites inferior y superior para metano en oxígeno son de 5.1% y 61%, respectivamente. No obstante, investigaciones recientes muestran que estos límites se desplazan al aumentar la presión, debido a cambios en la velocidad de reacción y en la difusión de especies, por lo que los valores a presión atmosférica no representan el comportamiento en condiciones presurizadas (Law, 2006; Turns, 2012).

En condiciones críticas y transcríticas, Singla *et al.* (Singla *et al.*, 2005) demostraron que la interfase entre fases tiende a desaparecer, favoreciendo la difusión molecular y modificando la estructura del flujo. De manera similar, estudios experimentales y numéricos evidencian alta difusividad y variaciones abruptas de densidad cerca del punto crítico, generando fuertes gradientes de presión y temperatura (Banuti, 2015; Dahms and Oefelein, 2015; Guo *et al.*, 2023). lo que representa un desafío para la modelación numérica.

Adicionalmente, estudios de chorros en condiciones transcríticas muestran que la interacción entre turbulencia y gradientes de densidad genera estructuras altamente inestables, con cizallamiento intenso y zonas de recirculación que incrementan la mezcla (Schmitt *et al.*, 2009; Gaballa, Habchi and de Hemptinne, 2023). A pesar de estos avances, persisten lagunas en configuraciones de inyección múltiple con un chorro central de oxígeno y chorros periféricos de

metano, donde la influencia de la presión sobre la estructura del flujo y la distribución de la mezcla no ha sido caracterizada sistemáticamente en condiciones transcríticas.

Esta limitación se reconoce como un problema abierto en la modelación numérica de sistemas de combustión a alta presión (Guo *et al.*, 2023; Wanstall, Bittle and Agrawal, 2023). En este sentido, es necesario integrar modelos de gas real, como Redlich–Kwong, junto con técnicas RANS para capturar las variaciones de densidad y los efectos turbulentos. Diversos estudios muestran que estas ecuaciones de estado mejoran la predicción termodinámica y la estabilidad numérica a presiones cercanas al punto crítico (Hasche *et al.*, 2023; Mishra *et al.*, 2023).

En conjunto, estos estudios establecen una base para analizar el proceso de mezcla en sistemas presurizados, con aplicaciones en generación de vapor en fondo de pozo, propulsión y turbinas de gas. La integración de modelos de gas real y CFD (Yan *et al.*, 2023) es esencial para comprender cómo la presión afecta la dinámica del mezclado y los límites de flamabilidad.

El presente estudio analiza numéricamente la mezcla de metano y oxígeno en un reactor presurizado bajo tres condiciones: 1 y 10 atm (subcríticas) y 50 atm (transcrítica). Se emplea RANS con el modelo $k-\epsilon$ realizable, bajo flujo isotérmico a 300 K, utilizando gas ideal en condiciones subcríticas y Redlich–Kwong en condición transcrítica para capturar las variaciones de densidad y su efecto en el flujo y el mezclado.

2. ECUACIONES GOBERNANTES

2.1 Ecuaciones de transporte

El flujo en el reactor a presión se modela mediante las ecuaciones de conservación de masa, cantidad de movimiento y transporte de especies bajo el enfoque RANS, el cual permite representar los efectos medios de la turbulencia en flujos tridimensionales. El sistema corresponde a una mezcla gaseosa no reactiva de metano y oxígeno, en régimen estacionario e isotérmico, con temperatura constante en todo el dominio. En estas condiciones, la densidad se considera variable y dependiente de la presión y de la composición de la mezcla, lo cual resulta necesario para describir el flujo en condiciones subcríticas y transcríticas. La conservación de masa para un flujo compresible se expresa mediante la ecuación de continuidad:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (1)$$

donde ρ representa la densidad del fluido y $\mathbf{u}$ es el vector velocidad.

El transporte de cantidad de movimiento se describe mediante las ecuaciones de Navier–Stokes promediadas de Reynolds:

$$\begin{aligned} \frac{\partial(\rho u_i)}{\partial t} + \frac{\partial(\rho u_i u_j)}{\partial x_j} = & -\frac{\partial p}{\partial x_i} + \frac{\partial}{\partial x_j} \left[\mu \left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) \right] \\ & - \frac{\partial}{\partial x_j} (\rho u'_i u'_j) \end{aligned} \quad (2)$$

donde u_i y u_j son las componentes de la velocidad en las direcciones espaciales, p es la presión, μ es la viscosidad dinámica y $\rho u'_i u'_j$ representa las tensiones de Reynolds asociadas a las fluctuaciones turbulentas de la velocidad. Estos términos adicionales constituyen el efecto de la turbulencia sobre el transporte de cantidad de movimiento y requieren un modelo de cierre, el cual se introduce mediante un modelo de turbulencia apropiado, descrito en la siguiente sección.

2.2 Transporte de especies

El proceso de mezcla entre metano y oxígeno en el reactor presurizado se describe mediante ecuaciones de transporte para cada una de las especies que componen la mezcla gaseosa. Estas ecuaciones permiten determinar la distribución espacial de las fracciones másicas en el dominio y caracterizar la interacción entre los chorros de combustible y oxidante bajo diferentes condiciones de presión. En este estudio se consideran dos especies principales, metano y oxígeno, cuya evolución se encuentra gobernada por los mecanismos de convección y difusión.

La ecuación general de transporte de especies se expresa como:

$$\frac{\partial(\rho Y_k)}{\partial t} + \nabla \cdot (\rho \mathbf{u} Y_k) = \nabla \cdot \left[\left(\rho D_k + \frac{\mu_t}{Sc_t} \right) \nabla Y_k \right] \quad (3)$$

donde Y_k representa la fracción másica de la especie k , D_k es el coeficiente de difusión molecular, μ_t es la viscosidad turbulenta y Sc_t es el número de Schmidt turbulento. Este término difusivo incorpora tanto los efectos de difusión molecular como el transporte adicional generado por la turbulencia. Debido a que el sistema es no reactivo, no se incluyen términos fuente por reacciones químicas, por lo que la evolución de las especies depende únicamente del transporte convectivo y difusivo.

El coeficiente de difusión efectiva se evalúa mediante un modelo de difusión turbulenta basado en la analogía entre el transporte de cantidad de movimiento y de masa, asumiendo un número de Schmidt turbulento constante en todo el dominio:

$$\frac{\partial \tilde{\rho} \tilde{Y}_{CH_4}}{\partial t} + \nabla \cdot (\tilde{\rho} \tilde{\mathbf{u}} \tilde{Y}_{CH_4}) = \nabla \cdot (\tilde{\rho} D_{eff} \nabla \tilde{Y}_{CH_4}) + \tilde{S}_{CH_4}, \quad (4)$$

$\tilde{Y}_{CH_4}$ es la fracción másica promedio de metano, D_{eff} es el coeficiente de difusión efectivo, que incluye la difusión molecular y turbulenta, $\tilde{S}_{CH_4}$ es el término fuente promedio de metano.

2.3 Modelo termodinámico

El sistema de ecuaciones se completa mediante una relación termodinámica que vincula la densidad del fluido con la presión, la temperatura y la composición de la mezcla. Para las condiciones subcríticas consideradas en este estudio, se emplea la ecuación de estado de gas ideal:

$$\rho = \frac{p}{RT} \quad (5)$$

donde R es la constante específica del gas y T es la temperatura absoluta. Sin embargo, cuando el sistema opera en condiciones cercanas al régimen transcrito, se emplea una ecuación de estado de gas real basada en el modelo de Redlich-Kwong:

$$p = \frac{RT}{v - b} - \frac{a}{T^{1/2}v(v + b)} \quad (6)$$

donde v representa el volumen específico del fluido y a y b son constantes características del gas que dependen de sus propiedades críticas.

En conjunto, las ecuaciones presentadas constituyen el marco matemático que gobierna el comportamiento del flujo en el reactor presurizado, describiendo el transporte de masa, cantidad de movimiento y especies en una mezcla gaseosa turbulenta bajo condiciones de presión variable.

2.4 Modelo de turbulencia $k - \varepsilon$ realizable

El cierre del sistema de ecuaciones promediadas de Reynolds se realiza mediante un modelo de turbulencia que represente los efectos de las fluctuaciones turbulentas sobre el transporte de cantidad de movimiento y de especies. En el presente estudio se emplea el modelo de turbulencia $k - \varepsilon$ realizable.

El modelo introduce ecuaciones de transporte para la energía cinética turbulenta y su tasa de disipación. La ecuación para k se expresa como:

$$\frac{\partial}{\partial t}(\rho \bar{k}) + \frac{\partial}{\partial x_j}(\rho \bar{k} u_j) = \frac{\partial}{\partial x_j} \left[\left(\mu + \frac{\mu_t}{\sigma_k} \right) \frac{\partial \bar{k}}{\partial x_j} \right] + G_k + G_b - \rho \bar{\varepsilon} - Y_M + S_k, \quad (7)$$

donde G_k representa la producción de energía cinética turbulenta debida a los gradientes de velocidad, σ_k es el número de Prandtl turbulento para la energía cinética y ε es la tasa de disipación turbulenta. G_b (producción por flotabilidad) Y_M (efecto de dilatación en flujos compresibles), S_k (término fuente adicional definido por el usuario)

La ecuación de transporte para la tasa de disipación turbulenta se expresa como:

$$\frac{\partial}{\partial t}(\rho \bar{\varepsilon}) + \frac{\partial}{\partial x_j}(\rho \bar{\varepsilon} u_j) = \frac{\partial}{\partial x_j} \left[\left(\mu + \frac{\mu_t}{\sigma_\varepsilon} \right) \frac{\partial \bar{\varepsilon}}{\partial x_j} \right] + \rho C_{1\varepsilon} S_\varepsilon - \rho C_{2\varepsilon} \frac{\bar{\varepsilon}^2}{\kappa + \sqrt{\nu \bar{\varepsilon}}} + C_{1\varepsilon} \frac{\bar{\varepsilon}}{\kappa} C_{3\varepsilon} G_b + S_\varepsilon, \quad (8)$$

donde S representa la magnitud del tensor de deformación medio, ν es la viscosidad cinemática del fluido y C_1 y C_2 son constantes del modelo que dependen de las propiedades locales del flujo. $C_1 = \max \left[0.43, \frac{\eta}{\eta + 5} \right]$, $\eta = S \frac{\bar{k}}{\varepsilon}$, $S = \sqrt{2 S_{ij} S_{ij}}$. La viscosidad turbulenta se determina mediante la relación:

$$\mu_t = \rho C_\mu \frac{\bar{k}^2}{\varepsilon}, \quad (9)$$

en la cual el coeficiente C_μ no es constante, sino que se calcula dinámicamente en función de las características locales del flujo.

La selección del modelo $k - \varepsilon$ realizable en este estudio se fundamenta en su desempeño comprobado en flujos turbulentos confinados y en configuraciones de chorros múltiples, donde la interacción entre corrientes de diferente densidad genera estructuras de flujo complejas y zonas de recirculación.

3. DETALLES NUMERICOS

El estudio numérico plantea el análisis del proceso de mezclado entre chorros de oxígeno y metano en dos etapas: condiciones subcríticas a 1 y 10 atm, y condiciones transcíticas a 50 atm. Para la primera etapa se calcula la densidad mediante la ecuación de estado de gas ideal, mientras que en la segunda se emplea tanto dicha ecuación como el modelo de Redlich-Kwong para representar el comportamiento real del gas. Se modela la inyección y el mezclado de los flujos en régimen turbulento-compresible, a una temperatura constante de 300 K y sin reacción química. La configuración consiste en un chorro central de oxígeno a través de una boquilla de 16.9 mm de diámetro y cuatro chorros periféricos de metano a través de boquillas de 5.86 mm, con una distancia radial entre inyectores de 33.87 mm, tal como se muestra en la Figura 1.

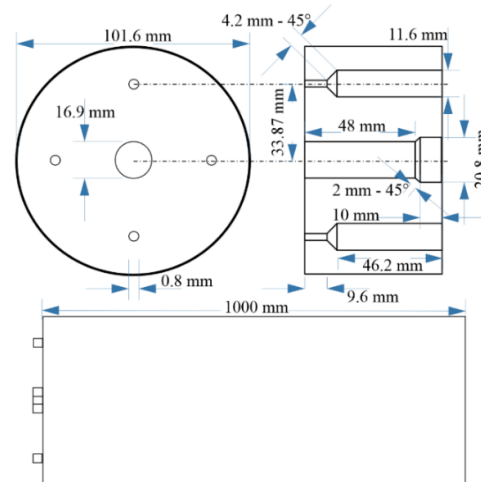


Figura 1. Esquema de configuración y dimensiones del quemador.

El objetivo de las simulaciones es evaluar cómo la presión en el reactor influye en el proceso de mezclado de ambas especies gaseosas, considerando los límites de flammabilidad. El dominio computacional se define a partir de la geometría física del sistema bajo estudio, considerando las dimensiones características necesarias para capturar el desarrollo hidrodinámico y térmico del flujo en condiciones subcríticas y transcíticas. La extensión del dominio se selecciona de manera que las condiciones de frontera no interfieran con la evolución natural de las estructuras del flujo ni introduzcan efectos artificiales en la solución numérica.

En la sección de entrada se prescribe una condición de flujo basada en variables termodinámicas medibles, típicamente presión total y temperatura total o, en su defecto, velocidad media y temperatura estática, dependiendo del régimen operativo del sistema.

En la sección de salida se establece una condición de presión estática, permitiendo que el flujo se ajuste dinámicamente a las variaciones internas del sistema sin imponer restricciones artificiales sobre la velocidad o la densidad.

Las superficies sólidas se modelan mediante condiciones de no deslizamiento para la velocidad, mientras que el tratamiento térmico de las paredes se define de acuerdo con las condiciones físicas del sistema, ya sea mediante flujo de calor impuesto o temperatura constante.

Las condiciones iniciales del cálculo se establecen de forma consistente con el estado termodinámico de operación, proporcionando un punto de partida estable para el proceso iterativo.

El dominio de cálculo del sistema es discretizado mediante una malla no estructurada conformada por 890,000 celdas en ambas etapas, como se muestra en la Figura 2. El estudio se enfoca en el transporte de especies sin premezcla y sin la presencia de reacción exotérmica (combustión). Los componentes de la mezcla, metano y oxígeno, son modelados como gases ideales e incompresibles inicialmente.

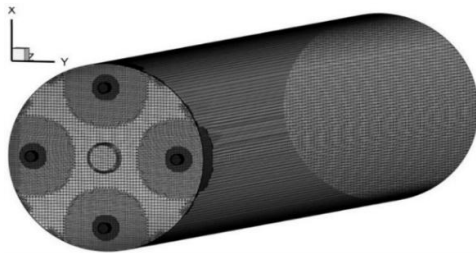


Figura 2. Mallado del dominio de cálculo.

La solución del sistema de ecuaciones gobernantes se realiza mediante el método de volúmenes finitos, el cual permite garantizar la conservación local de masa, cantidad de movimiento y energía dentro de cada celda del dominio computacional.

El acoplamiento entre presión y velocidad se resuelve mediante un esquema iterativo implícito, permitiendo mantener estabilidad numérica en presencia de gradientes fuertes de propiedades termodinámicas.

La discretización espacial de los términos convectivos y difusivos se implementa utilizando esquemas de segundo orden, con el objetivo de reducir la difusión numérica y preservar la precisión en la predicción de los gradientes.

El avance del cálculo se controla mediante un criterio de convergencia basado en la reducción de los residuos normalizados de las ecuaciones gobernantes, así como en la estabilización de variables globales del sistema.

El modelo $k-\epsilon$ realizable se seleccionó por su estabilidad numérica y bajo costo computacional en flujos compresibles con múltiples chorros, proporcionando una descripción adecuada del mezclado global; modelos más avanzados ($k-\omega$ SST, RSM, LES, DES) implican un costo computacional considerable sin mejoras significativas para el objetivo del estudio. Las paredes se modelaron con funciones de pared

estándar, manteniendo valores de y^+ en el rango 30–100. La convergencia se estableció con residuos menores a 10^{-5} y estabilización de variables globales, alcanzándose en ~2000 iteraciones. Se realizó un análisis de independencia de malla (gruesa, media y fina), seleccionando la malla de 890,000 celdas por presentar variaciones menores al 2%, con adecuada calidad ($skewness < 0.85$) y refinamiento en la zona de inyectores (De La Cruz *et al.*, 2015).

3.1 Condiciones de frontera y configuración de la mezcla

Las simulaciones numéricas se realizaron bajo condiciones de presión y temperatura que permiten reproducir el comportamiento del fluido en regímenes subcríticos y transcíticos, considerando la proximidad al punto crítico como el parámetro dominante en la variación de las propiedades termodinámicas.

Las condiciones de operación se definieron mediante la especificación directa de la presión y la temperatura de entrada, así como de la presión de salida, permitiendo controlar el nivel de compresión del fluido dentro del dominio computacional.

El rango de presiones considerado se seleccionó de manera que incluyera condiciones por debajo y por encima del valor crítico del fluido, permitiendo evaluar la influencia del estado termodinámico sobre la estabilidad del flujo y la evolución de las variables de interés. La temperatura de operación se mantuvo dentro de un intervalo que garantiza la sensibilidad de las propiedades termodinámicas a pequeñas variaciones de presión.

Los parámetros de simulación se mantuvieron constantes en todos los casos analizados, con el fin de aislar el efecto de la presión como variable principal del estudio y facilitar la comparación entre escenarios.

Las simulaciones numéricas consideran una condición de frontera de Dirichlet, imponiendo un perfil de velocidad uniforme de 10 m/s para los chorros de entrada, con una intensidad de turbulencia del 5%. En las paredes del reactor, se aplica la condición de no deslizamiento. Para la modelación de la mezcla, se emplea la base de datos predeterminada de especies de ANSYS Fluent, que incluye la mezcla metano-aire; sin embargo, para replicar el sistema metano-oxígeno, se excluye el nitrógeno, manteniendo únicamente el metano y el oxígeno. Esta modificación no introduce variaciones significativas, dado que el estudio se enfoca exclusivamente en el proceso de mezcla sin reacción química.

3.2 Límites de flamabilidad en el aire vs oxígeno

En términos generales, el límite inferior de flamabilidad (LFL, por sus siglas en inglés) de un gas presenta valores cercanos cuando se evalúa en atmósferas de aire y de oxígeno; sin embargo, la literatura reporta un número limitado de estudios comparativos estrictamente equivalentes.

Para el caso del metano, el LFL determinado bajo condiciones de ignición central o durante la propagación descendente de la llama en un conducto vertical resulta mayor en oxígeno puro que en aire. Este comportamiento se asocia a las diferencias en

las propiedades termo-físicas del medio oxidante, particularmente a la mayor capacidad calorífica del oxígeno en relación con el nitrógeno, lo que modifica el balance energético del sistema y conduce a un incremento en la temperatura requerida para sostener la propagación de la llama, situándose típicamente entre 1473 K y 1673 K.

Por el contrario, durante la propagación ascendente u horizontal, el límite inferior de flamabilidad del metano en oxígeno puede presentar valores ligeramente menores que los correspondientes en aire. De manera general, los límites de flamabilidad en atmósferas de oxígeno puro definen un intervalo de concentraciones combustibles más amplio, lo que implica una extensión del rango de mezclas potencialmente inflamables en comparación con sistemas que emplean aire como comburente, tal como se ilustra en la Figura 3.

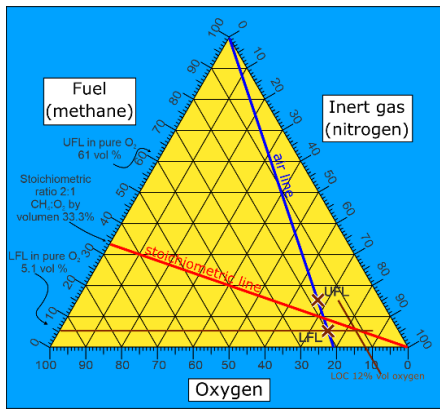


Figura 3. Límites de flamabilidad metano en aire fracción volumétrica.

3.3 Cálculo de los límites de flamabilidad CH_4 con O_2

Una vez determinados los límites de flamabilidad inferior (LFL) y superior (UFL) para la mezcla metano-oxígeno en función de la presión, es necesario evaluar la densidad de la mezcla correspondiente a dichas condiciones, ya que esta propiedad influye directamente en el comportamiento hidrodinámico del sistema presurizado y en la estabilidad del proceso de mezcla.

La mezcla entre el combustible y el oxígeno se genera como consecuencia de la turbulencia inducida en la corriente de oxígeno-metano a la salida del quemador, así como por las diferencias de densidad existentes entre el chorro central y el fluido circundante, lo que favorece el desarrollo de una capa de mezcla alrededor del núcleo potencial del chorro de oxígeno.

Tabla 1. Límites de flamabilidad de CH_4 a 1 atm y 300 K.

Porcentaje de CH_4 en por porcentaje de volumen de O_2	
Límite de la especie	Porcentaje de la especie
LFL del CH_4	5.228136882
UFL del CH_4	60.76388889

La formación de la mezcla combustible depende directamente de los mecanismos de convección y difusión turbulenta que controlan la distribución local de las especies en el dominio de flujo (Tabla 1). El LFL (Zabetakis and Richmond, 1953) se determina a partir de la siguiente relación:

$$LFL = \frac{0.55(100)}{4.76m + 1.19x - 2.38y + 1} \quad (11)$$

Y el UFL a través de:

$$ULF = \frac{3.50(100)}{4.76m + 1.19x - 2.38y + 1} \quad (12)$$

El modelo de Zabetakis (Zabetakis, 1965), se utiliza como referencia para evaluar los límites de flamabilidad de mezclas metano-oxígeno a presiones de 10, 50 y 100 atm. Este modelo propone una correlación empírica que permite estimar la variación de los límites de flamabilidad en función de la presión en sistemas presurizados. La relación empleada para determinar los límites de flamabilidad (FL) se expresa mediante la siguiente ecuación:

$$FL = FL_0 + k_{FL} \log P \quad (13)$$

Los límites de flamabilidad definen la concentración de combustible en el comburente necesaria para iniciar la combustión. Aunque existen diversos modelos desarrollados para combustibles en aire, como el modelo semiempírico de Jones (Coward and Jones, 1927), el presente trabajo considera una mezcla de metano en oxígeno puro. En este sentido, los datos reportados por Lewis y von Elbe (Lewis and Von Elbe, 1937) indican que los límites de flamabilidad inferior y superior del metano en oxígeno son aproximadamente 5.1% y 61% en fracción volumétrica, respectivamente (Tabla 2) (De La Cruz *et al.*, 2015).

Tabla 2. Límites de flamabilidad de CH_4 para varias atmósferas y 300 K.

% CH_4 por % O_2 (fracción volumétrica)						
atm	1	10	20	30	40	50
LFL _p	5.23	5.33	5.36	5.38	5.39	5.40
UFL _p	60.8	81.4	87.6	91.3	93.8	95.8

La densidad de la mezcla se calcula considerando la concentración volumétrica de cada especie en los límites de flamabilidad y su dependencia con la presión. Este cálculo permite establecer una relación directa entre el estado termodinámico del sistema y las condiciones necesarias para el inicio de la combustión. En particular, el incremento de la presión produce un aumento proporcional en la densidad de la mezcla, lo que modifica la dinámica de los chorros y la intensidad de la turbulencia generada en la región de mezcla.

$$\rho_m = \frac{P_m}{\sum_i (X_i R_{X_i}) * T_o p} \quad (14)$$

4. RESULTADOS

Los resultados obtenidos muestran que, para el límite inferior de flamabilidad, la densidad de la mezcla aumenta significativamente al incrementar la presión, pasando de valores cercanos a 1.23 kg/m^3 a presión atmosférica hasta valores superiores a 60 kg/m^3 a 50 atm. Este comportamiento refleja la fuerte dependencia de la densidad con la presión en sistemas presurizados y confirma la necesidad de considerar propiedades variables en el análisis del flujo (Tablas 3).

Tabla 3. Densidad de mezcla (metano-oxígeno) para los límites de flamabilidad inferiores.

Para el límite Inferior			
atm	% CH ₄ (mezcla)	% O ₂ (mezcla)	Densidad de LFF (kg/m ³)
1	5.23	94.77	1.235756692
5	5.3	94.7	6.174697033
10	5.33	94.67	12.34589472
50	5.4	94.6	61.68868647

De manera similar, en el ULF, la densidad de la mezcla también presenta un incremento notable con la presión, aunque con valores menores que los observados en el límite inferior debido a la mayor proporción de metano en la mezcla. Estas diferencias en densidad entre los límites inferior y superior permiten caracterizar el rango de operación seguro del sistema y proporcionan información relevante para el análisis de la estabilidad de la combustión en condiciones presurizadas (Tabla 4).

Tabla 4. Densidad de mezcla (metano-oxígeno) para los límites de flamabilidad superiores.

Para el límite Superior			
atm	% CH ₄ (mezcla)	% O ₂ (mezcla)	Densidad de mezcla ULF (kg/m ³)
1	60.88	39.12	0.8097
5	75.28	24.72	3.7170
10	81.48	18.52	7.1809
50	95.88	4.12	33.2723

Los resultados obtenidos permiten evaluar la influencia de la presión sobre el desarrollo del proceso de mezcla entre los chorros de metano y oxígeno en condiciones subcríticas y transcíticas. El análisis se realiza para tres niveles de presión: 1 y 10 atm, correspondientes al régimen subcrítico, y 50 atm, correspondiente al régimen transcítico, manteniendo constante la temperatura de operación en 300 K y una velocidad de inyección de 10 m/s para ambos fluidos (Figura 4).

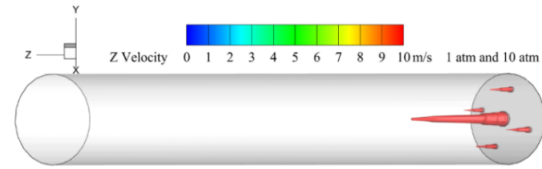


Figura 4. Núcleos potenciales de los chorros, 10m/s, 1 y 10 atm.

La distribución de fracciones de masa presenta gradientes suaves, indicando un proceso dominado por difusión turbulenta y por el intercambio de cantidad de movimiento entre los chorros.

Al aumentar la presión al régimen transcítico, se modifica la estructura del flujo, particularmente la longitud del núcleo potencial y la tasa de mezcla. Este comportamiento se asocia al incremento de la densidad y la reducción de la difusividad efectiva. El uso de una ecuación de estado de gas real permite capturar este efecto, mostrando diferencias frente al modelo de gas ideal.

El análisis de los campos de velocidad y fracción de masa indica que la presión intensifica los gradientes de concentración cerca de las boquillas, modificando la ubicación de las regiones donde se alcanzan los límites de flamabilidad. En conjunto, los resultados muestran que la presión afecta tanto las propiedades termodinámicas como la interacción entre chorros, condicionando la localización de zonas críticas dentro del dominio. 4.1 Comparación de los casos subcríticos a transcíticos

En condiciones subcríticas, la mezcla está dominada por el transporte turbulento y la diferencia de velocidad entre chorros, generando gradientes moderados y una transición gradual hacia la mezcla completa (Figura 5). Los contornos de densidad muestran que, cerca de la placa de las boquillas, la mezcla se aproxima a la condición estequiométrica, con una concentración de metano cercana al 34% en fracción volumétrica.

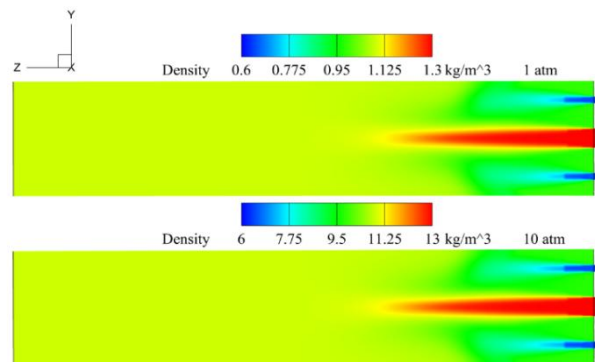


Figura 5. Contornos de densidad bajas presiones 1 y 10 atm.

El arrastre del metano permite alcanzar el límite superior de flamabilidad en el borde del chorro central de oxígeno a 1 atm y en la periferia de los chorros de metano a 10 atm.

En condiciones transcíticas, la variación no lineal de las propiedades termodinámicas modifica la estructura del flujo, aumentando la concentración de especies cerca del eje y

reduciendo la extensión de la región de mezcla. Las simulaciones con gas ideal a 50 atm muestran un comportamiento similar al subcrítico; sin embargo, dado que la presión crítica del metano es 45.95 atm, este régimen requiere el uso de una ecuación de estado de gas real.

Bajo estas condiciones, los chorros de metano presentan una contracción cercana al 7% al emplear gas real, mientras que el chorro de oxígeno, con presión crítica de ~50.93 atm, muestra una menor variación (2%).

Los contornos de densidad en la Figura 6 evidencian diferencias ligeras a 50 atm, indicando que la ecuación de estado de gas real mejora la predicción de la densidad en condiciones transcríticas.

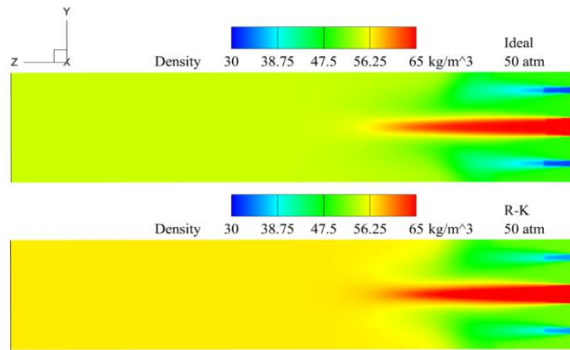


Figura 6. Contornos de densidad a presión transcrítica, 50 atm Gas ideal vs Redlich-Kwong.

La comparación entre el modelo de gas ideal y la ecuación de estado de gas real pone de manifiesto que las discrepancias en la predicción de la densidad se vuelven más relevantes a presiones elevadas, influyendo en la estimación del campo de velocidad y en la distribución de especies dentro del dominio (Figura 6).

Desde el punto de vista físico, estos resultados indican que el proceso de mezcla en sistemas presurizados no puede describirse adecuadamente mediante aproximaciones basadas en propiedades constantes, debido a que el comportamiento del fluido depende de forma directa de su estado termodinámico.

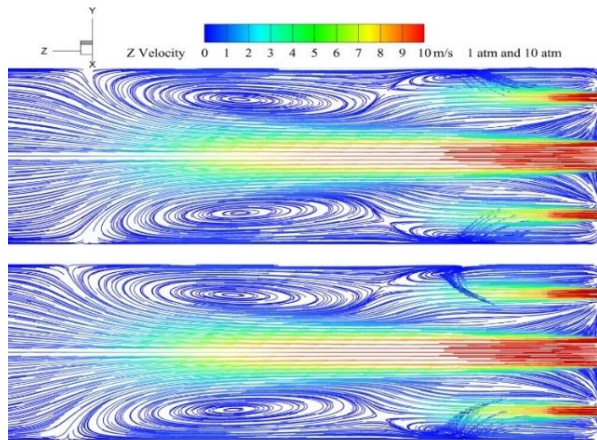


Figura 7. Líneas de corriente en contornos de velocidad: 1 y 10 atm.

El arrastre a bajas presiones (1 y 10 atm) y la velocidad de inyección permiten la dispersión y mezcla de los chorros de metano y oxígeno. Durante esta dispersión, los chorros de metano se mezclan por efecto de arrastre y por la recirculación debida al confinamiento (paredes del reactor ilustrados en la Figura 7).

En general, los resultados indican que la presión influye en la estructura del flujo y en la localización de las regiones donde se alcanzan los límites de flamabilidad, mientras que el proceso global de mezcla por difusión se mantiene dentro del rango de operación analizado. La recirculación está presente en dos zonas: la primera, ubicada entre los chorros de metano y las paredes del reactor, y la segunda, localizada entre los chorros de metano y el chorro de oxígeno. La primera zona presenta una mezcla pobre de oxígeno, aproximadamente de un 20 a un 10% según aumente la presión, permitiendo alcanzar el límite de flamabilidad superior a la salida de los inyectores de metano. La segunda zona, con baja concentración de metano (aproximadamente de un 11 a un 6%), permite alcanzar el límite inferior sobre la parte central del chorro de oxígeno.

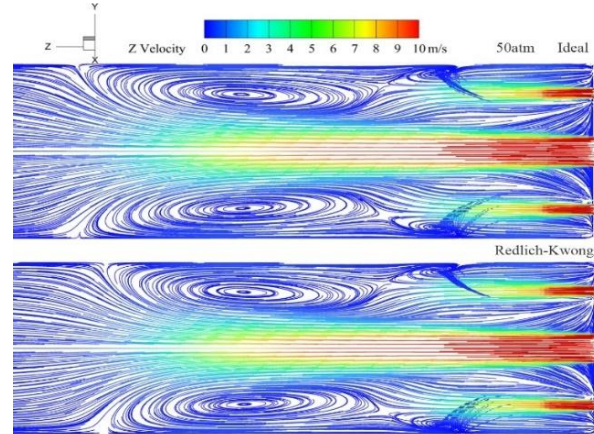


Figura 8. Líneas de corriente en contornos de velocidad G-I y R-K.

Por su parte la recirculación cerca de las paredes del reactor y los chorros de metano también presenta una contracción en sentido opuesto de la dirección de flujo como lo muestra la Figura 8. Esta imagen también exhibe las 2 dos zonas de recirculación encargadas del proceso de mezclado presente en todos los casos.

5. CONCLUSIONES

La mezcla entre el combustible y el oxígeno se produce principalmente como consecuencia de la turbulencia generada a la salida del quemador y de las diferencias de densidad entre el chorro central y el fluido circundante, lo que favorece la formación de una capa de mezcla alrededor del núcleo potencial del chorro de oxígeno. Las simulaciones muestran que el chorro central de oxígeno ejerce un efecto de arrastre sobre los chorros periféricos de metano debido a su mayor flujo másico, por lo que la región más activa del proceso de mezcla se localiza en la zona de interacción entre ambos chorros. En estas regiones se identifican zonas de recirculación, tanto entre los chorros de metano y las paredes



como en la interacción metano-oxígeno, las cuales favorecen la mezcla de las especies y permiten alcanzar concentraciones cercanas a los límites de flamabilidad, principalmente en los bordes de los núcleos potenciales.

El análisis numérico permitió evaluar la influencia de la presión sobre el proceso de mezcla en condiciones subcríticas y transcíticas, observándose que el incremento de la presión modifica la estructura del flujo, particularmente en la longitud de los núcleos potenciales y en la distribución espacial de las especies. En condiciones subcríticas, el proceso de mezcla está dominado por la turbulencia y la interacción hidrodinámica entre los chorros, generando una transición gradual hacia la mezcla completa. En condiciones transcíticas, la variación no lineal de las propiedades termodinámicas produce cambios en la densidad del fluido y en la dinámica del flujo, favoreciendo una mayor concentración de especies en regiones cercanas al eje del dominio.

REFERENCES

- Banuti, D.T. (2015) 'Crossing the Widom-line - Supercritical pseudo-boiling', *Journal of Supercritical Fluids*, 98, pp. 12–16. Available at: <https://doi.org/10.1016/j.supflu.2014.12.019>.
- Cioncolini, A., Thome, J.R. and Lombardi, C. (2009) 'Algebraic turbulence modeling in adiabatic gas-liquid annular two-phase flow', *International Journal of Multiphase Flow*, 35(6), pp. 580–596. Available at: <https://doi.org/10.1016/j.ijmultiphaseflow.2009.02.002>.
- Coward, H.F. and Jones, G.W. (1927) 'Limits of inflammability of gases and vapors', *Journal of the Franklin Institute*, p. 161. Available at: [https://doi.org/10.1016/s0016-0032\(27\)90111-2](https://doi.org/10.1016/s0016-0032(27)90111-2).
- Dahms, R.N. and Oefelein, J.C. (2015) 'Non-equilibrium gas-liquid interface dynamics in high-pressure liquid injection systems', *Proceedings of the Combustion Institute*, 35(2), pp. 1587–1594. Available at: <https://doi.org/10.1016/j.proci.2014.05.155>.
- Fossa, M. (1995) 'A simple model to evaluate direct contact heat transfer and flow characteristics in annular two-phase flow', *International Journal of Heat and Fluid Flow*, 16(4), pp. 272–279. Available at: [https://doi.org/10.1016/0142-727X\(95\)00027-N](https://doi.org/10.1016/0142-727X(95)00027-N).
- Gaballa, H., Habchi, C. and de Hemptinne, J.C. (2023) 'Modeling and LES of high-pressure liquid injection under evaporating and non-evaporating conditions by a real fluid model and surface density approach', *International Journal of Multiphase Flow*, 160(June 2022), p. 104372. Available at: <https://doi.org/10.1016/j.ijmultiphaseflow.2022.104372>.
- Guo, S. et al. (2023) 'A large-eddy simulation study of the transcritical mixing process in coaxial jet flow under supercritical condition', *Journal of Supercritical Fluids*, 203(September), p. 106080. Available at: <https://doi.org/10.1016/j.supflu.2023.106080>.
- Hasche, A. et al. (2023) 'Experimental and numerical assessment of the effects of hydrogen admixtures on premixed methane-oxygen flames', *Fuel*, 352(June), p. 128964. Available at: <https://doi.org/10.1016/j.fuel.2023.128964>.
- De La Cruz, M. et al. (2015) 'Estudio numérico del efecto de la presión en el proceso de mezcla metano-oxígeno en un arreglo de chorros 4-Lug Bolt', *Informacion Tecnologica*, 26(2), pp. 153–162. Available at: <https://doi.org/10.4067/S0718-07642015000200018>.
- Law, C.K. (2006) *Combustion Physics*. 1st edn. Edited by Cambridge University Press. New York, US.
- Lewis, B. and Von Elbe, G. (1937) 'Theory of flame propagation', *Chemical Reviews*, 21(2), pp. 347–358. Available at: <https://doi.org/10.1021/cr60069a011>.
- Ma, P.C. et al. (2017) 'Numerical framework for transcritical real-fluid reacting flow simulations using the flamelet progress variable approach', *AIAA SciTech Forum - 55th AIAA Aerospace Sciences Meeting*, (January), pp. 1–19. Available at: <https://doi.org/10.2514/6.2017-0143>.
- Mishra, A. et al. (2023) 'Flame structure and stability for oxygen enhanced non-premixed combustion of highly diluted methane with helium and nitrogen', *Fuel*, 347(March 2022), p. 128469. Available at: <https://doi.org/10.1016/j.fuel.2023.128469>.
- Schmitt, T. et al. (2009) 'Large-Eddy Simulation of transcritical flows', *Comptes Rendus - Mecanique*, 337(6–7), pp. 528–538. Available at: <https://doi.org/10.1016/j.crme.2009.06.022>.
- Singla, G. et al. (2005) 'Transcritical oxygen/transcritical or supercritical methane combustion', *Proceedings of the Combustion Institute*, 30 II(2), pp. 2921–2928. Available at: <https://doi.org/10.1016/j.proci.2004.08.063>.
- Turns, S.R. (2012) *An introduction to Combustion Concepts and Applications*. Third. McGraw-Hill.
- Wanstall, C.T., Bittle, J.A. and Agrawal, A.K. (2023) 'Phase diagram to demarcate supercritical, transcritical, and continuous phase regimes for binary fluid equilibrium mixing relevant to combustion applications', *Journal of Supercritical Fluids*, 199(March), p. 105935. Available at: <https://doi.org/10.1016/j.supflu.2023.105935>.
- Yan, Z. et al. (2023) 'A system-level multi-field coupling algorithm for regenerative cooling thrust chamber of a LOX/methane rocket engine', *Acta Astronautica*, 213(August), pp. 588–602. Available at: <https://doi.org/10.1016/j.actaastro.2023.09.042>.
- Zabetakis, M.G. (1965) 'Flammability characteristics of combustible gases and vapors', *Bureau of Mines*, 503(Bulletin 627).
- Zabetakis, M.G. and Richmond, J.K. (1953) 'The determination and graphic representation of the limits of flammability of complex hydrocarbon fuels at low temperatures and pressures', *Symposium (International) on Combustion*, 4(1), pp. 121–126.

Assessment of Thermal Comfort in Social Housing Through Occupant Perception And Environmental Monitoring In A Warm Subhumid Climate

Vázquez, Ch., H.*, Pachecano, L.*, Moya, S. L.*, Ascencio, J. A.**, Cruz-Sandoval, M.**

* *Tecnológico Nacional de México/Centro Nacional de Investigación y Desarrollo Tecnológico, Interior Internado Palmira S/N, Colonia Palmira, CP 62490, México.*

(e-mail: m24ce094@cenidet.tecnm.mx, luisana.cp@cenidet.tecnm.mx, sara.ma@cenidet.tecnm.mx).

** *Instituto Tecnológico y de Estudios Superiores de Monterrey, Campus Santa Fe, Av. Carlos Lazo 100, Santa Fe, La Loma, Álvaro Obregón, CP 01389, CDMX*

(e-mail: jascencio@tec.mx, cruzsandovalmarco@tec.mx)

Abstract: Thermal comfort in social housing located in warm subhumid climates is strongly influenced by indoor environmental conditions and occupant perception. This study integrates subjective thermal perception surveys with continuous environmental monitoring in a representative social dwelling located in Morelos, Mexico, during May 2025. A total of 81 occupants participated in a structured survey addressing thermal sensations, overheating periods, and adaptive behaviors. Simultaneously, indoor and outdoor temperature and relative humidity were recorded using IoT-based sensors installed in representative areas of the dwelling, following a previously validated monitoring methodology. Derived thermal indicators, including Heat Index (HI) and Vapor Pressure Deficit (VPD), were calculated to characterize thermal stress and hygrometric conditions. The results revealed strong agreement between occupant perception and monitored environmental conditions, particularly during afternoon periods identified as the most thermally critical. The upper floor exhibited persistent overheating and higher thermal stress levels compared to the ground floor, confirming the influence of vertical thermal stratification and heat accumulation within the building envelope. The integration of occupant perception with environmental monitoring provided a comprehensive interpretation of indoor thermal behavior, highlighting the limitations of passive cooling strategies based solely on natural ventilation. The proposed methodology demonstrates the potential of combining low-cost monitoring systems and perception-based analysis for thermal comfort assessment in social housing under warm subhumid climate conditions.

Keywords: Thermal comfort, occupant perception, environmental monitoring, social housing, Heat Index (HI), Vapor Pressure Deficit (VPD), indoor overheating, warm subhumid climate.

1. INTRODUCCIÓN

El confort térmico en viviendas de interés social ubicadas en climas cálidos subhúmedos representa uno de los principales desafíos asociados a la habitabilidad y bienestar de los ocupantes. En diversas regiones de México, gran parte de estas viviendas son construidas mediante sistemas convencionales de baja resistencia térmica, favoreciendo ganancias de calor a través de muros y cubiertas durante periodos de alta radiación solar. Como consecuencia, los espacios interiores presentan condiciones recurrentes de sobrecalentamiento, particularmente durante primavera y verano, reduciendo la efectividad de las estrategias pasivas de enfriamiento (ASHRAE, 2021).

Tradicionalmente, la evaluación del confort térmico se ha fundamentado en variables ambientales como temperatura del aire, humedad relativa, velocidad del aire y temperatura radiante media, siguiendo lineamientos establecidos por normas internacionales como ASHRAE Standard 55 e ISO 7726. Sin embargo, diversos estudios han señalado la importancia de integrar la percepción de los ocupantes y sus

respuestas adaptativas para una interpretación más representativa del ambiente interior (de Dear & Brager, 1998; Nicol & Humphreys, 2002).

En climas cálidos, la combinación de altas temperaturas y humedad relativa puede intensificar el estrés térmico fisiológico, especialmente en edificaciones naturalmente ventiladas. En este contexto, indicadores derivados como el Índice de Calor (Heat Index, HI) y el Déficit de Presión de Vapor (Vapor Pressure Deficit, VPD) han sido utilizados como herramientas complementarias para caracterizar condiciones higrotérmicas y niveles de incomodidad térmica percibida (NOAA, 2024; Allen et al., 1998).

A pesar del creciente interés en el análisis del confort térmico residencial, numerosos estudios continúan basándose únicamente en promedios climáticos históricos o mediciones ambientales aisladas, limitando la comprensión del comportamiento térmico real durante periodos críticos de sobrecalentamiento. Asimismo, la integración entre percepción térmica y monitoreo ambiental continuo sigue siendo limitada en estudios desarrollados en vivienda social bajo clima cálido subhúmedo en México.



En este contexto, el presente estudio evalúa el confort térmico en una vivienda de interés social ubicada en Morelos, México, mediante la integración de encuestas de percepción térmica y monitoreo ambiental continuo. Las variables de temperatura y humedad relativa fueron registradas mediante sensores IoT instalados en zonas representativas de la vivienda, siguiendo una metodología previamente validada experimentalmente (Vázquez et al., 2025). Adicionalmente, se calcularon indicadores derivados como Heat Index (HI) y Vapor Pressure Deficit (VPD) con el propósito de caracterizar las condiciones de estrés térmico y comportamiento higrotérmico interior. El objetivo principal consiste en establecer una relación entre la percepción térmica de los ocupantes y las condiciones ambientales monitoreadas para mejorar la interpretación del comportamiento térmico interior en viviendas de interés social bajo condiciones de clima cálido subhúmedo.

2. METODOLOGÍA

2.1 Caso de estudio y sistema constructivo

El caso de estudio corresponde a una vivienda dúplex de interés social ubicada en el municipio de Emiliano Zapata, Morelos, México, perteneciente a un conjunto habitacional de aproximadamente 4432 viviendas con características constructivas similares. Para el presente trabajo se seleccionó una vivienda representativa del modelo predominante, sin modificaciones respecto a su configuración original.

Con el propósito de caracterizar el inmueble, se realizó un levantamiento físico y dimensional para la elaboración de planos digitales y estimación del volumen interior, obteniéndose un valor aproximado de 138.293 m³ considerando ambos niveles de la vivienda.

El sistema constructivo predominante corresponde a concreto colado in situ, con muros de aproximadamente 10–11 cm de espesor y cubierta conformada por losa maciza de concreto armado. La vivienda cuenta con ventanas de vidrio sencillo y ventilación natural cruzada mediante aperturas laterales opuestas, configuración común en desarrollos habitacionales de bajo costo.

Adicionalmente, se realizaron pruebas experimentales de caracterización mecánica del concreto mediante esclerometría digital, siguiendo criterios de las normas ASTM C805 y ASTM C882, obteniéndose resistencias promedio cercanas a 280 kg/cm².

2.2 Instrumentación y sistema de adquisición de datos

El monitoreo ambiental se realizó mediante sensores digitales IoT para el registro continuo de temperatura del aire y humedad relativa en condiciones interiores y exteriores de la vivienda, siguiendo protocolos previamente validados para monitoreo termo-higrométrico en edificaciones residenciales (Vázquez et al., 2025).

Para la adquisición de datos se utilizaron sensores Wi-Fi Steren SHOME-180, con precisión aproximada de ± 0.3 °C para temperatura y ± 3 % para humedad relativa, de acuerdo con las especificaciones del fabricante. Los dispositivos fueron instalados en zonas representativas de la vivienda, incluyendo sala, recámara de planta baja, escaleras y habitaciones de planta alta, así como en una zona exterior protegida de

radiación solar directa. La colocación se realizó a una altura aproximada de 1.5 m sobre el nivel del piso, conforme a los criterios recomendados por ISO 7726 para estudios de confort térmico.

La adquisición de datos se efectuó de manera continua durante mayo de 2025, empleando una frecuencia de muestreo de un minuto para registrar la variabilidad higrotérmica diaria. Estudios previos reportaron errores promedio cercanos a ± 0.27 °C en temperatura y ± 0.17 % en humedad relativa respecto a sensores de referencia, respaldando la confiabilidad del sistema utilizado (Vázquez et al., 2025).

2.3 Variables ambientales monitoreadas

Las variables ambientales consideradas en el presente estudio fueron la temperatura del aire (T_a) y la humedad relativa (HR), debido a su influencia directa sobre las condiciones de confort térmico y percepción ambiental de los ocupantes. La selección de estas variables se fundamenta en los criterios establecidos por ASHRAE Standard 55 e ISO 7726, así como en metodologías previamente aplicadas en estudios de monitoreo higrotérmico (Vázquez et al., 2025).

A partir de los registros obtenidos mediante sensores IoT, se calcularon indicadores derivados como el Índice de Calor (Heat Index, HI) y el Déficit de Presión de Vapor (Vapor Pressure Deficit, VPD), con el propósito de complementar la interpretación de las condiciones higrotérmicas interiores y exteriores.

El monitoreo continuo permitió generar series temporales para analizar la evolución térmica diaria, las diferencias entre espacios interiores y exteriores, y la relación entre las condiciones ambientales medidas y la percepción térmica reportada por los ocupantes.

2.3.1 Índice de Calor (HI).

El Índice de Calor (Heat Index, HI) es un indicador empírico utilizado para estimar la temperatura aparente percibida por el cuerpo humano a partir de la combinación entre temperatura del aire y humedad relativa (NOAA, 2022; National Weather Service, 2023; Vázquez et al., 2025). Para su cálculo se empleó la formulación propuesta por el National Weather Service:

$$\begin{aligned} HI = & -42.379 + 2.04901523T + 10.14333127HR \\ & - 0.22475541T(HR) \\ & - 6.83783 \times 10^{-3}T^2 \\ & - 5.481717 \times 10^{-2}HR^2 \\ & + 1.22874 \times 10^{-3}T^2(HR) \\ & + 8.5282 \times 10^{-4}T(HR)^2 \\ & - 1.99 \times 10^{-6}T^2(HR)^2 \end{aligned} \quad (1)$$

donde:

- HI: Índice de calor
- T: temperatura del aire
- HR: humedad relativa

Su significado físico es directo: cuando la humedad relativa aumenta, el sudor pierde capacidad de evaporarse con rapidez, por lo que el cuerpo conserva más calor y “siente” una temperatura superior a la medida por el sensor. En este trabajo,

esa interpretación resulta especialmente útil para explicar por qué la percepción de bochorno, sudoración excesiva y dificultad para dormir aparece con tanta fuerza en la encuesta, aun cuando algunas temperaturas medias puedan parecer moderadas en términos absolutos.

2.3.2 Déficit de Presión de Vapor (VPD).

El Déficit de Presión de Vapor (Vapor Pressure Deficit, VPD) representa la diferencia entre la presión de vapor de saturación y la presión real de vapor presente en el aire, permitiendo evaluar el potencial evaporativo del ambiente (FAO, 1998; Monteith & Unsworth, 2013). Para su cálculo se estimaron primero la presión de saturación e_s y la presión real e_a :

$$e_s = 0.6108 \cdot \exp\left(\frac{17.27T_a}{T_a + 237.3}\right) \quad (2)$$

$$e_a = e_s \cdot \frac{HR}{100} \quad (3)$$

y posteriormente:

$$VPD = e_s - e_a \quad (4)$$

donde:

- VPD: Déficit de presión de vapor (kPa)
- e_s : presión de vapor de saturación (kPa)
- e_a : presión de vapor real (kPa)
- T_a : temperatura del aire (°C)
- HR: humedad relativa (%)

Desde el punto de vista fisiológico, un VPD elevado indica una atmósfera con alta demanda evaporativa, lo que, en condiciones de temperatura elevada, puede contribuir a procesos de deshidratación, fatiga térmica y aumento del malestar térmico, debido a la alteración del balance de calor del cuerpo humano (ASHRAE, 2020). En este sentido, el VPD no solo describe el estado del aire, sino también su interacción con los mecanismos de regulación térmica del organismo.

En el presente estudio, el VPD se emplea como complemento del Índice de Calor (HI). Mientras el HI permite estimar la carga térmica aparente percibida por el usuario en función de la temperatura y la humedad relativa (NOAA, 2022), el VPD aporta una interpretación física del potencial evaporativo del ambiente. La combinación de ambos indicadores permite obtener una evaluación más integral del confort térmico interior, al considerar tanto la percepción fisiológica como los procesos termodinámicos del aire.

2.3.3 Gradientes térmicos verticales.

Se definió el diferencial térmico vertical como:

$$\Delta T = T_{p2} - T_{pB} \quad (5)$$

donde valores positivos indican mayor acumulación de calor en planta alta, fenómeno asociado a estratificación térmica.

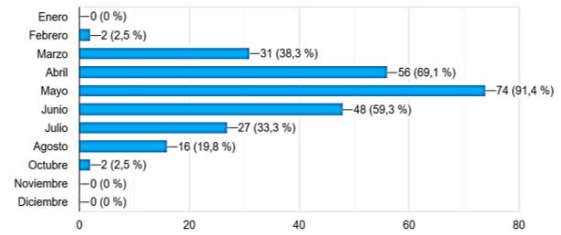
De manera análoga:

$$\Delta HR = HR_{p2} - HR_{pB} \quad (6)$$

permite analizar la distribución de humedad relativa en la vivienda.

En tu experiencia, ¿Cuáles son los meses más calurosos dentro de tu vivienda?

81 respuestas



El procesamiento y análisis de datos se realizó mediante RStudio®, utilizando procedimientos previamente validados en estudios anteriores de monitoreo higrotérmico desarrollados por los autores (Vázquez et al., 2025). Las rutinas implementadas permitieron el cálculo de índices derivados, correlaciones estadísticas y representación temporal de las variables ambientales monitoreadas.

2.4 Encuesta de percepción térmica

Con el propósito de complementar el monitoreo ambiental, se aplicó una encuesta de percepción térmica a 81 habitantes del conjunto habitacional estudiado. El instrumento permitió identificar percepción de confort térmico, periodos de mayor incomodidad y estrategias de adaptación utilizadas por los ocupantes frente a las condiciones ambientales de la vivienda. Las preguntas incluyeron variables relacionadas con sensación térmica percibida, horarios de mayor calentamiento, uso de ventilación natural y percepción general del comportamiento térmico interior. La integración entre encuesta y monitoreo permitió establecer una relación entre las condiciones ambientales registradas y la experiencia térmica reportada por los usuarios, favoreciendo una evaluación más integral del confort térmico (Humphreys & Nicol, 2000; ISO 7730, 2005).

3. RESULTADOS DE LA ENCUESTA

3.1 Percepción térmica de los ocupantes

Los resultados indican que el 100% de los participantes aceptaron voluntariamente formar parte del estudio, garantizando la validez de la muestra.

¿Aceptas participar de forma voluntaria en esta encuesta?

81 respuestas

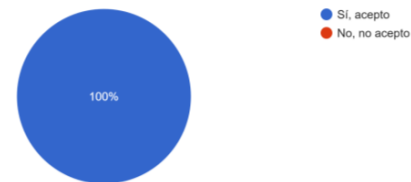


Figura 1. Participación en la encuesta

Figura 2. Selección de meses más calurosos en el área estudiada.

La Figura 2 muestra que mayo fue identificado como el mes más caluroso por el 91.4 % de los participantes, seguido de abril y junio. Este comportamiento coincide con las condiciones de sobrecalentamiento características de climas cálidos subhúmedos, asociadas a altas cargas de radiación y limitada disipación térmica nocturna (Santamouris, 2015). Respecto al horario, el 55.6 % de los encuestados identificó la tarde como el periodo de mayor calor dentro de la vivienda.



(Figura 3), comportamiento consistente con procesos de acumulación y desfase térmico asociados al almacenamiento de calor en la envolvente.

Durante esos meses más calurosos, ¿En qué horario se siente más el calor dentro de tu casa?

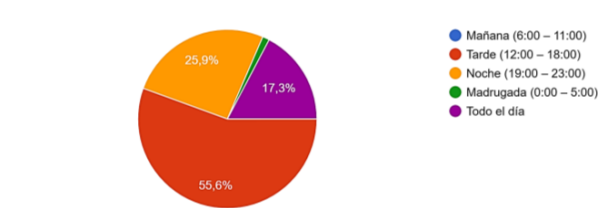


Figura 3. Horario de mayor calor dentro de la vivienda

La Figura 4 muestra que la percepción de intensidad térmica se concentró principalmente en niveles altos (4 y 5), indicando condiciones severas de incomodidad térmica. Estos resultados sugieren que las condiciones interiores excedieron el rango de confort adaptativo aceptable para los ocupantes (ASHRAE, 2020; Nicol & Humphreys, 2002).

Durante ese periodo más caluroso, ¿Qué tan intenso se vuelve el calor en el interior?

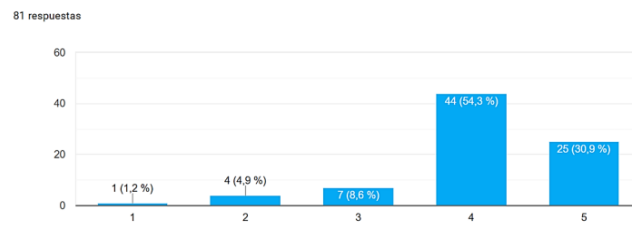


Figura 4. Evaluación de la intensidad térmica percibida en el interior de la vivienda

¿Qué zona de tu vivienda suele calentarse más?

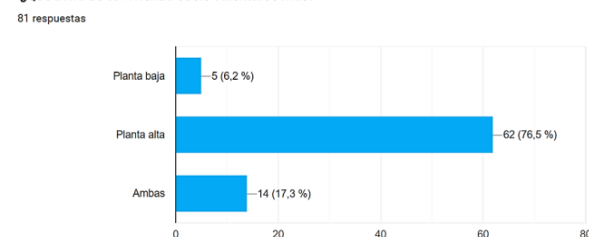


Figura 5. Percepción de la zona con mayor acumulación de calor en la vivienda

La Figura 5 muestra que el 76,5 % de los participantes identificó la planta alta como la zona de mayor acumulación de calor, mientras que únicamente el 6,2 % señaló la planta baja. Estos resultados sugieren la presencia de estratificación térmica vertical asociado a la acumulación de calor en niveles superiores y a la exposición de la cubierta a la radiación solar. Posteriormente, esta percepción fue contrastada con los resultados del monitoreo higrotérmico.

Durante el periodo más caluroso, ¿Qué situaciones has experimentado?

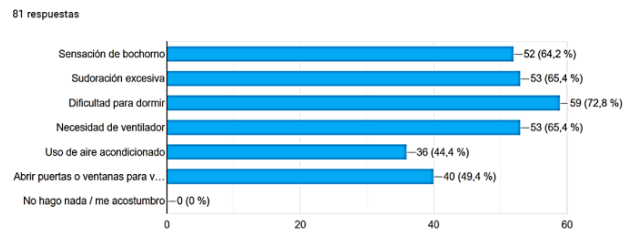


Figura 6. Situaciones de desconfort térmico y estrategias de adaptación reportadas por los ocupantes

La Figura 6 presenta las principales situaciones asociadas al desconfort térmico y las estrategias de adaptación reportadas por los ocupantes. La dificultad para dormir fue la condición más frecuente (72,8 %), seguida de sudoración excesiva y necesidad de ventilador (65,4 %), mientras que el 64,2 % reportó sensación de bochorno durante los periodos de mayor calentamiento.

Respecto a las estrategias de adaptación, el 49,4 % indicó recurrir a ventilación natural mediante apertura de puertas y ventanas, mientras que el 44,4 % señaló utilizar aire acondicionado. En conjunto, los resultados evidencian una percepción generalizada de desconfort térmico y la necesidad recurrente de estrategias de enfriamiento.

En conjunto, los resultados evidencian una percepción generalizada de desconfort térmico entre los ocupantes, así como la necesidad recurrente de implementar estrategias de ventilación o enfriamiento durante los periodos de mayor carga térmica.

4. MONITOREO HIGROTÉRMICO E ÍNDICES DERIVADOS: TEMPERATURA, HUMEDAD RELATIVA, HEAT INDEX Y DÉFICIT DE PRESIÓN DE VAPOR

Con el propósito de contrastar las respuestas obtenidas en la encuesta de percepción térmica con variables ambientales medibles, se realizó un monitoreo continuo de temperatura del aire y humedad relativa en distintos espacios de la vivienda durante mayo, identificado como el periodo de mayor severidad térmica del estudio. A partir de estas variables se estimaron adicionalmente el Heat Index (HI) y el Déficit de Presión de Vapor (VPD), utilizados para complementar la interpretación del comportamiento higrotérmico interior.

La evaluación conjunta de temperatura y humedad relativa resulta fundamental en estudios de confort térmico, ya que la percepción del ambiente interior depende de la interacción simultánea entre variables higrotérmicas y procesos de intercambio energético entre el cuerpo humano y el entorno (ASHRAE, 2020). La Figura 7 presenta la evolución diaria de temperatura, humedad relativa, HI y VPD para los sensores monitoreados durante mayo.



Evolución diaria de variables higrotérmicas e índices derivados durante mayo

Temperatura y humedad relativa como variables base para el cálculo de Heat Index y Déficit de Presión de Vapor

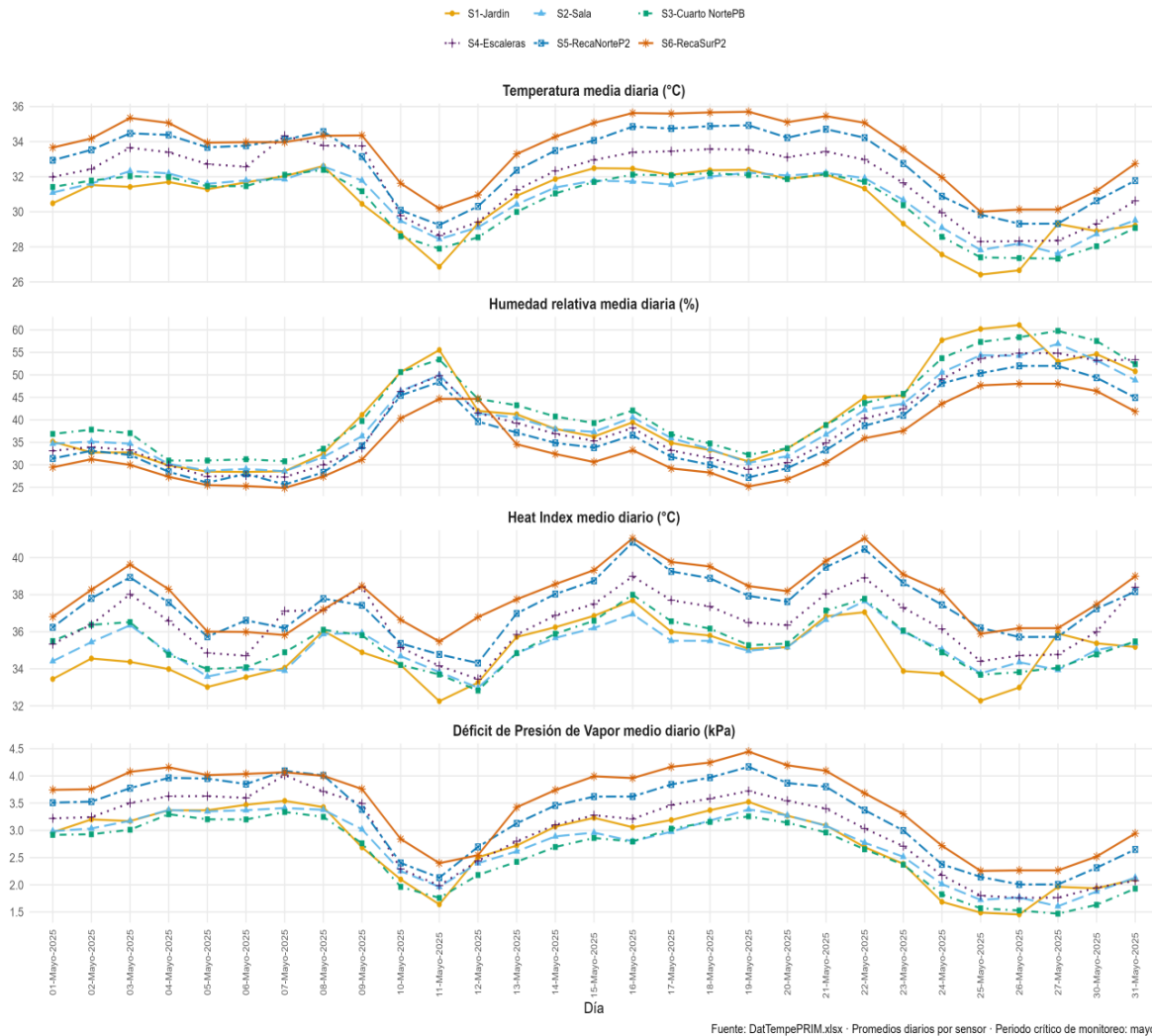


Figura 7. RESULTADOS DE MONITOREO HIGROTÉRMICO E ÍNDICES DERIVADOS: TEMPERATURA, HUMEDAD RELATIVA, HEAT INDEX Y DÉFICIT DE PRESIÓN DE VAPOR

Los resultados muestran que las mayores temperaturas se concentraron en los espacios de planta alta, particularmente en la recámara sur del segundo nivel (S6-RecaSurP2), donde se registraron valores medios superiores a 35 °C durante varios días consecutivos. En contraste, el sensor exterior presentó temperaturas relativamente menores, evidenciando procesos de acumulación y liberación gradual de calor en la envolvente de la vivienda.

La humedad relativa presentó un comportamiento inverso respecto a la temperatura, disminuyendo conforme aumentó el calentamiento interior, especialmente en los espacios superiores. Este comportamiento coincide con la relación psicrométrica reportada en estudios higrotérmicos previos (Vellei et al., 2017).

Los valores más altos de Heat Index (HI) y Déficit de Presión de Vapor (VPD) se registraron nuevamente en planta alta, alcanzando condiciones cercanas o superiores a 40 °C de temperatura aparente durante los periodos más cálidos. Estos resultados muestran correspondencia con las respuestas obtenidas en la encuesta, donde los ocupantes identificaron los

espacios superiores como las zonas de mayor incomodidad térmica (Rothfus, 1990; Zolnier et al., 2000).

En conjunto, los resultados muestran que las condiciones higrotérmicas interiores presentaron un comportamiento diferenciado respecto al ambiente exterior, particularmente en planta alta, donde se identificaron simultáneamente mayores temperaturas, menores humedades relativas y mayores valores de HI y VPD, evidenciando condiciones de mayor severidad térmica.

4.1 Estratificación térmica vertical entre planta alta y planta baja

Con el propósito de evaluar experimentalmente la percepción reportada por los ocupantes respecto a la existencia de zonas más calientes dentro de la vivienda, se analizó el diferencial térmico e higrométrico entre planta alta y planta baja. La Figura 8 muestra la evolución horaria del diferencial de temperatura entre ambos niveles, mientras que la Figura 9



Diferencial de temperatura P2 vs PB (por día + resumen)

Líneas finas: cada día · Línea negra: promedio por hora · Franja: p25-p75
 $\mu=1.90$, $\sigma=1.31$, mediana=1.83 [p25=0.87, p75=2.92] · banda de confort [-1, 1]

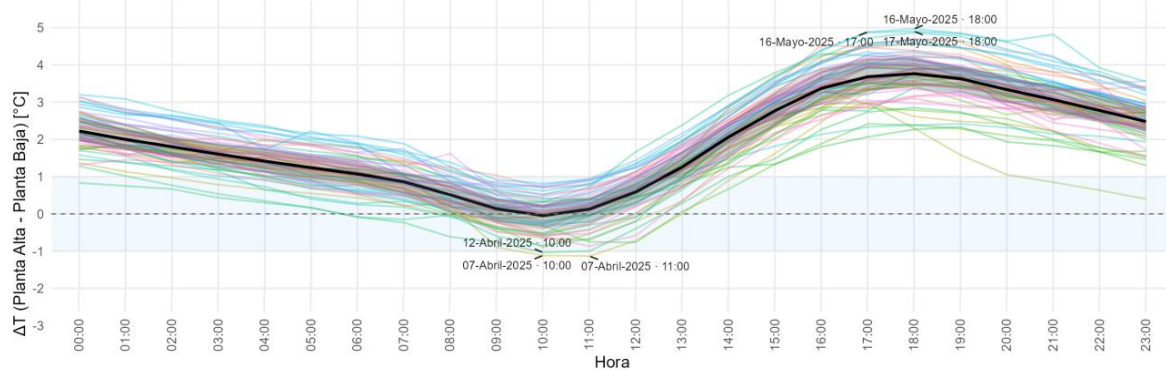


Figura 8. Diferencial vertical de humedad relativa entre planta alta y planta baja

Diferencial de HR P2 vs PB (por día + resumen)

Líneas finas: cada día · Línea negra: promedio por hora · Franja: p25-p75
 $\mu=-3.86$, $\sigma=2.15$, mediana=-3.67 [p25=-5.17, p75=-2.50] · banda de confort [-5, 5]

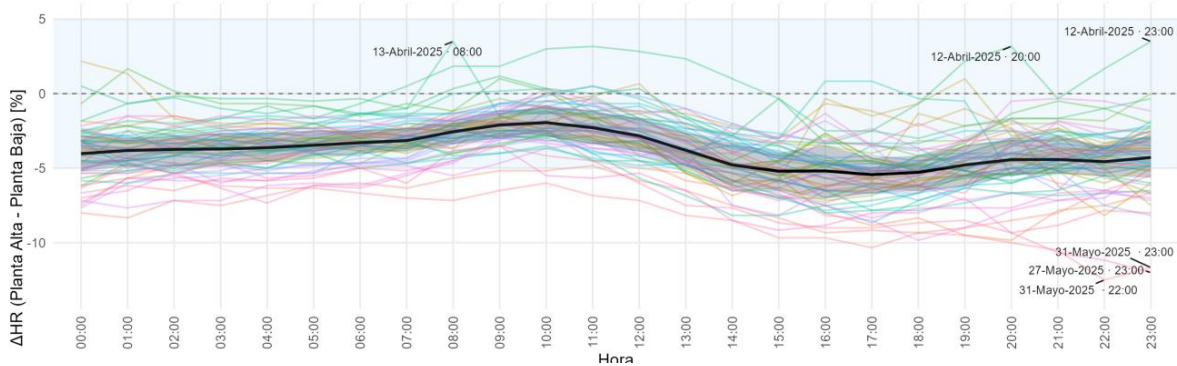
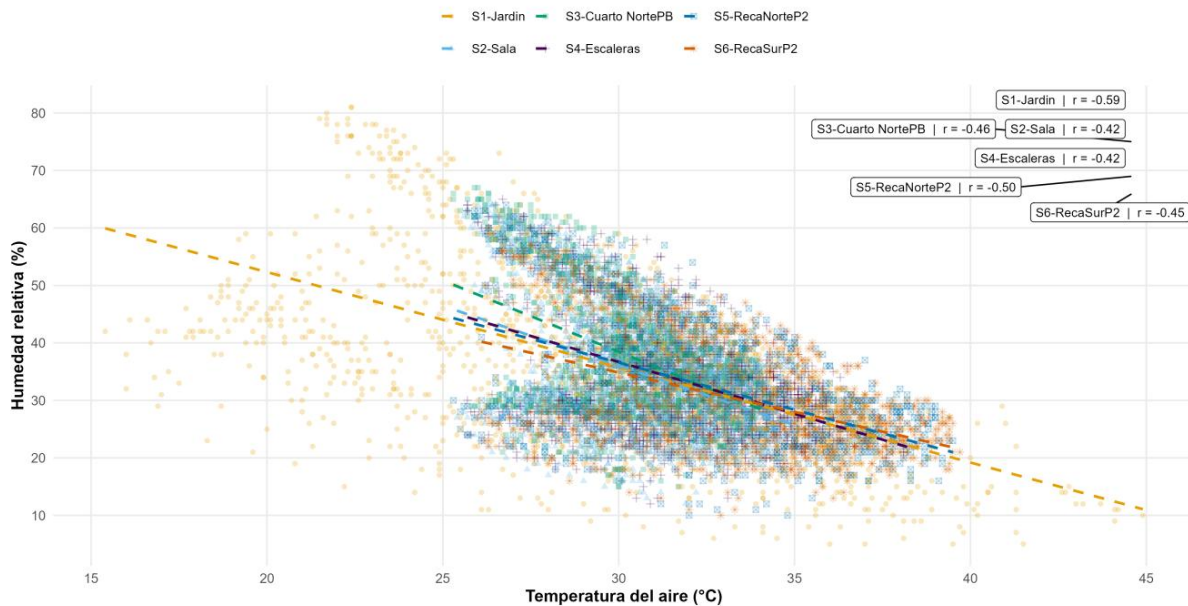


Figura 9. Diferencial vertical de humedad relativa entre planta alta y planta baja

Relación entre temperatura y humedad relativa

Análisis higrotérmico por sensor · Correlación de Pearson calculada individualmente



Fuente: DatTempePRIM.xlsx · Cada punto representa un registro horario monitoreado. Las líneas punteadas representan regresiones lineales por sensor.

Figura 10. Correlación higrotérmica entre temperatura del aire y humedad relativa.

presenta el comportamiento correspondiente para la humedad relativa.

Los resultados evidencian que el diferencial térmico entre planta alta y planta baja fue predominantemente positivo durante gran parte del ciclo diario, con una diferencia media cercana a 1.90 °C y máximos superiores a 4 °C durante las horas vespertinas. Este comportamiento indica que la planta alta mantuvo temperaturas sistemáticamente mayores respecto a los espacios inferiores, particularmente entre las 16:00 y las 19:00 h.

La persistencia de diferencias térmicas positivas constituye evidencia de estratificación térmica vertical, fenómeno asociado a la acumulación de calor en niveles superiores y ampliamente reportado en edificaciones con alta exposición solar y ventilación limitada (Baborska-Narożny et al., 2015). De manera complementaria, el diferencial de humedad relativa presentó valores predominantemente negativos, con una media aproximada de -3.86 %, indicando menores niveles de humedad relativa en planta alta. Este comportamiento resulta coherente con el incremento térmico observado en los espacios superiores.

La combinación de mayores temperaturas y menores humedades relativas en planta alta confirma la presencia de condiciones higrotérmicas más severas en los niveles superiores, resultado consistente con la percepción reportada por los ocupantes durante la encuesta aplicada.

4.1 Relación entre temperatura y humedad relativa mediante correlación de Pearson

Con el propósito de analizar la dependencia higrotérmica entre temperatura del aire y humedad relativa, se estimó la correlación de Pearson para cada uno de los sensores monitoreados. La Figura 10 presenta la distribución conjunta de ambas variables y las regresiones lineales correspondientes. Los resultados muestran correlaciones negativas consistentes en todos los sensores, con coeficientes comprendidos entre $r = -0.42$ y $r = -0.59$. Este comportamiento indica que, conforme aumenta la temperatura del aire, la humedad relativa tiende a disminuir, resultado físicamente coherente con la dinámica psicrométrica esperada en ambientes naturalmente ventilados y sometidos a procesos de calentamiento diurno.

El sensor exterior correspondiente al jardín (S1-Jardín) presentó la correlación negativa más alta ($r = -0.59$), sugiriendo una respuesta más directa a la dinámica atmosférica natural. En contraste, los sensores interiores mostraron correlaciones de menor magnitud ($r = -0.42$ a $r = -0.50$), comportamiento asociado al almacenamiento térmico de la envolvente y al desacoplamiento parcial entre ambiente interior y exterior.

Asimismo, los sensores localizados en planta alta presentaron un desplazamiento hacia regiones de mayor temperatura y menor humedad relativa, reforzando la identificación de estos espacios como las zonas de mayor severidad térmica dentro de la vivienda.

La estimación de correlaciones entre temperatura y humedad relativa ha sido utilizada previamente como herramienta para interpretar procesos de confort y comportamiento ambiental en edificaciones naturalmente ventiladas (Zamri et al., 2025; Vellei et al., 2017). En este estudio, la correlación de Pearson

permitió proporcionar evidencia cuantitativa sobre la interacción higrotérmica presente dentro de la vivienda.

5. DISCUSIÓN

Los resultados muestran una relación consistente entre la percepción térmica reportada por los ocupantes y las condiciones higrotérmicas registradas experimentalmente dentro de la vivienda. La integración entre encuesta y monitoreo permitió identificar que el discomfort térmico se encuentra asociado a procesos de sobrecalentamiento interior, acumulación energética de la envolvente y estratificación térmica vertical.

La percepción de la tarde como el periodo de mayor incomodidad coincidió con los máximos valores de temperatura, Heat Index y Déficit de Presión de Vapor registrados en planta alta. Este comportamiento es consistente con estudios desarrollados en climas cálidos, donde la envolvente funciona como un sistema de almacenamiento térmico capaz de prolongar el calentamiento interior después del máximo de radiación solar exterior (Santamouris, 2015). Los diferenciales térmicos entre planta alta y planta baja confirmaron la presencia de estratificación térmica vertical, fenómeno previamente reportado en edificaciones naturalmente ventiladas con alta exposición solar y limitada disipación energética (Baborska-Narożny et al., 2015). Asimismo, la combinación de mayores temperaturas, menores humedades relativas y mayores valores de VPD evidenció condiciones higrotérmicas más severas en planta alta.

Las correlaciones negativas entre temperatura y humedad relativa mostraron un comportamiento psicrométrico consistente con procesos reales de calentamiento interior (Vellei et al., 2017). La menor intensidad de correlación observada en los sensores interiores respecto al exterior sugiere un desacoplamiento parcial asociado a la capacidad de almacenamiento energético de la envolvente y a los efectos de inercia térmica previamente reportados en edificaciones con alta masa térmica (Orosa & Oliveira, 2011).

En conjunto, los resultados evidencian que el comportamiento térmico de la vivienda no depende exclusivamente de la ventilación natural, sino también de la interacción entre la envolvente, la acumulación energética y la dinámica higrotérmica interior. Bajo estas condiciones, fenómenos como persistencia térmica, acumulación energética y estratificación vertical pueden modificar significativamente la respuesta térmica interior aun cuando exista ventilación cruzada disponible.

Finalmente, los resultados muestran la importancia de complementar los análisis basados en promedios climáticos y registros meteorológicos históricos con monitoreos experimentales locales y de alta resolución temporal, permitiendo identificar fenómenos dinámicos asociados al comportamiento real de la vivienda bajo condiciones de clima cálido subhúmedo.

6. CONCLUSIONES

El presente estudio permitió evaluar el comportamiento higrotérmico de una vivienda de interés social mediante la integración de encuestas de percepción térmica y monitoreo



ambiental continuo. Los resultados mostraron una alta correspondencia entre la percepción reportada por los ocupantes y las condiciones ambientales registradas experimentalmente.

La planta alta presentó las condiciones de mayor severidad térmica, registrando mayores temperaturas del aire, menores humedades relativas y mayores valores de Heat Index y Déficit de Presión de Vapor. Los diferenciales térmicos verticales confirmaron además la presencia de estratificación térmica, con diferencias superiores a 4 °C durante las horas vespertinas. Los resultados evidenciaron que el comportamiento térmico interior estuvo influenciado por procesos de acumulación y liberación gradual de calor asociados a la envolvente de concreto, favoreciendo la persistencia térmica y un desacoplamiento parcial entre condiciones exteriores e interiores.

Finalmente, la integración entre percepción térmica y monitoreo higrotérmico continuo permitió obtener una representación más cercana del comportamiento ambiental real de la vivienda, proporcionando una base útil para futuras investigaciones relacionadas con estrategias pasivas y análisis térmico de edificaciones residenciales.

REFERENCIAS

- ASHRAE. (2020). *ANSI/ASHRAE Standard 55-2020: Thermal environmental conditions for human occupancy*. American Society of Heating, Refrigerating and Air-Conditioning Engineers.
- ASHRAE. (2021). *ASHRAE handbook—Fundamentals*. American Society of Heating, Refrigerating and Air-Conditioning Engineers.
- Allen, R. G., Pereira, L. S., Raes, D., & Smith, M. (1998). Crop evapotranspiration: Guidelines for computing crop water requirements (FAO Irrigation and Drainage Paper No. 56). Food and Agriculture Organization of the United Nations.
- Baborska-Narozny, M., Stevenson, F and Chatterton, P (2016) Temperature in housing: stratification and contextual factors. *Proceedings of the Institution of Civil Engineers: Engineering Sustainability*, 169 (4). pp. 125-137. ISSN: 1478-4629
- de Dear, Richard & Brager, Gail & D., Cooper. (1997). *Developing an Adaptive Model of Thermal Comfort and Preference - Final Report on RP-884..* ASHRAE Transactions. 104.
- Zamri, Ernisuhani & Ismail, Asmat & Md Ajis, Azizah. (2019). Thermal Comfort in Naturally Ventilated Classroom: A Literature Review. *International Journal of Property Sciences*. 9. 27-37. 10.22452/ijps.vol9no1.3.
- Givoni, B. (1998). *Climate considerations in building and urban design*. John Wiley & Sons.
- Humphreys, M. A., & Nicol, J. F. (2000). Outdoor temperature and indoor thermal comfort. *Building Research & Information*, 28(2), 92–101.
- International Organization for Standardization. (2005). ISO 7730: Ergonomics of the thermal environment—Analytical determination and interpretation of thermal comfort using calculation of the PMV and PPD indices.
- National Weather Service. (2023). Heat index equation. Weather Prediction Center. Heat Index Equation
- Nicol, F., & Humphreys, M. (2002). Adaptive thermal comfort and sustainable thermal standards for buildings. *Energy and Buildings*. Volume 34, Issue 6, July 2002, Pages 563-572
- Rothfus, L. P. (1990). The heat index equation (or, more than you ever wanted to know about heat index) (Technical Attachment SR 90-23). National Weather Service, Scientific Services Division.
- Orosa, J. A., & Oliveira, A. C. (2012). A field study on building inertia and its effects on indoor thermal environment. *Renewable Energy*, 37(1), 89–96.
- Santamouris, M. (2015). Regulating the damaged thermostat of the cities—Status, impacts and mitigation challenges. *Energy and Buildings*, 91, 43–56. <https://doi.org/10.1016/j.enbuild.2015.01.027>
- Vázquez-Chávez, H., Pachecano, L., Moya, S. L., Ascencio, J. A., & Cruz-Sandoval, M. (2025). Micro-scale wind assessment with commercial UAVs and IoT sensors for comfort and environmental management. *Revista Tecnología y Ciencia Aplicada*, 8(2), pag. 455 – 463 ISSN: 2992-8338.
- Vellei, M., Herrera, M., Fosas, D., & Natarajan, S. (2017). The influence of relative humidity on adaptive thermal comfort. *Building and Environment*. Volume 124, 1 November 2017, Pages 171-185
- Zolnier, Sérgio & Gates, Richard & Buxton, J & Mach, Chris. (2000). Psychometric and ventilation constraints vapor pressure deficit control. *Computers and Electronics in Agriculture* 26(3):343-359 DOI:10.1016/S0168-1699(00)00084-3

Parametric Numerical Analysis of an Earth-Air Heat Exchanger under Different Operating Conditions

B.A. Rosales-Daniel*, M.A. Chagolla-Aranda*, Y. Olazo-Gómez**, I. Hernández-López**, E. Sima*.

* *TecNM/CENIDET, Campus Apatzingan 2, Cuernavaca Morelos, México, C.P. 62490 (Tel. 5624450186, e-mail: m25ce025@cenidet.tecnm.mx, miguel.ca@cenidet.tecnm.mx, efrain.sm@cenidet.tecnm.mx).*

** *Universidad de Sonora, UNISON, Blvd. Luis Encinas y Rosales S/N, Col. Centro, Hermosillo, Sonora, CP. 83000, Mexico (e-mail: yessenia.olazo@unison.mx, irving.hernandez@unison.mx).*

Abstract: This study presents a parametric numerical analysis of an earth-air heat exchanger (EAHE) under the specific climatic conditions of Cuernavaca (Mexico), classified as Cwa (warm subhumid) according to the Köppen climate classification. A three-dimensional CFD model was developed in ANSYS Fluent under steady-state conditions. The numerical model was verified by comparing it with previously published CFD results, ensuring consistency with established numerical approaches described in the literature. Soil temperature was estimated using an analytical model based on local meteorological data collected over the course of a year, along with the thermophysical properties of different soil types obtained from the literature, which allowed for the evaluation of soil behavior at various depths and moisture conditions. Based on this analysis, a representative soil configuration was selected to perform simulations considering inlet air velocities from 1 to 6 m/s. The results show that lower air flow velocities improve heat transfer due to increased residence time inside the pipe. Under maximum cooling of approximately 6.08 °C at 1 m/s, while at 6 m/s the cooling effect decreased to 1.54 °C. In contrast, under minimum annual temperature conditions (9.46 °C), the EAHE operated in heating mode, increasing the outlet air temperature by up to 5.49 °C. The results also indicate that the outlet air temperature tends to approach the ground temperature, which remained within an approximate range of 22-25 °C at 3 m depth throughout the year. Furthermore, the analysis showed that thermal fluctuations decrease significantly with depth, identifying 3 m as a representative depth with stable thermal behavior. These findings demonstrate that the thermal performance of EAHE systems strongly depends on operating conditions and local climate characteristics, highlighting the importance of climate-specific analyses for the design and optimization of EAHE systems in warm subhumid regions.

Keywords: Earth-Air Heat Exchanger, Geothermal Cooling, Parametric Study, CFD, Heat Transfer.

1. INTRODUCCIÓN

Actualmente, la creciente demanda energética en sistemas de climatización ha impulsado el desarrollo de tecnologías basadas en fuentes renovables. Entre estas, el intercambiador de calor tierra-aire (Earth-Air Heat Exchanger, EAHE) el cual representa una alternativa eficiente para el acondicionamiento térmico. Este tipo de sistemas ha sido ampliamente utilizado en aplicaciones de climatización de edificaciones o en sistemas de enfriamiento de otros sistemas como lo son los paneles solares, con el fin de mejorar la eficiencia energética (Koshlak, 2025).

El principio de funcionamiento del EAHE se basa en el intercambio de calor entre el aire que circula a través de tuberías enterradas y el suelo circundante. Debido a que la temperatura del suelo permanece relativamente constante a profundidades mayores a 3 m por su alta inercia térmica, este actúa como disipador de calor en verano y como fuente de calor en invierno. Este comportamiento permite reducir la

carga térmica en sistemas de climatización y, por ende, el consumo energético.

Diversos estudios han evaluado el desempeño térmico de los EAHE considerando parámetros geométricos como longitud, diámetro y configuración del sistema. En particular, Hasan & Noori, (2019), desarrollaron un modelo numérico tridimensional para analizar diferentes configuraciones de EAHE, demostrando su potencial para reducir significativamente el consumo energético en aplicaciones residenciales.

Sin embargo, existe una limitada exploración del efecto de las condiciones de operación, tales como la velocidad del aire, la temperatura de entrada y la temperatura del suelo, sobre el desempeño térmico del sistema del artículo de Mihalakou et al. (2022). Además, el rendimiento de los sistemas EAHE puede variar significativamente en función de las condiciones climáticas locales, tales como la radiación solar, propiedades del suelo, etc. Por lo que su comportamiento aún continúa

siendo objeto de estudio en diferentes regiones y climas. Asimismo, la estimación precisa de la temperatura del suelo, considerando sus propiedades termofísicas y condiciones climáticas, representa un reto importante en este tipo de análisis.

Por lo tanto, el objetivo de este trabajo es realizar un análisis numérico paramétrico de un sistema EAHE, basado en la geometría de un sistema propuesto por Hasan & Noori, (2019), evaluando la influencia de variables como la velocidad del aire, la temperatura de entrada y la temperatura del suelo, esta última estimada mediante modelos analíticos dependientes de propiedades del suelo y condiciones meteorológicas.

2. FUNDAMENTOS TEORICOS

2.1 Principio de funcionamiento del EAHE.

El intercambiador de calor tierra-aire (Earth-Air Heat Exchanger, EAHE) es un sistema de climatización que aprovecha la estabilidad térmica del suelo para modificar la temperatura del aire que circula a través de tuberías enterradas. Su funcionamiento se basa en los principios de transferencia de calor por convección y conducción, en la *Figura 1* se observa una representación sencilla del sistema.

Como se mencionan en los artículos de Holmes et al. (2008) y Abdel-Ghany et al. (2022), a profundidades mayores a 3m, la temperatura del suelo permanece relativamente constante durante el año presentando valores menores que la temperatura ambiente en verano y mayores en invierno. Esta característica permite que el suelo actúe como un disipador o fuente de calor, dependiendo de las condiciones climáticas externas.

Cuando el aire es forzado a circular a través de las tuberías enterradas, se produce un intercambio de calor entre el aire y el suelo circundante. En condiciones de verano, el aire caliente transfiere calor al suelo, disminuyendo su temperatura antes de su salida del sistema. En invierno, el proceso ocurre en sentido contrario, calentando el aire frío.

Asimismo, el desempeño térmico del EAHE depende de diversos factores físicos y operativos. Entre los más relevantes se encuentran la velocidad del aire, la temperatura de entrada, la temperatura del suelo y las propiedades térmicas del medio.

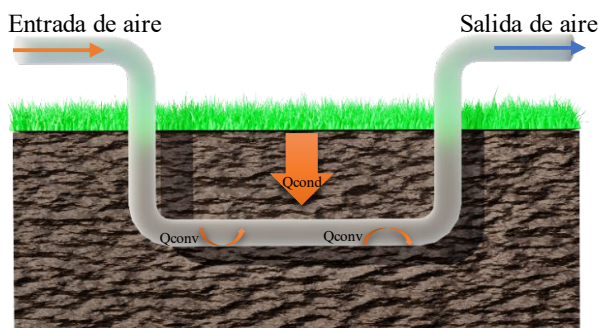


Figura 1. Representación de un sistema sencillo lineal de un EAHE.

2.2 Transferencia de calor en el EAHE.

El proceso de intercambio térmico en un EAHE involucra dos mecanismos principales:

- Convección entre el aire en movimiento y la pared interna de la tubería.
- Conducción a través de la pared de la tubería y el suelo circundante.

La cantidad de calor transferido puede expresarse como:

$$Q = \dot{m} C_p (T_{in} - T_{out}) \quad (1)$$

donde:

$\dot{m}$: flujo másico de aire.

C_p : calor específico del aire.

T_{in}, T_{out} : temperaturas de entrada y salida.

Esta ecuación permite evaluar el desempeño térmico del sistema.

2.3 Transferencia de calor en el EAHE.

El comportamiento del EAHE se describe mediante las ecuaciones de conservación de masa, cantidad de movimiento y energía. Estas ecuaciones, bajo los supuestos de flujo estacionario, propiedades constantes e incompresibilidad, permiten modelar el comportamiento del sistema mediante herramientas CFD como ANSYS Fluent.

Ecuación de continuidad

$$\nabla \cdot \vec{V} = 0 \quad (2)$$

Siendo $\vec{V}$ la velocidad.

Ecuaciones de cantidad de movimiento

(forma general de Navier-Stokes)

$$\rho(\vec{V} \cdot \nabla) \vec{V} = -\nabla p + \mu \nabla^2 \vec{V} \quad (3)$$

Siendo ρ la densidad, $\vec{V}$ la velocidad, p la presión, μ la viscosidad dinámica.

Ecuación de energía

$$\rho c_p (\vec{V} \cdot \nabla T) = k \nabla^2 T \quad (4)$$

Siendo c_p el calor específico, T la temperatura y k la conductividad térmica.

2.4 Modelo de temperatura del suelo.

La temperatura del suelo es un parámetro clave en el desempeño del EAHE, ya que determina el potencial de intercambio térmico entre el aire y el medio circundante.

En este trabajo, la temperatura del suelo se estimó mediante el modelo analítico propuesto por Ali, et al., (2023).

$$T_{soil}(z,t) = T_{mean} + A_0 e^{-z/d} \sin\left(\left(\frac{2\pi(t-t_0)}{365}\right) - \frac{z}{d} - \frac{\pi}{2}\right) \quad (5)$$

Donde:

T_{soil} : Temperatura del suelo.

T_{mean} : Temperatura media anual.

A_0 : Amplitud térmica.

t_0 : Desfase temporal

d : Profundidad de amortiguamiento.

z : Profundidad.

t : Día del año.

La profundidad de amortiguamiento depende de la difusividad térmica del suelo:

$$d = \sqrt{\frac{2\alpha}{\gamma}} \quad (6)$$

Donde:

α : Difusividad térmica.

γ : Frecuencia angular anual de la variación de temperatura.

Mientras que, la difusividad térmica se define como:

$$\alpha = \frac{k}{\rho C_p} \quad (7)$$

Donde:

C_p : Capacidad calorífica.

k : Conductividad térmica.

ρ : Densidad.

3. METODOLOGIA

En el presente trabajo se desarrolló un modelo numérico tridimensional de un intercambiador de calor tierra-aire (EAHE) con el objetivo de evaluar su desempeño térmico bajo diferentes condiciones de operación. Las simulaciones se realizaron mediante el software de dinámica de fluidos computacional (CFD) ANSYS Fluent. Asimismo, el modelo EAHE se basa en la configuración del sistema 1 propuesta por Hasan & Noori, (2019).

3.1 Descripción del dominio físico.

El modelo geométrico corresponde a un sistema de tuberías enterradas por donde circula aire, representando un intercambiador tierra-aire típico. El sistema está compuesto por tuberías horizontales enterradas a una profundidad constante, rodeadas por un medio sólido que representa el suelo.

El dominio computacional incluye tanto la región fluida (aire) como la región sólida (suelo), permitiendo el análisis del fenómeno de transferencia de calor conjugada entre ambos medios. En la *figura 2* se observa el dominio físico en el software de ANSYS.

El sistema EAHE cuenta con una profundidad de enterramiento de 3 m y consiste en 7 tuberías horizontales con una longitud de 3 m, un diámetro de 0.1016 m, y una separación de 1 m, además cuenta con dos colectores de 0.152 m de diámetro y 7 m de longitud.

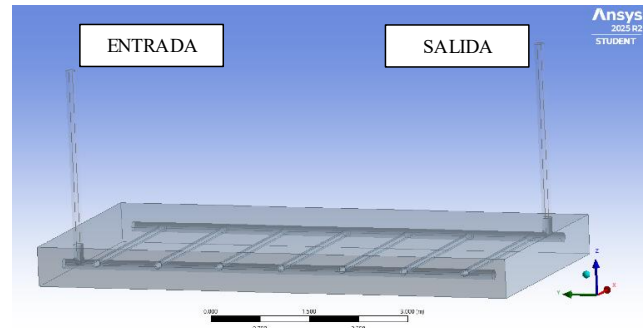


Figura 2. Dominio físico del sistema EAHE.

3.2 Consideraciones del modelo

Se consideró que el flujo se encuentra en estado estacionario debido a que el análisis se enfoca en condiciones representativas de operación donde las variaciones temporales son despreciables en comparación con los gradientes espaciales de temperatura. Además, el aire fue modelado como un fluido incompresible, ya que las velocidades de operación son relativamente bajas (del orden de 1 a 6 m/s), lo que implica variaciones de densidad insignificantes.

Asimismo, las propiedades termofísicas de los materiales se asumieron constantes, considerando que las variaciones de temperatura en el sistema no son lo suficientemente elevadas como para generar cambios significativos en dichas propiedades. En cuanto al suelo, fue modelado como un medio homogéneo e isotrópico, permitiendo representar su comportamiento térmico de manera efectiva sin introducir complejidad adicional asociada a las características del terreno.

También se asumió la ausencia de generación interna de calor dentro del dominio dado que no existen fuentes activas que contribuyan al incremento de energía térmica en el sistema. Además, el mecanismo de transferencia de calor se modeló considerando conducción en el suelo y convección en el aire, lo cual representa los fenómenos físicos dominantes, donde el intercambio térmico ocurre principalmente entre el aire en movimiento y el terreno circundante.

3.3 Generación de la malla

La geometría fue discretizada mediante una malla no estructurada compuesta por elementos tetraédricos, adecuada para representar geometrías complejas y capturar la interacción entre el flujo de aire y el medio sólido circundante. Para el desarrollo del modelo numérico se empleó un nivel de refinamiento y un número de elementos cercano al utilizado en el trabajo de referencia de Trilok Singh et al. (2014), el cual previamente había mostrado una adecuada concordancia con datos experimentales. Con base en ello, se consideró que la discretización adoptada proporcionaba una resolución espacial adecuada y suficiente para la obtención de resultados satisfactorios y numéricamente confiables.

La calidad del mallado fue evaluada mediante parámetros geométricos como asimetría y calidad ortogonal, obteniéndose valores promedio de 0.19 y 0.83, respectivamente, con la



intención de tener una estabilidad numérica. El refinamiento empleado, junto con la calidad del mallado, permitió reducir los errores asociados a la discretización espacial. Asimismo, la confiabilidad del modelo numérico fue respaldada mediante un proceso de verificación, comparando los resultados obtenidos con resultados numéricos reportados para una geometría y condiciones de operación equivalentes, obteniéndose una adecuada concordancia entre ambos casos. La configuración numérica adoptada también es consistente con estudios previos de sistemas EAHE en los que se han empleado modelos CFD tridimensionales, transferencia de calor conjugada y modelos de turbulencia para reproducir satisfactoriamente el comportamiento térmico del sistema, como se reporta en el trabajo de Trilok Singh et al. (2014).

3.4 Modelo físico y configuración en ANSYS Fluent

Las simulaciones numéricas se llevaron a cabo en el entorno de ANSYS Fluent mediante un modelo tridimensional con transferencia de calor conjugada, considerando la interacción térmica entre el flujo de aire en la tubería y el suelo circundante. Este enfoque ha sido utilizado en estudios CFD de sistemas EAHE debido a su capacidad para representar simultáneamente la conducción en el suelo y la convección en el flujo del aire (Hasan & Noori, 2019).

El flujo de aire fue modelado como turbulento, empleando el modelo estándar $k-\epsilon$, utilizado en estudios de flujos internos con transferencia de calor debido a su robustez, estabilidad numérica y bajo costo computacional (Mahdi et al. 2024). Además, considerando que para las condiciones de operación analizadas (velocidades del orden de 1 a 6 m/s y diámetros característicos del sistema), el flujo se encuentra en régimen turbulento ($Re > 4000$), por lo que el uso de un modelo de turbulencia resulta adecuado para capturar el transporte de cantidad de movimiento y energía.

Por otra parte, para el acoplamiento presión-velocidad se utilizó el algoritmo SIMPLE, mientras que las ecuaciones de transporte fueron discretizadas mediante esquemas de segundo orden con el fin de mejorar la precisión numérica de la solución.

La convergencia del modelo se estableció con base en la reducción de los residuales escalados, considerándose alcanzada cuando los residuales de continuidad, momento y variables turbulentas fueron inferiores a 10^{-4} , y el residual de energía inferior a 10^{-6} . Adicionalmente, se verificó la estabilidad de las variables de interés, como la temperatura del aire a la salida, asegurando que no presentaran variaciones significativas durante las iteraciones finales.

3.5 Condiciones de frontera

En la entrada del dominio se estableció una condición de velocidad uniforme, junto con una temperatura de entrada específica, permitiendo controlar de manera directa los parámetros operativos del flujo de aire que ingresa al sistema.

En la salida se impuso una condición de presión atmosférica, asumiendo que el aire descarga hacia el ambiente sin

restricciones, lo cual es una aproximación común en el modelado de sistemas de ventilación y ductos.

En las paredes internas de la tubería se aplicó una condición de no deslizamiento (no slip), considerando que la velocidad del fluido en contacto con la superficie sólida es nula, lo cual es consistente con el comportamiento físico de fluidos viscosos y permite capturar adecuadamente los efectos de la capa límite.

Finalmente, en el dominio externo del suelo se estableció una condición de temperatura constante, representando la temperatura del suelo no perturbado a la profundidad de enterramiento del sistema, ya que a partir de cierta profundidad la temperatura del suelo presenta variaciones mínimas a lo largo del tiempo, permitiendo modelar el terreno como una fuente térmica estable.

3.6 Modelado de la temperatura del suelo.

La temperatura del suelo fue estimada mediante el modelo analítico propuesto por Ali et al. (2023), el cual describe la variación de la temperatura en función de la profundidad y el tiempo, dependiente de la temperatura media anual, la amplitud térmica, la difusividad térmica del suelo y las propiedades termofísicas.

Dado que son necesarios los datos de las propiedades del suelo se utilizaron datos experimentales reportados por Sailor et al. (2008), para distintas composiciones de suelo y niveles de saturación. En particular, se consideraron la densidad (ρ), conductividad térmica (k) y capacidad calorífica específica (Cp), las cuales determinan la difusividad térmica del suelo y su comportamiento térmico.

Se analizaron ocho tipos de suelo (DH01–DH08) compuestas por diferentes proporciones volumétricas de agregado inorgánico (pumita), arena y materia orgánica (compost), lo que permite representar una amplia variabilidad en condiciones reales de suelo. Estas composiciones junto con el contenido de humedad influyen directamente en las propiedades térmicas y en la distribución de temperatura del suelo.

Tabla 1 Composición (% en volumen) y capacidad de humedad de muestras de suelo.

No. de muestra	Pómez (%)	Lutita expandida (%)	Composta (%)	Arena (%)	Capacidad de humedad (g/g)
DH01	50	0	10	40	0.32
DH02	50	0	0	50	0.30
DH03	75	0	0	25	0.43
DH04	75	0	10	15	0.44
DH05	0	50	10	40	0.23
DH06	0	50	0	50	0.23
DH07	0	75	0	25	0.22
DH08	0	75	10	15	0.24

3.7 Verificación del modelo numérico.

Con el fin de evaluar la confiabilidad del modelo numérico desarrollado, se realizó una verificación mediante la comparación de los resultados obtenidos con aquellos reportados por Hasan & Noori (2019), utilizando las mismas condiciones de operación.

En la *Tabla 2* y la *Figura 3* se presenta la comparación entre las temperaturas de salida del aire obtenidas en ambos estudios. Se observa una buena concordancia entre los resultados, con diferencias absolutas menores a 1 °C y errores relativos bajos.

Estos resultados indican que el modelo numérico implementado reproduce adecuadamente el comportamiento térmico del sistema reportado en la literatura, proporcionando confianza en su uso para el análisis paramétrico bajo diferentes condiciones de operación.

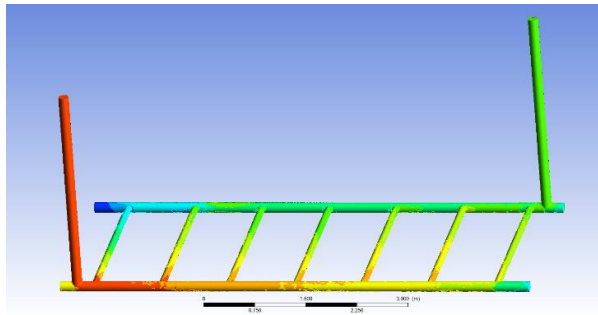


Figura 3. Verificación del sistema 1 para $T_{in}=50^{\circ}$ y $v=6$ m/s.

Tabla 2 Resultados de la verificación del artículo de Hasan et al., 2019.

No. de sistema	Temperatura de entrada (°C)	Temperatura de salida (°C)	Temperatura de salida del trabajo presente	Diferencia abs.	Diferencia (%)
Sistema 1	50.0	39.6	40.4	0.8	2.0

3.8 Selección de condiciones térmicas.

Con el fin de definir condiciones de frontera representativas para el análisis paramétrico, se seleccionaron distintos días del año a partir de la base de datos meteorológicos del año 2024 correspondiente a Cuernavaca, Morelos. La selección se realizó buscando cubrir escenarios térmicos extremos, promedios estacionales y condiciones estadísticas representativas del comportamiento anual de la temperatura ambiente.

En primer lugar, se consideró el día 131 (10 de mayo) como correspondiente a la temperatura máxima anual, ya que representa la condición más severa de calentamiento del aire de entrada y, por tanto, el mayor potencial de enfriamiento del sistema EAHE. De manera complementaria, se tomó el día 269 (25 de septiembre) como asociado a la temperatura mínima anual, con el propósito de representar la condición opuesta dentro del ciclo anual.

Además de los extremos absolutos, se seleccionaron condiciones medias de los meses más representativos del comportamiento térmico anual. El día 135 (15 de mayo) se tomó como día medio del mes más cálido (con promedio de temperatura mensual de 28.8 °C), mientras que el día 349 (14 de diciembre) correspondió al día medio del mes más frío (con promedio de temperatura mensual de 20.3 °C). Esta elección permite evaluar el sistema bajo condiciones más estables y representativas que los extremos puntuales, evitando que el análisis dependa únicamente de eventos aislados.

Asimismo, se consideró el día 183 (1 de julio) para representar la temperatura media anual, ya que constituye una condición de referencia del comportamiento térmico promedio del sitio de estudio. Esta condición resulta útil como caso base para comparar el efecto de los escenarios extremos y de los escenarios promedio mensuales.

Finalmente, con el fin de incorporar un criterio estadístico adicional, se seleccionaron los percentiles 0.05 y 0.95, correspondientes a los días 343 (8 de diciembre) y 135 (15 de mayo), respectivamente. Estos percentiles permiten representar condiciones cercanas a los valores bajos y altos de la distribución anual de temperatura, pero menos sensibles a valores atípicos que los extremos absolutos. De esta forma, su inclusión aporta robustez al análisis paramétrico al considerar no solo máximos y mínimos puntuales, sino también condiciones térmicas estadísticamente representativas.

En conjunto, la selección de estos días permitió construir un conjunto de condiciones de entrada capaz de representar el comportamiento térmico anual desde diferentes enfoques: extremos absolutos, promedios estacionales, promedio anual y percentiles de la distribución.

3.9 Estudio paramétrico.

Con el fin de analizar la influencia de las condiciones de operación sobre el desempeño del sistema, se realizó un estudio paramétrico considerando la velocidad del aire, la temperatura de entrada y la temperatura del suelo. Los parámetros evaluados y sus respectivos intervalos se presentan en la *Tabla 3*.

Tabla 3 Parámetros evaluados.

Parámetro	Valores evaluados	Unidad
Velocidad del aire	1, 2, 3, 4, 5 y 6	m/s
Temperatura de entrada del aire	9 – 38	°C
Temperatura del suelo	21 – 24	°C
Profundidad	1 – 5	m
Tipo de suelo	DH01-DH08	—

Los valores de temperatura de entrada y temperatura de suelo fueron los obtenidos a partir del modelo analítico basado en propiedades termofísicas del suelo y datos meteorológicos locales. A diferencia de otros estudios, en este trabajo la temperatura del suelo no se asume constante, sino que se

obtiene mediante un modelo físico basado en propiedades reales del suelo y datos meteorológicos.

Asimismo, previo al modelado se consideró que las propiedades termofísicas entre las distintas composiciones de suelo muestran un comportamiento térmico similar, no representando cambios significativos en la temperatura del suelo. A partir de 3 m de profundidad, la temperatura se estabiliza y el rango de temperatura del suelo se mantiene cercano a 25 °C. Con base a ello se seleccionó la muestra DH01, con saturación 0 y una profundidad de 3 m, como condición representativa para las simulaciones numéricas del sistema EAHE.

La selección de la composición DH01 se realizó considerando, en primer lugar, que sus propiedades representan de forma razonable una condición de suelo granular-mineral con contenido orgánico, la cual puede considerarse cercana a configuraciones de suelo comúnmente encontradas en la zona de estudio. Además, las diferencias térmicas entre las distintas composiciones analizadas no fueron suficientemente significativas como para modificar de forma importante la tendencia general del comportamiento del suelo.

Por otra parte, se eligió la condición de saturación 0 debido a que permite establecer un caso base de referencia bajo condiciones secas, evitando introducir variaciones adicionales asociadas al contenido de humedad. Dado que el efecto de la saturación sobre la temperatura del suelo tendió a reducirse a mayores profundidades, esta condición resultó adecuada para simplificar el análisis paramétrico y facilitar la comparación entre escenarios.

Respecto a la profundidad, los resultados mostraron que las capas de suelo cercanas a la superficie presentan mayores fluctuaciones térmicas, mientras que a partir de 3 m la temperatura se estabiliza notablemente a lo largo del año. Esto indica que dicha profundidad ofrece un comportamiento térmico suficientemente estable para la operación del EAHE. Al mismo tiempo, profundidades mayores implican mayores requerimientos de excavación e instalación, lo que incrementa el costo del sistema sin aportar mejoras térmicas significativas. Por ello, la profundidad de 3 m fue seleccionada como un punto de equilibrio entre estabilidad térmica y viabilidad económica de instalación. Adicionalmente, los resultados obtenidos permiten evidenciar el comportamiento térmico del suelo bajo condiciones climáticas específicas de un clima cálido subhúmedo (Cwa), proporcionando información que puede ser utilizada como base para el análisis y diseño de sistemas EAHE en estudios futuros bajo condiciones similares.

4. RESULTADOS

4.1 Comportamiento térmico del suelo y la influencia de la profundidad.

Se realizó un análisis de la variación de la temperatura del suelo a lo largo del año considerando diferentes profundidades y con ocho composiciones de suelo (DH01–DH08), que corresponden a distintas mezclas de agregado inorgánico,

materia orgánica y arena en distintas proporciones y distintos niveles de saturación.

En primer lugar, se estudió el efecto de la profundidad sobre la temperatura del suelo para todas las composiciones, aunque por motivos de claridad y espacio a continuación se presenta únicamente en la *Figura 4* el caso correspondiente a la composición DH01 bajo una condición de saturación fija ($sat = 0$). En ella se observa que al incrementar la profundidad hay una disminución progresiva en la amplitud térmica, así como un desfase temporal en la ocurrencia de los máximos y mínimos de temperatura. A profundidades mayores (3-5 m), la temperatura tiende a estabilizarse alrededor de un valor casi constante, evidenciando el efecto amortiguador del suelo frente a las variaciones de la temperatura ambiente.

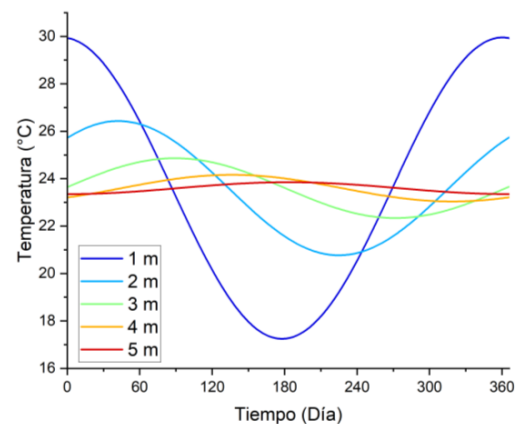


Figura 4. Comportamiento de la temperatura del suelo del compuesto DH01 a diferentes profundidades con saturación 0.

Posteriormente, en la *figura 5* se observa la evolución anual de la temperatura del suelo para las distintas composiciones bajo una condición de saturación representativa ($sat = 0$) a una profundidad de 3 m. Los resultados muestran que, a pesar de las diferencias en sus propiedades termofísicas, todas las composiciones presentan una tendencia térmica similar a lo largo del año, con temperaturas que oscilan en un intervalo aproximado de 22 °C a 25 °C.

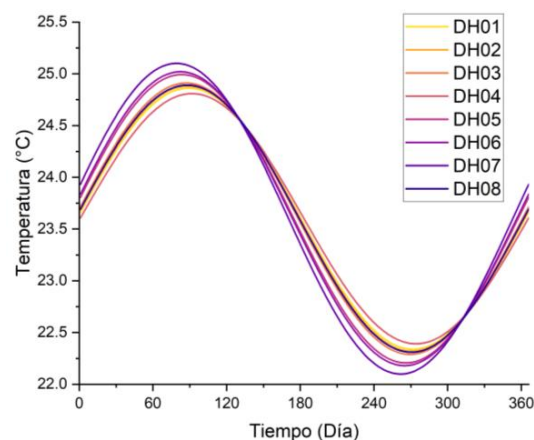


Figura 5. Comportamiento de la temperatura del suelo a diferentes composiciones, a una saturación de 0 y una profundidad de 3 m.



Los resultados muestran que la composición del suelo tiene un efecto secundario en comparación con la profundidad, por lo que es posible emplear una composición representativa sin afectar la tendencia general del sistema.

Además, el análisis de la temperatura del suelo en función de la profundidad mostró que las capas cercanas a la superficie presentan mayores fluctuaciones térmicas debido a la influencia directa de las condiciones ambientales. Sin embargo, a medida que aumenta la profundidad, la amplitud de estas variaciones disminuye significativamente.

4.2 Influencia de cantidad de saturación.

Se analizó la influencia del contenido de saturación sobre el comportamiento térmico del suelo a lo largo del año, considerando distintas composiciones y diferentes profundidades. El incremento en el contenido de humedad modifica las propiedades termofísicas del suelo, ya que al incrementarse el contenido de humedad los poros del suelo contienen mayor cantidad de agua en lugar de aire. Debido a que el agua posee mayor conductividad térmica y mayor capacidad de almacenamiento de calor que el aire, un suelo más húmedo tiende a transferir calor con mayor facilidad y presentar una mayor inercia térmica, como puede apreciarse en el comportamiento del gráfico de la *figura 6*.

Esto implica que el aumento de la saturación puede favorecer una distribución más uniforme de la temperatura dentro del suelo y reducir parcialmente las variaciones térmicas, especialmente en comparación con el suelo seco. Sin embargo, este efecto depende de la profundidad. En capas cercanas a la superficie, donde la temperatura está más influenciada por las condiciones ambientales, las diferencias entre niveles de saturación son más perceptibles.

Conforme aumenta la profundidad, la amplitud de la variación térmica disminuye y la temperatura del suelo tiende a estabilizarse. Por ello, a profundidades cercanas o mayores a 3 m, las curvas correspondientes a distintos niveles de saturación presentan comportamientos similares, con diferencias reducidas entre sí. Esto indica que, aunque la saturación modifica la respuesta térmica del suelo, su influencia sobre la temperatura se vuelve secundaria frente al efecto amortiguador de la profundidad.

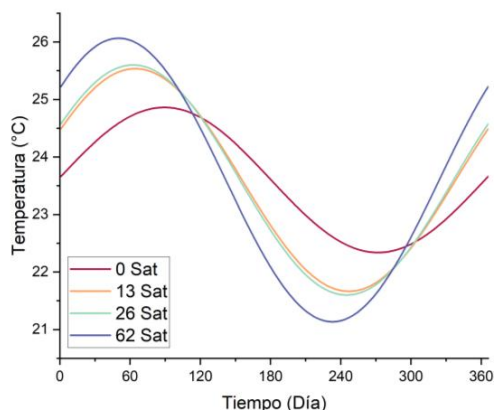


Figura 6. Comportamiento térmico anual de la composición DH01 del suelo a diferentes saturaciones a una profundidad de 3 m.

4.3 Efecto de la velocidad del aire.

Para el análisis paramétrico se seleccionó el suelo tipo DH01, con saturación 0 y una profundidad de 3 m, debido a que en secciones anteriores se observó que estas condiciones representan adecuadamente el comportamiento térmico general del suelo. Se evaluaron velocidades de entrada de 1 a 6 m/s con el fin de analizar el efecto del tiempo de residencia del aire dentro de la tubería sobre la temperatura de salida del sistema EAHE.

Los resultados muestran que la velocidad del aire influye directamente en el intercambio térmico. A bajas velocidades, el aire permanece más tiempo dentro de la tubería, lo que favorece el contacto térmico con el suelo y permite mayores variaciones de temperatura entre la entrada y la salida. En condiciones cálidas, este efecto se traduce en una mayor reducción de la temperatura del aire, mientras que, en condiciones frías, el sistema puede incrementar la temperatura del aire a la salida. Por el contrario, al aumentar la velocidad, el tiempo de residencia disminuye, por lo que el aire intercambia menos calor con el suelo. Esto provoca que la temperatura de salida se aproxime más a la temperatura de entrada y que el valor de ΔT disminuya. Este comportamiento se observa de forma consistente en los diferentes escenarios climáticos evaluados y se cuantifica en la *Tabla 4*.

Tabla 4 Resultados del modelado con variación de velocidad y temperatura.

Condición	Velocidad (m/s)	Temp. entrada (°C)	Temp. suelo (°C)	Temp. salida (°C)	ΔT (°C)	Aprovechamiento térmico (%)
Temperatura máxima anual.	1	37.9	24.2	31.8	6.0	44.3
	2	37.9	24.2	34.2	3.7	27.5
	3	37.9	24.2	35.2	2.7	20.0
	4	37.9	24.2	35.7	2.1	15.8
	5	37.9	24.2	36.1	1.7	13.0
	6	37.9	24.2	36.4	1.5	11.2
Promedio del mes más cálido.	1	28.8	24.1	26.7	2.0	44.4
	2	28.8	24.1	27.5	1.2	27.5
	3	28.8	24.1	27.8	0.9	19.9
	4	28.8	24.1	28.0	0.7	15.9
	5	28.8	24.1	28.1	0.6	13.1
	6	28.8	24.1	28.2	0.5	11.2
Percentil 0.95	1	29.1	24.1	26.9	2.2	44.4
	2	29.1	24.1	27.7	1.3	27.5
	3	29.1	24.1	28.1	1.0	19.9
	4	29.1	24.1	28.3	0.8	15.9
	5	29.1	24.1	28.4	0.6	13.2
	6	29.1	24.1	28.5	0.5	11.2
Media anual	1	23.6	22.5	23.1	0.4	44.3
	2	23.6	22.5	23.3	0.3	27.3
	3	23.6	22.5	23.4	0.2	20.0
	4	23.6	22.5	23.4	0.1	15.8
	5	23.6	22.5	23.4	0.1	13.1
	6	23.6	22.5	23.5	0.1	11.2
Temperatura mínima anual.	1	9.4	21.8	14.9	-5.4	44.4
	2	9.4	21.8	12.8	-3.3	27.3
	3	9.4	21.8	11.9	-2.4	20.0
	4	9.4	21.8	11.4	-1.9	15.9
	5	9.4	21.8	11.0	-1.6	13.1
	6	9.4	21.8	10.8	-1.3	11.2
Promedio del mes más frío.	1	20.2	24.1	22.0	-1.7	44.4
	2	20.2	24.1	21.3	-1.0	27.4
	3	20.2	24.1	21.0	-0.7	20.0
	4	20.2	24.1	20.9	-0.6	15.8
	5	20.2	24.1	20.8	-0.5	13.1
	6	20.2	24.1	20.7	-0.4	11.2
Percentil 0.05	1	19.5	23.9	21.5	-1.9	44.4
	2	19.5	23.9	20.7	-1.2	27.4
	3	19.5	23.9	20.4	-0.8	20.0
	4	19.5	23.9	20.2	-0.7	15.8
	5	19.5	23.9	20.1	-0.5	13.2
	6	19.5	23.9	20.0	-0.5	11.3

4.4 Efecto de la temperatura de entrada.

La temperatura de entrada del aire mostró una influencia en el desempeño térmico del sistema, ya que determina el gradiente térmico inicial entre el aire y el suelo. Cuando la temperatura de entrada es mayor que la temperatura del suelo, el sistema funciona en modo de enfriamiento, transfiriendo calor desde el aire hacia el suelo. En cambio, cuando la temperatura de entrada es menor que la temperatura del suelo, el sistema actúa en sentido inverso, calentando el aire a su paso por la tubería.

Los mayores valores de ΔT se obtuvieron bajo condiciones de temperatura máxima anual, donde la diferencia entre la temperatura del aire y la del suelo fue más elevada. Para este caso, la reducción máxima fue de aproximadamente de 6.08 °C a 1 m/s. En contraste, para condiciones cercanas a la temperatura media anual, el gradiente térmico fue menor y, por lo tanto, la variación de temperatura a la salida fue más reducida.

Esto indica que el desempeño del EAHE depende de las condiciones climáticas de entrada. El sistema presenta mayor potencial de enfriamiento durante periodos cálidos, mientras que en periodos fríos puede contribuir al precalentamiento del aire.

4.5 Efecto de la temperatura del suelo.

La temperatura del suelo actúa como referencia térmica del sistema, por lo que el intercambio de calor depende del gradiente entre el aire de entrada y el suelo. Cuando $T_{suelo} < T_{aire}$, el sistema enfría el aire; cuando $T_{suelo} > T_{aire}$, lo calienta. A 3 m de profundidad, la temperatura del suelo varía en un intervalo reducido, pero incluso pequeñas diferencias afectan el valor de ΔT , lo que resalta la importancia de su correcta estimación.

5. CONCLUSIONES

Los resultados muestran que el desempeño térmico del EAHE está influenciado por las condiciones de operación. A bajas velocidades del aire (1 m/s), se alcanzó una reducción máxima de temperatura de aproximadamente 6.08 °C bajo condiciones de temperatura máxima anual, mientras que a velocidades altas (6 m/s) la reducción disminuyó hasta 1.54 °C, evidenciando la importancia del tiempo de residencia en el intercambio térmico.

La temperatura de entrada del aire también mostró un efecto significativo. Para condiciones cálidas (≈ 37.97 °C), el sistema presentó mayor capacidad de enfriamiento, mientras que en condiciones frías (≈ 9.46 °C), operó en modo de calentamiento, incrementando la temperatura del aire hasta en 5.29 °C, lo que confirma su comportamiento bidireccional.

Asimismo, se observó que la temperatura del aire a la salida tiende a aproximarse a la temperatura del suelo, la cual se mantiene en un intervalo aproximado de 22-25 °C a 3 m de profundidad. A partir de esta profundidad, las variaciones

térmicas del suelo son mínimas, lo que permite una operación estable del sistema.

En términos de operación, se identificó que profundidades cercanas a 3 m y bajas velocidades del aire maximizan el intercambio térmico. Para el clima analizado, el sistema EAHE demostró ser una alternativa viable para el acondicionamiento pasivo del aire, permitiendo reducciones térmicas en condiciones cálidas y contribuyendo al precalentamiento del aire en condiciones frías.

REFERENCIAS

- Abdel-Ghany, A. M., Al-Helal, I. M., Alsadon, A. & Shady, M., 2022. Measuring and predicting the in-ground temperature profile for geothermal energy systems in the desert of arid regions. *Energies*, 15(19), p. 7268.
- Ali, M. H., Kurjak, Z. & Beke, J., 2023. Investigation of earth air heat exchangers functioning in arid locations using Matlab/Simulink. *Renewable Energy*, 209(10.1016/j.renene.2023.04.042), pp. 632-643.
- Hasan, M. I. & Noori, S. W., 2019. A study of the potential of using the Earth to air heat exchanger for cooling and heating of residential buildings in Iraq. *Heat Transfer—Asian Research*, 48(8), pp. 3202-3927.
- Holmes, T. R. H., Owe, M., De Jeu, R. A. M. & Kooi, H., 2008. Estimating the soil temperature profile from a single depth observation: A simple empirical heatflow solution. *Water Resources Research*, 44(2).
- Koshlak, H., 2025. Earth-Air Heat Exchangers: A Comprehensive Review. *Preprints.org*.
- Mahdi, E. J., Algburri, S., Al-Abadi, N., Ahmed, O. K., Ahmed, A. K., 2024. Photovoltaic panel cooling using ground source energy: CFD simulation. *Results in Engineering*, Volumen 22, p. 102144.
- Mihalakakou, G., Souliotis, M., Papadaki, M., Halkos, G., Paravantis, J., Makridis, S., Papaefthimiou, S., 2022. Applications of earth-to-air heat exchangers: A holistic review. *Renewable and Sustainable Energy Reviews*, 155(111921).
- Sailor, D. J., Hutchinson, D. & Bokovoy, L., 2008. Thermal property measurements for ecoroof soils common in the western U.S.. *Energy and Buildings*, 40(7), pp. 1246-1251.
- Trilok Singh, B., Anil, K. & Prashant, B., 2014. Parametric Analysis of Earth-Air Heat Exchanger System Based on CFD Modelling. *International Journal of Power and Renewable Energy Systems*, Volumen 1.

Effect of WAAM-CMT Process Parameters on Microstructure and Electrical Conductivity of Pure Aluminum

Adalberto Guzmán Náfata*. Arturo Gómez Ortega*. Juan M. Salgado López*. Christian Félix Martínez*.

**Centro de Ingeniería y Desarrollo Industrial, 76125, Santiago de Querétaro, Querétaro, MX.*

(e-mail: a.guzman@posgrado.cidesi.edu.mx / arturo.gomez@cidesi.edu.mx).

Abstract: The effect of WAAM-CMT process parameters — boost current, boost duration, and wait current — on the microstructure and electrical conductivity of pure aluminum (Al-1100) walls was investigated using a factorial design of experiments. A 70-layer wall was fabricated and the bottom, middle, and top regions were characterized. Results reveal a predominantly columnar and homogeneous microstructure throughout the wall, with slight variations in electrical conductivity between regions: 34.27 MS/m (bottom), 34.09 MS/m (middle), and 33.94 MS/m (top). These findings demonstrate that electrical conductivity constitutes a reliable indicator of metallurgical quality in WAAM-CMT of pure aluminum.

Palabras clave: Additive manufacturing, WAAM-CMT, pure aluminum, electrical conductivity, microstructure, process parameters.

1. INTRODUCCIÓN

La manufactura aditiva (MA), actualmente se posiciona como una tecnología emergente para la fabricación de componentes geométricamente complejos y/o personalizados mediante la deposición capa por capa, ofreciendo ventajas significativas frente a los procesos convencionales en términos de flexibilidad de diseño, eficiencia en el uso de material, reducción de tiempos de producción y bajo costo (DebRoy et al., 2018).

Dentro de estas tecnologías, la manufactura aditiva por arco y alambre (WAAM) utiliza alambre metálico y arco eléctrico por lo cual puede auxiliarse de tres procesos convencionales de soldadura por arco: soldadura por arco metálico con gas (GMAW), soldadura por arco de tungsteno con gas (GTAW) y soldadura por arco de plasma (PAW), esta tecnología ha ganado especial interés debido a su alta tasa de depósito, bajo costo operativo y capacidad para fabricar estructuras o componentes metálicos de mediano y gran tamaño, lo que la posiciona como una alternativa viable para aplicaciones industriales en diversos sectores como el aeroespacial, automotriz, electrónica, entre otros. (Williams et al., 2016).

De manera particular, la variante basada en Transferencia de Metal en Frío (CMT) permite la transferencia controlada de gotas metálicas, reduciendo significativamente el aporte térmico durante el proceso, característica que la convierte en la opción ideal para su integración con el proceso WAAM, especialmente cuando se emplean materiales de aporte sensibles al calor. Asimismo, el proceso CMT ha demostrado ventajas relevantes en el procesamiento de aluminio, debido a su bajo aporte térmico, estabilidad del arco y reducción de salpicaduras. Estas cualidades contribuyen a mejorar tanto la calidad geométrica como la integridad metalúrgica de los depósitos. Lo anterior resulta particularmente importante en

materiales como el aluminio, ampliamente utilizado en el sector industrial debido a su baja densidad, alta conductividad eléctrica y buena resistencia a la corrosión. (Fang et al., 2021).

No obstante, uno de los principales desafíos del proceso WAAM-CMT radica en el control de los parámetros de operación, ya que estos influyen directamente en la estabilidad del arco, la transferencia metálica y la solidificación del material. En consecuencia, afectan la microestructura, la presencia de defectos, las propiedades finales y en consecuencia la calidad del material depositado. Estudios previos han demostrado que variables como la corriente, el voltaje y la velocidad de alimentación del alambre influyen significativamente en la formación de porosidad, falta de fusión y en la morfología del grano (Gu et al., 2016; Xiong & Zhang 2021).

En este contexto, la conductividad eléctrica se presenta como un parámetro relevante para la evaluación del material depositado, ya que es altamente sensible a la composición química, la microestructura, la presencia de discontinuidades y de impurezas. Diversos estudios han señalado que variaciones en el tamaño de grano, la distribución de fases y la presencia de discontinuidades internas pueden reflejarse directamente en cambios en la conductividad eléctrica, lo que la convierte en un indicador indirecto de la calidad metalúrgica del material (ASM International, 1990; Davis, 1999). Sin embargo, su aplicación como herramienta de evaluación en procesos WAAM-CMT aún es limitada, particularmente en aluminio puro, donde la relación entre parámetros de proceso, microestructura y propiedades funcionales no está completamente establecida (Williams et al., 2016; Fang et al., 2021).

El presente estudio tiene como objetivo analizar el efecto de los parámetros del proceso WAAM-CMT sobre la microestructura y la conductividad eléctrica de depósitos de

aluminio puro (Al-1100), evaluando la uniformidad microestructural y la estabilidad de la conductividad eléctrica a lo largo de las diferentes zonas del muro fabricado, con el fin de establecer correlaciones entre las condiciones de procesamiento y el comportamiento homogéneo observado.

2. METODOLOGÍA

2.1 Material y equipo

Se realizaron depósitos de capas unitarias y multicapas de aluminio puro mediante un sistema de alimentación continua de alambre de 1.2 mm de diámetro (Al-1100, conforme a la norma ASTM B209), cuya composición química se presenta en la Tabla 1.

Tabla 1. Composición química de aluminio puro.

Elemento	Al	Fe + Si	Cu	Mn	Zn	Otros
Wt-%	> 99	< 0.95	0.05-0.20	< 0.05	< 0.10	< 0.05

El proceso se llevó a cabo con movimiento automatizado mediante un brazo robótico FANUC 120iC, el cual opera bajo una arquitectura de control de lazo cerrado que permite el monitoreo y la corrección en tiempo real de la posición de sus seis ejes. El sistema cuenta con un alcance lineal máximo de 1,407 mm, restringido por las dimensiones de la mesa de trabajo de 900 × 500 mm (Figura 1). Como sustrato se empleó una aleación de aluminio Al-6061 (base Al-Mg-Si) con un espesor de 6.35 mm.



Figura 1. Proceso WAAM-CMT

Se realizó una exploración preliminar de parámetros a través de un diseño factorial completo 3^2 con 2 réplicas para evaluar el acabado superficial y salpicadura mediante inspección visual, con los siguientes parámetros: velocidades de alimentación de alambre (WFS) de 5.0, 5.5 y 6.0 m·min⁻¹ y velocidades de desplazamiento (TS) de 6.0, 10.0 y 14.0 mm·s⁻¹, la tabla 2 muestra los resultados.

Tabla 2. Diseño experimental exploratorio 3^2 de Al-1100

Prueba	WFS (m/min)	TS (mm/s)	Acabado Superficial	Salpicadura
1	5.0	10	Buena	Baja
2	6.0	14	Buena	Baja
3	5.5	6	Regular	Alta
4	5.0	14	Buena	Baja
5	5.5	10	Buena	Baja
6	5.5	14	Buena	Baja
7	6.0	10	Regular	Baja
8	5.0	6	Mala	Media
9	6.0	6	Mala	Media
10	5.5	6	Mala	Baja
11	5.0	14	Regular	Baja
12	6.0	6	Mala	Alta
13	5.5	10	Regular	Baja
14	6.0	14	Buena	Media
15	5.0	10	Buena	Baja
16	6.0	10	Regular	Alta
17	5.0	6	Regular	Alta
18	5.5	14	Buena	Alta

Basado en la respuesta del diseño experimental anterior, se seleccionan 4 combinaciones “WFS” y “TS” para ser probadas con depósitos tipo muro (10 capas) y evaluar el crecimiento vertical como respuesta, maximizando de esta manera la principal ventaja del proceso WAAM (Figura 2).



Figura 2. Depósitos tipo muro (10 capas) de Al-1100. (1) WFS:5; TS:10. (2) WFS:6; TS:14. (3) WFS:5.5; TS:10 (4) WFS:5.5. TS: 14

Los resultados del diseño exploratorio mostraron que velocidades de desplazamiento bajas ($TS = 6 \text{ mm} \cdot \text{s}^{-1}$) generaron acabados superficiales deficientes independientemente del WFS empleado, debido al exceso de aporte térmico por unidad de longitud. Por otro lado, velocidades de desplazamiento altas ($TS = 14 \text{ mm} \cdot \text{s}^{-1}$) produjeron cordones con menor sección transversal, limitando el crecimiento vertical por capa.

A partir de estas observaciones, se seleccionaron cuatro combinaciones con $TS = 10 \text{ mm} \cdot \text{s}^{-1}$ para la etapa de evaluación de crecimiento vertical en depósitos multicapa, equilibrando calidad superficial y estabilidad del proceso (Figura 2). La combinación $WFS = 5.5 \text{ m} \cdot \text{min}^{-1}$ y $TS = 10 \text{ mm} \cdot \text{s}^{-1}$ fue seleccionada por presentar el mayor crecimiento vertical multicapa (21.88 mm) junto con la mejor calidad de depósito entre las cuatro combinaciones evaluadas. Se descartó $WFS = 5.0 \text{ m} \cdot \text{min}^{-1}$ por menor aporte de material y $WFS = 6.0$



$\text{m} \cdot \text{min}^{-1}$ por mayor inestabilidad observada en el arco durante la fabricación. El valor de $\text{WFS} = 5.5 \text{ m} \cdot \text{min}^{-1}$ se estableció como punto de partida para la siguiente etapa del estudio.

Los parámetros adicionales de operación se definieron con base en valores reportados en la literatura y recomendaciones del fabricante, se estableció un tiempo de espera entre capas de 90 segundos para controlar la temperatura y evitar acumulación térmica (Giulio, M., et al. 2026), se empleó argón 5.0 de ultra alta pureza como gas de protección con un flujo de $15 \text{ L} \cdot \text{min}^{-1}$, y se fijó una distancia entre boquilla y sustrato de 19 mm, conforme a las recomendaciones técnicas de Lincoln Electric para procesos MIG/MAG.

2.2 Parámetros del proceso CMT

El proceso WAAM-CMT se rige por quince parámetros sinérgicos fundamentales, definidos a partir de la WFS seleccionada. Estos parámetros pueden clasificarse en tres fases: ignición, depósito o plasma, y término del ciclo (Gómez Ortega et al., 2021). El presente trabajo se centra en la fase de plasma, la cual comprende dos etapas: fase de impulso, en la que la corriente eléctrica es alta y funde la punta del alambre de aporte; y la fase de espera, en la que la corriente es baja y el alambre avanza hacia el baño de fusión.

Para una WFS fija de $5.5 \text{ m} \cdot \text{min}^{-1}$, se identificaron tres parámetros de la fase de plasma con mayor influencia sobre la energía lineal del proceso — calculada según la Ecuación (1) — y sobre la dinámica de solidificación durante el ciclo CMT (Figura 3), conforme a lo reportado en la literatura (Gómez Ortega et al., 2021). Los demás parámetros sinérgicos permanecen determinados por la curva característica del proceso para la WFS seleccionada y no son ajustables de forma independiente bajo la configuración del equipo utilizado.

$$E = \frac{VI}{v} \quad (1)$$

donde V es el voltaje, I la corriente eléctrica y v la velocidad de desplazamiento.

Los tres parámetros seleccionados son:

- **Corriente de espera (I_{wait} , parámetro 6):** Define el valor de la corriente eléctrica durante la fase de espera, posterior a la fase de impulso. Se mantiene hasta la inmersión del alambre metálico de aporte en el baño de fusión, lo que provoca la extinción del arco eléctrico y el inicio de la fase de cortocircuito.
- **Corriente de impulso (I_{boost} , parámetro 11):** Establece el valor de la corriente eléctrica durante la fase de plasma, permitiendo la fusión controlada de la punta del alambre metálico. Al variar este parámetro se modifica la energía térmica aportada en el proceso y a su vez el volumen de la gota transferida durante cada ciclo, se calcula mediante la siguiente Ecuación (2):

$$P = \frac{1}{t_f} \int_0^{t_f} UI dt \quad (2)$$

donde U es el voltaje instantáneo, I la corriente eléctrica instantánea y t_f la duración total del ciclo.

- **Tiempo de impulso (t_{boost} , parámetro 12):** Define la duración en milisegundos de la corriente eléctrica de impulso durante la fase de plasma.

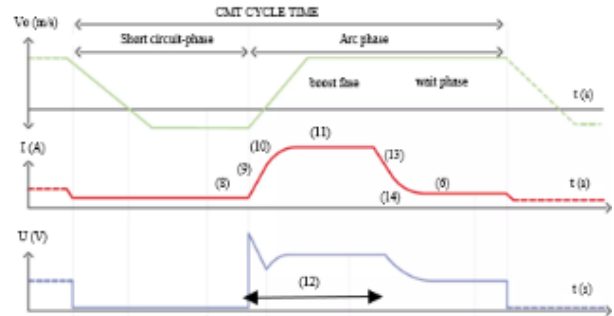


Figura 3. Tiempo y variables durante el ciclo CMT

2.3 Diseño experimental y fabricación del muro

Durante el proceso WAAM-CMT existe una fuerte interdependencia entre variables eléctricas, térmicas y cinemáticas, lo que dificulta aislar sus efectos individuales. Por ello, se requiere una estrategia experimental robusta que permita analizar no solo el impacto de cada variable, sino también las posibles interacciones entre ellas.

En este contexto, se seleccionó un diseño experimental factorial 2^k como estrategia metodológica para el presente estudio ya que permite evaluar de forma sistemática la influencia de múltiples parámetros sinérgicos del proceso CMT sobre las características geométricas de la capa depositada, se consideraron los siguientes tres parámetros: I_{wait} (70–81 A), I_{boost} (150–165 A) y t_{boost} (2.5–3.0 ms), se estableció como variable de respuesta la maximización de la conductividad eléctrica Tabla 3.

Tabla 3. Diseño experimental factorial 2^3 de Al-1100.

Orden	I_{wait} (A)	I_{boost} (A)	t_{boost} (ms)	Réplica	Conductividad (IACS)
1	81	150	2.5	3	46.62
2	70	150	2.5	1	46.48
3	81	165	3.0	2	45.14
4	70	150	3.0	2	45.68
5	70	165	2.5	3	42.46
6	70	165	3.0	1	45.14
7	81	150	3.0	3	49.93
8	81	165	2.5	1	50.64
9	70	165	2.5	2	42.20
10	70	150	3.0	1	40.45
11	81	165	3.0	3	49.12
12	81	150	2.5	2	46.75
13	70	150	2.5	3	49.79
14	70	165	3.0	3	40.36
15	70	150	2.5	2	43.54
16	81	165	2.5	3	45.23
17	70	165	3.0	2	41.26
18	81	150	3.0	2	52.77
19	81	165	3.0	1	49.36
20	70	150	3.0	3	48.60
21	81	150	2.5	1	47.96
22	70	165	2.5	1	45.50
23	81	165	2.5	2	37.91
24	81	150	3.0	1	51.78

Se utilizaron herramientas como análisis de varianza (ANOVA), diagramas de Pareto y gráficos de efectos principales para conocer la significancia estadística de cada factor. De acuerdo con el ANOVA con un nivel de significancia $\alpha = 0.05$, se identificaron 2 factores que presentan significancia durante el proceso I_wait e I_boost, esto puede apreciarse de mejor manera en el diagrama de Pareto y gráfico de interacciones mostrados a continuación en la Figura 4.

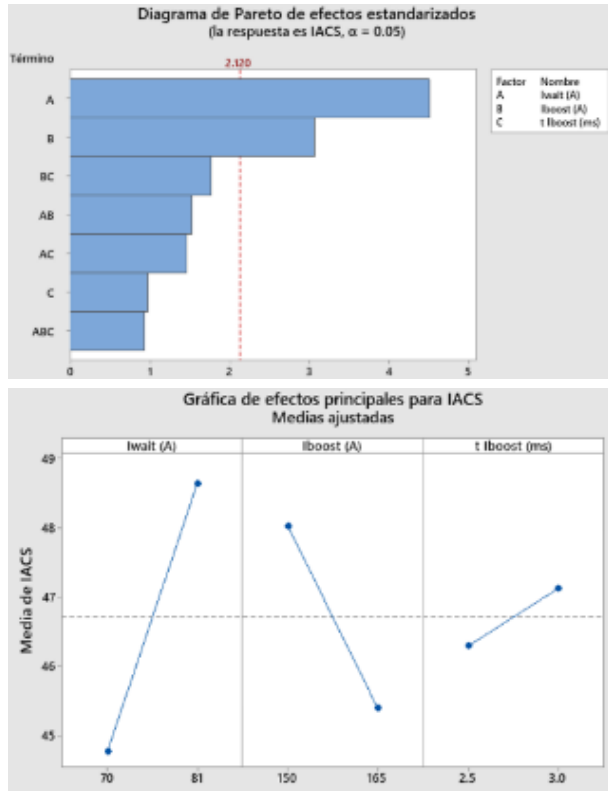


Figura 4. Diagrama de Pareto y gráfico de interacciones

De acuerdo con el objetivo del diseño experimental, la solución óptima se determinó mediante Minitab Statical Software, versión 22.x (Minitab LLC, 2024), ver Figura 5.

Solución

Solución	Iwait (A)	Iboost (A)	t Iboost (ms)	IACS Ajuste	Deseabilidad compuesta
1	81	150	3	51.4933	0.897126

Figura 5. Solución de respuesta múltiple

Con base en los parámetros óptimos obtenidos, se llevó a cabo la fabricación de un muro de aluminio puro Al-1100 mediante el proceso WAAM-CMT. Los parámetros empleados se presentan en la Tabla 4.

Tabla 4. Parámetros de fabricación Al-1100.

Layer	WFS (m/min)	TS (mm/s)	Iboost (A)	t Iboost (ms)	I wait (A)
1-70	5.5	10	150	3.0	81

El muro fabricado se obtuvo mediante el depósito de 70 capas, alcanzando dimensiones finales de 150 mm en la dirección longitudinal y 125 mm en la dirección de construcción. La configuración geométrica final se presenta en la Figura 6.

2.4 Caracterización microestructural

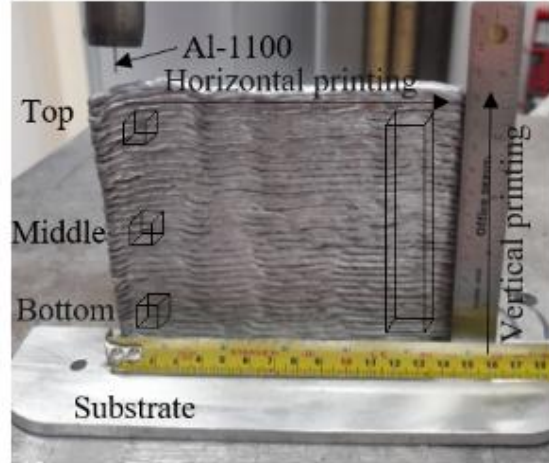


Figura 6. Fabricación tipo muro de Al-1100

La microestructura de las zonas inferior, media y superior del muro fabricado se caracterizó mediante microscopía óptica (Nikon EPIPHOT 200) y microscopía electrónica de barrido (AMETEK APOLLO X). Las probetas se extrajeron de cada zona mediante corte con disco abrasivo, obteniendo muestras con orientación longitudinal (10X10 mm) y transversal (10X6.35 mm) al depósito, como se muestra en la Figura 6. La preparación metalográfica consistió en el lijado con papel abrasivo de carburo de silicio (SiC) con granulometría progresiva de 300 a 3000 grit, seguido de pulido con pasta de diamante de 3 a 0.5 μm hasta obtener un acabado superficial tipo espejo. El ataque químico se realizó con reactivo de Keller, cuya composición es: 5 mL de ácido nítrico (HNO_3), 3 mL de ácido clorhídrico (HCl), 2 mL de ácido fluorhídrico (HF) y 190 mL de agua destilada.

2.5 Medición de conductividad eléctrica

La conductividad eléctrica se determinó mediante un medidor de conductividad por corrientes de Eddy (Eddy Current Conductivity Tester FD-101). Para ello se extrajo una muestra vertical del muro con dimensiones de 10 x 120 mm (ancho x altura), diseñada para abarcar de manera continua las tres zonas de interés: inferior, media y superior de acuerdo al diagrama mostrados en la Figura 6.

Previo a las mediciones, la superficie de la muestra fue sometida a un proceso de preparación metalográfica hasta obtener un acabado tipo espejo, con el fin de minimizar la rugosidad superficial y asegurar un acoplamiento adecuado entre el sensor del medidor de corriente y el espécimen preparado de Al-1100. Este procedimiento permitió garantizar la fiabilidad y repetibilidad de las mediciones de conductividad eléctrica a lo largo de toda la probeta. La caracterización microestructural y eléctrica se muestra en la figura 7.



Figura 7. Caracterización microestructural y eléctrica

3. RESULTADOS Y DISCUSIÓN

3.1 Porosidad

Se observó la presencia de porosidad aislada en todas las regiones inspeccionadas. El poro de mayor diámetro se localizó en la muestra transversal de la zona inferior, con un diámetro de 278.4 μm , identificado a un aumento de $\times 250$ (Figura 8). Estas discontinuidades son características del proceso WAAM cuando se emplea aluminio puro como material de aporte, y se atribuyen principalmente a la alta solubilidad del hidrógeno en el aluminio en estado líquido, la cual disminuye significativamente al solidificarse, favoreciendo la formación de poros (Gou et al., 2015).

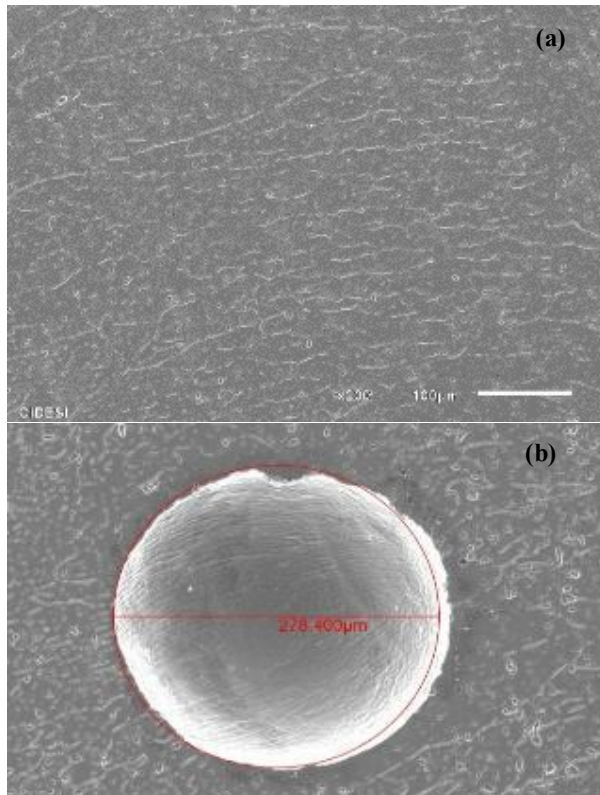


Figura 8. (a) Porosidad aislada (b) Poro mayor en el depósito

3.2 Microestructura

La microestructura del aluminio puro fabricado por WAAM-CMT es característica de un proceso de solidificación bajo altas velocidades de enfriamiento, como ocurre en la manufactura aditiva por depósito de energía directa por arco. Se distingue por el crecimiento de granos columnares con orientación aproximadamente perpendicular a las capas depositadas, siguiendo la dirección del flujo de calor durante la solidificación.

Significativamente, esta morfología de crecimiento de grano se mantiene relativamente constante desde la zona inferior hasta las zonas media y superior del muro, sin evidenciar cambios microestructurales relevantes entre regiones (Figura 9). Este comportamiento se atribuye principalmente a la alta conductividad térmica del aluminio, que favorece una disipación eficiente del calor y mantiene condiciones de solidificación similares a lo largo de la dirección de construcción.

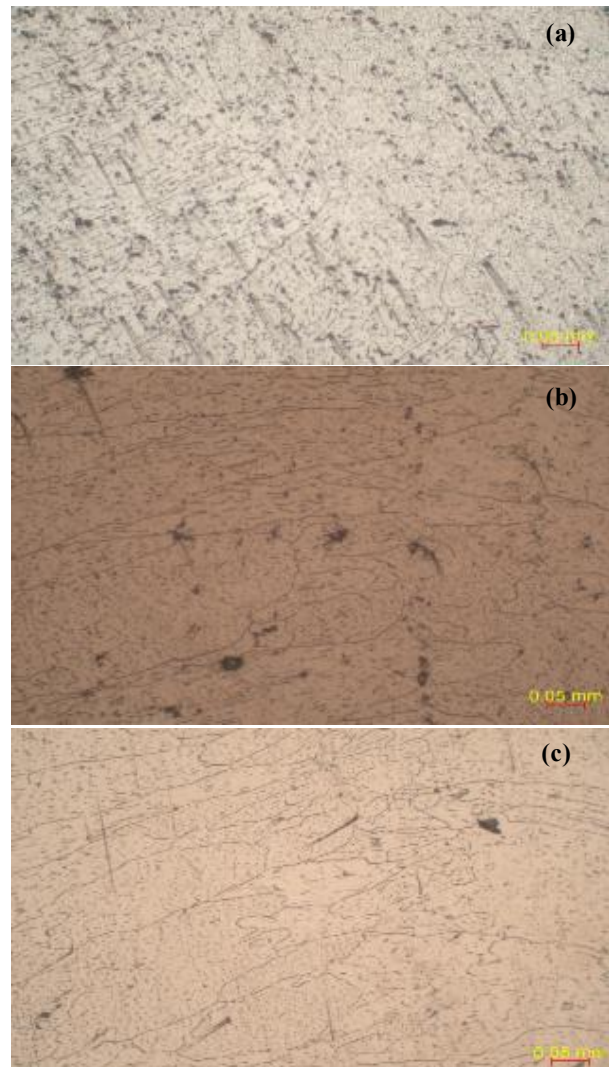


Figura 9. Morfología de grano columnar en zona (a) inferior, (b) media y (c) superior Al-1100 fabricado por MA.

3.3 Conductividad eléctrica

La conductividad eléctrica presentó ligeras variaciones al evaluarse en las distintas regiones del muro. Se realizaron mediciones tanto en la probeta vertical continua como en las muestras preparadas para microscopía. La zona inferior registró los valores más altos, con 34.27 MS/m (59.10% IACS), seguida de la zona media con 34.09 MS/m (58.79% IACS) y la zona superior con el valor más bajo de 33.94 MS/m (58.53% IACS), este efecto se aprecia de mejor manera en el gráfico mostrado en la Figura 10.

La baja dispersión de valores entre regiones — con una variación máxima de 0.26 MS/m entre la zona inferior y la superior — es consistente con la homogeneidad microestructural observada en la sección 4.1, lo que sugiere que las condiciones térmicas similares durante la fabricación favorecieron tanto un crecimiento columnar uniforme como una distribución consistente de defectos internos a lo largo del muro.

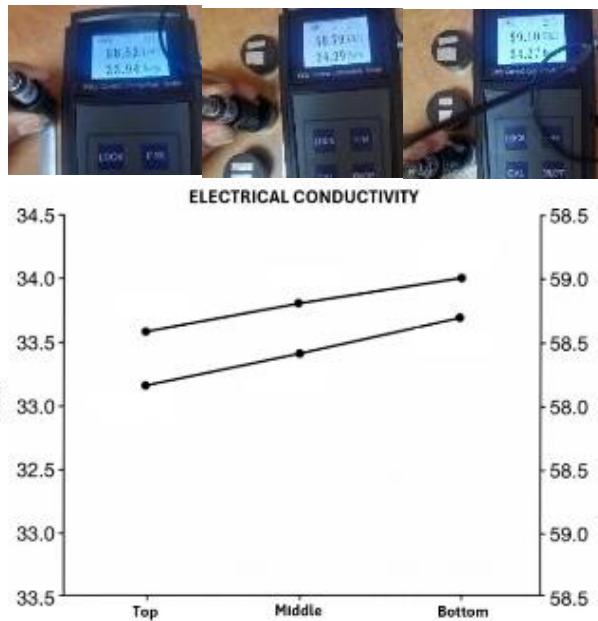


Figura 10. Conductividad eléctrica en las distintas regiones

3.4. Discusión

Los resultados obtenidos demuestran que el proceso WAAM-CMT permite la fabricación de muros de aluminio puro Al-1100 con una microestructura predominantemente columnar y homogénea a lo largo de la dirección de construcción. Este comportamiento contrasta con lo reportado en aleaciones de aluminio de mayor complejidad, donde la acumulación térmica progresiva entre capas suele generar gradientes microestructurales significativos entre las zonas inferior, media y superior del muro (Fang et al., 2021).

Aunque no se realizaron mediciones directas de temperatura durante el proceso, la evidencia experimental disponible sugiere que no se desarrollaron gradientes térmicos pronunciados durante la fabricación. En particular, la baja

variación de conductividad eléctrica entre regiones ($< 0.8\%$) y la similitud microestructural observada a lo largo del muro constituyen indicadores indirectos de condiciones de solidificación relativamente estables en cada capa. Este comportamiento es consistente con la alta conductividad térmica del aluminio puro ($\approx 220 \text{ W} \cdot \text{m}^{-1} \cdot \text{K}^{-1}$), que favorece la disipación del calor acumulado entre capas, limitando el desarrollo de gradientes térmicos significativos a lo largo de la dirección de construcción (Fang et al., 2021). El tiempo de espera de 90 s entre capas, establecido con base en criterios reportados en la literatura (Giulio et al., 2026), contribuyó adicionalmente al control de la temperatura entre depósitos.

La selección de parámetros del proceso CMT mediante un diseño factorial 2^k permitió identificar combinaciones con efectos significativos sobre la conductividad eléctrica del material depositado. En particular, la condición óptima — $I_{\text{wait}} = 81 \text{ A}$, $I_{\text{boost}} = 150 \text{ A}$ y $t_{\text{boost}} = 3.0 \text{ ms}$ — regula de manera directa el aporte térmico durante las fases de impulso y espera del ciclo CMT. El incremento controlado de la corriente de impulso (I_{boost}) y su duración (t_{boost}) favorece una transferencia de material más estable y una adecuada fusión, mientras que una corriente de espera moderada (I_{wait}) contribuye a la disipación térmica entre ciclos, evitando la acumulación excesiva de calor.

Este balance térmico promueve condiciones de solidificación más uniformes, reduciendo la formación de porosidad, favoreciendo una microestructura más homogénea. En consecuencia, se obtiene una ligera mejora en la conductividad eléctrica, al minimizar los obstáculos al movimiento de electrones. Estos resultados son consistentes con lo reportado en la literatura, donde se establece que el control del aporte térmico en procesos CMT influye directamente en la estabilidad del arco, la calidad del depósito y la integridad metalúrgica del material (Gu et al., 2016).

La conductividad eléctrica medida en las tres zonas del muro — 34.27 MS/m en la zona inferior, 34.09 MS/m en la zona media y 34.01 MS/m en la zona superior — mostró una variación máxima de 0.26 MS/m, lo que equivale a una diferencia relativa menor al 0.8% entre regiones. Esta baja dispersión confirma que la conductividad eléctrica es un indicador sensible y confiable de la homogeneidad microestructural en este tipo de proceso. Esto es consistente con lo reportado en la literatura (ASM International, 1990 y Davis, 1999), donde se destaca la alta sensibilidad de esta propiedad frente a variaciones microestructurales y a la presencia de discontinuidades internas.

No obstante, los valores de conductividad eléctrica obtenidos son inferiores a los reportados para el aluminio puro Al-1100 en estado recocido. Para contextualizar cuantitativamente esta diferencia, el Al-1100 recocido presenta una conductividad de aproximadamente 61% IACS (ASM International, 1990), mientras que los valores obtenidos en el presente estudio oscilan entre 58.53% y 59.10% IACS, representando una reducción de entre 1.9 y 2.5 puntos porcentuales IACS. Esta reducción puede descomponerse en contribuciones específicas atribuibles a mecanismos bien documentados en la literatura:



En primer lugar, la porosidad aislada observada en todas las regiones del muro actúa como barrera física al flujo de electrones, reduciendo la sección transversal efectiva del conductor e incrementando la resistividad local del material (Davis, 1999).

En segundo lugar, la microestructura columnar fuera de equilibrio, característica del proceso WAAM-CMT, introduce una mayor densidad de fronteras de grano con orientación preferencial. Estas fronteras actúan como centros de dispersión electrónica, incrementando la resistividad del material respecto a una microestructura equiaxial en equilibrio (ASM International, 1990).

En tercer lugar, las altas velocidades de enfriamiento características del proceso WAAM-CMT favorecen la retención de soluto en solución sólida, generando distorsiones en la red cristalina que reducen la movilidad electrónica. Adicionalmente, los ciclos térmicos repetitivos entre capas promueven la generación de tensiones residuales, las cuales también contribuyen como centros de dispersión de electrones (Davis, 1999; Fang et al., 2021).

En conjunto, estos mecanismos explican la reducción observada en la conductividad eléctrica respecto al material de referencia en estado recocido, y son consistentes con lo reportado en la literatura para materiales de aluminio fabricados mediante procesos de solidificación rápida (Fang et al., 2021). Este resultado pone de manifiesto el impacto de las condiciones de procesamiento sobre las propiedades eléctricas del material y sugiere como línea de trabajo futura la evaluación de tratamientos térmicos post-proceso orientados a promover la recristalización, reducir la densidad de dislocaciones y homogeneizar la microestructura, con el objetivo de incrementar la conductividad eléctrica del material fabricado.

La porosidad aislada observada en todas las regiones del muro constituye un fenómeno inherente al procesamiento de aluminio puro mediante WAAM, asociado a la alta solubilidad del hidrógeno en el aluminio en estado líquido y a su posterior precipitación durante la solidificación (Gou et al., 2015). Sin embargo, la distribución relativamente uniforme de estas discontinuidades a lo largo del muro — sin evidencia de concentraciones preferenciales en ninguna zona — resulta consistente con la homogeneidad microestructural y la estabilidad de conductividad eléctrica previamente observadas, este comportamiento sugiere que los parámetros de proceso seleccionados favorecen condiciones de solidificación estables en cada ciclo de depósito.

Finalmente, los resultados de este estudio representan una contribución al conocimiento del procesamiento de aluminio puro por WAAM-CMT, un material poco reportado en la literatura en comparación con aleaciones de aluminio más complejas como la serie 4xxx y 5xxx (Fang et al., 2021; Williams et al., 2016). La integración de un diseño factorial para la selección de parámetros, en combinación con la caracterización microestructural y la medición de la conductividad eléctrica como indicador de calidad constituye

una metodología sólida y replicable para la optimización de parámetros en procesos de manufactura aditiva por arco.

4. CONCLUSIONES

Se demostró que el proceso WAAM-CMT permite fabricar muros de aluminio puro Al-1100 con microestructura predominantemente columnar y homogénea a lo largo de la dirección de construcción, sin evidenciar cambios microestructurales significativos entre las zonas inferior, media y superior del muro. Este comportamiento se atribuye a la alta conductividad térmica del aluminio puro, que favorece condiciones de solidificación similares en cada capa durante la fabricación.

Mediante un diseño factorial 2^k se identificaron combinaciones óptimas de los parámetros CMT — corriente de espera, corriente de impulso y tiempo de impulso — que maximizan la conductividad eléctrica del material depositado. Los parámetros óptimos seleccionados promovieron un aporte térmico controlado en cada ciclo CMT, contribuyendo a una microestructura más uniforme y a una menor densidad de discontinuidades internas.

La conductividad eléctrica medida en las tres zonas del muro presentó una variación máxima de 0.26 MS/m — equivalente a menos del 0.8% de diferencia relativa — confirmando la homogeneidad del material fabricado y validando el uso de la conductividad eléctrica como un indicador confiable y práctico de la calidad metalúrgica en procesos WAAM-CMT de aluminio puro.

La porosidad observada, distribuida de manera uniforme a lo largo del muro, es característica del procesamiento de aluminio puro por arco y no representó una fuente de variabilidad significativa entre las distintas regiones. Este comportamiento es consistente con los resultados obtenidos tanto en el análisis microestructural como en la medición de la conductividad eléctrica.

Como líneas de trabajo futuro, se proponen los siguientes tópicos de investigación orientadas a ampliar y consolidar los resultados obtenidos:

- En primer lugar, evaluar el efecto de tratamientos térmicos post-proceso sobre la microestructura y la conductividad eléctrica del material fabricado, con el fin de comprender los mecanismos de transformación y su impacto en las propiedades finales.
- Asimismo, se plantea extender el estudio a otras aleaciones de aluminio, considerando condiciones de carga térmica más complejas, lo que permitirá validar y generalizar la metodología propuesta en distintos escenarios de procesamiento.
- Finalmente, se propone orientar la fabricación de aleaciones de aluminio hacia aplicaciones específicas en manufactura aditiva por arco, integrando criterios de diseño, desempeño funcional y requerimientos industriales para el desarrollo de componentes.



REFERENCIAS

- ASM International (1990). Electrical conductivity and resistivity of metals. In ASM Handbook, Volume 2: Properties and Selection: Nonferrous Alloys and Special-Purpose Materials. ASM International, Materials Park, OH.
- Davis, J. R. (1999). Corrosion of aluminum and aluminum alloys. ASM International, Materials Park, OH.
- DebRoy, T., Wei, H. L., Zuback, J. S., Mukherjee, T., Elmer, J. W., Milewski, J. O., and Zhang, W. (2018). Additive manufacturing of metallic components – Process, structure and properties. *Progress in Materials Science*, 92, 112–224. <https://doi.org/10.1016/j.pmatsci.2018.06.001>
- Fang, X., Zhang, L., Chen, G., Dang, X., and Huang, K. (2021). Microstructure and mechanical properties of aluminum alloys fabricated by WAAM-CMT. In *Journal of Materials Processing Technology*, 291, 117044. Elsevier, Amsterdam. <https://doi.org/10.1016/j.jmatprotec.2020.117044>
- Giulio, M., et al. (2026). Influence of interlayer cooling strategies and contact tip to work distance on layer geometry, microstructure, hardness, and corrosion properties of WAAM ER70S-6 steel. In *Welding in the World*. Springer. <https://doi.org/10.1007/s40194-026-02385-3>
- Gómez Ortega, A., Corona Galván, L., Deschaux-Beaume, F., Mezrag, B., and Rouquette, S. (2021). Effect of process parameters on the quality of aluminium alloy Al5Si deposits in wire and arc additive manufacturing using a cold metal transfer process. *International Journal of Advanced Manufacturing Technology* Springer. <https://doi.org/10.1007/s00170-021-07876-8>
- Gou, G., Zhang, M., Chen, H., et al. (2015). Effect of humidity on porosity, microstructure, and fatigue strength of A7N01S-T5 aluminum alloy welded joints in high speed trains. *Materials & Design*, 85, 309–317. <https://doi.org/10.1016/j.matdes.2015.06.137>
- Gu, J., Ding, J., Williams, S. W., Gu, H., Bai, J., Zhai, Y., and Ma, P. (2016). The effect of inter-layer cold working and post-deposition heat treatment on porosity in additively manufactured aluminium alloys. In *Journal of Materials Processing Technology*, 230, 26–34. Elsevier, Amsterdam. <https://doi.org/10.1016/j.jmatprotec.2015.11.006>
- Williams, S. W., Martina, F., Addison, A. C., Ding, J., Pardal, G., and Colegrove, P. (2016). Wire + Arc Additive Manufacturing. In *Materials Science and Technology*, 32(7), 641–647. Taylor & Francis, Abingdon. <https://doi.org/10.1080/02670836.2015.1115753>
- Xiong, J., and Zhang, G. (2021). Effect of process parameters on the quality of aluminium alloy Al5Si deposits in wire and arc additive manufacturing using a cold metal transfer process. *International Journal of Advanced*

Manufacturing Technology Springer.
<https://doi.org/10.1007/s00170-021-07876-8>

Thermal Matching and Annual Performance of PCM Retrofit in Commercial Opaque Envelopes: A Whole-Building Parametric Study

Samanta López Salazar*, Dong Li**, Ruitong Yang***, Wanyu Hu****

*School of Architecture and Civil Engineering, Northeast Petroleum University
China (Tel: 777 1797523; e-mail: samanta.lopez.salazar@gmail.com)

**School of Architecture and Civil Engineering, Northeast Petroleum University
China (e-mail: lidonglvyan@126.com)

***School of Architecture and Civil Engineering, Northeast Petroleum University
China (e-mail: yangruitong17@163.com)

****School of Architecture and Civil Engineering, Northeast Petroleum University
China (e-mail: 136976906@qq.com)

Abstract: PCM performance in opaque envelopes depends on thermal match between latent storage and the dynamic thermal regime of each envelope component. This condition was examined through annual EnergyPlus simulations of a stand-alone retail prototype in Denver, Colorado, with common geometry, envelope definition, internal load profile, schedule set, HVAC system, and climate file. PCM retrofit was considered in external walls, roof, and combined wall–roof application, with transition temperatures from 12 to 35 °C, three layer positions, and thin retrofit layers of 12, 15, and 18 mm. Total site energy was used to rank the parametric cases, while HVAC end uses, peak electric and gas demand, thermal discomfort, energy use intensity, and source-energy indicators were used as complementary performance indicators. Annual PCM performance was governed by the coupling between the latent layer and the indoor-side thermal response. The strongest response was obtained with interior placement, 18 mm thickness, and transition temperatures near 23–25 °C. Total site energy was reduced by 1.77%, 7.44%, and 10.37% in the best wall, roof, and wall–roof cases, respectively. In the selected wall–roof case, gas heating demand and electric cooling demand were reduced by 48.68% and 25.92%, thermal discomfort by 3.85%, and electric and gas peak demand by 0.44% and 12.30%. Fan electricity increased under the preserved HVAC control structure, indicating that the main PCM benefit was concentrated at coil-load level. These results define a controlled Denver retail reference case for PCM selection in commercial opaque envelopes.

Keywords: Phase change material; Opaque envelope; Building energy simulation; Thermal matching; Parametric analysis.

1. INTRODUCTION

Transient heat transfer through the building envelope mediates the exchange of energy between the outdoor environment and the conditioned indoor space (Wijesuriya et al., 2025). This response arises from conduction across multilayer assemblies, convective and long-wave radiative exchange at exposed surfaces, and heat storage within construction materials (Han et al., 2023). Hourly variations in outdoor air temperature, solar irradiance, effective sky temperature, and wind impose time-dependent boundary conditions that the envelope transforms into attenuated and phase-shifted heat fluxes delivered to the conditioned zone (Guo et al., 2024). The resulting sensible heat gains and losses define the temporal profile of heating and cooling demand, and the HVAC capacity and operating schedule required to maintain indoor thermal conditions (Cui et al., 2024). Recent sectoral assessments place buildings at 32% of global energy demand and 34% of energy-related CO₂ emissions in 2023. Building operations remained near 30% of global final energy consumption in 2024, and commercial and public buildings

accounted for about 30% of total building energy demand (IEA, 2025).

Research on building envelopes has examined how material thermophysics, layer topology, and external boundary conditions control heat transport, thermal storage, and the dynamic transmission of loads across the building skin (Ghazwani et al., 2024; Kitsopoulou et al., 2024; Laasri et al., 2024; Dabous and Hosny, 2025; Cuce, 2025; Zhang et al., 2026). Phase change materials are treated in this literature as thermally activated latent-storage media that augment the effective heat capacity of opaque envelope assemblies over a prescribed temperature interval. Their role lies in buffering peak thermal loads, delaying heat transmission, and redistributing heat gains and losses over time, provided that the phase-change interval remains aligned with the thermal trajectory imposed by climate and building operation (Jha et al., 2024; Dora et al., 2026). The performance of PCM-integrated envelopes is controlled by the coupling among phase-transition temperature, storage capacity per unit area, layer placement, envelope component, and the boundary conditions imposed by climate and building operation (Sun et

al., 2021; Yang et al., 2023; Zhao et al., 2024; Gutiérrez et al., 2025). Across comparative studies, effective PCM layers are frequently reported in the 10–30 mm range, while wall and roof applications often place the active layer near the interior side of the assembly (Jia et al., 2021). Dardouri et al. (2023) linked 21 °C to heating periods and 29 °C to cooling periods, and reported demand reductions of up to 41.6% for combined wall–roof applications with thicknesses up to 40 mm. Austin et al. (2022) found a stronger annual response for roof integration than for wall integration in Panama and identified 27 °C as the transition temperature required for completion of the annual charge–discharge cycle. Chan et al. (2025) reported annual savings of 35.24–35.47% in Riverside and 18.52–22.13% in Palm Springs. Hou et al. (2023) reported annual reductions of 20.7% in traditional dwellings, while Alsagri et al. (2025) reported reductions of 20.72% in energy use and 21.05% in CO₂ emissions for a zero-energy building with PCM29 at 50 mm.

Previous studies have clarified the variables that govern PCM performance in opaque envelopes and have reported high-performing configurations for wall, roof, and combined deployment. What remains unresolved is the comparability of those selections across envelope subsystems. Published optima are usually embedded in study-specific combinations of building typology, internal loads, schedules, control assumptions, climate, and evaluation metric. Under these conditions, a configuration selected for one study cannot be interpreted as a directly comparable design outcome for another. Thus, the open issue is more specific: whether PCM selection is preserved across wall, roof, and combined deployment when the comparison is conducted on a common and controlled basis, or whether the preferred configuration shifts with the integration domain itself.

This study addresses that issue through annual whole-building simulation in EnergyPlus for a code-based stand-alone retail reference prototype in Denver, Colorado. External walls, roof, and combined wall–roof deployment are examined under the same geometry, internal loads, schedules, and operating conditions. For each domain, the parametric space is defined by nominal phase-transition temperature, active-layer thickness, and layer position within the opaque assembly. This formulation enables a direct comparison of domain-specific selections and establishes whether the preferred PCM configuration is intrinsic to the material choice or contingent on the envelope subsystem in which it is implemented.

2. PARAMETRIC SETUP AND SELECTION CRITERIA

This section defines the parametric space adopted for PCM deployment in opaque envelope components and the basis used to select the preferred configuration in each application domain.

2.1 PCM deployment parameters

The analysis was structured around the variables that directly govern PCM behavior within the opaque assembly: nominal phase-transition temperature, active-layer thickness, and layer position within the stratigraphy. These variables were examined for three deployment cases: external walls, roof, and combined wall–roof application.

Nominal phase-transition temperature defines the thermal interval over which latent storage is activated. Active-layer thickness determines the storage capacity incorporated into the component. Layer position governs the thermal coupling between PCM activation and the indoor-side temperature field. External walls and roofs were examined separately because each one is characterized by a distinct exposure pattern and heat-transfer path. Combined wall–roof application was included to evaluate the response of simultaneous integration under the same building-scale conditions.

2.2 Selection metric and evaluation indicators

Configuration selection was performed at full-model scale with annual total site energy as the primary metric. This metric ranks each PCM variant according to the energy use at the model boundary, before source-energy factors or emissions coefficients alter the relative weight of electricity and gas. All variants shared the same prototype, climate file, HVAC system, internal loads, schedules, thermostat settings, and fuel structure; differences in total site energy therefore represent the annual response to transition temperature, layer position, and thickness under a common basis. The selected cases were then interpreted through HVAC end-use disaggregation, annual maximum electric demand, annual maximum gas demand, thermal discomfort, energy use intensity, and total source energy. This hierarchy keeps the selection criterion separate from the complementary impact indicators and maintains a consistent comparison across wall, roof, and wall–roof retrofit domains.

3. ENERGYPLUS IMPLEMENTATION

Figure 1 shows the methodological workflow adopted for the EnergyPlus assessment, structured as a continuous sequence from reference-model definition to comparative evaluation of the PCM cases.

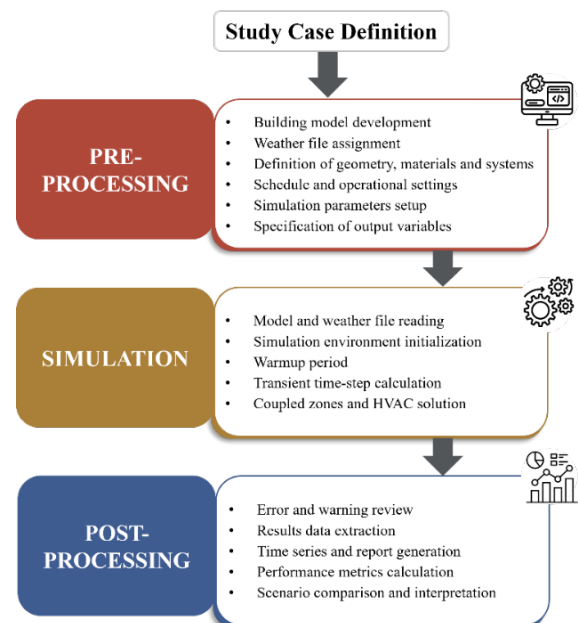


Figure 1. Methodological workflow used for the EnergyPlus-based assessment.

3.1 Reference building and climatic condition

A stand-alone retail building from the DOE/PNNL prototype series developed for ASHRAE 90.1 was adopted as the reference case for the present study (U.S. Department of Energy, Building Energy Codes Program, 2023). This prototype provided a standardized basis for geometry, envelope construction, internal loads, schedules, ventilation, infiltration, and HVAC configuration, allowing the PCM assessment to be conducted within a consistent whole-building framework. Its selection also placed the analysis in a commercial building type with sustained operational loads and a representative interaction among envelope heat transfer, HVAC response, and internal gains.

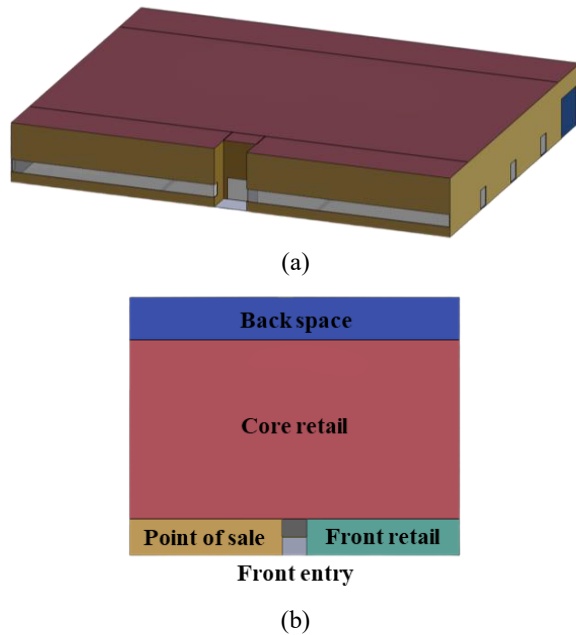


Figure 2. Reference stand-alone retail prototype: (a) building model and (b) thermal zoning layout.

Denver, Colorado, corresponding to ASHRAE climate zone 5B, was selected as the climatic condition. This location represents a cold-dry climate with a pronounced heating season and an appreciable cooling requirement, providing a suitable basis for examining the annual response of PCM-integrated opaque envelopes under mixed thermal demand. Figure 2 shows the simulated retail prototype and its thermal zoning layout, while the main model specifications are summarized in Table 1.

The reference building has a total floor area of 2294 m² distributed over a single story, with a floor-to-ceiling height of 6.096 m. The plan comprises Back Space, Core Retail, Point of Sale, Front Retail, and Front Entry zones. The envelope consists of a concrete wall with continuous insulation and gypsum board in the external walls, and a built-up roofing membrane and rigid insulation in the roof assembly. The glazing system is characterized by a window U-factor of 1.4082 Wm⁻²K⁻¹, a solar heat gain coefficient of 0.21, and a visible transmittance of 0.29. The HVAC system follows the reference prototype definition shown in Table 2, with

packaged single-zone air-conditioning units with gas heating in the main retail zones and a gas unit heater in the front entry. Heating and cooling setpoints, setback values, and outdoor air rates were also preserved from the prototype specification in order to maintain a common operational basis throughout the parametric analysis.

Table 1. Geometric and envelope specifications of the stand-alone retail reference building (García et al., 2022).

Characteristic	Description
	2294.00 (54.27 m × 42.27 m)
	- Back Space 379.90 (54.27 m × 7.00 m) - Core Retail 1600.50 (54.27 m × 29.49 m) - Point of Sale 150.80 (25.14 m × 6.00 m) - Front Retail 150.80 (25.14 m × 6.00 m) - Front Entry 12.00 (4.00 m × 3.00 m)
Floor area, m ² (length × width)	
Number of floors	1
Floor to ceiling height, m	6.096
Window-to-wall ratio	7.13 %
South-facing façade glazing ratio	25.13 %
Building envelope (outer layer, intermediate layer, inner layer)	
Walls	concrete wall (0.2032 m) + continuous insulation (R=1.4031 m ² K ⁻¹ W ⁻¹) + gypsum board (0.01270 m)
Roof	Built-up roofing membrane (0.0095 m) + rigid thermal insulation (R=5.3066 m ² K ⁻¹ W ⁻¹)
Window U-factor	1.4082 Wm ⁻² K ⁻¹
Solar heat gain coefficient (SHGC)	0.21
Visible transmittance	0.29
Foundation	F-factor: 0.90012 Wm ⁻¹ K ⁻¹

Table 2. HVAC system and control specifications of the stand-alone retail reference building (García et al., 2022).

HVAC	
Heating type	PSZ-AC with gas heating in Back space, Core retail, Point of sale and Front retail; gas unit heater in Front entry

Cooling type	Air-cooled DX cooling in Back space, Core retail, Point of sale and Front retail; none in Front entry
Size	Autosized
Heating efficiency	0.88 – 0.891
Cooling efficiency	COP 3.57 – 4.32
Thermostat setpoint, °C	Heating: 21.11 °C in the main zones, 16.0 °C in the front entry; Cooling: 23.89 °C.
Thermostat setback, °C	Heating: 15.56 °C; Cooling: 29.44 °C in the zone assigned to the optimum cooling schedule; no cooling setback in the remaining retail zones.
Ventilation	OA by area/person: 0.00061 m ³ /s·m ² (Back space), 0.00354 m ³ /s·person + 0.00061 m ³ /s·m ² (main retail zones), 0.00118 m ³ /s·m ² (Front entry).

3.2 Simulation framework and parametric implementation

The parametric study was implemented in EnergyPlus through systematic reassignment of PCM-integrated constructions to the envelope elements associated with each application case. For each run, the baseline model was preserved and only the construction definitions of the corresponding opaque surfaces were modified according to the combinations established for the screening stage. This procedure produced the complete simulation set required for the parametric assessment and enabled the annual response of each construction variant to be evaluated within the same modelling environment.

3.3 Parametric case definition

Figure 3 presents the PCM deployment matrix examined for the external wall and roof assemblies. For each component, the PCM layer was placed at the outdoor side, within the assembly core, or at the indoor side, so that the parametric screening could resolve the effect of layer placement on the annual building response. The analysis was restricted to external walls and the roof, as these opaque envelope elements are directly exposed to external boundary conditions and govern the thermal load transmitted to the conditioned zones. Interior partitions were outside the scope of the screening stage, since they do not act as primary interfaces between the building and the outdoor environment. The PCM layer was added as a retrofit layer without reducing the baseline insulation thickness. Consequently, the reported response includes the combined effect of latent storage, added sensible mass, and the additional thermal resistance introduced by the PCM layer.

The PCM set comprised nine materials, identified as P12, P16, P19, P23, P25, P27, P31, P33, and P35 according to nominal

phase-change temperature. Each material was examined at three layer positions and three thicknesses of 12, 15, and 18 mm. The screening matrix included 81 PCM configurations for each retrofit domain, corresponding to nine transition temperatures, three layer positions, and three thicknesses. The full parametric set comprised 243 PCM cases plus the baseline model. For the combined wall–roof domain, the same PCM identifier, layer position, and thickness were assigned simultaneously to the wall and roof assemblies in each case.

The selected thicknesses define a thin retrofit band below 20 mm, consistent with practical wall and roof insertion and with previous PCM envelope work that identifies 10 mm as a cost-effective lower reference for latent layers (Austin et al., 2022). For the material properties used here, these thicknesses provide latent storage capacities of 1.94–2.47 MJm⁻² at 12 mm, 2.43–3.09 MJm⁻² at 15 mm, and 2.92–3.71 MJm⁻² at 18 mm. The phase-change interval was prescribed as $T_{pc,nom} \pm 1$ °C. A common material-property basis was adopted across the PCM set, with density, specific heat capacity, and thermal conductivity fixed at 1330 kgm⁻³, 1930 Jkg⁻¹K⁻¹, and 0.30 Wm⁻¹K⁻¹, respectively. Latent heat varied from 121.80 to 155.00 kJkg⁻¹ according to the nominal transition temperature assigned to each PCM, and the enthalpy–temperature curves defined the latent response solved by CondFD. This material definition frames the PCM set as an idealized family for controlled comparison of transition temperature, latent capacity, layer position, and thickness. Product-level application would require measured enthalpy curves and temperature-dependent properties for the selected PCM formulation.

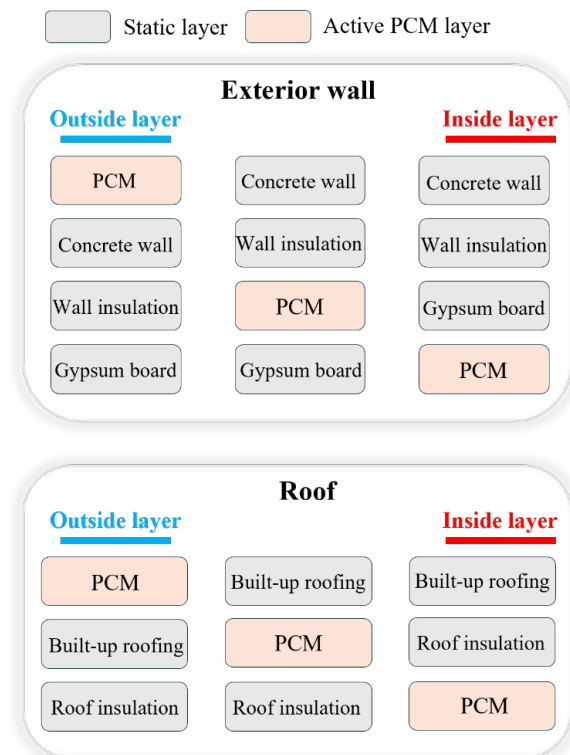


Figure 3. Simplified PCM insertion schemes considered in the parametric analysis for wall and roof assemblies.

3.4 Model configuration and numerical settings

Annual simulations covered the period from 1 January to 31 December with 20 time steps per hour. At each time step, EnergyPlus solved the transient thermal response of the building envelope through the inside-surface heat-balance formulation (Mazidi et al., 2025). For the interior face of an opaque element, Eq. (1) accounts for long-wave radiative exchange, short-wave radiation, long-wave radiation from internal sources, absorbed transmitted solar radiation, conduction through the construction, and convection with the zone air.

$$q''_{LWX} + q''_{SW} + q''_{LWS} + q''_{k,i} + q''_{sol} + q''_{conv} = 0 \quad (1)$$

The convective heat flux in Eq. (1) is evaluated from the temperature difference between the inside surface and the zone air, as given in Eq. (2).

$$q''_{conv} = h_c(T_s - T_a) \quad (2)$$

where T_s is the inside surface temperature, T_a is the zone-air temperature, and h_c is the convective heat-transfer coefficient. Interior and exterior surface convection were modeled with the TARP algorithm. For interior surfaces, EnergyPlus evaluates h_c from correlations that depend on surface orientation and on the surface-to-air temperature difference. For a vertical wall, the corresponding relation is given in Eq. (3), with $\Delta T = T_s - T_a$. For exterior surfaces, the convective heat transfer coefficient includes natural and forced contributions within the same convection model.

$$h_c = 1.31|\Delta T|^{1/3} \quad (3)$$

Heat conduction through the opaque envelope assemblies was solved using the Conduction Finite Difference (CondFD) algorithm. The heat-balance method was set to CondFD for the annual simulations so that the enthalpy–temperature response of the PCM layers could be resolved within the wall, roof, and combined wall–roof constructions (Feng et al., 2022). For an internal node, the finite-difference energy balance is written as Eq. (4).

$$C_p \rho \Delta x \frac{T_i^{j+1} - T_i^j}{\Delta t} = k_w \frac{T_{i+1}^{j+1} - T_i^{j+1}}{\Delta x} + k_E \frac{T_{i-1}^{j+1} - T_i^{j+1}}{\Delta x} \quad (4)$$

Here, i and j denote the spatial node and time level, respectively; ρ is the density, C_p the specific heat, Δx the nodal thickness, Δt the time step, and k_w and k_E the conductivities at the adjacent interfaces. For PCM layers, latent storage enters the CondFD formulation through the enthalpy–temperature function given in Eq. (5).

$$h_i = HTF(T_i) \quad (5)$$

where HTF denotes the material enthalpy–temperature function. Under the hysteresis formulation implemented in EnergyPlus, Eq. (5) is evaluated along separate melting and freezing branches, so the material state depends on the current temperature and on the preceding phase path. All simulations were performed under the reference schedules and control settings defined for the RetailStandalone prototype. Equations (1)–(5) define the numerical scheme adopted for the annual evaluation of the PCM configurations.

The prototype HVAC schedules and thermostat settings were retained across all simulations. HVAC availability followed the reference operating schedule, with system operation from 06:00 to 21:00 on weekdays, from 06:00 to 22:00 on Saturdays, and from 08:00 to 19:00 on Sundays and holidays. During occupied hours, the main conditioned zones were controlled to a heating setpoint of 21.11 °C, whereas the front entry was maintained at 16.0 °C; the heating setback temperature was 15.56 °C. The occupied cooling setpoint was 23.89 °C. A cooling setback of 29.44 °C was applied in the zone assigned to the optimum cooling schedule, while the remaining retail zones operated without cooling setback (U.S. Department of Energy, Building Energy Codes Program, 2023; Goel et al., 2023). Numerical settings and control inputs were kept unchanged throughout the parametric campaign, and the same output structure was used for all cases.

3.5 Performance metrics and comparison basis

The comparison was based on annual EnergyPlus outputs for the reference model and the PCM cases. Total site energy defined the parametric rank order as the direct annual energy use at the model boundary under a common prototype, climate file, HVAC system, schedules, and fuel structure (Sangwan et al., 2022). For the selected configurations, HVAC end-use disaggregation, annual peak electric and gas demand, thermal discomfort, energy use intensity, and total source energy were examined as complementary indicators on a consistent whole-model basis. Annual peak demand characterized the peak-load response at model scale (Laafer et al., 2026), whereas energy use intensity provided an area-normalized view and total source energy added a primary-energy conversion view of the annual energy balance (Kulumkanov et al., 2024). Thermal discomfort was quantified as the annual sum of heating and cooling unmet setpoint hours reported by EnergyPlus for the conditioned zones. Therefore, this value represents cumulative zone-level discomfort hours rather than a single-zone annual duration. All cases were evaluated from the same annual output set, with the rank order tied to total site energy and the complementary indicators used to interpret energy, demand, and comfort impacts.

3.6 Numerical basis and PCM material check

The PCM layer definition was reviewed against the EnergyPlus CondFD formulation and the enthalpy–temperature data assigned in the IDF. CondFD is the EnergyPlus heat-balance method for envelope layers with phase-change behavior or variable thermal properties; latent storage enters the solution through the enthalpy–temperature relation, from which an equivalent specific heat is evaluated at each time step. The model used 20 time steps per hour, equal

to a 3-min step, consistent with the EnergyPlus Reference for finite-difference models. This time-step choice also agrees with Tabares-Velasco et al. (2012), who verified and validated the EnergyPlus PCM model for opaque wall assemblies and recommended time steps of 3 min or less for PCM cases. Figure 4 complements this numerical check with the enthalpy–temperature curves obtained from the IDF material inputs. The curves confirm that P12, P16, P19, P23, P25, P27, P31, P33, and P35 follow the prescribed transition interval of $T_{pc,nom} \pm 1\text{ }^{\circ}\text{C}$ and preserve the latent heat range assigned to the PCM set, from 121.80 to 155.00 kJkg⁻¹.

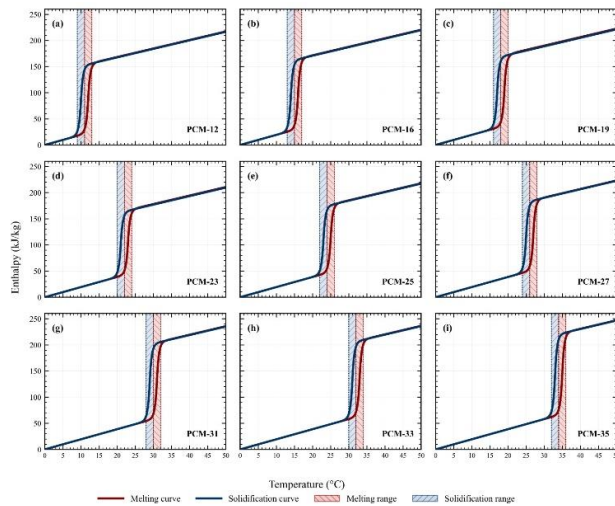


Figure 4. Enthalpy–temperature curves generated from the EnergyPlus PCM material definitions.

4. RESULTS

This section presents the results of the parametric PCM analysis under the baseline operating schedule. First, the baseline building performance is described. Then, the effects of transition temperature, PCM position, and thickness are examined for exterior walls, the roof, and the combined wall–roof configuration.

4.1 Energy performance of the baseline building

Figure 5a presents the annual end-use distribution of the baseline building. For a conditioned floor area of 2,293.99 m², total site energy reached 285,928.25 kWh, equivalent to 124.64 kWhm⁻², and total source energy reached 600,692.05 kWh, equivalent to 261.85 kWhm⁻². Space heating represented the largest individual end use, with 85,595.19 kWh, corresponding to 29.94% of total site energy. Non-heating end uses represented a substantial fraction of the annual site energy, which limits the relative impact of envelope-based PCM storage on whole-building energy use. Under this baseline condition, PCM assessment must therefore be interpreted in relation to the heat-transfer-driven share of HVAC demand, since envelope-based latent storage acts through transient heat exchange across the external walls and roof.

Figure 5b presents the monthly HVAC energy profile of the baseline case. Heating dominated the cold season, whereas

cooling and fan energy gained relevance during the warm period. This seasonal split defines the operating context for PCM retrofit: reduction of heat losses during cold months and attenuation of heat gains during warm months.

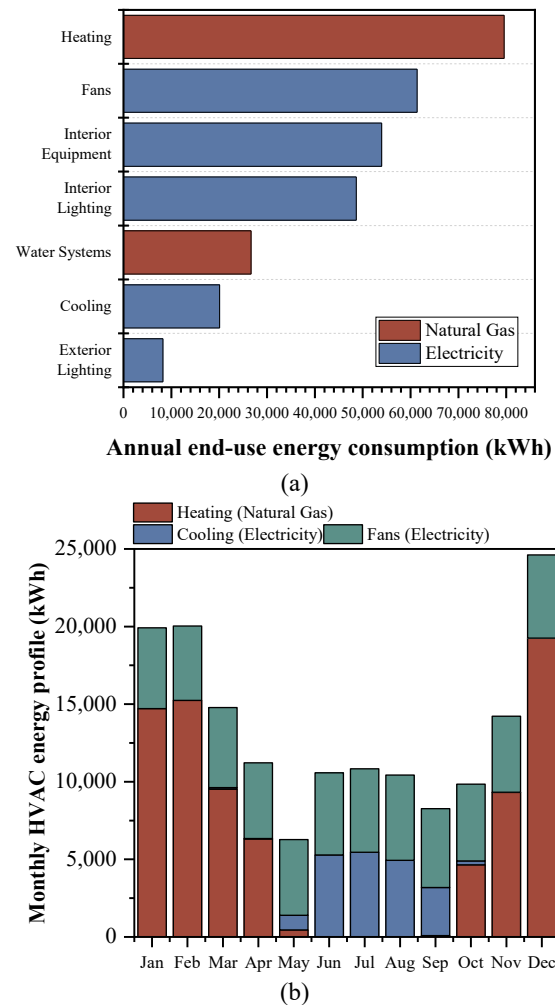


Figure 5. Baseline building energy characterization through (a) annual end-use energy consumption and (b) monthly HVAC energy profile.

4.2 Annual energy ranking of PCM retrofit configurations

The annual energy results for wall integration, presented in Figure 6, remain close to the reference case, with Site EUI from 122.44 to 134.57 kWhm⁻² and total site energy from a 1.77% reduction to a 7.96% increase. P23 records the lowest annual demand in this set. At 18 mm, interior P23 reaches 122.44 kWhm⁻² and middle-layer P23 123.36 kWhm⁻², equivalent to reductions of 1.77% and 1.03%, respectively. Exterior P23 and interior P16 remain in the penalty range.

A broader annual-energy spread appears in the roof results presented in Figure 7, where Site EUI extends from 115.37 to 148.06 kWhm⁻² and total site energy from a 7.44% reduction to an 18.79% increase. At 18 mm, the lowest annual demand corresponds to interior P23 and interior P25, with Site EUI values of 115.37 and 115.39 kWhm⁻². Interior P19 reaches the highest Site EUI, 148.06 kWhm⁻², and records an 18.79%

increase. Middle-layer and exterior placement remain above the baseline.

The annual energy results for the combined wall–roof cases, shown in Figure 8, extend this spread further. Site EUI ranges from 111.72 to 150.75 kWhm⁻², while total site energy ranges from a 10.37% reduction to a 20.94% increase. At 18 mm, interior P23 reaches the lowest Site EUI, 111.72 kWhm⁻², followed by interior P25 with 112.32 kWhm⁻²; these cases reduce total site energy by 10.37% and 9.89%, respectively. Interior P19 reaches the highest Site EUI, 150.75 kWhm⁻², and records a 20.94% increase. Mid-layer placement crosses the baseline, whereas exterior placement remains within a 7.33–9.39% increase.

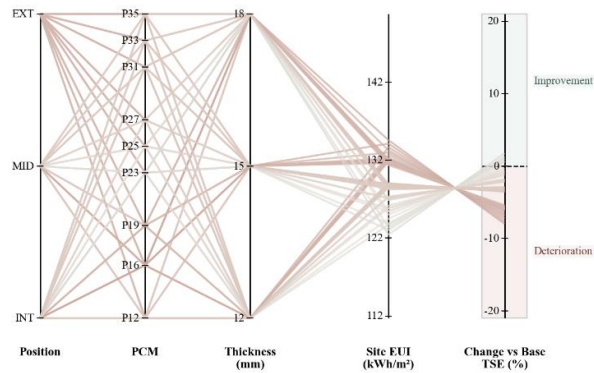


Figure 6. Wall: Multivariate relationship between PCM design parameters and energy performance.

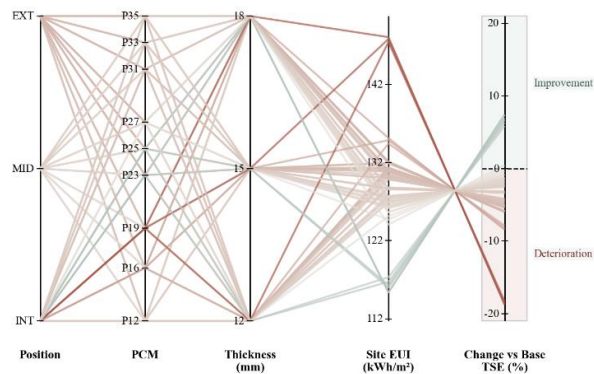


Figure 7. Roof: Multivariate relationship between PCM design parameters and energy performance.

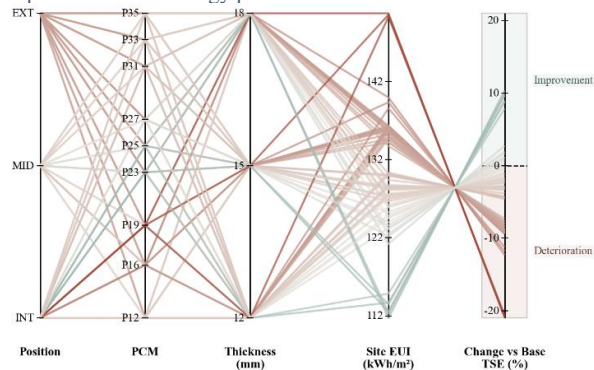


Figure 8. Wall–roof: Multivariate relationship between PCM design parameters and energy performance.

Annual energy ranking places combined wall–roof retrofit first, roof retrofit second, and wall retrofit third. The same order applies to the design variables: interior placement ranks first, middle-layer placement second, and exterior placement last; P23 ranks ahead of P25, and 18 mm corresponds to the best annual-energy response within the evaluated range.

4.3 Annual performance comparison of the selected PCM retrofit against the baseline case

The annual comparison summarized in Figure 9 shows that the principal effect of PCM retrofit appears in the space-conditioning end uses. Among the selected cases, the wall–roof configuration records the lowest gas heating and electric cooling requirements, 43,930.15 and 15,967.64 kWh, equivalent to reductions of 48.68% and 25.92% relative to the baseline, while thermal discomfort also reaches its lowest level, 3382.00 h, although with a smaller reduction of 3.85%.

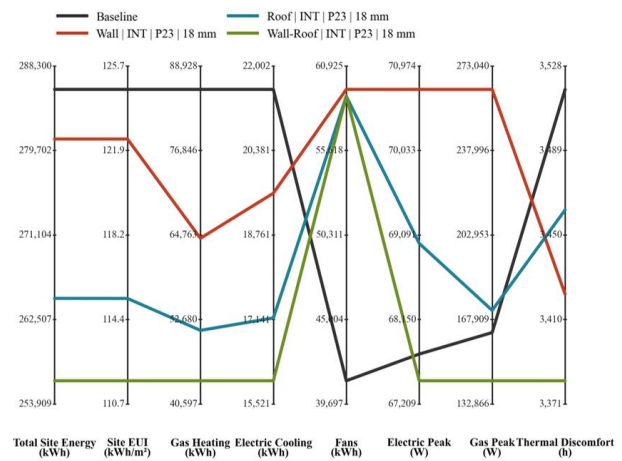


Figure 9. Annual comparison of end uses, peak demand, and thermal discomfort for the selected PCM retrofit cases and the baseline case.

Fan electricity follows a different annual pattern: consumption rises from 41,161.25 kWh in the baseline to 59,460.69, 59,030.51, and 59,053.93 kWh for the wall, roof, and wall–roof cases, respectively, corresponding to increases of 44.46%, 43.41%, and 43.47%, which suggests that the attenuation of envelope-driven loads modifies HVAC operation without a parallel reduction in air-side electricity use. Peak demand reinforces the distinction among retrofit strategies, since wall and roof applications raise electric and gas peaks, the wall case reaching the highest gas peak, 263.37 kW, whereas the wall–roof configuration remains the only option with lower values for each peak, 67.47 kW for electricity and 142.53 kW for gas, equivalent to reductions of 0.44% and 12.30%. Under this annual assessment, the simultaneous retrofit of wall and roof provides the most balanced performance among the selected cases.

At full-model scale, total site energy decreases by 10.37%, a value affected by fan electricity, internal electric loads, and other end uses beyond direct envelope control. At the HVAC end-use scale, gas heating demand and electric cooling demand decrease by 48.68% and 25.92%, respectively; these



components maintain the closest physical link with transient heat flow through walls and roof. Recent PCM-envelope studies report comparable end-use benefits under phase-change intervals aligned with seasonal demand. Dardouri et al. (2023) found demand reductions up to 41.6% for combined wall–roof PCM cases in Mediterranean climates and associated PCM21 with heat-demand reduction and PCM29 with cool-demand reduction. Chan et al. (2025) reported annual reductions of 35.24–35.47% at Riverside and 18.52–22.13% at Palm Springs. The wall–roof end-use reductions fall within that performance range, whereas the aggregate site-energy response remains shaped by the broader retail energy structure, including fan electricity, internal electric loads, and other end uses beyond direct envelope control.

4.4 Discussion

Three annual patterns define the parametric response. Retrofit effect rises with the thermal weight of the envelope path under latent control; roof application exceeds wall application, and simultaneous wall–roof retrofit gives the strongest response. Layer position establishes a second order. The response weakens as the PCM plane moves away from the room side; interior placement preserves thermal contact with the zone-side wave, while middle and exterior locations shift a larger fraction of the charge–discharge cycle toward outdoor-driven fluctuations. Transition temperature provides a third order. The response is non-monotonic, with the highest-performing range concentrated near P23–P25; lower transition temperatures entered the penalty range because their activation interval was poorly aligned with the annual indoor-side surface-temperature field.

The wall–roof retrofit affects each HVAC component according to its physical link with opaque-envelope heat transfer. Gas heating demand decreases from 85,595.19 to 43,930.15 kWh, and electric cooling demand reaches 15,967.64 kWh, as the latent layer absorbs part of the conductive heat wave when wall and roof temperatures cross the phase-change range, then releases that energy after the peak exchange with the zone. This phase delay reduces the heat that the coils must add or remove when the transition range remains close to the indoor-side surface-temperature field. Fan electricity follows a different control path and rises from 41,161.25 to 59,053.93 kWh, as air flow remains tied to system availability, outdoor-air rates, fan power curves, minimum-flow requirements, and part-load operation, instead of the conductive envelope load alone. Thermal discomfort decreases from about 3517 to 3382 h, equivalent to 135 h or 3.85%, with residual hours still constrained by thermostat setpoints, setbacks, HVAC schedules, internal gains, ventilation, and zone-level operation. Thus, the PCM retrofit reduces and shifts coil loads, whereas fan electricity and discomfort hours remain governed by the HVAC control structure preserved from the reference prototype. A source-energy or carbon metric would assign different weights to gas heating demand, electric cooling demand, and fan electricity, with those weights set by regional fuel factors and grid composition.

The selected configuration represents a Denver retail reference case for a mixed heat–cool demand climate. PCM performance

depends on the match between the phase-change interval and the local envelope temperature field; the preferred transition range can shift with climate severity, solar exposure, occupancy profile, ventilation rate, thermostat schedule, and HVAC control. Hot climates can favor higher transition temperatures as roof solar gains and cool demand gain weight, while cold climates can require latent release closer to heat-demand periods. A direct extension would apply the parametric matrix to representative ASHRAE climate zones, alternative commercial prototypes, and modified operation schedules to assess the stability of the P23–P25, 18 mm, interior-placement pattern. Transfer to a commercial PCM formulation would also require measured enthalpy curves and temperature-dependent properties, since the present material set represents an idealized thermophysical basis for controlled comparison of transition temperature, latent capacity, position, and thickness.

5. CONCLUSIONS

Annual whole-building EnergyPlus simulations were used to compare PCM retrofit in external walls, roof, and combined wall–roof deployment for a stand-alone retail building in Denver under the same building definition and operating conditions. By screening transition temperature, layer thickness, and layer position on a common basis, the analysis determined whether the preferred PCM solution remains stable across retrofit domains or shifts with the envelope subsystem.

- The annual benefit concentrated in the combined wall–roof application, where total site energy decreased by about 10%, while the best roof and wall solutions remained near 7% and 2%.
- A common design pattern emerged across the three retrofit domains: interior placement, 18 mm thickness, and transition temperatures close to P23–P25.
- The selected wall–roof case reduced gas heating demand by 48.68% and electric cooling demand by 25.92%, with a 3.85% decrease in thermal discomfort. Fan electricity increased under the preserved HVAC control structure; therefore, the main PCM benefit was associated with coil-load reduction rather than a proportional response across air-side electricity and comfort hours.

These results define a controlled Denver retail reference case, with annual total site energy used as the selection basis. The shift in PCM selection across retrofit domains supports full-model evaluation before transfer to another envelope component, climate, prototype, or control scheme. Future work should apply the same matrix across representative ASHRAE climate zones and alternative commercial prototypes to assess transferability under different operation profiles.



REFERENCES

- Alsagri, A. S. (2025). Approaching a Nearly Zero Energy Building Integrated with PCM by Optimization of Energy Sources. *Buildings*, 15(13), 2205.
- Austin, M. C., Araúz, J., & Mora, D. (2022). Numerical Assessment of Different Phase Change Materials as a Passive Strategy to Reduce Energy Consumption in Buildings under Tropical Climates. *Buildings*, 12(7), 906.
- Chan, Y., Hoke, T., Meredith, K., & Chen, X. (2025). Annual Simulation of Phase Change Materials for Enhanced Energy Efficiency and Thermal Performance of Buildings in Southern California. *Energies*, 18(4), 847.
- Cuce, P. M. (2025). Sustainable Insulation Technologies for Low-Carbon Buildings: From Past to Present. *Sustainability*, 17(11), 5176.
- Cui, W., Liu, G., Wang, Y., & Li, K. (2024). Integrated optimization of the building envelope and the HVAC system in office building retrofitting. *Case Studies in Thermal Engineering*, 62, 105185.
- Dabous, S. A., & Hosny, F. (2025). A review of building envelope retrofitting methods for improving energy efficiency, aesthetic, and indoor environmental quality. *Energy Nexus*, 18, 100407.
- Dardouri, S., Tunçbilek, E., Khaldi, O., Arıcı, M., & Sghaier, J. (2023). Optimizing PCM Integrated Wall and Roof for Energy Saving in Building under Various Climatic Conditions of Mediterranean Region. *Buildings*, 13(3), 806.
- Dora, S., Kuznik, F., & Madhavan, M. K. (2026). Phase change materials for buildings' thermal energy management: an overview. *Thermal Science and Engineering Progress*, 71, 104521.
- Feng, F., Fu, Y., Yang, Z., & O'Neill, Z. (2022). Enhancement of phase change material hysteresis model: A case study of modeling building envelope in EnergyPlus. *Energy and Buildings*, 276, 112511.
- García, M., Vera, S., Rouault, F., Gironás, J., & Bustamante, W. (2022). Cooling potential of greenery systems for a stand-alone retail building under semiarid and humid subtropical climates. *Energy and Buildings*, 259, 111897.
- Ghazwani, K., Beach, T., & Rezgui, Y. (2024). Energy retrofitting using advanced building envelope materials for sustainable housing: A review. *Building and Environment*, 267, 112243.
- Goel, S., Rosenberg, M., Mets, E., & Eley, C. (2023). ANSI/ASHRAE/IES Standard 90.1-2019 Performance Rating Method Reference Manual (PNNL-29674). Pacific Northwest National Laboratory.
- Guo, T., Sang, G., Zhang, Y., Cui, X., & Zhang, L. (2024). Study on the dynamic thermal response of phase change material integrated wall under indoor-outdoor dual thermal disturbances. *Journal of Energy Storage*, 85, 111059.
- Gutiérrez, R. A., Audenaert, A., & Verbeke, S. (2025). Passive phase change materials integration in office buildings: Review on impact of HVAC operation controls on thermal performance. *Journal of Building Engineering*, 119, 115026.
- Han, Y., Zhou, Z., Li, W., Feng, J., & Wang, C. (2023). Exploring building component thermal storage performance for optimizing indoor thermal environment – A case study in Beijing. *Energy and Buildings*, 304, 113834.
- Hou, J., Liu, Z., Zhang, L., Zhang, T., Hou, C., & Fukuda, H. (2023). Parametric and economic analysis of incorporating phase change material (PCM) into exterior walls to reduce energy demand for traditional dwellings in northeast of Sichuan hills, China. *Applied Thermal Engineering*, 223, 119982.
- IEA. (2025). Buildings. In: *Energy Efficiency 2025*. International Energy Agency, Paris.
- Jha, S. K., Sankar, A., Zhou, Y., & Ghosh, A. (2024). Incorporation of Phase Change Materials in Buildings. *Construction Materials*, 4(4), 676–703.
- Jia, J., Liu, B., Ma, L., Wang, H., Li, D., & Wang, Y. (2021). Energy saving performance optimization and regional adaptability of prefabricated buildings with PCM in different climates. *Case Studies in Thermal Engineering*, 26, 101164.
- Kitsopoulou, A., Bellos, E., & Tzivanidis, C. (2024). An Up-to-Date Review of Passive Building Envelope Technologies for Sustainable Design. *Energies*, 17(16), 4039.
- Kulumkanov, N., Memon, S. A., & Khawaja, S. A. (2024). Evaluating future building energy efficiency and environmental sustainability with PCM integration in building envelope. *Journal of Building Engineering*, 93, 109413.
- Laafer, A., Hammouma, T., Hmida, A., & Bourouis, M. (2026). Comprehensive Review of Phase Change Materials for Building Applications: Passive, Active,



- and Hybrid Systems (2022–2025). *Energies*, 19(5), 1151.
- Laasri, I. A., Es-Sakali, N., Charai, M., Mghazli, M. O., & Outzourhit, A. (2024). Recent progress, limitations, and future directions of macro-encapsulated phase change materials for building applications. *Renewable and Sustainable Energy Reviews*, 199, 114481.
- Mazidi, H. A., Wijesuriya, S., Medina, M. M., Booten, C., Dicker, B., & Kishore, R. A. (2025). Thermal performance and energy consumption validation of an occupied local government office building outfitted with ceiling tile phase change materials. *Journal of Building Engineering*, 106, 112593.
- Sangwan, P., Mehdizadeh-Rad, H., Ng, A. W. M., Tariq, M. A. U. R., & Nnachi, R. C. (2022). Performance Evaluation of Phase Change Materials to Reduce the Cooling Load of Buildings in a Tropical Climate. *Sustainability*, 14(6), 3171.
- Sun, X., Zhang, Y., Xie, K., & Medina, M. A. (2021). A parametric study on the thermal response of a building wall with a phase change material (PCM) layer for passive space cooling. *Journal of Energy Storage*, 47, 103548.
- Tabares-Velasco, P. C., Christensen, C., & Bianchi, M. (2012). Verification and validation of EnergyPlus phase change material model for opaque wall assemblies. *Building and Environment*, 54, 186-196.
- U.S. Department of Energy, Building Energy Codes Program. (2023). 90.1 Prototype Building Models 90.1-2019 - Stand-alone Retail.
- Wijesuriya, S., Kishore, R. A., Mitchell, M., & Booten, C. (2025). Enhancing EnergyPlus capabilities to model dynamic building envelopes using python plugin. *Energy and Buildings*, 339, 115776.
- Yang, T., Ding, Y., Li, B., & Athienitis, A. K. (2023). A review of climate adaptation of phase change material incorporated in building envelopes for passive energy conservation. *Building and Environment*, 244, 110711.
- Zhang, J., Rasdi, M. T. M., Zainordin, N., & Qin, Y. (2026). A comprehensive review of advanced energy-efficient technologies for building envelopes: Focus on walls, windows, and roofs. *Energy Reports*, 15, 108981.
- Zhao, S., Ming, T., Wu, Y., Cai, C., Yin, K., Fang, Y., De Richter, R., Chen, Y., & Zhou, N. (2024). Application research and effectiveness analysis of phase change materials in building envelope: A review. *Energy and Buildings*, 324, 114923.

CFD analysis of the effect of a wind catcher on cross ventilation and thermal comfort in a room of a house in the state of Morelos

E. G. Ortiz Salvador * S. L. Moya Acosta **
R. Campos Amezcua ***

* Departamento de mecánica, Tecnológico Nacional De México/CENIDET, Morelos, México,
(e-mail: m21ce002@cenidet.tecnm.mx).
** (e-mail: sara.ma@cenidet.tecnm.mx)
*** (e-mail: rafael.ca@cenidet.tecnm.mx)

Abstract: This paper presents a CFD analysis of the implantation of wind catcher (WC) as a passive ventilation strategy to improve thermal comfort in a residential room located in a warm climate in Morelos, Mexico. A previous study demonstrated that cross ventilation alone is insufficient to achieve thermal comfort conditions. Therefore, this work evaluates and compares two configurations: a base case with cross ventilation and a modified case incorporating a wind tower. Numerical simulations were performed using ANSYS Fluent under realistic boundary conditions obtained from meteorological data. The results show a significant improvement in air change per hour (ACH), increasing from 31.37 to 65.23 ACH, as well as an improvement in thermal comfort, with the predicted percentage to dissatisfied people (PPD) reduced from 13.86 % to 6.12 % and the predicted mean vote (PMV) decreased from 0.66 to 0.26. The findings indicate that the wind tower is an effective passive strategy for improving indoor environmental conditions in warm climates.

Keywords: Natural ventilation, wind tower, thermal comfort, CFD, PMV, PPD.

1. INTRODUCCIÓN

El incremento sostenido en la demanda energética del sector de la edificación, junto con el impacto ambiental derivado del uso intensivo de sistemas de climatización mecánica, ha impulsado el desarrollo de estrategias pasivas orientadas a mejorar las condiciones interiores con un menor consumo energético (Santamouris y Vasilakopoulou, 2021). En este contexto, la ventilación natural se ha consolidado como una alternativa viable para reducir la demanda energética y mejorar la calidad del aire interior. Entre sus distintas modalidades, la ventilación cruzada destaca por su simplicidad y capacidad para incrementar los cambios de aire por hora (ACH). No obstante, su desempeño depende en gran medida de factores como la geometría del espacio, la disposición de las aberturas y las condiciones del viento exterior (Zhang et al., 2022). En climas cálidos, diversos estudios han evidenciado que la ventilación cruzada, por sí sola, resulta insuficiente para garantizar condiciones adecuadas de confort térmico (Nasrollahi y Ghobadi, 2022). Ante estas limitaciones, las torres de viento (TV) han sido ampliamente estudiadas como una estrategia pasiva capaz de mejorar la ventilación y reducir la temperatura interior. Investigaciones recientes han analizado su comportamiento aerodinámico mediante enfoques numéricos y experimentales, reportando mejoras en la distribución del flujo de aire, la ventilación interior y la reducción de zonas de estancamiento (Sadeghi et al., 2019; Nejat et al., 2021). Asimismo, se han propuesto di-

seños optimizados mediante modificaciones geométricas y la incorporación de elementos adicionales para incrementar la eficiencia del sistema (Sotelo et al., 2023). La investigación ha evolucionado hacia el desarrollo de sistemas adaptativos, incluyendo configuraciones multicanal con mecanismos capaces de mantener un desempeño eficiente bajo distintas direcciones del viento, así como su integración con otras estrategias pasivas de bajo consumo energético (Li et al., 2024; Li et al., 2025). Investigaciones recientes han analizado su comportamiento aerodinámico mediante enfoques numéricos y experimentales, reportando mejoras en la distribución del flujo de aire, la ventilación interior y la reducción de zonas de estancamiento (Sadeghi et al., 2019; Nejat et al., 2021). Asimismo, se han propuesto diseños optimizados mediante modificaciones geométricas y la incorporación de elementos adicionales para incrementar la eficiencia del sistema (Sotelo et al., 2023). La investigación ha evolucionado hacia el desarrollo de sistemas adaptativos, incluyendo configuraciones multicanal con mecanismos capaces de mantener un desempeño eficiente bajo distintas direcciones del viento, así como su integración con otras estrategias pasivas de bajo consumo energético (Li et al., 2024; Li et al., 2025). Diversos estudios han evaluado la aplicación de TV en distintos contextos climáticos, confirmando su potencial como estrategia de enfriamiento pasivo y su contribución a la reducción del consumo energético en edificaciones (Carreto et al., 2025, 2026). Asimismo, se han explorado configuraciones que incorporan técnicas complementarias, como la humidificación, para

mejorar el desempeño térmico en condiciones específicas (Carreto et al., 2022). Particularmente en el Estado de Morelos, estudios previos han evidenciado que, si bien la ventilación cruzada permite alcanzar niveles adecuados de ACH, las condiciones térmicas permanecen fuera de los rangos de confort establecidos por normas internacionales (Ortiz et al., 2022). Estos resultados ponen de manifiesto la necesidad de incorporar estrategias pasivas adicionales que permitan mejorar simultáneamente la ventilación y el confort térmico en este tipo de edificaciones. A pesar de los avances en el estudio de las TV, aún es limitada la investigación que evalúa su integración con sistemas de ventilación cruzada en viviendas ubicadas en climas cálidos de México, bajo condiciones representativas de operación. En este contexto, el presente trabajo tiene como objetivo evaluar numéricamente el efecto de la implementación de una TV en el desempeño de la ventilación cruzada y el confort térmico en una habitación (ver Fig. 1) de una casa en el estado de Morelos. Para ello, se realiza un análisis comparativo entre un caso base con ventilación cruzada y un caso modificado con TV, utilizando herramientas de dinámica de fluidos computacional (CFD) ANSYS Fluent (ANSYS U., 2021).

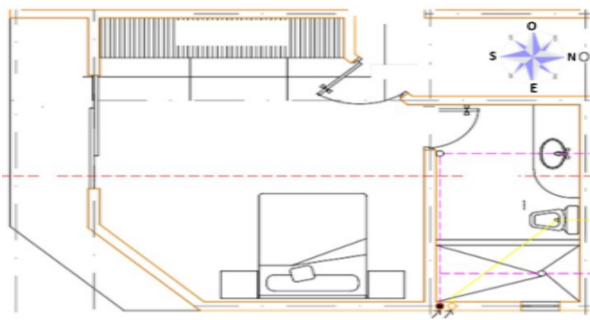


Fig. 1. Plano de la habitación (Ortiz Salvador et al. 2022)

1.1 Alcances y limitaciones

Alcances

- Evaluar numéricamente el efecto de una TV en la ventilación cruzada y el confort térmico de una habitación.
- Analizar indicadores de desempeño como ACH, PMV y PPD bajo condiciones climáticas de verano.
- Determinar la viabilidad de la TV como estrategia pasiva para mejorar la ventilación y el confort térmico.

Limitaciones

- El estudio se limita a una sola habitación y a condiciones climáticas de verano.
- Las simulaciones se realizaron en régimen estacionario y bajo condiciones simplificadas.
- No se consideran cargas internas de calor ni variaciones dinámicas de ocupación.
- Los resultados provienen únicamente de simulaciones CFD, sin validación experimental directa.

- El desempeño analizado depende de una dirección predominante específica del viento.

2. METODOLOGÍA

Los datos meteorológicos provienen de la estación ESOLM ET-IER-UNAM, ubicada en Temixco, Morelos ($18^{\circ}50'25.62''$ N, $99^{\circ}14'10.49''$ O, 1253 m s. n. m.), la cual es la más cercana al área de estudio y proporcionó los registros en bruto del periodo 2018–2020. Se consideraron variables de temperatura del aire, humedad relativa, velocidad y dirección del viento, analizadas a escala estacional, mensual y horaria para identificar condiciones críticas. Se determinó el verano y el mes de julio como representativos del caso de estudio. Las velocidades de viento fueron ajustadas mediante la ley de potencia del perfil vertical para estimar condiciones a la altura de las aberturas, en consecuencia determinar las condiciones de frontera para las simulaciones CFD y comparativa de resultados sin/con TV.

2.1 Periodos de analisis de datos meteorológicos

Se analizaron registros correspondientes al periodo 2018–2020, considerando variables como temperatura, humedad relativa, velocidad y dirección del viento. El análisis inicial por estaciones del año permitió identificar el verano como el periodo crítico, ya que concentra las condiciones térmicas más severas (ver Fig. 2). Posteriormente, se realizó un análisis a escala mensual y diaria para determinar condiciones representativas de operación. A partir de este análisis, se identificó julio como el mes crítico, ya que presenta la mayor amplitud térmica, con valores entre 17.57°C y 29.30°C , así como velocidades de viento elevadas que alcanzan hasta 2.1 m/s (ver Fig. 3). En comparación, junio muestra condiciones ligeramente más estables, mientras que agosto presenta temperaturas máximas moderadas (26.14°C), aunque con velocidades de viento similares (hasta 2.15 m/s) (ver Tabla 1).

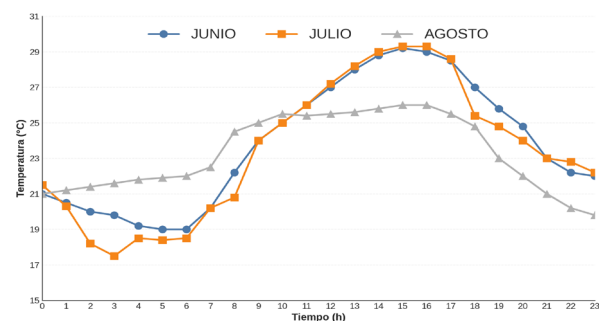


Fig. 2. Comportamiento de la temperatura ambiente en la ESOLMET-IER-UNAM de Verano 2018-2020

El análisis horario evidencia una relación directa entre la temperatura y la velocidad del viento, donde ambas variables incrementan durante las horas de mayor radiación solar (aproximadamente entre las 12:00 y 16:00 hrs). Este comportamiento sugiere la influencia de fenómenos convectivos que favorecen el aumento del flujo de aire en los periodos más cálidos del día (ver las Figs. 2 y 3).

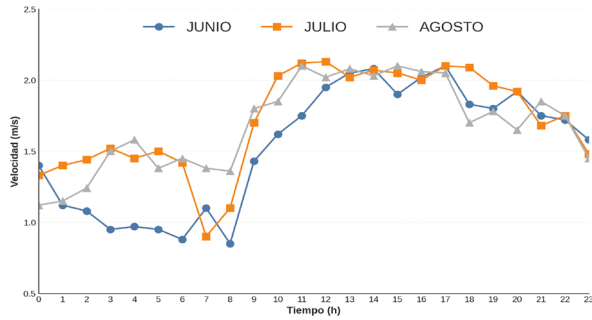


Fig. 3. Comportamiento de la velocidad de viento en la ESOLMET-IER-UNAM de Verano 2018-2020

Por otra parte, el análisis de la dirección del viento muestra una clara predominancia de flujos provenientes del sur, con mayor frecuencia en rangos de velocidad entre 1 y 2 m/s (ver Fig. 4). En particular, se identificó una velocidad máxima representativa de aproximadamente 2.1 m/s en el mes de julio, durante la estación de verano, la cual se considera una condición crítica (ver Tabla 1). Este valor fue adoptado como condición de entrada en la abertura correspondiente a la TV (ver Fig. 6), para la definición de las condiciones de frontera en la simulación CFD.

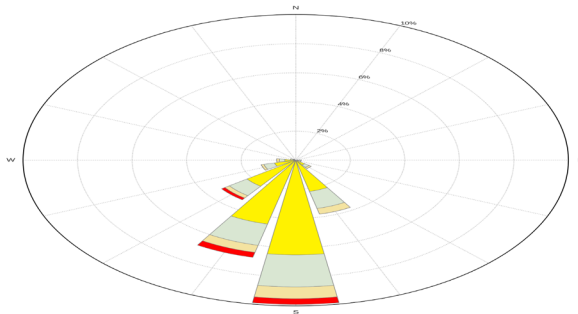


Fig. 4. Dirección predominante promedio del viento en la ESOLMET-IER-UNAM de Verano 2018-2020

Adicionalmente, los valores de humedad relativa, que alcanzan hasta 86.31 % en junio y 84.47 % en agosto (ver Tabla 1), indican condiciones ambientales con alta carga de humedad, lo que puede influir en la percepción de confort térmico y en la eficiencia de las estrategias de ventilación natural.

Tabla 1. Promedios de datos meteorológicos en verano (2018-2020)

Mes	Variable	Mín	Máx
Junio	Temperatura (°C)	19.18	29.05
	Humedad (%)	42.22	86.31
	Velocidad (m/s)	0.85	2.14
Julio	Temperatura (°C)	17.57	29.30
	Humedad (%)	35.73	73.51
	Velocidad (m/s)	0.89	2.10
Agosto	Temperatura (°C)	20.02	26.14
	Humedad (%)	58.13	84.47
	Velocidad (m/s)	1.12	2.15

Para estimar la velocidad del viento en diferentes aberturas, se empleó la ley de potencia empírica, considerando un exponente de rugosidad $\alpha=0.13$, valor típico para terrenos ligeramente accidentados. En el ámbito de la ingeniería y la mecánica de fluidos, esta relación se conoce como perfil vertical de velocidad del viento mediante ley de potencia (power law wind profile) (Justus y Mikhail, 1976), y describe la variación de la velocidad del viento con la altura:

$$v_2 = v_1 \left(\frac{h_2}{h_1} \right)^\alpha \quad (1)$$

Donde n es el coeficiente de rugosidad del terreno, h_1 la altura de referencia de la estación meteorológica, h_2 la altura media de la abertura, v_1 la velocidad del viento medida en la estación y v_2 la velocidad estimada a la altura de interés. Con base en los datos analizados, se obtuvo una velocidad de viento de entrada en la ventana de la habitación ($Inlet_1$) de 1.89 m/s.

3. MODELO FÍSICO Y MATEMÁTICO

3.1 Modelo físico

El modelo físico corresponde a una habitación representativa de una casa ubicada en el estado de Morelos, México, la cual presenta características constructivas típicas de climas cálidos, tales como materiales convencionales, geometría simple y un esquema de ventilación natural basado en aberturas. Este tipo de configuración es común en la región y resulta adecuada para evaluar estrategias pasivas de enfriamiento. El dominio computacional de la simulación CFD incluye el volumen interior del recinto (aire), así como las aberturas destinadas a la ventilación natural, a través de las cuales se permite la entrada ($Inlet_1$) y salida del flujo de aire ($Outlet_1$), bajo condiciones de frontera representativas del entorno exterior (ver Fig. 5).

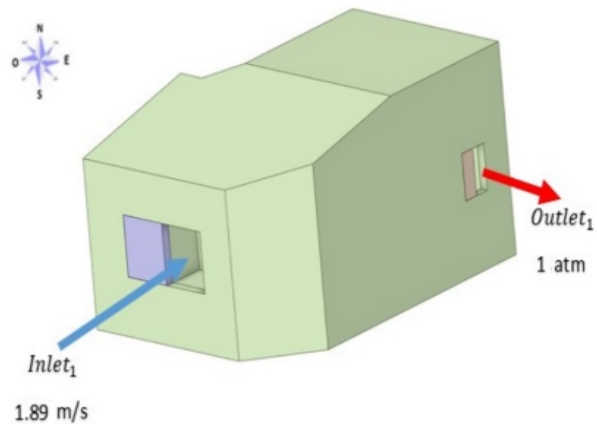


Fig. 5. Modelo físico desarrollado sin la implementación de una TV

Con el objetivo de analizar el efecto de la incorporación de dispositivos pasivos en el desempeño térmico y de ventilación, se evaluaron dos configuraciones:

- Caso base: Ventilación cruzada mediante dos aberturas opuestas ($Inlet_1$ y $Outlet_1$) (ver Fig. 5).

- Caso con TV: Ventilación cruzada asistida mediante una TV que capta el aire exterior y lo dirige al interior del recinto. (ver Fig. 6).

El modelo CAD del caso base reproduce las dimensiones reales de la habitación representativa (ver Fig. 1) y constituye la base de la simulación numérica CFD. Sobre esta geometría se incorporaron las condiciones de frontera necesarias para representar el comportamiento del flujo de aire dentro del dominio, definiendo una abertura de entrada ($Inlet_1$) en la pared frontal izquierda, por donde el fluido ingresa con una velocidad de 1.89 m/s, de acuerdo con el análisis de los datos obtenidos de la estación ESOLMET-IER-UNAM, y una abertura de salida ($Outlet_1$) en la pared lateral derecha, donde se estableció una presión atmosférica de 1 atm para permitir la descarga del flujo de aire (ver Fig. 5). De esta manera, la configuración del modelo permite analizar la distribución y el comportamiento del flujo en el interior del recinto bajo condiciones iniciales controladas. El diseño de la TV fue seleccionado a partir de las condiciones climáticas locales y de la evidencia reportada en la literatura. Su desempeño depende principalmente del número de ventanas, su orientación respecto a los vientos predominantes y su integración con el espacio interior. En presencia de una dirección predominante del viento (ver Fig. 4), se recomienda el uso de una TV de dos ventanas para mejorar la captación y distribución del flujo de aire (ver Fig. 6) (Reyes et al., 2013).

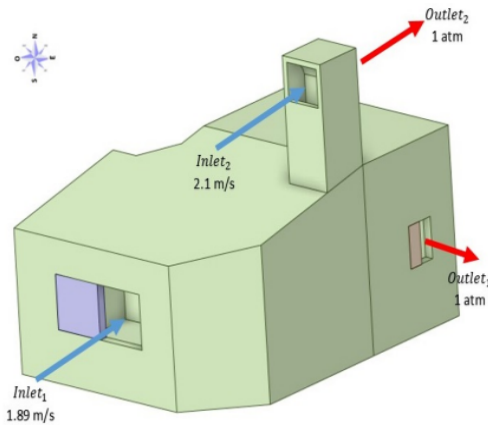


Fig. 6. Representación del modelo físico equipado con una TV

La orientación del sistema se definió con base en el análisis de los vientos predominantes registrados en la estación ESOLMET-IER-UNAM, donde se identificó una dirección principal proveniente del sur (ver Fig. 4), por lo que la TV se orientó hacia dicha dirección. Asimismo, la altura de la TV se determinó considerando la ubicación de la estación meteorológica y el perfil de velocidad del viento, estableciéndose una elevación de 2.3 m sobre el nivel del techo, representativa del viento incidente. Finalmente, se adoptó un diseño con dos ventanas de dimensiones iguales (0.84 m × 0.6 m) y una curva de inducción de 39° (ver Fig. 6), configuración que favorece la alineación del flujo y mejora la ventilación natural en el interior del recinto (Carreto et al., 2022).

3.2 Modelo matemático

El software ANSYS Fluent emplea el método de volúmenes finitos para resolver las ecuaciones gobernantes del flujo de fluidos. Estas incluyen la ecuación de conservación de la masa (Ec. 2), las ecuaciones de cantidad de movimiento (Ecs. 3–5) y la ecuación de energía (Ec. 6), formuladas bajo el enfoque RANS (Reynolds-Averaged Navier–Stokes) (ANSYS U., 2021). Dichas ecuaciones se aplican a un flujo turbulento tridimensional en régimen estacionario, bajo la aproximación de Boussinesq. En este enfoque, los promedios se realizan en el tiempo y cada variable instantánea del campo turbulento se descompone en una componente media y otra fluctuante, expresada como: $u = \bar{u} + u'$

$$\nabla \cdot (\rho \bar{\vec{V}}) = 0 \quad (2)$$

$$\nabla \cdot (\rho \bar{u}_i \bar{\vec{V}}) = -\frac{\partial \bar{P}}{\partial x} + \nabla \cdot [(\mu \nabla \bar{u}_i) (\rho \bar{u}'_i \bar{u}'_j)] \quad (3)$$

$$\nabla \cdot (\rho \bar{u}_j \bar{\vec{V}}) = -\frac{\partial \bar{P}}{\partial y} + \nabla \cdot [(\mu \nabla \bar{u}_j) + (\rho \bar{u}'_i \bar{u}'_j)] - \rho g \beta (\bar{T} - \bar{T}_o) \quad (4)$$

$$\nabla \cdot (\rho \bar{u}_k \bar{\vec{V}}) = -\frac{\partial \bar{P}}{\partial z} + \nabla \cdot [(\mu \nabla \bar{u}_k) + \rho \bar{u}'_i \bar{u}'_j] \quad (5)$$

$$\nabla \cdot (\rho C_p \bar{T} \bar{\vec{V}}) = \nabla \cdot (\lambda \nabla \bar{T}) + \rho \bar{u}'_i \bar{T}' \quad (6)$$

$\bar{\vec{V}}$ representa el vector de velocidad, cuyas componentes corresponden a: u_i, u_j, u_k . Asimismo, ρ indica la densidad del fluido y la presión. Por otra parte $g, \mu, \beta, C_p, \lambda, \nabla \bar{T}$ y T_o representan, respectivamente, la aceleración gravitacional, la viscosidad molecular, el coeficiente de expansión térmica volumétrica de aire, el calor específico, la conductividad térmica, el gradiente de temperatura y la temperatura de referencia. El tensor de esfuerzos turbulentos de Reynolds, expresado como $\rho \bar{u}'_i \bar{u}'_j$, que aparece en las ecuaciones (3)–(5), origina un problema de cierre en el modelo matemático. Este tensor puede escribirse como:

$$\rho \bar{u}'_i \bar{u}'_j = \mu_t \frac{\partial \bar{u}_i}{\partial x_j} + \frac{\partial \bar{u}_j}{\partial x_i} \quad (7)$$

En esta relación, la viscosidad turbulenta vincula los esfuerzos turbulentos de Reynolds con los gradientes de la velocidad media. A diferencia de la viscosidad molecular μ , la viscosidad turbulenta μ_t no constituye una propiedad intrínseca del fluido, ya que depende considerablemente del estado local de la turbulencia y puede variar de un punto a otro dentro del flujo. En este estudio se empleó el modelo $k - \varepsilon$ realizable, ampliamente utilizado en investigaciones relacionadas con ventilación natural. Para este modelo, la viscosidad turbulenta se calcula mediante:

$$\mu_t = \rho C_\mu \frac{k^2}{\varepsilon} \quad (8)$$

donde la energía cinética turbulenta (k) y la tasa de disipación de dicha energía (ε) se obtiene resolviendo las ecuaciones de transporte que se presentan a continuación:

$$\frac{\partial}{\partial x_j} (\rho k \bar{u}_j) = \frac{\partial}{\partial x_j} \left[\left(\mu + \frac{\mu_t}{\sigma_t} \right) \frac{\partial k}{\partial x_j} \right] + G_k + G_b - \rho \varepsilon - Y_M + S_K \quad (9)$$

$$\frac{\partial}{\partial x_j}(\rho \varepsilon \bar{u}_j) = \frac{\partial}{\partial x_j} \left[\left(\mu + \frac{\mu_t}{\sigma_t} \right) \frac{\partial \varepsilon}{\partial x_j} \right] + \rho C_1 S_\varepsilon - \rho C_2 \frac{\varepsilon^2}{k + \sqrt{\nu \varepsilon}} + C_{1\varepsilon} \frac{\varepsilon}{K} C_{3\varepsilon} G_b + S_\varepsilon \quad (10)$$

Las variables relacionadas de las Ecs. (9) y (10) están definidas en la guía de usuario de ANSYS Fluent (Fluent U., 2021).

4. RESULTADOS

4.1 Establecimiento de condiciones de frontera

Las condiciones de frontera dinámicas se definieron a partir del análisis de los datos obtenidos de la estación ESOLMET-IER-UNAM (ver Tabla 3). En cuanto a la transferencia de calor, se asumió que el suelo y los muros interiores son adiabáticos, lo que implica que no hay intercambio de calor entre el sistema y el entorno. Los materiales considerados para los muros y la cubierta corresponden a ladrillo rojo y concreto, con conductividades térmicas de 0.78 y 1.74 W/(m·K), respectivamente (ver Tabla 2). Asimismo, se adoptaron las condiciones de muro masivo establecidas en la norma NOM-020-ENER-2011 (Estados Unidos Mexicanos – Secretaría de Energía, 2011, p. 12), la cual regula la eficiencia energética en edificaciones de uso habitacional mediante criterios de diseño térmico, con el objetivo de garantizar el confort de los ocupantes con el menor consumo de energía posible.

Tabla 2. Propiedades térmicas de los materiales

Muro	Mat.	ρ (kg/m ³)	c_p (J/kg·K)	k (W/m·K)
Sur y Este	Ladrillo	1600	840	0.78
Techo	Concreto	2300	1000	1.74

Tabla 3. Condiciones del sistema: volumen de la habitación, temperatura, velocidad del flujo y presión atmosférica.

Habitación	
Habitación	64.15 m ³
Habitación con TV	65.35 m ³
Muros	
Orientación	Temperatura (°C)
Sur	28
Este	27
Techo	39
Ventanas	
Inlet ₁	1.89 m/s
Inlet ₂	2.10 m/s
Outlet ₁ , Outlet ₂	1 atm

4.2 Resultados sin TV

Para la simulación numérica se asumieron condiciones de ventilación natural con ventanas abiertas, considerando una entrada de aire en la ventana sur (Inlet₁) y una salida a presión (Outlet₁), con velocidades de 1.89 m/s y presión de 1 atm, respectivamente. Los resultados muestran altas velocidades en la entrada; sin embargo, el flujo se disipa rápidamente dentro de la habitación, generando amplias zonas de baja velocidad y evidenciando una distribución de aire no homogénea (ver Fig. 7). Asimismo, se identifican

regiones de recirculación en paredes y esquinas que limitan la renovación efectiva del aire, por lo que un aumento en la velocidad de entrada no garantiza necesariamente una mejora en la ventilación. En términos de desempeño, la habitación alcanza 31.37 ACH, valor inferior al rango recomendado de 60–70 ACH (ASHRAE S., 2020), lo cual concuerda con los patrones de flujo observados y la presencia de zonas de aire estancado.

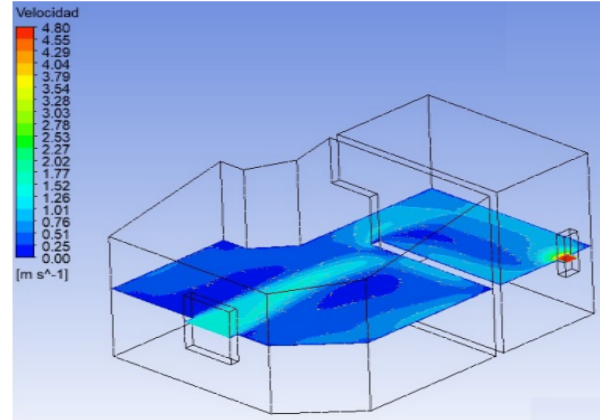


Fig. 7. Campos de velocidad dentro de la habitación

La simulación numérica permitió evaluar la distribución del confort térmico mediante los índices de voto medio predicho (PMV) y porcentaje de personas insatisfechas (PPD), siguiendo los estándares de la norma establecida ($-0.5 \leq \text{PMV} \leq 0.5$ y $\text{PPD} < 10\%$) (ASHRAE S., 2020). Los contornos de PMV muestran una distribución no uniforme, con valores predominantemente positivos (mínimo de 0.59), lo que indica una sensación térmica cálida en todo el recinto. Las zonas con mayores valores de PMV se concentran en el centro y en regiones con menor ventilación, mientras que los valores más bajos se ubican cerca de las aperturas; no obstante, en ningún caso se alcanzan condiciones de confort térmico (ver Fig. 8).

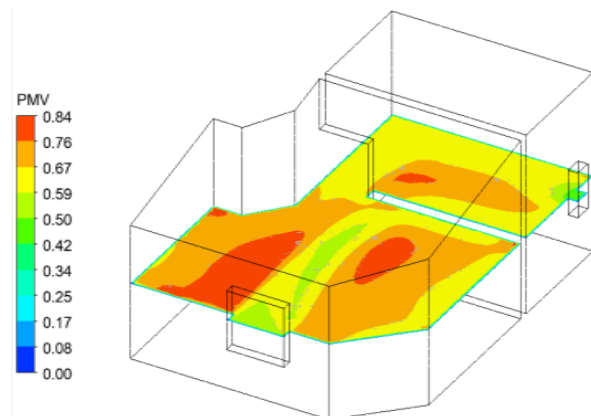


Fig. 8. Campos de PMV

De manera consistente, los contornos de PPD presentan un valor mínimo de 11.96 %, superando el límite recomendado de <10 %, con mayores niveles de insatisfacción en las mismas zonas de alto PMV. Esto confirma que el disconfort térmico es generalizado en la habitación (ver Fig. 9).

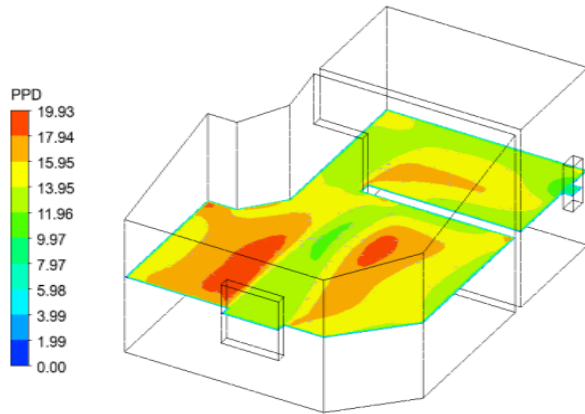


Fig. 9. Campos de PPD

El cálculo de estos índices térmicos se realizó considerando un conjunto específico de condiciones ambientales y fisiológicas. En primer lugar, se utilizó una velocidad del aire de 0.056 m/s, se tomó una temperatura del aire de 31.2 °C y una humedad relativa de 28.17%. En cuanto a las variables fisiológicas, se asumió un factor de ropaje de 0.5 clo, correspondiente a ropa ligera, lo que implica una resistencia térmica baja pero no despreciable. Adicionalmente, se adoptó una tasa metabólica de 1.2 met, asociada a una actividad sedentaria o de reposo ligero, es decir, una producción interna de calor relativamente baja.

4.3 Resultados con TV

La simulación CFD con TV muestra un comportamiento del flujo de aire claramente influenciado por los gradientes de presión y la geometría de los canales de entrada y salida del sistema (ver Fig. 10). En las regiones correspondientes a *Inlet₂*, *Outlet₂* y *Outlet₁* se observa un incremento de la velocidad del aire, lo cual se debe a la aceleración del flujo al atravesar las secciones de captación y extracción asociadas al funcionamiento de la TV. Este aumento de velocidad es consecuencia directa de la diferencia de presiones entre el exterior y el interior del sistema, lo que genera un flujo dirigido y continuo hacia el espacio analizado. Al ingresar el aire a la habitación (*Inlet₁* y *Inlet₂*), el flujo presenta un proceso de expansión y disipación progresiva, reduciendo su velocidad conforme aumenta la distancia respecto a las aberturas. En el volumen interior, las velocidades se mantienen en un intervalo comprendido entre 0.25 m/s y 1.76 m/s. Este comportamiento indica que el aire se distribuye de manera relativamente uniforme, sin concentraciones excesivas de velocidad en zonas específicas. La transición del flujo desde regiones de alta velocidad hacia zonas de menor velocidad permite una renovación de aire gradual y estable dentro de la habitación. La ausencia de corrientes intensas dentro del volumen ocupado sugiere que el sistema de ventilación basado en TV mantiene condiciones adecuadas para la ocupación humana, minimizando el riesgo de discomfort asociado a velocidades elevadas o turbulencias localizadas. En términos de desempeño global, el sistema alcanza un valor de 65.23 ACH, este valor se encuentra dentro del rango recomendado de 60-70 ACH (ASHRAE, 2020).

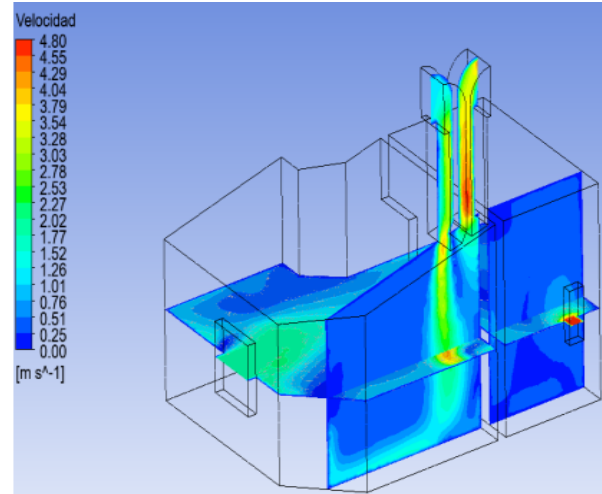


Fig. 10. Campos de velocidad con TV

La simulación numérica permitió obtener los contornos de PMV y PPD en la habitación (ver Figs. 11 y 12). El valor máximo de PMV es de 0.45, dentro del rango de confort normativo, lo que indica una sensación térmica cercana a la neutralidad. De manera consistente, el PPD máximo es de 8%, cumpliendo con el límite aceptable y reflejando baja insatisfacción térmica.

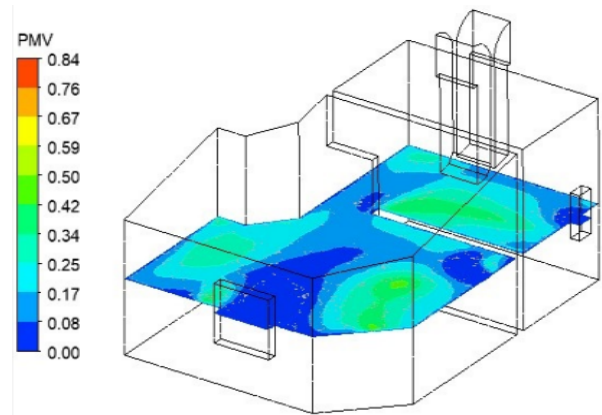


Fig. 11. Campos de PMV con TV

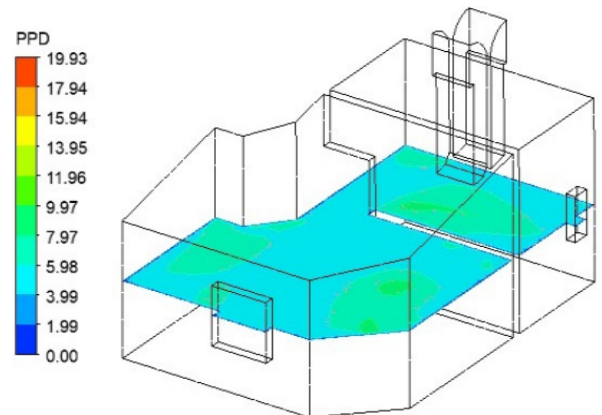


Fig. 12. Campos de PPD con TV

5. CONCLUSIONES

Los resultados sin la implementación de la TV muestran una ventilación de 31.37 ACH y valores de confort térmico de 0.66 para el PMV y 13.86 % para el PPD, lo que indica condiciones inadecuadas tanto de ventilación como de confort (ver Figs. 7, 8 y 9). En contraste, al incorporar la TV, la ventilación aumenta a 65.23 ACH, ubicándose dentro del rango recomendado de 60–70 ACH. De igual forma, los índices de confort mejoran, alcanzando un PMV de 0.26 y un PPD de 6.12 %, cumpliendo con los límites normativos establecidos (ver Tabla 4) (ASHRAE S., 2020). Estos resultados evidencian que la implementación de una TV permite lograr condiciones adecuadas de ventilación y confort térmico en la habitación.

Tabla 4. Comparación de ACH, PMV y PPD con y sin TV.

Variable	Sin TV	Con TV	Confort térmico y ventilación
ACH	31.37	65.23	60–70
PMV	0.66	0.26	-0.5 a 0.5
PPD (%)	13.86	6.12	<10

REFERENCIAS

- ASHRAE. (2020). *ASHRAE Standard 55: Thermal environmental conditions for human occupancy*. Atlanta: ASHRAE.
- ANSYS Fluent. (2021). *ANSYS Fluent 12.0/12.1 Documentation*. <https://www.afs.enea.it/project/neptunius/docs/fluent/index.html>.
- Carreto Hernandez, L.G., Báez García, W.G., García Merino, J.A., Cárdenas Hernández, A.J., Gómez Parra, A.Y. and Ramírez Sánchez, P. (2026). 'Study on thermal comfort and ventilation performance of a wind catcher in hot-dry climate of Mexico: A sustainable energy-saving strategy'. *Building Services Engineering Research and Technology*, pp. 1–20. <https://doi.org/10.1177/01436244261426152>.
- Carreto Hernandez, L.G., Hernandez, A.F., García Merino, J.A., Báez García, W.G., Juárez Sosa, I. and Reyes Portillo, I.A. (2025). 'CFD evaluation of wind catcher performance in residential ventilation in New York: towards a sustainable architecture'. *Advances in Building Energy Research*. <https://doi.org/10.1080/17512549.2025.2607572>.
- Carreto Hernandez, L.G., Moya, S.L., Varela-Boydo, C.A. and Francisco-Hernández, A. (2022). 'Studies of ventilation and thermal comfort in different wind tower-room configurations considering humidification for a warm climate of Mexico'. *Journal of Building Engineering*, 46. <https://doi.org/10.1016/j.jobe.2021.103675>.
- Secretaría de Energía. (2011). *NOM-020-ENER-2011: Eficiencia energética en edificaciones – Envoltante de edificios para uso habitacional*. <https://www.gob.mx/cms/uploads/attachment/data/file/181660/NOM020ENER2011.pdf>
- Justus, C.G. and Mikhail, A. (1976). 'Height variation of wind speed and wind distributions statistics'. *Geophysical Research Letters*, 3(5), pp. 261–264. <https://doi.org/10.1029/GL003i005p00261>
- Li, J., Calautit, J. and Jimenez-Bescos, C. (2024). 'Parametric analysis of a novel rotary scoop dual-channel windcatcher for multi-directional natural ventilation of buildings'. *Energy and Built Environment*. <https://doi.org/10.1016/j.enbenv.2024.07.002>
- Li, J., Calautit, J., Jimenez-Bescos, C., Song, W., Riffat, S. and Chen, Q. (2025). 'Climate-adaptive windcatcher natural ventilation integrated with passive and low-energy technologies: A review of current and future developments'. *Building and Environment*, 284, p. 113436. <https://doi.org/10.1016/j.buildenv.2025.113436>
- Nejat, P., Ferwati, M.S., Calautit, J., Ghahramani, A. and Sheikhshahrokhdehkhordi, M. (2021). 'Passive cooling and natural ventilation by the windcatcher (Badgir): An experimental and simulation study of indoor air quality, thermal comfort and passive cooling power'. *Journal of Building Engineering*, 41, 102436. <https://doi.org/10.1016/j.jobe.2021.102436>
- Nasrollahi, N. and Ghobadi, P. (2022). 'Field measurement and numerical investigation of natural cross-ventilation in high-rise buildings: Thermal comfort analysis'. *Applied Thermal Engineering*, 211, 118500. <https://doi.org/10.1016/j.applthermaleng.2022.118500>
- Ortiz Salvador, E.G., Ricoy Zárate, E.G., Carreto Hernández, L.G., Sotelo Rodríguez, P., Moya Acosta, S.L. and Campos Amezcua, R. (2022). 'Estudio numérico y experimental de ventilación cruzada en una habitación de una casa del Estado de Morelos'. *JCYTA, TecNM/CENIDET, Cuernavaca, Morelos, México*.
- Reyes, V., Moya, S., Morales, J. and Sierra-Espinosa, F. (2013). 'A study of air flow and heat transfer in building-wind tower passive cooling systems applied to arid and semi-arid regions of Mexico'. *Energy and Buildings*, 66, pp. 211–221. <https://doi.org/10.1016/j.enbuild.2013.07.032>
- Sadeghi, M., Samali, B., Wood, G. and De Dear, R. (2019). 'Comfort cooling by wind towers in the Australian residential context: Experimental wind tunnel study of comfort'. *Journal of Wind Engineering and Industrial Aerodynamics*, 196, p. 104014. <https://doi.org/10.1016/j.jweia.2019.104014>
- Santamouris, M. and Vasilakopoulou, K. (2021). 'Present and future energy consumption of buildings: Challenges and opportunities towards decarbonisation'. *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, 1, 100002. <https://doi.org/10.1016/j.prime.2021.100002>
- Sotelo Rodríguez, P., Carreto Hernández, L.G., Ortiz Salvador, E.G., Campos Amezcua, R. and Moya Acosta, S.L. (2023). 'Numerical study of the air capture effectiveness of a wind tower without/with inlet extensions'. *I CITCA (2023)*, TecNM/CENIDET, Cuernavaca, México.
- Zhang, X., Weerasuriya, A.U., Wang, J., Li, C.Y., Chen, Z., Tse, K. and Hang, J. (2022). Cross-ventilation of a generic building with various configurations of external and internal openings. *Building and Environment*, 207, 108447. <https://doi.org/10.1016/j.buildenv.2021.108447>

From Static Navigation to Dynamic Risk Assessment: Analysis of Assistive Technologies for Visual Impairment

Bustos Ramírez, Maximiliano * Blanco Ortega, Andrés*

**Tecnológico Nacional de México/Centro Nacional de Investigación y Desarrollo Tecnológico, Interior Internado Palmira s/n, C.P. 62490, Morelos, México (e-mail: m25ce094@cenidet.tecnm.mx)*

Abstract: Autonomous urban mobility poses a critical challenge for the visually impaired, especially during street crossings where assessing vehicular risk is vital. This paper presents a comparative review of state-of-the-art assistive technologies, classifying current solutions into smart infrastructure (IoT), mobile applications, instrumented canes, and guide robots. The technical limitations of each approach are analyzed, revealing that mobile apps lack the required robustness in adverse lighting conditions, whereas most contemporary smart canes are restricted to detecting static obstacles rather than evaluating dynamic traffic. Through an evaluation of latency, usability, and accuracy, a significant gap is identified: the absence of portable devices capable of real-time vehicle distance estimation without reliance on external infrastructure. The study concludes that integrating embedded computing with lightweight detection algorithms (such as YOLOv8) and haptic interfaces represents the most promising path to ensure user safety in complex environments.

Keywords: Assistive Technologies, Visual Impairment, Object Detection, Edge Computing, YOLOv8.

1. INTRODUCTION

La movilidad autónoma en entornos urbanos constituye uno de los desafíos más críticos para las personas con discapacidad visual (PDV). Según estimaciones globales, cientos de millones de personas viven con deficiencias visuales que limitan su capacidad de desplazamiento seguro (Flaxman et al., 2017). Si bien el bastón blanco ha sido históricamente la herramienta estándar para la navegación, su funcionalidad se limita estrictamente a la detección de obstáculos estáticos inmediatos y a la identificación de superficies táctiles (Mai et al., 2024), dejando al usuario vulnerable ante peligros dinámicos fuera de su alcance físico.

El escenario de mayor riesgo para este colectivo es el cruce peatonal no semaforizado o complejo. En este contexto, la seguridad del usuario depende de dos factores críticos: la capacidad de percibir brechas seguras en el flujo vehicular y la interacción con los conductores. Estudios de comportamiento vial han evidenciado que la “cultura de protección” es insuficiente; investigaciones observacionales indican que los conductores ceden el paso a peatones con discapacidad visual en tan solo un 37% de las interacciones (Geruschat & Hassan, 2005). Esta imprevisibilidad del comportamiento humano, sumada a la incapacidad de las PDV para estimar la velocidad y distancia de los vehículos mediante el oído – especialmente en presencia de vehículos eléctricos silenciosos–, genera un entorno de alta incertidumbre y riesgo vital.

Actualmente, existen diversas soluciones tecnológicas orientadas a mitigar estos riesgos, desde aplicaciones móviles basadas en GPS hasta sistemas de infraestructura conectada (V2P o Vehicle-to-Pedestrian). Sin embargo, la literatura reciente sugiere que estas aproximaciones presentan limitaciones fundamentales para la seguridad en tiempo real.

Las soluciones basadas en smartphones enfrentan problemas de latencia, dependencia de la conectividad y una baja eficiencia en condiciones de iluminación adversa (Budrionis et al., 2022; Khan & Khusro, 2021), mientras que los sistemas de infraestructura inteligente requieren inversiones urbanas masivas que limitan su escalabilidad.

Por lo tanto, existen una necesidad técnica no resuelta: el desarrollo de dispositivos de asistencia portátiles (wearables) con capacidad de cómputo, capaces de realizar una evaluación de riesgo dinámico sin depender de infraestructura externa.

Este artículo presenta una revisión comparativa de las tecnologías de asistencia actuales y propone una arquitectura de sistemas basada en visión artificial estéreo y algoritmos de detección ligera (YOLOv8). El objetivo es demostrar que la integración de hardware embebido de alto rendimiento con interfaces de retroalimentación háptica ofrece una solución superior para la estimación de riesgo vehicular en tiempo real, superando las limitaciones de latencia y usabilidad de las aplicaciones móviles convencionales.

2. TRABAJOS RELACIONADOS

Con el objetivo de establecer un marco comparativo riguroso y formalizar el análisis del estado del arte, esta revisión se delimita a la evaluación de tecnologías orientadas a la movilidad peatonal bajo condiciones dinámicas. Se priorizó el análisis de investigaciones recientes que abordan específicamente la interacción del usuario con el tráfico vehicular, diferenciándolas de aquellas limitadas a la evasión estática de obstáculos. Para contrastar la viabilidad, eficacia y seguridad de cada enfoque tecnológico, se establecieron cuatro criterios metodológicos fundamentales:

- **Autonomía Operativa y Dependencia de Infraestructura (P2I/V2P):**

Se evalúa la escalabilidad de la solución tecnológica frente a entornos urbanos heterogéneos. El criterio penaliza aquellos sistemas que condicionan su funcionamiento a inversiones gubernamentales previas (Smart City), priorizando tecnologías capaces de operar de manera independiente en cruces no semaforizados o sin conectividad de red.

- **Latencia Computacional y Procesamiento Temporal:**

Dado que la estimación de riesgo vehicular requiere tiempos de reacción inmediatos, se analiza la eficiencia del hardware y software para emitir alertas en tiempo real. Este criterio evalúa los cuellos de botella computacionales, particularmente en el procesamiento de visión artificial, y la capacidad de los sistemas para mantener una alta tasa de cuadros por segundo (FPS) sin incurrir en sobrecalentamiento.

- **Robustez Perceptual y Tolerancia Ambiental:**

Se contrasta la fidelidad de los arreglos sensoriales utilizados (cámaras de smartphones, LiDAR, cámaras RGB-D) frente a la alta variabilidad del mundo real. Se examina específicamente la degradación del rendimiento algorítmico ante escenarios de iluminación adversa (noche o deslumbramiento solar) y la capacidad de mantener el seguimiento de objetos a diferentes escalas.

- **Interacción Humano-Computadora (HCI) y Carga Cognitiva:**

Se analiza la eficacia ergonómica y comunicativa del sistema. Este parámetro evalúa las fricciones operativas impuestas al usuario (como la dificultad para alinear un dispositivo de forma manual) y la idoneidad de las interfaces de salida (auditivas frente a hápticas) para garantizar la transmisión efectiva de la alerta en entornos con alta contaminación acústica, minimizando la saturación sensorial.

A partir de la aplicación transversal de estos cuatro parámetros de evaluación, la literatura actual se ha estructurado en cuatro grandes categorías tecnológicas. A continuación, se detalla el desempeño de cada enfoque frente a las métricas establecidas.

2.1 Sistemas Basados en Infraestructura (IoT)

Esta categoría agrupa soluciones que dependen de la comunicación Person-to-Infraestructura (P2I) para asistir el cruce. La premisa fundamental es externalizar la percepción; en lugar de que el usuario “vea” el semáforo, la infraestructura le comunica su estado digitalmente.

- **Comunicación Directa y Nube:**

Investigaciones recientes han implementado sistemas bidireccionales; en (Huang et al., 2022) desarrollaron en Taiwán una arquitectura IoT donde los semáforos transmiten su estado directamente al smartphone del usuario, eliminando

la necesidad de señales auditivas ambientales que pueden ser opacadas por el ruido del tráfico. En una aproximación similar pero centralizada (Ihejimba & Wenkster, 2020) propusieron “DetectSignal”, un sistema basado en la nube que utiliza la geolocalización (GPS) del usuario para consultar el estado de la intersección en tiempo real sin requerir línea de vista visual.

- **Percepción desde la Infraestructura:**

Llevando el concepto P2I un paso más allá (Montanha et al., 2022) invirtieron la carga de procesamiento. Su sistema detecta la presencia del peatón mediante señales Wi-Fi y utiliza cámaras IP instaladas en el semáforo para analizar el tráfico con algoritmos Deep Learning. Este enfoque libera al dispositivo del usuario de la carga computacional (ver Fig. 1).



Figura 1. Sistema basado en infraestructura.

Si bien estos sistemas mitigan problemas de visión artificial (como la oclusión o la baja iluminación), presentan una barrera de implementación crítica: su dependencia de una infraestructura urbana (Smart City) altamente digitalizada. La viabilidad de estas soluciones está condicionada a inversiones gubernamentales masivas para equipar cada intersección, lo que las hace imprácticas para una adopción universal, especialmente en ciudades de países en desarrollo donde la infraestructura básica es heterogénea o inexistente. Además, estos sistemas no protegen al usuario en cruces no regulados ni ante conductores que ignoran la señal de alto, una variable de riesgo impredecible que la infraestructura por sí sola no puede gestionar.

2.2 Aplicaciones Móviles y Visión Computacional

Esta categoría explora el uso de smartphones como dispositivos de asistencia primarios, aprovechando su ubicuidad y capacidad de procesamiento creciente. Sin embargo, a pesar de la popularidad de herramientas como Seeing AI o Lookout, revisiones sistemáticas recientes identifican limitaciones críticas de fiabilidad para tareas de seguridad vital.

- Robustez y Condiciones Ambientales:

Estudios exhaustivos (Khan & Khusro, 2021) concluyen que las aplicaciones basadas en visión artificial presentan una degradación significativa de rendimiento en escenarios de iluminación adversa (noche o contraluz) y generan un consumo energético insostenible para el monitoreo continuo, reduciendo drásticamente la autonomía del dispositivo móvil y, por ende, la seguridad del usuario.

- Barreras de Usabilidad (Evaluación Cuantitativa):

Más allá del rendimiento técnico, la interacción humano-computadora presenta fricciones severas, por ejemplo, (Palma-Mecias et al., 2025) aplicaron la metodología IPDE (Identificación, Planificación, Diseño y Evaluación) para cuantificar la eficiencia de estas herramientas. Sus resultados experimentales revelaron que las aplicaciones de visión artificial obtuvieron puntuaciones de usabilidad significativamente inferiores (<91 puntos) en comparación con herramientas de accesibilidad nativas como Google Talkback (149 puntos). Esta disparidad sugiere que la interfaz táctil impone una carga cognitiva excesiva durante la navegación.

- El Desafío de la Alineación de Cámara:

Una limitación operativa fundamental, validada por (Bhagat et al., 2023) en su evaluación de las soluciones líderes del mercado, es el problema de la “Alineación de la Cámara”. Los usuarios con discapacidad visual enfrentan dificultades sustanciales para encuadrar y estabilizar el dispositivo manual hacia objetos de interés dinámicos sin una referencia visual previa. Esta limitación convierte la detección de vehículos en un proceso lento y frustrante, restringiendo la utilidad de las apps a entornos estáticos y controlados.

La literatura evidencia que, aunque el software móvil ha avanzado, el factor de forma del smartphone es inherentemente inadecuado para la percepción ambiental activa. La necesidad de sostener y apuntar el dispositivo anula la ventaja de “manos libres” y la falta de sensores dedicados (como cámaras de profundidad o visión nocturna) impide una evaluación de riesgo robusta. Esto justifica la transición hacia hardware dedicado (wearables o bastones instrumentados) que garantice una percepción pasiva, continua y alineada con la perspectiva del usuario, como se ejemplifica en la Fig. 2.



Figura 2. Smartphone y visión por computadora.

2.3 Hardware Dedicado: Bastones Inteligentes y Robots

Para superar las limitaciones del smartphone como sensor primario, una tercera línea de investigación se ha centrado en el desarrollo de hardware dedicado que integra la percepción directamente en el dispositivo de asistencia o en plataformas autónomas.

- Bastones Inteligentes y Navegación Estática:

El desarrollo de bastones instrumentados ha evolucionado desde la incorporación de sensores ultrasónicos básicos hasta soluciones integradas con sistemas operativos móviles. Un ejemplo representativo es el bastón basado en la arquitectura de Android que utiliza sensores para el reconocimiento de obstáculos acoplado a alertas de voz (Bhatnagar et al., 2022). Sin embargo, el alcance de estas soluciones se limita predominantemente a la navegación en aceras y la detección de obstáculos estáticos, como postes o paredes (Bhatnagar et al., 2022) (ver Fig. 3).



Figura 3. Smart Canes.

Sistemas de vanguardia han intentado robustecer esta percepción mediante la fusión de sensores complejos, integrando, por ejemplo, un LiDAR 2D con una cámara de profundidad RGB-D para generar un reconocimiento complejo del pavimento y del entorno inmediato (Mai et al., 2024). A pesar de su alta precisión técnica para identificar características estáticas del entorno, estos dispositivos no están diseñados para el análisis del tráfico dinámico a distancia, dejando una brecha crítica en la seguridad durante el cruce de calles (Mai et al., 2024).

Dando un paso más allá de la navegación estática, investigaciones recientes abordan directamente el problema del cruce dinámico mediante sistemas portátiles (wearables) basados en cámara RGB-D y algoritmos de Deep Learning (Tian et al., 2021). Estos enfoques se especializan en la comprensión de la escena del cruce, rastreando vehículos en movimiento y estimando su intención para realizar una evaluación de riesgo en tiempo real (Tian et al., 2021). Adicionalmente, estudios sobre las modalidades de retroalimentación en estas situaciones han demostrado que el feedback háptico (vibraciones) es significativamente más efectivo, rápido e intuitivo para transmitir el riesgo de vehículos en aproximación que las alertas puramente visuales o sonoras (Ren et al., 2025).

• Robots Guía Autónomos y Limitaciones de la Plataforma:

Paralelamente a los dispositivos portátiles, se ha explorado el desarrollo de robots guía autónomos que actúan como acompañantes activos. Un sistema destacado en esta categoría es CaBot, un robot con forma de maleta que implementa una avanzada fusión de sensores combinando LiDAR para localización y evasión de obstáculos con una cámara estéreo (ZED), la cual es procesada por redes neuronales convolucionales (CNN) para añadir comprensión semántica del entorno (Guerreiro et al., 2019), cuyo concepto se ilustra en la Fig. 4.



Figura 4. Robots Guía autónomos.

Recientemente, la investigación ha evaluado el uso de robots cuadrúpedos (perros-robot) debido a su forma biomimética y potencial agilidad en terrenos complejos. No obstante, estudios experimentales con usuarios ciegos han revelado que el ruido excesivo y los movimientos bruscos inherentes a la

locomoción robótica actual limitan significativamente la usabilidad y la confianza del usuario, siendo preferidas las plataformas de ruedas convencionales (Wang et al., 2023). Para mitigar estos problemas de interacción humano-robot, nuevas propuestas se enfocan en un diseño basado en el confort, utilizando, por ejemplo, correas motorizadas y elásticas guiadas por modelos de movimiento humano para garantizar una tracción suave durante la navegación (Chen et al., 2023).

A pesar de la sofisticación de estas plataformas electromecánicas, para resolver el problema específico de cruzar una intersección vehicular, el factor limitante sigue siendo la capacidad del hardware embebido para analizar las amenazas de forma inmediata. Esta necesidad de procesar video estéreo en fracciones de segundo para evitar accidentes justifica la transición de la investigación hacia arquitecturas que integren aceleración por hardware optimizada y algoritmos ligeros de visión artificial, asegurando la baja latencia requerida en situaciones de riesgo vital.

2.4 Algoritmos de Percepción y Procesamiento Embebido

Un desafío crítico en el desarrollo de dispositivos portátiles de asistencia es la limitada capacidad de cómputo del hardware embebido. En el contexto de la movilidad urbana en México, investigaciones previas desarrollaron sistemas de detección de tráfico vehicular y semáforos implementando la arquitectura YOLOv3 sobre placas de propósito general como la Raspberry Pi 4 (Cleofas-Sánchez et al., 2023). Aunque el sistema logró identificar elementos del entorno local, reportó una tasa de procesamiento ineficiente de apenas 0.4 cuadros por segundo (FPS) y problemas significativos de sobre calentamiento (Cleofas-Sánchez et al., 2023). Este antecedente evidencia que el uso de arquitecturas de detección clásicas sobre hardware genérico resulta insuficiente para garantizar la seguridad en tiempo real, lo que justifica la transición hacia modelos de inteligencia artificial altamente optimizados y aceleradores gráficos embebidos. Un ejemplo de esta capacidad de detección se muestra en la Fig. 5.

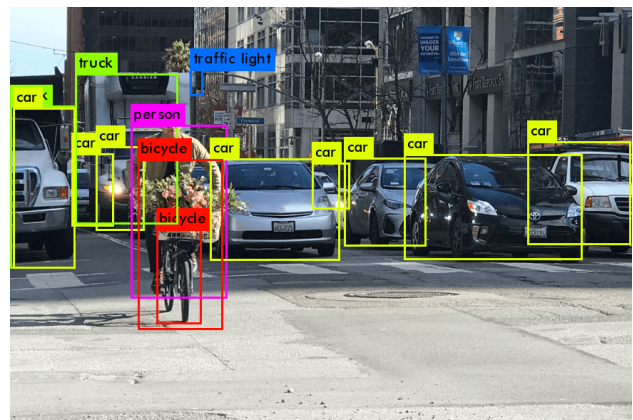


Figura 5. Detección vehicular con software YOLOv8.



- Evolución hacia la Arquitectura YOLOv8:

Frente a las limitaciones de latencia, la literatura reciente ha validado la eficiencia de los modelos de detección de última generación. Un estudio comparativo exhaustivo evaluó el desempeño de la familia YOLOv8 (en sus variantes Nano, Small y Medium) para el reconocimiento de elementos viales (Zanevych, 2024). Los resultados demostraron que YOLOv8 supera significativamente a sus predecesores en precisión, destacando la variante YOLOv8n (Nano) por ofrecer un equilibrio excepcional entre eficiencia computacional y rendimiento, alcanzando una precisión del 94.7% (Zanevych, 2024). Este avance es determinante, ya que sugiere que la adopción de redes ligeras resuelve los cuellos de botella de latencia sin sacrificar la confiabilidad operativa.

Avanzando en la optimización para escenarios urbanos complejos, investigaciones recientes introdujeron arquitecturas mejoradas como YOLO-BS, basadas en el marco de YOLOv8 (Zhang et al., 2025). Al abordar la dificultad de detectar objetos pequeños o a gran distancia, este modelo integra redes piramidales de características bidireccionales (BiFPN), logrando incrementar la sensibilidad del sistema hasta alcanzar una precisión media del 90.1% manteniendo una tasa operativa de 78 FPS (Zhang et al., 2025). Esto demuestra que es posible obtener un rendimiento de grado de seguridad en hardware moderno mediante la optimización estructural de la red neuronal.

- Análisis Temporal y Entornos Dinámicos:

Más allá de la detección estática en fotogramas individuales, la frontera actual de la percepción artificial para la movilidad requiere la comprensión de dinámicas temporales para resolver ambigüedades. El desarrollo de enfoques como FlashLightNet, que integra arquitecturas ligeras (YOLOv10n) con redes de memoria a corto y largo plazo (LSTM), ha permitido resolver problemas complejos como la interpretación de señales de tráfico parpadeantes (Khaled et al., 2025). Alcanzando precisiones superiores al 99.2%, este enfoque metodológico valida la necesidad de analizar secuencias continuas de video en lugar de imágenes aisladas para comprender la dinámica del entorno y calcular el riesgo real a través de los vectores de movimiento de los vehículos (Khaled et al., 2025).

- Sinergia de Hardware y Contexto Geográfico:

Consolidando estas tendencias algorítmicas, revisiones sistemáticas sobre la detección de tráfico mediante la familia YOLO identifica que las plataformas NVIDIA Jetson (como Xavier NX y Nano) se han posicionado como el estándar predominante para estas aplicaciones debido a su arquitectura de procesamiento en paralelo (Flores-Calero et al., 2024). De igual forma, la literatura subraya una escasez de investigaciones que aborden las disparidades geográficas en los conjuntos de datos, los cuales provienen mayoritariamente de entornos europeos o sudamericanos (Flores-Calero et al., 2024). Esto valida de manera imperativa la necesidad de generar conjuntos de datos locales para garantizar que la

estimación de riesgo sea aplicable y segura en el diseño de la infraestructura vial mexicana.

4. ANÁLISIS COMPARATIVO

Tabla 1. Comparación de tecnologías para asistencia.

Enfoque Tecnológico	Percepción Principal	Ventajas Principales	Limitaciones Críticas
Infraestructura IoT (P2I)	Estado del semáforo e intersección conectada	Evita problemas de visión artificial (oclusión, baja luz).	Dependencia total de inversión gubernamental (Smart City). Inútil en cruces no regulados.
Aplicaciones Móviles (Smartphone)	Visión computacional mediante cámara del teléfono	Alta disponibilidad y nulo costo de hardware adicional.	Pobre rendimiento en baja iluminación, alto consumo de batería y dificultad de alineación manual.
Bastones Básicos (Sensores/LiD AR)	Obstáculos estáticos a corto alcance (postes, paredes).	Autonomía de red, alta precisión para evasión inmediata en banquetas.	Incapacidad para detectar vehículos a distancia o evaluar tráfico dinámico.
Edge Computing (Cámaras + YOLOv8)	Vehículos en movimiento, intención y riesgo dinámico.	Percepción pasiva continua, procesamiento en tiempo real con baja latencia.	Requiere hardware especializado (GPUs embebidas) y optimización de consumo energético.

5. CONCLUSIONES

La movilidad autónoma en cruces peatonales constituye un desafío crítico que expone a las personas con discapacidad visual a riesgos vitales, exacerbados por la naturaleza dinámica y a menudo impredecible del tráfico vehicular. A través de la presente revisión comparativa, se ha evidenciado que las tecnologías de asistencia actuales presentan limitaciones estructurales significativas que impiden una mitigación integral de este riesgo.

Por un lado, los sistemas basados en comunicación con la infraestructura (IoT) ofrecen información precisa sobre las intersecciones, pero su viabilidad está estrictamente condicionada a inversiones urbanas masivas, lo que limita severamente su escalabilidad y uso universal. Por otro lado, las aplicaciones móviles basadas en visión computacional



enfrentan barreras críticas de usabilidad, latencia de procesamiento y una drástica degradación del rendimiento en condiciones de iluminación adversa. Asimismo, se identificó que la mayoría de los dispositivos físicos dedicados, como los bastones inteligentes y los robots guía convencionales, se han centrado tradicionalmente en la navegación estática y la evasión de obstáculos inmediatos, omitiendo la necesidad algorítmica de analizar el flujo vehicular a distancia.

Por lo tanto, se concluye que la evolución de las tecnologías debe transitar de la simple percepción pasiva de obstáculos hacia la evaluación activa del riesgo dinámico. La evidencia analizada apunta a que la integración de arquitecturas de Edge Computing de alto rendimiento, en sinergia con algoritmos de detección ligeros y altamente optimizados como YOLOv8, representa la ruta tecnológica más prometedora. Al combinar esta capacidad de procesamiento en tiempo real con interfaces de retroalimentación háptica, es posible desarrollar sistemas verdaderamente portátiles y autónomos. Este enfoque descentralizado elimina la dependencia de la infraestructura externa y minimiza la carga cognitiva, garantizando una toma de decisiones segura y fomentando una autonomía real para los usuarios en entornos urbanos complejos.

REFERENCIAS

- S. R. Flaxman et al., "Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis," *Lancet Glob. Health*, vol. 5, no. 12, 2017, doi: 10.1016/S2214-109X(17)30393-5.
- C. Mai et al., "A Smart Cane Based on 2D LiDAR and RGB-D Camera Sensor-Realizing Navigation and Obstacle Recognition," *Sensors*, vol. 24, no. 3, 2024, doi: 10.3390/s24030870.
- Smith, S.E. and Jones, L.Q. (2008).
- D. R. Geruschat and S. E. Hassan, "Driver behavior in yielding to sighted and blind pedestrians at roundabouts," *J. Vis. Impair. Blind.*, vol. 99, no. 5, 2005, doi: 10.1177/0145482x0509900504.
- A. Budrionis, D. Plikynas, P. Daniušis, and A. Indrulionis, "Smartphone-based computer vision travelling aids for blind and visually impaired individuals: A systematic review," 2022. doi: 10.1080/10400435.2020.1743381.
- A. Khan and S. Khusro, "An insight into smartphone-based assistive solutions for visually impaired and blind people: issues, challenges and opportunities," 2021. doi: 10.1007/s10209-020-00733-8.
- C. Y. Huang, C. K. Wu, and P. Y. Liu, "Assistive technology in smart cities: A case of street crossing for the visually-impaired," *Technol. Soc.*, vol. 68, 2022, doi: 10.1016/j.techsoc.2021.101805.
- C. Ihejimba and R. Z. Wenkstern, "DetectSignal: A Cloud-Based Traffic Signal Notification System for the Blind and Visually Impaired," in *2020 IEEE International Smart Cities Conference, ISC2 2020*, 2020. doi: 10.1109/ISC251055.2020.9239004.
- A. Montanha, A. M. Oprescu, and Mc. Romero-Ternero, "A Context-Aware Artificial Intelligence-based System to Support Street Crossings For Pedestrians with Visual Impairments," *Applied Artificial Intelligence*, vol. 36, no. 1, 2022, doi: 10.1080/08839514.2022.2062818.
- J. Palma-Mecias, J. Herrera-Tapia, D. Muñoz-Verduga, L. Fernández-Capestany, and H. Herrera-Tapia, "Guía de Evaluación de Aplicaciones Móviles de Aprendizaje para Personas con Discapacidad Visual," *Revista Científica de Informática ENCRIPtar*, vol. 8, no. 16, pp. 227–240, Oct. 2025, doi: 10.56124/encriptar.v8i16.012.
- S. Bhagat, P. Joshi, A. Agarwal, and S. Gupta, "Accessibility evaluation of major assistive mobile applications available for the visually impaired," *ITU Journal on Future and Evolving Technologies*, vol. 4, no. 4, 2023, doi: 10.52953/tnrv4696.
- A. Bhatnagar, A. Ghosh, and S. M. Florence, "Android Integrated Voice based walking stick for blind with obstacle recognition," in *4th International Conference on Emerging Research in Electronics, Computer Science and Technology, ICERECT 2022*, 2022. doi: 10.1109/ICERECT56837.2022.10060809.
- S. Tian, M. Zheng, W. Zou, X. Li, and L. Zhang, "Dynamic Crosswalk Scene Understanding for the Visually Impaired," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, 2021, doi: 10.1109/TNSRE.2021.3096379.
- G. Ren, Z. Huang, W. Lin, T. Huang, G. Wang, and J. H. Lee, "Enhancing Street-Crossing Safety for Visually Impaired Pedestrians with Haptic and Visual Feedback," *Applied Sciences*, vol. 15, no. 7, p. 3942, Apr. 2025, doi: 10.3390/app15073942.
- J. Guerreiro, D. Sato, S. Asakawa, H. Dong, K. M. Kitani, and C. Asakawa, "Cabot: Designing and evaluating an autonomous navigation robot for blind people," in *ASSETS 2019 - 21st International ACM SIGACCESS Conference on Computers and Accessibility*, 2019. doi: 10.1145/3308561.335377.
- L. Wang et al., "Can Quadruped Guide Robots be Used as Guide Dogs?," in *IEEE International Conference on Intelligent Robots and Systems*, 2023. doi: 10.1109/IROS55552.2023.10341792.
- Y. Chen et al., "Quadruped Guidance Robot for the Visually Impaired: A Comfort-Based Approach," in *Proceedings - IEEE International Conference on Robotics and Automation*, 2023. doi: 10.1109/ICRA48891.2023.10160854.
- L. Cleofas-Sánchez, J. P. F. Posadas-Durán, P. Martínez-Ortiz, G. Loyo-Desiderio, E. A. Ruvalcaba-Hernández, and O. G. Brito, "Automatic Detection of Vehicular Traffic Elements based on Deep Learning for Advanced Driving Assistance Systems," *Computacion y Sistemas*, vol. 27, no. 3, 2023, doi: 10.13053/CyS-27-3-4508.



- Y. Zanevych, "TRAFFIC LIGHT DETECTION WITH YOLO MODELS," Grail of Science, no. 38, pp. 194–199, Apr. 2024, doi: 10.36074/grail-of-science.12.04.2024.033.
- H. Zhang, M. Liang, and Y. Wang, "YOLO-BS: a traffic sign detection algorithm based on YOLOv8," Sci. Rep., vol. 15, no. 1, Dec. 2025, doi: 10.1038/s41598-025-88184-0.
- L. B. Khaled, M. Rahman, I. A. Ebu, and J. E. Ball, "FlashLightNet: An End-to-End Deep Learning Framework for Real-Time Detection and Classification of Static and Flashing Traffic Light States," Sensors, vol. 25, no. 20, Oct. 2025, doi: 10.3390/s25206423.
- M. Flores-Calero et al., "Traffic Sign Detection and Recognition Using YOLO Object Detection Algorithm: A Systematic Review," 2024. doi: 10.3390/math12020297.

Analysis and Design of an Didactic Platform for Monitoring and Fault Detection in Hydraulic Systems

Cesar R. Tovar-Salazar* Romina G. Quintero-Flores*
Ian M. Muñoz-Chairez* Montserrat Y. Santos-Rojas*
Andrik Jarillo-Mata* Iván A. Reyes-Portillo*

** Departamento de Electrónica, Centro de Formación Profesional (CFP)/ Bachillerato Tecnológico Salesiano Carlos Gómez, 78240, San Luis Potosí, S.L.P., México. (e-mail: ia.reyes@ieee.org).*

Abstract: This paper presents the design and implementation of an educational platform for real-time monitoring and fault detection in hydraulic systems. The system integrates flow, current, and vibration sensors, along with an architecture based on Arduino and ESP32 microcontrollers, enabling local and remote data acquisition, processing, and visualization via LCD screens and a mobile app. A methodology is implemented that includes sensor calibration, data acquisition, and the definition of operational thresholds to identify normal and abnormal conditions. Experimental validation is conducted under various operating scenarios, including load variations and fault emulation. The results demonstrate the system's ability to detect changes in system behavior in a timely manner. The proposed platform offers an accessible, low-cost solution for educational settings, facilitating the understanding of monitoring techniques and predictive maintenance strategies in hydraulic systems. Furthermore, it provides a practical foundation for integrating instrumentation, data processing, and real-time visualization into engineering education.

Keywords: Hydraulic systems, fault detection, monitoring, ESP32, vibration analysis.

1. INTRODUCTION

Hydraulic systems are among the most widely used technologies in the industrial sector due to their high power density, robustness, and ability to transmit energy efficiently and in a controlled manner (Tian et al., 2024; Bhola and Wratt, 2025). In this context, industrial processes require highly efficient systems designed to optimize energy use at the various stages of production (Reyes-Portillo et al., 2025). In particular, these systems are widely used in sectors such as manufacturing, automation, heavy machinery, and various industrial processes, where the application of large forces with high precision is required. Their operation is based on the transmission of energy through pressurized fluids, which allows for precise control of variables such as force, speed, and position (Tian et al., 2024). Despite their advantages, hydraulic systems are prone to various failures that can compromise their performance, safety, and service life (Gheisi et al., 2016). Among the most common failures are leaks in lines and connections, hydraulic fluid contamination, cavitation, component wear, and abnormal variations in pressure and flow (Sharma et al., 2024). These conditions not only reduce system efficiency but can also lead to unscheduled downtime and high maintenance costs (Dunscombe et al., 2000).

In this context, the implementation of monitoring and early-detection strategies becomes essential for improving operational reliability and facilitating the transition to predictive maintenance schemes (Adebisi et al., 2021). Advances in instrumentation and data acquisition technologies have enabled the development of real-time monitoring systems capable of tracking critical variables such as pressure, flow rate, temperature, and vibrations. By analyzing these variables, it is possible to identify anomalous behavior patterns associated with incipient failures, which facilitates timely decision-making (Zheng et al., 2023).

However, in the educational setting, access to such systems is often limited, which hinders the practical understanding of concepts related to fault diagnosis in hydraulic systems. In this regard, there is a need to develop educational platforms that integrate both theoretical principles and hands-on experimentation (Liuta et al., 2019). These platforms allow students and professionals to interact directly with real or semi-real systems, understand their operational dynamics, and analyze the impact of different failure conditions. In addition, they foster the development of skills in key areas such as instrumentation, signal processing, data acquisition, and interpretation (Canut-Montalva et al., 2025; Mirauda et al., 2019). This paper presents the analysis and design of an educational platform for monitoring and detecting faults in hydraulic systems. The proposal integrates sensors for measuring critical variables, a data acquisition system, and a real-time visualization interface that enables the analysis of system behavior under various

* The authors thank the Centro de Formación Profesional (CFP) / Bachillerato Tecnológico Salesiano Carlos Gómez for providing the resources used in this work.



operating conditions. In addition, the platform is designed to simulate common failures, which facilitates the experimental study and validation of diagnostic techniques. The aim is to help strengthen the teaching-learning process in engineering by providing a tool that combines theoretical analysis with practical application.

2. DESIGN OF THE DIDACTIC PLATFORM

This work is conducted using a quantitative approach, based on the measurement, acquisition, and analysis of physical variables associated with the operation of a hydraulic pumping system. The research is applied and technological in nature, as it aims to design and implement an educational platform focused on monitoring and fault detection, integrating concepts of electronics, instrumentation, programming, and data analysis. Furthermore, the study has an experimental scope, as controlled tests were conducted on a real system, which allowed for the evaluation of the equipment's behavior under various operating conditions and the validation of the proposed platform's performance. The platform is conceived as a comprehensive system capable of acquiring, processing, and displaying information in real time. To this end, a hierarchical architecture comprising three functional levels is proposed:

- **Physical (sensing) level:** responsible for measuring variables using appropriate sensors.
- **Processing level:** responsible for the acquisition, conditioning, and processing of signals.
- **Visualization level:** oriented toward the presentation and interpretation of data by the user.

Three key variables were selected to characterize the system:

- **Flow rate:** an indicator of the system's hydraulic performance.
- **Vibration:** associated with the mechanical condition and potential failure conditions.
- **Electric current:** related to the energy behavior of the drive system.

The selection of these variables provides a comprehensive representation of the pumping system's operational status, facilitating the early identification of anomalies and contributing to the analysis of both normal and failure operating conditions. The following components were used to implement the educational platform; they were selected based on their performance, availability, and suitability for the experimental purpose:

- 1 HP water pump, used as a test system for the controlled generation of hydraulic flow, allowing for the reproduction of various operating conditions.
- ESP32 microcontroller, used as the main data acquisition and processing unit due to its computing power, wireless connectivity (Wi-Fi/Bluetooth), and support for real-time applications.
- Hall-effect flow sensor (FS300A G3/4), selected for its ease of integration, low cost, and ability to provide reliable flow measurements in low- and medium-pressure applications.
- MPU6050 sensor (accelerometer and gyroscope), used to measure mechanical vibrations, enabling the iden-

tification of patterns associated with abnormal conditions in the system.

- ACS712 current sensor, used to monitor the motor's power consumption, facilitating the detection of load variations and potential electrical faults.
- 16x2 LCD screens with I2C interface, used for local display of variables, reducing the number of connections and simplifying system integration.
- Mobile app developed in Flutter, designed for remote monitoring and real-time visualization of variables, leveraging the ESP32's connectivity.
- Arduino IDE development environment, used for programming and integrating devices due to its compatibility and ease of use.
- Electrical startup system, consisting of a switch, contactor, and relay, which ensures safe and reliable control of system operation.
- Auxiliary components, such as wiring, a breadboard, and a power supply, necessary for system interconnection and operation.

3. CHARACTERISATION OF THE SENSORS

The sensors integrated into the platform were selected on the basis of low cost, commercial availability, ease of integration and compatibility with the data acquisition system. This selection ensures the platform's reproducibility in educational settings, whilst maintaining reliability in the measurement of key hydraulic variables. Table 1 summarizes the main specifications of the FS300A sensor, used for flow rate measurement in the hydraulic system. This device operates based on the Hall-effect principle, with a measurement range of 1–60 L/min, a 5 V DC supply voltage, and pulse output, enabling direct integration with the data acquisition and processing system.

Table 1. FS300A flow sensor specifications.

Characteristic	Description
Model	FS300A (Hall-effect)
Measurement range	1–60 L/min
Operating voltage	5 V DC
Output type	Pulse signal
Operating pressure	0–1.75 MPa

Fig. 1 shows the relationship between flow rate (Q , in L/min) and the sensor output frequency (F , in Hz). The obtained linear fit is $F = 7.50Q + 0.12$, with a coefficient of determination $R^2 = 0.9997$. This R^2 value (very close to 1) indicates that the frequency varies in an almost perfectly linear manner with flow rate, confirming the sensor's high linearity across its entire operating range. Fig. 2 shows the frequency-to-voltage conversion. The fitted equation is $V = 0.0375Q + 0.0062$, with $R^2 = 0.9996$. This demonstrates that the conversion preserves the linearity required for accurate flow rate estimation.

Table 2 presents the specifications of the ACS712 sensor, used for current monitoring in electrical actuators. Its analog output and bidirectional measurement range of ± 30 A enable straightforward integration with microcontrollers.

Fig. 3 shows the ACS712 response using a 10-bit ADC. The measurement error remains within ± 0.5 A, confirming its suitability for electrical load monitoring applications.

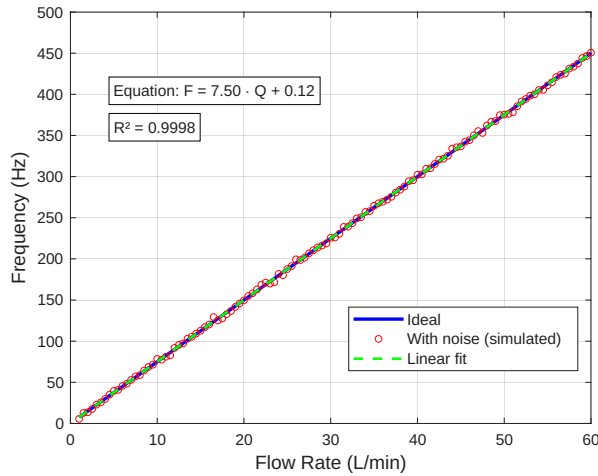


Fig. 1. Characterisation of the FS300A sensor: output frequency as a function of flow rate.

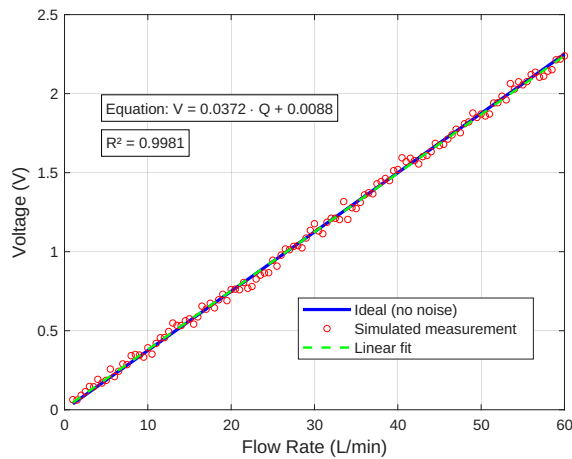


Fig. 2. Characterisation of the FS300A with F/V converter: output voltage as a function of flow rate.

Table 2. ACS712 Current sensor specifications.

Characteristic	Description
Model	ACS712
Measurement range	±30 A
Operating voltage	5 V DC
Output type	Analog signal
Sensitivity	66–185 mV/A

The MPU6050 integrates an accelerometer and gyroscope in a single module, enabling mechanical vibration monitoring associated with hydraulic faults. Its main specifications are summarized in Table 3.

Table 3. MPU6050 Sensor specifications.

Characteristic	Description
Model	MPU6050
Functions	Accelerometer and gyroscope
Operating voltage	3.3–5 V
Communication interface	I2C
Gyroscope range	±250 to ±2000 °/s
Accelerometer range	±2g to ±16g

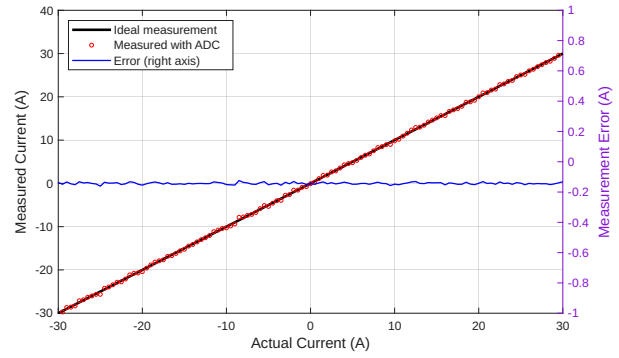


Fig. 3. ACS712 sensor characterization: current measurement and associated error.

Fig. 4 shows the vibration-to-voltage transfer characteristic, exhibiting a linear response with a sensitivity of 0.103 V/(m/s²).

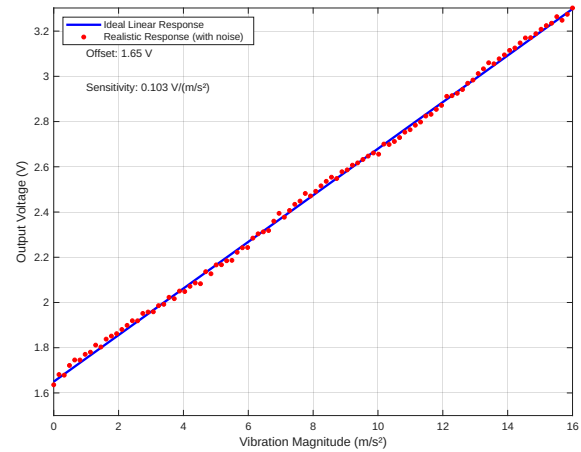


Fig. 4. MPU6050 characterization: output voltage as a function of vibration magnitude.

Fig. 5 shows the temporal response under normal and fault conditions, where the increase in RMS amplitude enables early anomaly detection.

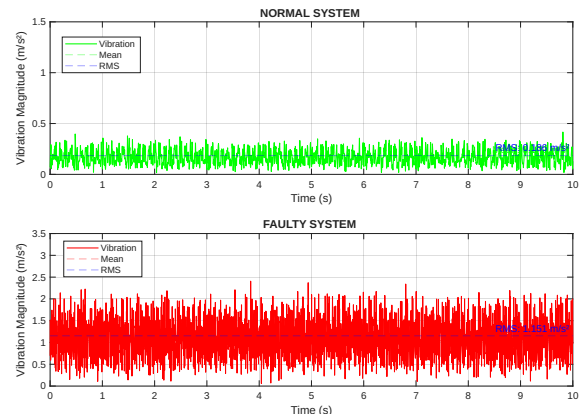


Fig. 5. Temporal response of the MPU6050 under normal and fault conditions.

Overall, the experimental results validate that the selected sensors satisfy the range, sensitivity, and linearity requirements necessary for reliable hydraulic system monitoring.

4. INTEGRATION OF THE PROPOSED PLATFORM

Fig. 6 shows the block diagram of the proposed architecture for the integration of the hydraulic monitoring platform. In this configuration, the system is divided into two distributed subsystems: the first, based on Arduino, acquires and processes signals from the FS300A and ACS712 sensors for flow and current measurement, respectively, displaying the information locally on an LCD screen. The second subsystem uses an ESP32 to acquire vibration data from the MPU6050 sensor, enabling its wireless transmission to a mobile application developed in Flutter. Additionally, an interlock module is incorporated for the safe control of pump start-up and shutdown, ensuring reliable operation of the hydraulic system.

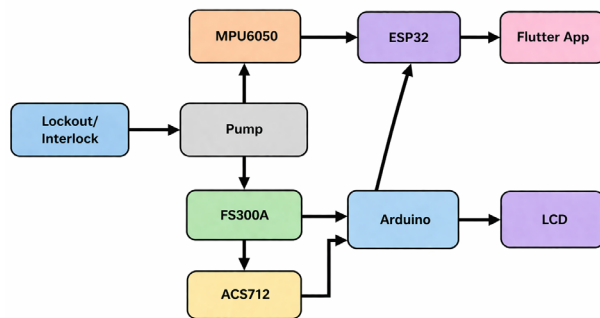


Fig. 6. Block diagram of the proposed architecture for monitoring flow, current and vibration in a hydraulic system.

The proposed monitoring architecture is organized into two subsystems. The first subsystem, based on an Arduino board, acquires signals from the FS300A flow sensor and the ACS712 current sensor, enabling local monitoring of hydraulic and electrical variables through an LCD interface. The second subsystem employs an ESP32 to process signals from the MPU6050 sensor, dedicated to the detection of mechanical vibrations associated with incipient faults. Wireless data transmission is achieved via Wi-Fi to a mobile application developed in Flutter, which provides real-time remote visualization. Fig. 7 shows the digital platform for real-time monitoring connected via Wi-Fi.

Additionally, the system incorporates an electrical interlocking module for the safe start and stop of the hydraulic pump, ensuring reliable operating conditions. The implementation of this scheme guarantees operational continuity through contactor self-holding and immediate disconnection upon a stop signal, following standard industrial control practices. Fig. 8 shows the experimental implementation of the hydraulic monitoring teaching platform during its real-time operation. It shows two hydraulic pumps integrated into the test bench: Pump 1, in optimal operating condition, used to monitor vibrations under normal conditions; and Pump 2, in a degraded condition,



Fig. 7. Digital platform for real-time monitoring connected via Wi-Fi

used to measure vibrations associated with failure scenarios. The figure also shows the electronic data acquisition and processing module, which integrates the flow, current and vibration sensors, enabling the simultaneous monitoring of critical process variables. The local interface displays the measured parameters in real time, validating the coordinated operation between the instrumentation, the hydraulic system and the embedded control unit.

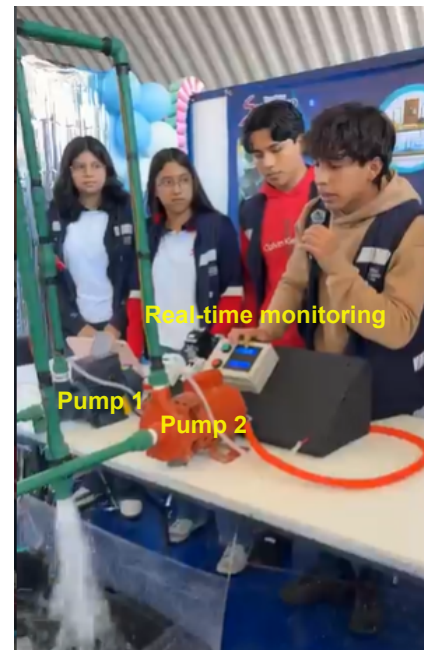


Fig. 8. Experimental implementation of the educational platform for hydraulic monitoring during its real-time operation.



Table 4. Detection performance for different fault conditions.

Fault Condition	Flow (L/min)	Vibration (m/s ²)	Current (A)	Efficiency (%)	Time (s)
No fault	9.92 ± 0.19	0.87 ± 0.08	5.58 ± 0.13	—	—
Partial blockage	4.85 ± 0.22	0.94 ± 0.10	5.85 ± 0.15	92.5	2.1
Cavitation	8.20 ± 0.33	1.95 ± 0.12	5.35 ± 0.14	94.3	1.8
Unbalance	9.65 ± 0.26	5.25 ± 0.20	6.05 ± 0.17	95.0	1.5
Worn bearing	9.48 ± 0.29	7.90 ± 0.23	5.78 ± 0.16	94.6	1.2
Overload	2.10 ± 0.48	1.18 ± 0.11	9.15 ± 0.22	91.8	3.2
Undervoltage	8.75 ± 0.30	1.05 ± 0.09	6.85 ± 0.18	90.5	2.5
Multiple fault	4.48 ± 0.40	5.50 ± 0.25	6.65 ± 0.20	90.0	2.7

Table 4 summarizes the performance of the monitoring system under seven operating conditions, including normal operation and six types of faults. Under normal conditions, the system recorded an average flow rate of 9.92 L/min, a vibration of 0.87 m/s², and a current consumption of 5.58 A. The experimentally induced faults included partial blockage, cavitation, mechanical imbalance, worn bearing, overload due to a closed valve, low voltage (110 V rms), and a multiple-fault condition. The results show detection efficiencies ranging from 90.0% to 95.0%. The mechanical imbalance fault exhibited the highest detection efficiency (95.0%), attributable to its high signal-to-noise ratio in the vibration axis. The multiple fault was the most difficult to detect (90.0% efficiency) due to overlapping fault signatures and partial flow rate cancellation. The shortest response time was obtained for the worn bearing (1.2 s), while overload exhibited the highest latency (3.2 s), explained by the need to integrate the current over multiple line cycles to avoid false alarms. In addition, the low-voltage condition (110 V rms) was identified with an efficiency of 90.5% through a combined analysis of high current (6.85 A) and low flow rate (8.75 L/min), demonstrating the system's ability to diagnose electrical faults in addition to mechanical and hydraulic ones. In addition, the low-voltage condition (110 V rms) was identified with an efficiency of 90.5% through a combined analysis of high current (6.85 A) and low flow rate (8.75 L/min), demonstrating the system's ability to diagnose electrical faults in addition to mechanical and hydraulic ones.

5. CONCLUSION

This paper presents the design and implementation of an educational platform for monitoring and fault detection in hydraulic systems, integrating flow, current, and vibration sensors with architectures based on Arduino and ESP32. Experimental results demonstrated that the system is capable of timely identifying various operating conditions, including hydraulic, mechanical, and electrical faults, with detection efficiencies exceeding 90% and response times suitable for real-time applications. Characterization of the sensors revealed linear and accurate behavior, ensuring reliability in data acquisition. Furthermore, the proposed architecture enabled local and remote visualization of variables, facilitating system analysis and fault interpretation. Overall, the developed platform constitutes a low-cost, easy-to-implement solution suitable for educational environments, as it enhances practical learning in areas such as instrumentation, monitoring, and predictive maintenance. For future work, we propose the incorporation of advanced diagnostic algorithms based on artificial intelligence, as

well as the expansion of the system to include a greater number of variables and failure scenarios, in order to improve the accuracy and robustness of the detection system.

ACKNOWLEDGEMENTS

The authors thank the Centro de Formación Profesional (CFP) / Bachillerato Tecnológico Salesiano Carlos Gómez for providing the resources used in this work.

REFERENCES

- Adebisi, B., Aigbedion, E., Ayorinde, O.B., and Onukwulu, E.C. (2021). A conceptual model for predictive asset integrity management using data analytics to enhance maintenance and reliability in oil & gas operations. *International Journal of Multidisciplinary Research and Growth Evaluation*, 2(1), 534–54. doi: <https://doi.org/10.54660/IJMRGE.2021.2.1.534-541>.
- Bhola, M. and Wratt, G. (2025). Energy-efficient hydraulics in heavy machinery: Technologies, challenges, and future directions. *Sustainability*, 18(1), 302. doi: <https://doi.org/10.3390/su18010302>.
- Canut-Montalva, A., Rizo-Maestre, C., and Martínez-López, J. (2025). Using pilot plants as a didactic resource to facilitate the acquisition of technical competences in the urban water cycle in vocational and higher education. *Water*, 17(5), 675. doi: <https://doi.org/10.3390/w17050675>.
- Duncombe, P., Roberts, G., and Valiquette, L. (2000). Preventative maintenance and unscheduled downtime from an economic perspective. *Journal of Applied Clinical Medical Physics*, 1(2), 68–75. doi: <https://doi.org/10.1120/jacmp.v1i2.2647>.
- Gheisi, A., Forsyth, M., and Naser, G. (2016). Water distribution systems reliability: A review of research literature. *Journal of Water Resources Planning and Management*, 142(11), 04016047. doi: [https://doi.org/10.1061/\(ASCE\)WR.1943-5452.0000690](https://doi.org/10.1061/(ASCE)WR.1943-5452.0000690).
- Liuta, A.V., Perig, A.V., Afanasieva, M.A., and Skyrach, V.M. (2019). Didactic games as student-friendly tools for learning hydraulics in a technical university's undergraduate curriculum. *Industry and higher education*, 33(3), 198–213. doi: <https://doi.org/10.1177/0950422218824>.
- Miranda, D., Capece, N., and Erra, U. (2019). Streamflowvl: A virtual fieldwork laboratory that supports traditional hydraulics engineering learning. *Applied Sciences*, 9(22), 4972. doi: <https://doi.org/10.3390/app9224972>.



- Reyes-Portillo, I.A., Claudio-Sánchez, A., Méndez Elizondo, S.R., Castro-López, D.L., and Carreto-Hernández, L.G. (2025). Analysis of kinetic energy recovery system based on four-phase interleaved buck converter for vehicle verification processes. *IEEE Latin America Transactions*, 23(6), 527–536. doi:10.1109/TLA.2025.11007189.
- Sharma, A.K., Kumar, N., and Das, A.K. (2024). A review on wear failure of hydraulic components: existing problems and possible solutions. *Engineering Research Express*, 6(1), 012502. doi:10.1088/2631-8695/ad299f.
- Tian, X., Guo, X., Stump, P., Dessy, G., Vacca, A., Fiorati, S., and Pintore, F. (2024). New hydraulic control technologies for improving the energy efficiency of the hydraulic system of agricultural tractors and their implements. *International Journal of Fluid Power*, 203–224. doi:10.13052/ijfp1439-9776.2525.
- Zheng, X., Liu, Y., and Shao, B. (2023). Causal analysis of fall accidents in hydraulic engineering based on text mining and decision-making trial and evaluation laboratory and interpretative structural modeling. *Water*, 15(21), 3810. doi:https://doi.org/10.3390/w15213810.

Meshing Contour Detection-based Triangulation for 2D Topology Optimization

Carlos Uriel Jiménez Monzón*. Arturo Gómez Ortega*. Mariela Gómez Castañeda*.

*Centro de Ingeniería y Desarrollo Industrial, 76125, Santiago de Querétaro, Querétaro, MX. (e-mail: cu.jimenez@posgrado.cidesi.edu.mx / arturo.gomez@cidesi.edu.mx).

Abstract: A 2D Python-based algorithm is presented for the generation of triangular finite element meshes on image-defined domains. The method integrates contour detection via the Suzuki-Abe algorithm, equidistant sampling with closure correction, and adaptive curvature-based sampling for complex geometries. Internal node distribution is achieved through Lloyd's algorithm and Delaunay triangulation. The result is a triangular mesh containing all the essential data for direct implementation in topology optimization algorithms.

Keywords: Topology optimization, additive manufacturing, Delaunay triangulation, Lloyd's algorithm, finite element method, Python.

1. INTRODUCCIÓN

La optimización topológica es una herramienta computacional ampliamente utilizada para el diseño estructural eficiente, permitiendo determinar la distribución óptima de material dentro de un dominio dado bajo restricciones de carga y desplazamiento (Bendsøe & Kikuchi, 1988). Su objetivo principal es minimizar el *compliance* — definido como la flexibilidad global de la estructura, o equivalentemente, la energía de deformación total bajo las cargas aplicadas — lo que se traduce en una distribución de material que maximiza la rigidez estructural con la menor cantidad de material posible. Entre sus principales ventajas destacan la reducción de material en la manufactura y la disminución de tiempos de producción, características que pueden aprovecharse directamente con la manufactura aditiva (Wang et al., 2023). Esta última ha ganado terreno como proceso de fabricación estable en diversas industrias, ya que al basarse en la fabricación capa por capa permite producir geometrías complejas que serían difíciles de obtener mediante procesos convencionales.

La mayoría de las implementaciones de optimización topológica trabajan sobre dominios rectangulares discretizados con mallas estructuradas de elementos cuadriláteros, lo que limita su aplicación a geometrías simples y puede introducir errores de escalonado en los bordes del dominio, como se ilustra en la Figura 1 (Zhuang et al., 2023). El uso de mallas triangulares no estructuradas sobre dominios arbitrarios representa una alternativa con mayor fidelidad geométrica, capaz de representar con precisión contornos curvos y geometrías complejas (Li et al., 2022). Sin embargo, su generación automatizada a partir de geometrías definidas por el usuario sigue siendo un reto en implementaciones de código abierto accesibles (Islam & Nayem, 2024).

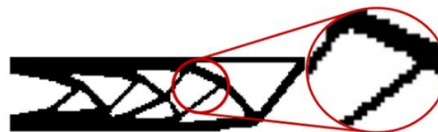


Figura 1. Efecto "escalonado" presente en los bordes del modelo alcanzado por OT mediante el algoritmo BESO.

En este contexto, el presente trabajo propone un algoritmo 2D desarrollado en Python (Python Software Foundation, 2025) para la generación automática de mallas triangulares de elementos finitos sobre dominios definidos por imagen. El método integra la detección de contornos mediante el algoritmo de Suzuki & Abe (1985) implementado en OpenCV (Bradski, 2000), el muestreo equidistante con corrección de cierre, el muestreo por curvatura para contornos complejos, el algoritmo de Lloyd y la triangulación de Delaunay mediante SciPy (Virtanen et al., 2020). Adicionalmente, se desarrolló una interfaz gráfica de usuario (GUI) que permite al usuario, señalar cargas, nodos fijos y exportar toda la información de la malla en formato JSON. Este archivo contiene la información completa necesaria para ejecutar análisis estructurales estáticos por elemento finito, siendo su implementación en algoritmos de optimización topológica una de sus aplicaciones principales.

2. METODOLOGÍA

2.1 Descripción general del algoritmo

El algoritmo propuesto genera mallas triangulares de elementos finitos sobre dominios definidos por imagen, siguiendo un flujo de trabajo secuencial que comprende cinco etapas principales: detección de contornos, discretización de la frontera, distribución de puntos internos, triangulación y exportación de resultados. El proceso se ejecuta completamente en Python (Python Software Foundation, 2025), haciendo uso de las bibliotecas OpenCV (Bradski, 2000), NumPy, SciPy (Virtanen et al., 2020) y Matplotlib. La Figura 2 ilustra el enfoque convencional de mado

rectangular, mientras que la Figura 3 muestra un ejemplo del tipo de dominio arbitrario que acepta el algoritmo propuesto, en este caso una viga con orificio central.

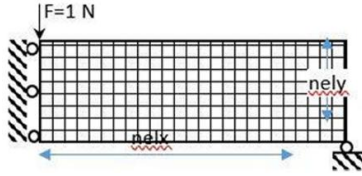


Figura 2. Subdivisión del dominio en elementos finitos mediante geometría cuadrada.

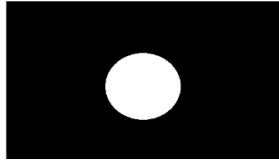


Figura 3. Ejemplo de figura aceptada por el código.

El resultado final es un archivo JSON que contiene toda la información necesaria para ejecutar un análisis estructural estático por elemento finito — posición de nodos, conectividad de elementos, cargas y condiciones de frontera — siendo su implementación en el algoritmo de optimización topológica SERA una de sus aplicaciones principales.

2.2 Definición del dominio

El algoritmo propuesto permite trabajar con cualquier geometría que el usuario proporcione como imagen, con la única condición de que el área del dominio sea de un color uniforme y el fondo sea blanco. A diferencia de los dominios rectangulares tradicionales en optimización topológica, esta característica permite analizar geometrías arbitrarias más representativas de casos de ingeniería reales.

La imagen es procesada mediante una lectura de la matriz de píxeles en espacio de color RGB, identificando el color predominante y generando una máscara binaria mediante un filtro de umbralización basado en rangos de color, descrito por la función de intensidad de la Ecuación (1):

$$f(x, y) = \begin{cases} 1 & \text{si } L \leq I(x, y) \leq U \\ 0 & \text{en caso contrario} \end{cases} \quad (1)$$

donde $I(x, y)$ es la intensidad del píxel y $[L, U]$ representa el intervalo de color detectado como predominante.

2.3 Detección de contornos

Sobre la máscara binarizada, el algoritmo ejecuta el método de seguimiento de bordes de Suzuki & Abe (1985), el cual detecta y analiza contornos en imágenes binarias, transformando la información de píxeles en vectores de coordenadas y permitiendo pasar de una representación discreta a una representación geométrica continua.

Para distinguir entre contornos externos e internos — es decir, huecos o cavidades dentro de la pieza — se emplea RETR_TREE, que construye una jerarquía de contornos en forma de árbol genealógico de la geometría. Posteriormente, la Fórmula de Shoelace (Ecuación 2) permite clasificar cada

contorno según su orientación: área positiva indica contorno exterior y área negativa indica un hueco o cavidad interior.

$$A_{\text{signo}} = \frac{1}{2} \sum_{i=0}^{n-1} (x_i y_{i+1} - x_{i+1} y_i) \quad (2)$$

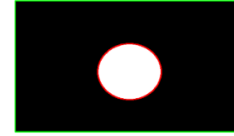


Figura 4. Detección de contorno externo (verde) e interno (rojo).

2.4 Distribución de puntos en el contorno

Una vez clasificados los contornos, el siguiente paso es discretizarlos colocando nodos sobre la frontera. Estos nodos actúan como límites inamovibles que confinan los puntos internos generados posteriormente. El algoritmo implementa dos estrategias de muestreo según la complejidad geométrica del contorno, como se ilustra en la Figura 5.

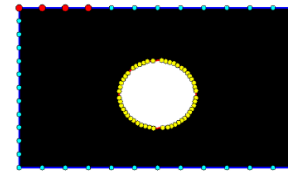


Figura 5. Contorno discretizado. Los puntos azules están posicionados mediante muestreo equidistante. Puntos rojos con la variación equidistante para cierre continuo. Puntos amarillos con muestreo adaptativo por curvatura.

Muestreo equidistante con corrección de cierre

El usuario define una distancia objetivo d_{obj} entre puntos. El sistema calcula el perímetro total P del contorno mediante la Ecuación (3) y determina el número de puntos N mediante la Ecuación (4). Para garantizar el cierre continuo del contorno se recalcula una distancia ajustada d_{real} mediante la Ecuación (5), logrando la distribución correcta de los puntos, señalados de color rojo en la Figura 5.

$$P = \sum_{i=0}^{m-1} \sqrt{(x_{i+1} - x_i)^2 + (y_{i+1} - y_i)^2} \quad (3)$$

$$N = \left(\frac{P}{d_{obj}} \right) \quad (4)$$

$$d_{real} = \frac{P}{N} \quad (5)$$

Adicionalmente, para preservar la geometría en contornos ortogonales, el algoritmo evalúa el ángulo θ entre vectores incidentes en cada píxel mediante la Ecuación (6).

$$\theta = \cos^{-1} \left(\frac{v_1 \cdot v_2}{|v_1| |v_2|} \right) \quad (6)$$

Si el ángulo evidencia un cambio de dirección drástico, ese punto se bloquea como obligatorio, asegurando que la topología original no pierda su forma ortogonal, problema ilustrado en la Figura 6.

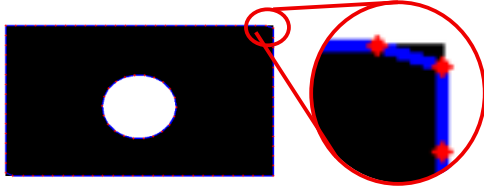


Figura 6. Problema en las esquinas por la distancia equidistante.

Muestreo adaptativo por curvatura

Para contornos con geometría compleja — curvas orgánicas u orificios pequeños — el algoritmo emplea muestreo adaptativo por curvatura. La distancia entre puntos se ajusta dinámicamente mediante la curvatura local κ , concentrando mayor densidad de nodos en zonas de alta curvatura sin saturar las zonas rectas, como describe la Ecuación (7).

$$S = \frac{d_b}{1 + f_c \cdot \kappa} \quad (7)$$

En tramos rectos la curvatura κ es cercana a cero, por lo que $S \approx d_b$, colocando pocos puntos. En esquinas o curvas cerradas, κ aumenta significativamente, reduciendo el espaciado S y concentrando más puntos para capturar fielmente la geometría. El cambio en la distribución de puntos de esta estrategia comparado con el muestreo equidistante se presenta en la Figura 7.

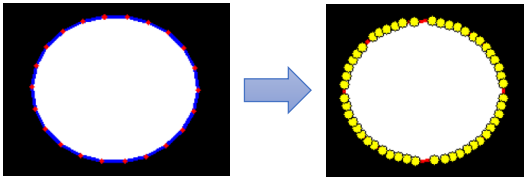


Figura 7. Comparación con el método de distancia equidistante (izquierda) con el método de muestreo adaptativo por curvatura.

2.5 Distribución de puntos internos

Una vez definida la frontera del dominio, el algoritmo emplea una función de generación aleatoria (*random*) para situar los nodos internos. Para garantizar y asegurar la repetibilidad y un comportamiento determinista, se implementó una semilla fija (*valor 42*). De lo contrario, la variación en este parámetro derivaría en distribuciones espaciales distintas en cada ejecución, lo que generaría resultados cualitativamente similares pero carentes de una reproducibilidad exacta. Independientemente del número de la semilla, el uso de una semilla fija garantiza la consistencia total de las pruebas experimentales.

Posteriormente a esto, se verifica su pertenencia mediante la función de trayectoria de Shapely (Gillies et al., 2007). Estos puntos se refinan iterativamente mediante el algoritmo de Lloyd, que desplaza cada punto hacia el centroide de su celda de Voronoi correspondiente, logrando una distribución más uniforme y adecuada para el análisis de elemento finito. La

Figura 8 ilustra la evolución de las celdas de Voronoi durante este proceso.

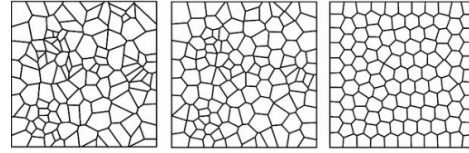


Figura 8. Evolución de las celdas de Voronoi durante cada iteración.

2.6 Triangulación y generación de la malla

Con el conjunto final de puntos — frontera más interiores suavizados — se aplica la triangulación de Delaunay mediante SciPy (Virtanen et al., 2020), obteniendo la conectividad de los elementos triangulares. La triangulación de Delaunay garantiza la calidad de los elementos generados al maximizar el ángulo mínimo de cada triángulo, evitando elementos altamente distorsionados que afectarían la precisión del análisis estructural. La Figura 9 muestra el resultado del mallado triangular sobre el dominio de la Figura 3.

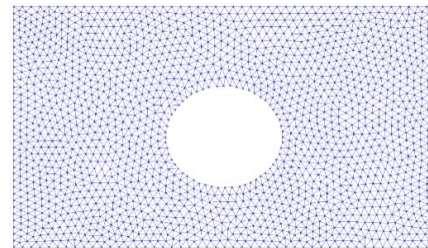


Figura 9. Malla triangular de elementos finitos.

2.7 Interfaz gráfica y exportación JSON

El algoritmo incluye una interfaz gráfica de usuario (GUI) desarrollada en Python que permite al usuario definir interactivamente las condiciones de frontera: selección de nodos fijos y nodos con carga aplicada, como se muestra en la Figura 10. Una vez completada la definición, toda la información de la malla — coordenadas de nodos, conectividad de elementos, cargas y condiciones de frontera — se exporta en formato JSON para su uso en análisis estructurales estáticos por elemento finito o en algoritmos de optimización topológica como SERA.

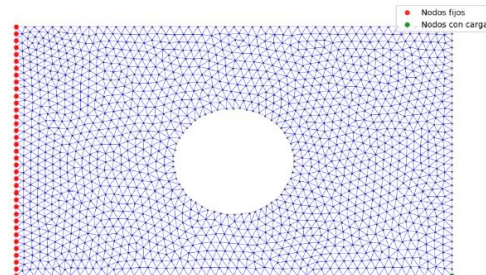


Figura 10. Visualización en tiempo real en la interfaz de los nodos señalados por el usuario. Fijos (rojos) y con carga (verdes).

3 RESULTADOS

3.1 Desempeño computacional del algoritmo de Lloyd

La Figura 11 muestra el comportamiento del tiempo de cómputo en función del número de iteraciones del algoritmo de Lloyd. Se observa un crecimiento exponencial conforme aumenta el número de iteraciones, lo que representa una limitación práctica en la generación de mallas con alta densidad de nodos. Este comportamiento es inherente al cálculo iterativo de los diagramas de Voronoi y debe considerarse al definir el número de iteraciones en función del tamaño del dominio y la precisión requerida. Por lo tanto, para fines de este trabajo, se utilizarán 100 iteraciones de Lloyd en todos los casos de estudio presentados, al demostrar un equilibrio entre tiempo y funcionalidad (Figura 16).

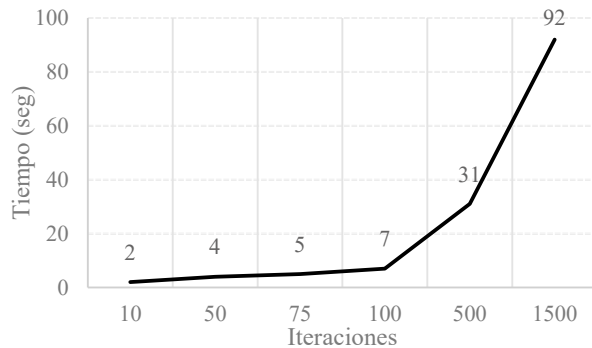


Figura 11. Gráfica que muestra el crecimiento exponencial del tiempo de las iteraciones del algoritmo de Lloyd.

3.2 Validación con el algoritmo SERA

Para validar la integración del algoritmo de mallado con un proceso de optimización topológica, el archivo JSON generado por el algoritmo propuesto se implementó como entrada del algoritmo SERA. Se utilizó como caso de prueba el dominio de la Figura 3 — una viga con orificio central — con condiciones de frontera definidas mediante la GUI: nodos fijos en un extremo y carga aplicada en el extremo opuesto. Donde se estableció una cantidad de 3500 elementos, los cuales se alcanzaron mediante 1638 puntos internos, un umbral de distancia de puntos de 20 y 100 iteraciones de Lloyd.

La Figura 12 muestra el resultado de la optimización topológica obtenida con un radio mínimo de filtro (R_{min}) = 2.5 y una fracción de volumen (V_{frac}) = 0.5. El resultado es coherente con los patrones esperados para las condiciones de carga y frontera definidas, demostrando que la malla triangular generada por el algoritmo propuesto es compatible con el algoritmo SERA y que el archivo JSON contiene correctamente toda la información necesaria para ejecutar el análisis.

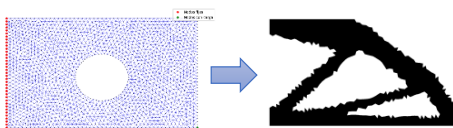


Figura 12. Optimización topológica lograda con la técnica de mallado de este trabajo. R_{min} : 2.5, V_{frac} : 0.5

3.3 Convergencia de la malla

Una vez validada la integración con SERA, se realizó un análisis de convergencia para determinar el número mínimo de elementos necesarios para obtener resultados consistentes. Para ello, se corrió repetidamente el mismo problema de optimización topológica del caso de prueba descrito en la sección 3.2, variando el número de elementos de la malla en cada ejecución. Dado que el objetivo de SERA es minimizar el *compliance*, este parámetro se empleó como indicador de convergencia: se registró el valor final de *compliance* de cada ejecución y se identificó el punto a partir del cual este valor dejó de cambiar significativamente entre ejecuciones.

La Figura 13 muestra los resultados de este análisis. Se observa que a partir de aproximadamente 3000 elementos — señalados por las líneas rojas — el *compliance* final se estabiliza, lo que indica que incrementar el número de elementos más allá de este umbral no mejora la precisión del resultado pero sí aumenta el costo computacional. Este valor de 3000 elementos se establece por tanto como referencia para la generación de mallas en dominios de complejidad similar al caso de prueba. Cabe mencionar que los valores de *compliance* reportados corresponden a una carga de control de -0.001 N aplicada durante las pruebas.

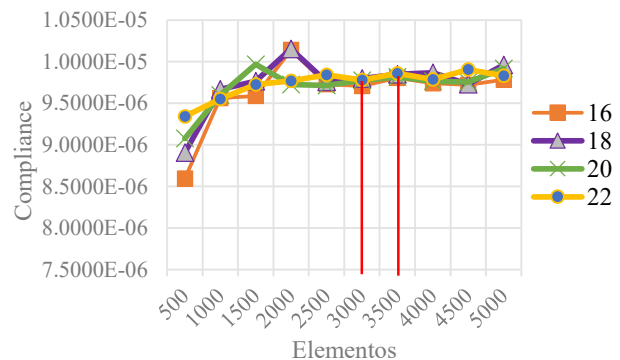


Figura 13. Gráfica que muestra la estabilización en el *compliance* al alcanzar 3000 elementos (líneas rojas). Nota: Los valores del *compliance* son causados debido a una carga -0.001 N durante las pruebas, siendo este un valor de control.

3.4 Casos de estudio

Ya con el algoritmo validado y el número de elementos establecidos, se optó por realizar pruebas con diferentes geometrías para comprobar la robustez del código. Para ello se dibujaron tres tipos de viga y se guardaron como imagen. Estas se muestran en la Figura 14, se buscó específicamente que estas vigas tuvieran geometrías más complejas y condiciones de frontera más específicas que ejemplos anteriores.



Figura 14. Geometrías más complejas dibujadas por el usuario.

Dentro del algoritmo de mallado, se señaló en el código la ruta del archivo y se establecieron los parámetros para la construcción de la malla, donde se establecieron 100 iteraciones de Lloyd — garantizando una buena distribución de los puntos internos — y la cantidad de puntos necesarios para alcanzar los 3500 elementos.

Los resultados del mallado se pueden apreciar en la Figura 15, las cuales demuestran que el algoritmo de mallado es lo suficientemente robusto para detectar correctamente cualquier geometría ingresada.

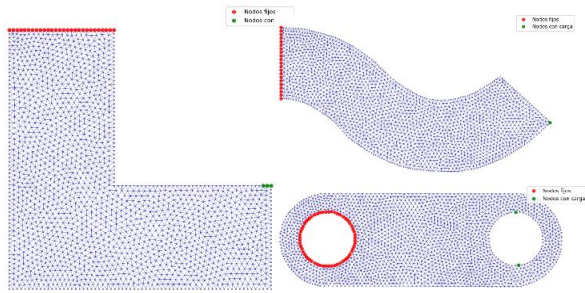
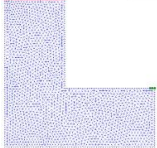
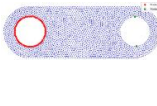
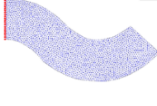


Figura 15. Mallado triangular logrado de las tres geometrías, junto con los nodos de carga (verdes) y fijos (rojos).

Durante las pruebas, se analizaron las mallas generadas en estos casos de estudio con el fin de evaluar la eficacia del algoritmo con base en las métricas estándar de los elementos finitos. Cabe destacar que dicha información se calcula directamente dentro del código; por lo tanto, constituye información de fácil acceso para el usuario. Estas métricas se presentan en la Tabla 1:

Tabla 1. Métricas de las mallas de los casos de estudio.

Malla de los casos de estudio	Métricas de calidad de la malla
	Numero de elementos: 3500 Area promedio de los elementos: 157.244534 Aspect ratio promedio: 1.1721 Aspect ratio maximo: 1.6277 Angulo minimo promedio: 52.80° Angulo minimo: 37.86°
	Numero de elementos: 3500 Area promedio de los elementos: 39.157939 Aspect ratio promedio: 1.1973 Aspect ratio maximo: 1.9484 Angulo minimo promedio: 51.89° Angulo minimo: 28.34°
	Numero de elementos: 3500 Area promedio de los elementos: 64.783457 Aspect ratio promedio: 1.2087 Aspect ratio maximo: 2.2867 Angulo minimo promedio: 51.54° Angulo minimo: 23.67°

En la tabla se puede observar cómo en ningún caso los elementos presentan ángulos menores de 20°, habiendo en promedio ángulos mínimos de 50° —validando la distribución lograda con las 100 iteraciones de Lloyd explicada en la

sección 3.1—. En el caso del *aspect ratio*, se busca que este sea lo más cercano a 1. Durante los casos de estudio no se presentó un *aspect ratio* promedio que superara el 1.20, obteniendo así mallas que están cerca del valor 1 de *aspect ratio*.

Siguiendo con el proceso, con ayuda del algoritmo SERA — señalado en la sección 3.2 — se realizaron las pruebas correspondientes para demostrar el correcto funcionamiento de la carga y lectura del archivo JSON en el algoritmo de optimización topológica. Los resultados obtenidos a raíz de las geometrías mostradas en la Figura 14 fueron las siguientes:



Figura 16. Resultados de Optimización Topológica lograda con la técnica de mallado.

El resultado de la optimización topológica se obtuvo con los mismos parámetros mostrados en la sección 3.2 (radio mínimo de filtro = 2.5 y una fracción de volumen = 0.5). De esta manera se demostró que el algoritmo no solamente malla la geometría de manera correcta, si no que, utiliza satisfactoriamente la información guardada en el JSON dando unos resultados de optimización topológica coherentes a lo señalado en las condiciones de frontera.

3.5 Comparación

A efectos de validar el desempeño y la eficacia del algoritmo propuesto, se llevó a cabo una fase de evaluación experimental utilizando un trabajo anterior basado en el método SERA (Sequential Element Rejection and Admission) con elementos finitos cuadráticos, desarrollado en el Centro de Ingeniería y Desarrollo Industrial (CIDESI).

El propósito de este análisis comparativo consiste en examinar la evolución del *compliance* y la morfología topológica resultante bajo dos esquemas de discretización distintos. Para tal fin, el estudio se centró en casos de estudio basados en vigas bajo condiciones de contorno estándar: viga empotrada (*Fixed*), viga en voladizo (*Left Cantilever*) y voladizo con carga central (*Left Cantilever Middle Load*).

Los parámetros de optimización se fijaron en una densidad de 3500 elementos, con una fracción de volumen objetivo de 0.5, un radio de filtro (*Rmin*) de 2.5 y una magnitud de carga unitaria de 1 N. En cuanto a las propiedades constitutivas del material, integradas mediante el archivo JSON, se definieron un módulo de Young (*E*) de 1.7 GPa y un coeficiente de Poisson (*ν*) de 0.33. Los hallazgos derivados de esta comparativa se detallan en la tabla a continuación:

Tabla 2. Tabla comparativa de los distintos tipos de viga con sus respectivos resultados.

Algoritmo SERA (malla cuadrada)		Algoritmo SERA (malla triangular – Algoritmo propuesto)	
	Viga Fixed Iter = 30 Volfrac = 0.5 Comp = 1.17		Viga Fixed Iter = 29 Volfrac = 0.5 Comp = 1.42
	Left Cantilever Iter = 35 Volfrac = 0.5 Comp = 47.55		Left Cantilever Iter = 29 Volfrac = 0.5 Comp = 14.58
	Middle Load Iter = 30 Volfrac = 0.5 Comp = 44.89		Middle Load Iter = 27 Volfrac = 0.5 Comp = 12.69

Tal como se ilustra en la Tabla 2, el comportamiento del *compliance* derivado del algoritmo propuesto muestra una tendencia decreciente en comparación con el esquema de malla cuadrática convencional. Específicamente, se cuantificaron reducciones significativas del 30.66% para el caso *Left Cantilever* y del 28.27% en la configuración *Middle Load*. Adicionalmente, el proceso de optimización exhibió una mayor eficiencia en la convergencia, logrando una disminución en el número de iteraciones necesarias en cada uno de los casos de estudio analizados.

Es importante acotar que, si bien el flujo de trabajo propuesto requiere una configuración manual de las condiciones de carga y la densidad de elementos —lo que podría incrementar marginalmente el tiempo de pre-procesamiento total—, el tiempo de ejecución por iteración dentro del algoritmo SERA resulta sensiblemente inferior al observado en mallas cuadráticas. Se reconoce, no obstante, la pertinencia de realizar estudios estadísticos de mayor profundidad para determinar un porcentaje de eficiencia efectivo con mayor precisión en el tiempo.

La reproducibilidad de los resultados experimentales expuestos en las secciones 3.4 y 3.5 está garantizada mediante la implementación de una semilla fija, tal como se fundamentó en la sección 2.5. La integridad de la información necesaria para la reconstrucción de las topologías obtenidas se encuentra centralizada en el archivo con formato JSON, el cual encapsula la totalidad de parámetros y condiciones de contorno empleadas. Para información detallada de la arquitectura lógica del software, se remite al lector a la sección final del Apéndice A.

4. DISCUSIÓN

El algoritmo propuesto demostró ser capaz de generar mallas triangulares de elementos finitos sobre dominios arbitrarios definidos por imagen, superando la limitación principal de los enfoques convencionales que restringen el análisis a dominios rectangulares con mallas estructuradas de elementos cuadriláteros (Zhuang et al., 2023). La eliminación del efecto de escalonado en los bordes del dominio, ilustrada en la comparación entre las Figuras 1 y 9, representa una mejora

directa en la fidelidad geométrica de la representación, lo que se traduce en un análisis estructural más cercano a la geometría real de la pieza.

La integración de dos estrategias de muestreo — equidistante con corrección de cierre y adaptativo por curvatura — permite al algoritmo manejar tanto geometrías simples como contornos complejos con orificios internos, sin necesidad de intervención manual adicional por parte del usuario. Esta característica lo distingue de implementaciones más simples que aplican una única estrategia de discretización independientemente de la complejidad geométrica del dominio (Islam & Nayem, 2024).

El uso del algoritmo de Lloyd para la distribución interna de nodos garantiza una mayor uniformidad en el tamaño de los elementos triangulares, lo cual es relevante para la precisión del análisis de elemento finito. Sin embargo, el crecimiento exponencial del tiempo de cómputo observado en la Figura 11 representa una limitación práctica que debe considerarse al escalar el algoritmo a dominios más grandes o con mayor densidad de elementos. En este sentido, el análisis de convergencia presentado en la sección 3.3 proporciona una referencia útil: con aproximadamente 3000 elementos el *compliance* se estabiliza, lo que permite al usuario definir un criterio objetivo para la generación de la malla sin incurrir en costos computacionales innecesarios.

La validación con el algoritmo SERA demostró que el archivo JSON generado contiene correctamente toda la información necesaria para ejecutar un análisis de optimización topológica, produciendo resultados coherentes con los esperados para las condiciones de carga y frontera definidas. Esto confirma que el formato de exportación propuesto es funcional y extensible a otros algoritmos de análisis estructural estático por elemento finito, no limitándose exclusivamente a SERA.

Con base en lo anteriormente expuesto, se establece en la Tabla 3 la información técnica de la malla y los parámetros obtenidos en la optimización topológica final de los cuatro casos de estudio presentados.

Tabla 3. Tabla de desempeño global de los casos de estudio presentados.

Caso de estudio	Métrica de la malla	Parámetros de la OT
	Número de elementos: 3500 Iteraciones de Lloyd: 100 Carga: 1 N	Rmin: 2.5 Volfrac: 0.5 Comp: 28.13 Iter: 29
		Rmin: 2.5 Volfrac: 0.5 Comp: 73.63 Iter: 29
		Rmin: 2.5 Volfrac: 0.5 Comp: 184.67 Iter: 30
		Rmin: 2.5 Volfrac: 0.5 Comp: 122.83 Iter: 28



Dicha tabla tiene como fin el proporcionar una comparativa cuantitativa que valide la eficiencia del algoritmo propuesto tanto en la calidad de la discretización como en la convergencia de la solución estructural.

En comparación con herramientas comerciales como SolidWorks (Dassault Systèmes, 2023), que también ofrecen capacidades de mallado y optimización topológica, el algoritmo propuesto presenta la ventaja de ser completamente de código abierto, personalizable y ejecutable sin licencias comerciales, lo que lo hace accesible para entornos académicos y de investigación. Su desarrollo en Python y funcionamiento modular facilita además su integración con otras bibliotecas científicas y su adaptación a flujos de trabajo específicos (Wang et al., 2023).

Como limitación principal, el algoritmo trabaja actualmente en dominios bidimensionales, lo que restringe su aplicación a problemas planos. Asimismo, el tamaño de elemento en la malla es uniforme — determinado por la distancia entre nodos de frontera — sin posibilidad de refinamiento local en zonas de alto gradiente de esfuerzos, característica presente en malladores más avanzados (Li et al., 2022).

5. CONCLUSIÓN

Se desarrolló un algoritmo 2D en Python para la generación automática de mallas triangulares de elementos finitos sobre dominios definidos por imagen, integrando detección de contornos, muestreo equidistante con corrección de cierre, muestreo adaptativo por curvatura, algoritmo de Lloyd y triangulación de Delaunay. El algoritmo demostró ser capaz de trabajar con geometrías arbitrarias, superando la limitación de los enfoques convencionales basados en dominios rectangulares y eliminando el efecto de escalonado en los bordes del dominio.

La validación con el algoritmo SERA confirmó que el archivo JSON generado contiene correctamente toda la información necesaria para ejecutar análisis estructurales estáticos por elemento finito, siendo la optimización topológica una de sus aplicaciones principales. El análisis de convergencia estableció que a partir de aproximadamente 3000 elementos el *compliance* se estabiliza, proporcionando un criterio objetivo para la definición del tamaño de malla en dominios de complejidad similar al caso de prueba.

El crecimiento exponencial del tiempo de cómputo del algoritmo de Lloyd representa la principal limitación del trabajo actual, restringiendo su aplicación práctica a dominios de tamaño moderado. Adicionalmente, el algoritmo opera exclusivamente en dominios bidimensionales y no contempla refinamiento local de malla en zonas de alto gradiente de esfuerzos.

Como trabajo futuro se propone abordar estas limitaciones mediante las siguientes líneas de desarrollo:

- Optimización del algoritmo de Lloyd para reducir su costo computacional, explorando estrategias de convergencia anticipada o implementaciones paralelizadas.

- Incorporación de refinamiento local de malla en zonas de alto gradiente de esfuerzos, mejorando la precisión del análisis sin incrementar innecesariamente el número total de elementos.
- Extensión del algoritmo a dominios tridimensionales, ampliando su aplicabilidad a problemas de ingeniería más representativos.
- Validación del algoritmo con casos de prueba adicionales de mayor complejidad geométrica y distintas condiciones de carga y frontera.
- Integración directa con otros algoritmos de optimización topológica además de SERA, aprovechando la flexibilidad del formato JSON como interfaz de comunicación.

REFERENCIAS

- Bendsøe, M.P. & Kikuchi, N. (1988). Generating optimal topologies in structural design using a homogenization method. *Computer Methods in Applied Mechanics and Engineering*, 71(2), 197–224.
- Bradski, G. (2000). The OpenCV Library. *Dr. Dobbs's Journal of Software Tools*. Versión utilizada: 4.2.0.34.
- Dassault Systèmes. (2023). *SolidWorks*. Dassault Systèmes, Vélizy-Villacoublay, Francia. Disponible en: <https://www.solidworks.com>
- Gillies, S., et al. (2007). *Shapely: Manipulation and analysis of geometric objects in the Cartesian plane* [Software de computación]. PyPI. <https://github.com/shapely/shapely>
- Islam, M.S. & Nayem, A.Z. (2024). Finite element-based optimization procedure for an irregular domain with unstructured mesh. *Heliyon*, 10(4), e25994. <https://doi.org/10.1016/j.heliyon.2024.e25994>
- Li, J., Gao, L., Ye, M., Li, H. & Li, L. (2022). Topology optimization of irregular flow domain by parametric level set method in unstructured mesh. *Journal of Computational Design and Engineering*, 9(1), 100–113. <https://doi.org/10.1093/jcde/qwab071>
- Python Software Foundation. (2025). *Python*, versión 3.13. Disponible en: <https://www.python.org>
- Suzuki, S. & Abe, K. (1985). Topological structural analysis of digitized binary images by border following. *Computer Vision, Graphics, and Image Processing*, 30(1), 32–46.
- Virtanen, P. et al. (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, 17, 261–272. <https://doi.org/10.1038/s41592-019-0686-2>
- Wang, Y., Li, X., Long, K. & Wei, P. (2023). Open-source codes of topology optimization: A summary for beginners to start their research. *Computer Modeling in Engineering & Sciences*, 137(1), 1–34. <https://doi.org/10.32604/cmescs.2023.027603>



Zhuang, Z., Xie, Y.M., Li, Q. & Zhou, S. (2023). A 172-line Matlab code for structural topology optimization in the body-fitted mesh. *Structural and Multidisciplinary Optimization*, 66, 11. <https://doi.org/10.1007/s00158-022-03464-x>

Apéndice A. CÓDIGO FUENTE PARA LA RELAJACIÓN DE LLOYD Y TRIANGULACIÓN DE DELAUNAY

Como detalle complementario, se presenta secciones del código explicadas durante este trabajo, secciones cuales fueron ligeramente condensadas en su estructura.

Tras un muestreo aleatorio inicial mediante técnica de rechazo, los nodos internos se someten a un proceso iterativo de suavizado. En cada ciclo, se calcula la región de Voronoi asociada a cada nodo y se realiza una intersección con el dominio Ω donde el nodo es desplazado al centroide de la región resultante.

```
def generar_puntos_voronoi(puntos_contorno_fijos,
    num_puntos_internos, dominio_poligono,
    iteraciones_lloyd=10, seed=42):

    if dominio_poligono is None or
    dominio_poligono.is_empty: return np.array([])

    minx, miny, maxx, maxy = dominio_poligono.bounds

    puntos_internos = []

    np.random.seed(seed)

    while len(puntos_internos) < num_puntos_internos:

        x, y = np.random.uniform(minx, maxx),
        np.random.uniform(miny, maxy)

        if dominio_poligono.contains(Point(x, y)):

            puntos_internos.append([x, y])

    puntos_internos = np.array(puntos_internos)

    for _ in range(iteraciones_lloyd):

        todos = np.vstack([puntos_contorno_fijos,
            puntos_internos])

        try:

            vor = Voronoi(todos)

            nuevos = []

            for i in range(len(puntos_internos)):

                idx = i + len(puntos_contorno_fijos)

                if idx in vor.point_region and
                vor.point_region[idx] != -1:

                    vertices =
                    vor.vertices[vor.regions[vor.point_region[idx]]]

                    recorte =
                    ShapelyPolygon(vertices).intersection(dominio_poligono)

                    nuevos.append([recorte.centroid.x,
                        recorte.centroid.y] if not recorte.is_empty else
                    puntos_internos[i])
```

```
        else: nuevos.append(puntos_internos[i])

    puntos_internos = np.array(nuevos)

    except: break

    return puntos_internos
```

Durante la fase parcialmente final de esta etapa, a partir de la función *crear_malla_delaunay*, se establece la conectividad del mallado logrado mediante la Triangulación de Delaunay. Se evalúa el centroide de cada simplex y se mantienen únicamente aquellos que estén contenidos en Ω . De esta manera el código asegura que la malla final represente con precisión la geometría analizada.

```
def crear_malla_delaunay(puntos_contorno_fijos,
    puntos_internos, dominio_poligono):

    if len(puntos_contorno_fijos) == 0: return None, []

    todos = np.vstack([puntos_contorno_fijos,
        puntos_internos])

    try:

        tri = Delaunay(todos)

        validos = [s for s in tri.simplices if
            dominio_poligono.contains(Point(np.mean(todos[s],
                axis=0)))]

        return tri, validos

    except: return None, []

def visualizar_malla_final(puntos_contorno_fijos,
    puntos_internos, triangulos_validos, imagen_original, ax):

    if imagen_original is not None:

        ax.imshow(cv2.cvtColor(imagen_original,
            cv2.COLOR_BGR2RGB), origin='upper',

            extent=[0, imagen_original.shape[1], 0,
                imagen_original.shape[0]])

        todos = np.vstack([puntos_contorno_fijos,
            puntos_internos])

        for t in triangulos_validos:

            ax.add_patch(plt.Polygon(todos[t], fill=None,
                edgecolor='green', linewidth=0.8, alpha=0.8))

            ax.plot(todos[:, 0], todos[:, 1], 'ro', markersize=1)

            ax.set_aspect('equal')

            ax.axis('off')
```

El código fuente íntegro y sin condensar, que comprende todas las funciones de los pasos revisados en este trabajo, está disponible bajo petición formal para fines de revisión académica. Asimismo, se informa que se está trabajando activamente en la creación de un repositorio público (GitHub) para facilitar la reproducibilidad de los resultados y fomentar la colaboración dentro de la comunidad de investigación en optimización estructural.

Mechanical Design of an ASRS Prototype with Artificial Vision as a Learning Resource for Industrial Engineering

GF Pérez-García* SM Marquez-Marina*
KM Velázquez-Nataren* JA Fragoso-Mandujano*
EN Escobar-Gómez* MC Salgado-Gutiérrez*

* *Tecnológico Nacional de México / Instituto Tecnológico de Tuxtla
Gutiérrez, Tuxtla Gutiérrez, Chiapas, México.*

Abstract

This work presents the mechanical design of a didactic-scale Automated Storage and Retrieval System (ASRS) prototype intended to support the training of Industrial Engineering students. The project focuses on the CAD modeling and assembly of the main mechanical elements of the system, including the structural base, rack, storage containers, guide elements, transmission components, and positioning mechanism. The objective is to provide a physical and visual learning resource that allows students to analyze ASRS principles from a structural, dimensional, and functional perspective.

The proposed design was developed in SolidWorks and evaluated through digital assembly review, dimensional compatibility analysis, and verification of the relationship between the rack, containers, guide elements, and motion components. In addition, preliminary design criteria were defined to support future manufacturing, automation, and integration of control and artificial vision modules. The prototype is not intended to compete with industrial ASRS solutions, but to serve as a scalable educational platform for understanding storage organization, mechanical configuration, component integration, and future automation stages. The results provide a mechanical basis for subsequent implementation using Python, ESP32-based control, and computer vision.

Keywords: ASRS, mechanical design, SolidWorks, Industrial Engineering, didactic prototype, ESP32, computer vision

1. INTRODUCCIÓN

Los sistemas automáticos de almacenamiento y recuperación, conocidos como ASRS, se han convertido en una referencia importante en logística, manufactura y manejo de materiales. De acuerdo con el informe anual de MHI de 2025, el 74 % de los líderes de cadena de suministro está aumentando su inversión en tecnología e innovación, y el 90 % planea invertir más de un millón de dólares en soluciones para su operación [MHI, 2025]. A esto se suma que la presión por automatizar también está relacionada con el problema de personal: Descartes reportó que el 76 % de los responsables de operaciones logísticas percibe escasez relevante de fuerza laboral [Descartes Systems Group, 2024b], y otro estudio del mismo grupo señala que el 54 % está priorizando la automatización de tareas repetitivas para mejorar la productividad [Descartes Systems Group, 2024a]. En este contexto, comprender el funcionamiento de un ASRS deja de ser un tema opcional y se vuelve parte de la formación técnica de quienes trabajarán en operaciones, logística, manufactura y mejora de procesos.

Desde el punto de vista operativo, los ASRS ofrecen ventajas ampliamente reconocidas: mejor aprovechamiento del espacio, reducción de recorridos, mayor control de inventario, menor error de surtido y una operación más ordenada bajo esquemas *goods-to-person*. También pueden favorecer mejores condiciones ergonómicas al acercar el material al operador y reducir desplazamientos innecesarios. Sin embargo, estos sistemas también presentan restricciones importantes: requieren inversión inicial, integración con software, una adecuada definición del layout y, sobre todo, una base mecánica bien resuelta. Si estas condiciones no se atienden correctamente, el sistema pierde eficiencia o se vuelve difícil de mantener [Kardex, 2022]. Por ello, estudiar un ASRS no solo implica conocer sus ventajas operativas, sino también comprender los fundamentos estructurales y mecánicos que hacen posible su funcionamiento.

Para un estudiante de Ingeniería Industrial, el estudio de un ASRS resulta especialmente valioso porque en un mismo sistema convergen distintos temas relevantes de la formación profesional. Permite analizar layout y aprovechamiento del espacio, flujo y manejo de materiales, organización de inventarios, diseño de estaciones, ergonomía, integración mecánica, automatización y planeación de mejoras. También ayuda a comprender que la automatización no inicia únicamente en el software o en el control

* Mechanical Design of an ASRS Prototype with Artificial Vision as a Learning Resource for Industrial Engineering



electrónico, sino en una estructura física cuya geometría, configuración y compatibilidad entre componentes deben estar correctamente definidas desde la etapa de diseño. Si el rack está mal planteado, si los contenedores no se ajustan adecuadamente o si la relación dimensional entre los elementos no es consistente, cualquier etapa posterior de automatización presentará limitaciones.

Con base en esta necesidad, el objetivo general de este trabajo es desarrollar el diseño mecánico de un prototipo ASRS de escala didáctica para apoyar la enseñanza de sistemas automatizados de almacenamiento y recuperación en Ingeniería Industrial. El alcance se centra en la etapa mecánica del proyecto, por lo que incluye la definición de la arquitectura general, el modelado de componentes y el ensamble digital en SolidWorks, antes de avanzar hacia fases posteriores de manufactura, control electrónico, sensores o visión artificial.

De manera específica, el trabajo busca definir la configuración general del prototipo, modelar sus componentes principales, establecer criterios dimensionales para el rack, los contenedores, las guías y los elementos de movimiento, y dejar una plataforma preparada para futuras etapas de integración tecnológica. El aporte no se plantea como una solución industrial, sino como un recurso didáctico escalable que permite relacionar conceptos de *layout*, almacenamiento, movimiento, compatibilidad dimensional y diseño mecánico dentro de un mismo sistema físico.

2. ESTADO DEL ARTE

Los sistemas automáticos de almacenamiento y recuperación han evolucionado desde configuraciones rígidas de carga unitaria hasta soluciones modulares, compactas y conectadas con software. Su aplicación principal se relaciona con el aumento de densidad de almacenamiento, la reducción de recorridos manuales, la mejora en precisión de surtido y la operación bajo esquemas *goods-to-person* [Aslam et al., 2025, SSI Schaefer, 2026, Kardex, 2025, Honeywell, 2026]. Sin embargo, su implementación requiere inversión inicial, planeación del *layout*, integración con sistemas de control y una selección adecuada de la tecnología según el tipo de producto, el volumen de operación y la capacidad de mantenimiento [Honeywell Intelligranted, 2021, Aslam et al., 2025].

Dentro de las soluciones comerciales, los sistemas de carga unitaria y *miniloat* continúan utilizándose en aplicaciones donde se requiere almacenamiento vertical, manejo repetitivo y operación con pasillos definidos [Dematic, 2026, SSI Schaefer, 2026, Honeywell, 2026]. En operaciones con menor superficie disponible, los módulos verticales de almacenamiento, como los *vertical lift modules*, permiten concentrar material en altura y entregar bandejas al operador en una estación ergonómica [Kardex, 2025, 2026]. Por otro lado, los sistemas *shuttle* y las configuraciones modulares para cajas o charolas ofrecen mayor velocidad y escalabilidad, aunque demandan una integración más cuidadosa entre mecánica, control y software [Honeywell, 2026, Dematic, 2026, SSI Schaefer, 2026].

Más recientemente, los sistemas de almacenamiento cúbico y robótica móvil han ampliado las alternativas de automatización. AutoStore utiliza una rejilla compacta con

contenedores apilados y robots que operan sobre la parte superior del sistema [AutoStore, 2025b,a]. En paralelo, soluciones *shelf-to-person*, como las de Geek+, emplean robots móviles para trasladar estanterías o unidades de almacenamiento hacia estaciones de trabajo [Geek+, 2026]. Estas tecnologías muestran una tendencia hacia arquitecturas más flexibles y escalables, pero también incrementan la dependencia de software, integración especializada y soporte técnico.

Desde el punto de vista educativo, esta evolución plantea una limitación práctica. Las soluciones comerciales permiten estudiar aplicaciones reales de alto desempeño, pero no siempre facilitan comprender desde cero la relación entre estructura, rack, contenedores, guías, sistema de movimiento y lógica de almacenamiento. Además, su costo y dependencia tecnológica pueden dificultar su adopción como recurso de laboratorio en programas de Ingeniería Industrial.

En este contexto, los prototipos didácticos de escala reducida permiten estudiar los principios básicos de un ASRS sin reproducir la complejidad completa de una solución industrial. Su utilidad se encuentra en representar físicamente la organización de posiciones, la compatibilidad dimensional, el movimiento entre ejes y la integración gradual con control, sensores o visión artificial. Por ello, el diseño mecánico de un prototipo académico constituye una etapa necesaria antes de avanzar hacia pruebas de automatización, reconocimiento visual o validación experimental.

3. METODOLOGÍA

La metodología se enfocó en desarrollar el diseño mecánico de un prototipo ASRS de escala didáctica, partiendo de los requerimientos de enseñanza y de las restricciones propias de un sistema de laboratorio. El procedimiento se organizó en cinco etapas: definición de requerimientos, selección de arquitectura, dimensionamiento preliminar, modelado y ensamble en SolidWorks, y revisión técnica del ensamble CAD.

3.1 Definición de requerimientos del prototipo

El prototipo se planteó como un sistema didáctico para representar operaciones básicas de almacenamiento y recuperación de contenedores. A diferencia de un ASRS industrial, el objetivo no fue maximizar capacidad, velocidad o densidad de almacenamiento, sino construir una plataforma clara para explicar la relación entre estructura, movimiento, posiciones de almacenamiento y futura automatización.

A partir de este enfoque, los requerimientos de diseño se definieron considerando la función mecánica y didáctica del prototipo. El sistema debía representar una estructura tipo rack con posiciones de almacenamiento visibles, utilizar contenedores de tamaño reducido compatibles con impresión 3D e integrar una base estructural capaz de soportar el conjunto de movimiento. También se consideró necesario incorporar guías lineales, soportes y elementos de transmisión para permitir desplazamientos horizontales y verticales. Como base de accionamiento se propuso el uso de motores paso a paso, manteniendo una configuración



abierta para futuras etapas de control mediante ESP32, interfaz en Python y visión artificial.

Estos requerimientos permitieron delimitar el diseño mecánico antes del modelado CAD. También sirvieron para evitar una propuesta sobredimensionada respecto a su finalidad académica.

3.2 Selección de arquitectura mecánica

Con base en los requerimientos anteriores, se seleccionó una arquitectura tipo rack con nueve posiciones de almacenamiento distribuidas en una configuración 3×3 . Esta elección permite representar de manera sencilla la lógica de localización de posiciones, el desplazamiento hacia una columna y el posicionamiento en un nivel determinado.

El sistema se dividió en cuatro subsistemas mecánicos principales:

- **Base estructural:** elemento de soporte del conjunto y referencia para la instalación de guías, motores y soportes.
- **Sistema de movimiento:** conjunto formado por varillas guía, soportes SK8, poleas dentadas y motores Nema 17.
- **Rack de almacenamiento:** estructura destinada a organizar las posiciones donde se colocan los contenedores.
- **Contenedores y plataforma:** elementos manipulados por el sistema y superficie de apoyo para operaciones de almacenamiento y recuperación.

Esta división permitió modelar cada parte de forma independiente y posteriormente integrarla en un ensamble general. La secuencia de funcionamiento considerada para el prototipo se muestra en la Fig. 1.

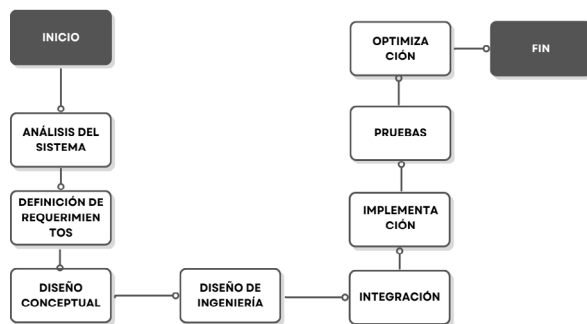


Figura 1. Diagrama general de funcionamiento del sistema ASRS.

3.3 Criterios de diseño y dimensionamiento

El dimensionamiento del prototipo se realizó considerando compatibilidad entre piezas, facilidad de fabricación y claridad visual para su uso en prácticas académicas. Para el rack se definieron dimensiones generales de $60 \times 24 \times 40$ cm,

con capacidad para nueve contenedores. Cada contenedor se propuso con dimensiones aproximadas de $15 \times 10 \times 7$ cm, de modo que pudiera fabricarse mediante impresión 3D y manipularse con facilidad dentro del sistema.

Para el sistema de guiado se consideraron varillas de acero de 8 mm de diámetro, montadas sobre soportes SK8. Esta selección permite representar un sistema de desplazamiento lineal sencillo y de bajo costo. Para el accionamiento se consideraron motores paso a paso Nema 17, ya que su eje de 5 mm es compatible con poleas dentadas comerciales y permite una futura implementación de control por posición.

Los criterios aplicados durante esta etapa se resumen en la Tabla 1.

Tabla 1. Criterios de diseño considerados en el prototipo ASRS

Criterio	Aplicación en el diseño
Estabilidad estructural	Uso de una base rígida y elementos de soporte para mantener alineación durante el desplazamiento.
Compatibilidad dimensional	Relación entre rack, contenedores, plataforma, guías, poleas y motores dentro del ensamble CAD.
Claridad didáctica	Distribución visible de posiciones para explicar almacenamiento, recuperación y movimiento entre ejes.
Factibilidad de fabricación	Uso de piezas compatibles con impresión 3D, acrílico, perfiles comerciales y componentes mecánicos disponibles.
Escalabilidad	Configuración preparada para integrar posteriormente control electrónico, interfaz en Python y visión artificial.
Mantenimiento y ajuste	Selección de elementos desmontables que permitan modificar o sustituir piezas durante prácticas de laboratorio.

3.4 Modelado y ensamble en SolidWorks

El modelado se realizó en SolidWorks a partir de los subsistemas definidos. Primero se modelaron los componentes individuales: base, rack, contenedores, soportes SK8, varillas guía, motores, poleas y plataforma. Posteriormente, los elementos se integraron en un ensamble general para revisar su relación geométrica.

El uso del entorno CAD permitió verificar la posición relativa de las piezas, la correspondencia entre ejes, el espacio disponible para el desplazamiento y la ubicación de los contenedores dentro del rack. Esta etapa fue importante porque cualquier error dimensional en el rack, la plataforma o el sistema de guías podría afectar la futura automatización del prototipo.

La Tabla 2 presenta los componentes principales considerados en el diseño.

3.5 Ensamble CAD

Una vez integrado el ensamble general, se realizó una revisión preliminar del modelo en SolidWorks. Esta revisión se centró en cuatro aspectos: compatibilidad rack-contenedor, alineación del sistema de guías, integración motor-polea y disposición general de la plataforma.

Tabla 2. Componentes principales del prototipo ASRS

Componente	Material o referencia	Dimensión principal
Base estructural	Lámina tipo Lamitec	Perfiles de 450 mm de largo y 73.2 mm de ancho.
Soportes SK8	Aluminio	Compatibles con varillas de 8 mm.
Varillas guía	Acero	Diámetro de 8 mm.
Motores Nema 17	Motor paso a paso	Eje de 5 mm y longitud de 20 a 21 mm.
Poleas dentadas	Polea para eje de 5 mm	Diámetro exterior aproximado de 12 a 13 mm.
Plataforma	Acrílico	Propuesta para corte láser.
Rack	Estructura CAD	60 × 24 × 40 cm.
Contenedores	Impresión 3D	15 × 10 × 7 cm.

La compatibilidad rack-contenedor se revisó comparando las dimensiones del contenedor con los espacios disponibles en el rack. La alineación del sistema de guías se verificó mediante la ubicación de las varillas de 8 mm respecto a los soportes SK8. La integración motor-polea se revisó considerando la compatibilidad del eje de 5 mm del motor Nema 17 con las poleas seleccionadas. Finalmente, la disposición de la plataforma se revisó para comprobar que pudiera funcionar como elemento de apoyo durante las operaciones de almacenamiento y recuperación.

Esta propuesta permite detectar problemas geométricos antes de fabricar el prototipo, ayuda a identificar interferencias, falta de holgura, desalineaciones o incompatibilidad entre piezas. Con ello, el diseño queda preparado para una siguiente etapa de manufactura física y validación mediante pruebas de desplazamiento, repetibilidad y control.

4. RESULTADOS

Los resultados obtenidos corresponden al diseño mecánico y al ensamble CAD del prototipo ASRS. Se presentan en cuatro apartados: configuración general del sistema, sistema de movimiento, integración rack-contenedor y criterios de evaluación preliminar del ensamble.

4.1 Configuración mecánica general

El prototipo quedó conformado por una base estructural, un sistema de guías lineales, una estructura vertical, una plataforma móvil, un rack de almacenamiento y contenedores. La base se diseñó como elemento de soporte del conjunto y como referencia para la ubicación de guías, soportes y elementos de movimiento. Para esta parte se consideraron perfiles de lámina tipo Lamitec con dimensiones aproximadas de 450 mm de longitud y 73.2 mm de ancho, como se muestra en la Fig. 2.

El guiado lineal se planteó mediante varillas de acero de 8 mm de diámetro y ocho soportes tipo SK8. Cada soporte presenta una base aproximada de 14.4 mm × 41.8 mm y una altura total de 32.8 mm. La Fig. 3 muestra el soporte modelado y la Fig. 4 muestra su integración dentro del ensamble. Esta configuración permite establecer una referencia mecánica para el desplazamiento de los componentes móviles y facilita la explicación de alineación, soporte y guiado dentro de una práctica de laboratorio.

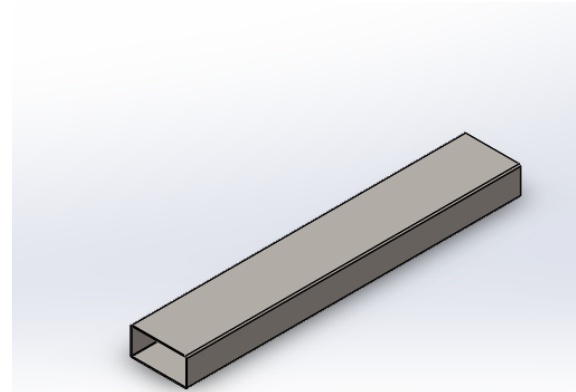


Figura 2. Modelo de la base estructural propuesta para el sistema ASRS.

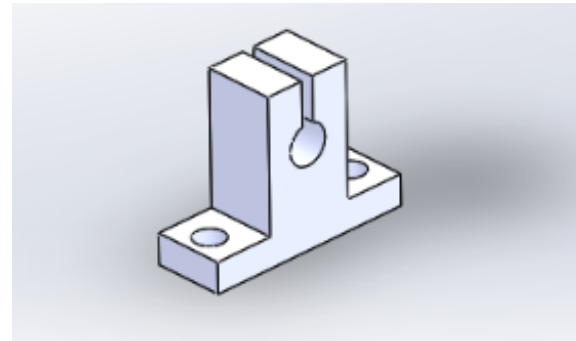


Figura 3. Modelo del soporte tipo SK8 para guía lineal de 8 mm.

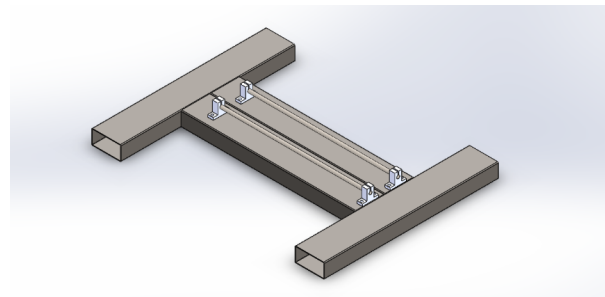


Figura 4. Integración de varillas de acero al sistema de guía lineal dentro del ensamble digital.

4.2 Sistema de movimiento

El sistema de movimiento se planteó con tres motores paso a paso Nema 17: dos para el desplazamiento horizontal en el eje X y uno para el movimiento vertical. Esta configuración permite representar las dos acciones básicas de un ASRS: desplazarse hacia una columna de almacenamiento y posicionarse en el nivel correspondiente. El motor considerado cuenta con eje de 5 mm de diámetro y longitud aproximada de 20 a 21 mm, como se observa en la Fig. 5.

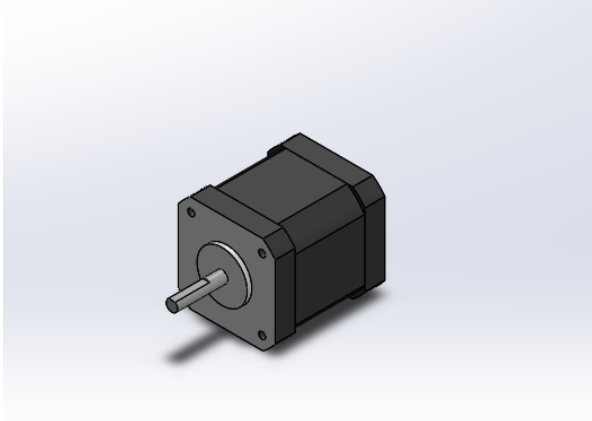


Figura 5. Motor a pasos Nema 17 considerado en el diseño del prototipo.

Para la transmisión se consideraron poleas dentadas compatibles con eje de 5 mm, con diámetro exterior aproximado de 12 a 13 mm y altura de 15 a 16 mm. La Fig. 6 presenta la polea modelada. La selección final del conjunto polea-banda deberá definirse en una etapa posterior, ya que de ella dependerá la resolución de posicionamiento del sistema.

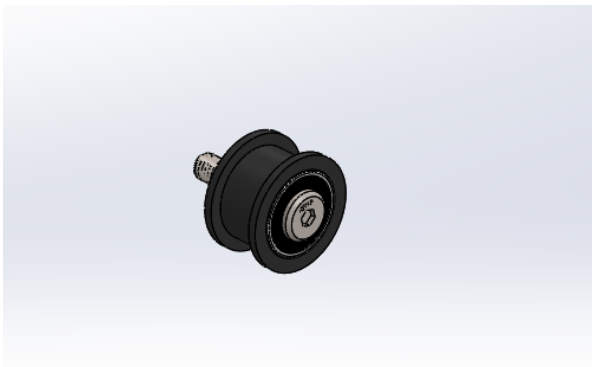


Figura 6. Polea dentada de eje de 5 mm considerada en el diseño del sistema.

Como criterio preliminar, la resolución lineal del sistema de transmisión puede estimarse mediante:

$$\Delta x = \frac{pN_d}{N_s} \quad (1)$$

donde Δx representa el desplazamiento lineal por paso del motor, p el paso de la banda, N_d el número de dientes de la polea y N_s el número de pasos por revolución del motor. Esta relación no se usa todavía para reportar un valor final, porque falta seleccionar la banda definitiva; sin embargo, deja establecido el criterio que deberá emplearse para validar la precisión de posicionamiento.

La estructura vertical se integró como soporte del desplazamiento en altura y como elemento de unión entre la base, el sistema de transmisión y la plataforma. La Fig. 7 muestra la configuración del eje vertical, mientras que la Fig. 8 presenta el ensamble general del sistema.

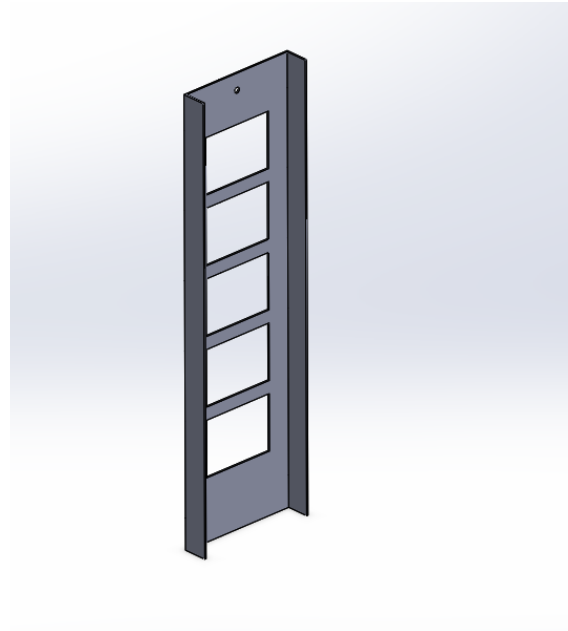


Figura 7. Configuración estructural del eje vertical dentro del diseño del sistema.

La plataforma de almacenamiento, mostrada en la Fig. 9, se propuso como una pieza ligera para apoyar el manejo de contenedores. Su fabricación futura puede realizarse en acrílico cortado con láser, lo que facilita ajustes dimensionales y reemplazo de la pieza durante prácticas de laboratorio.

4.3 Integración del rack y contenedores

El rack se definió con dimensiones de $60 \times 24 \times 40$ cm y capacidad para nueve contenedores, organizados en una configuración 3×3 . Esta distribución permite representar operaciones básicas de localización, almacenamiento y recuperación sin aumentar la complejidad mecánica del prototipo. La Fig. 10 muestra el modelo del rack propuesto.

Los contenedores se modelaron con dimensiones aproximadas de $15 \times 10 \times 7$ cm, considerando compatibilidad con el rack y fabricación mediante impresión 3D. El modelo se muestra en la Fig. 11.

La capacidad nominal del rack puede expresarse como:

$$N = n_f n_c \quad (2)$$

donde N representa el número total de posiciones, n_f el número de filas y n_c el número de columnas. Para la configuración propuesta:

$$N = 3(3) = 9 \quad (3)$$

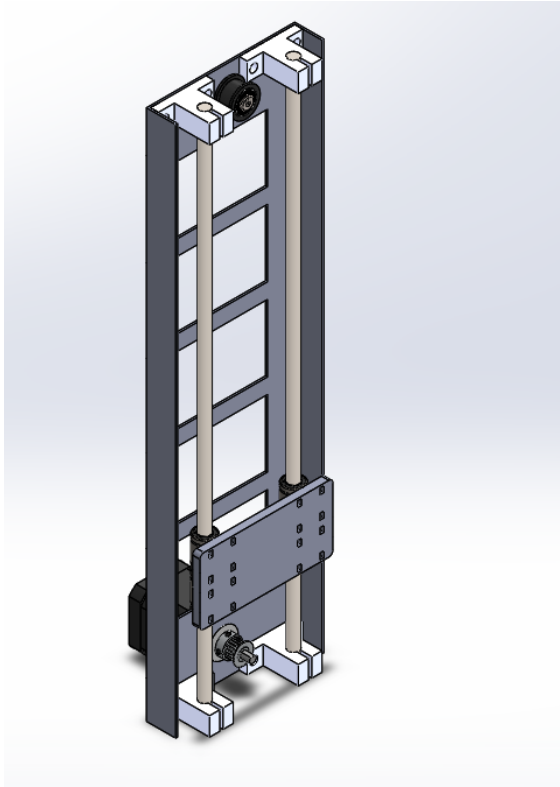


Figura 8. Ensamble general del sistema desarrollado en SolidWorks.

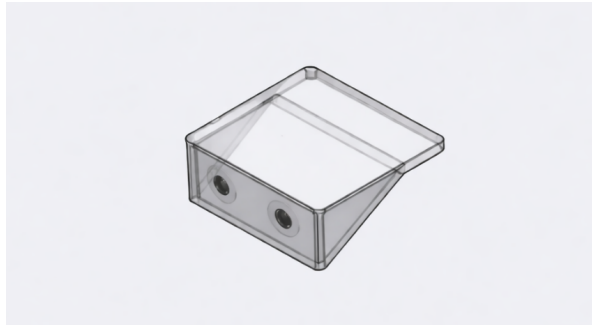


Figura 9. Plataforma propuesta para el almacenamiento y recuperación de contenedores.

También se estimó una relación volumétrica preliminar entre el volumen ocupado por los contenedores y el volumen geométrico del rack:

$$\eta_v = \frac{NV_c}{V_r} \times 100 \quad (4)$$

donde V_c es el volumen geométrico de un contenedor y V_r el volumen geométrico del rack. Con $V_c = 15 \times 10 \times 7 = 1050 \text{ cm}^3$, $V_r = 60 \times 24 \times 40 = 57600 \text{ cm}^3$ y $N = 9$, se obtiene:

$$\eta_v = \frac{9(1050)}{57600} \times 100 = 16.4\% \quad (5)$$

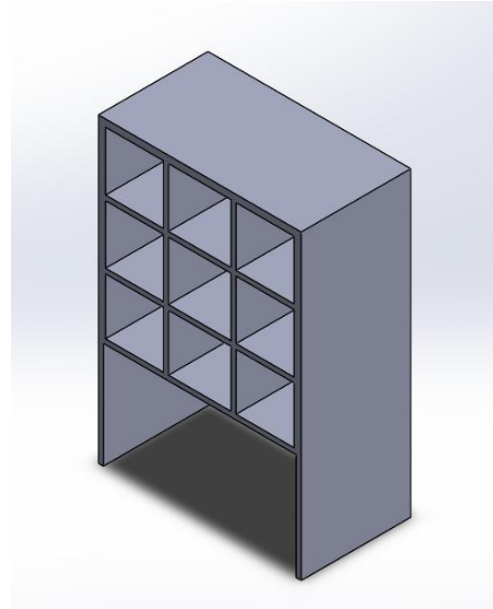


Figura 10. Modelo del rack para el almacenamiento de contenedores en el sistema ASRS.

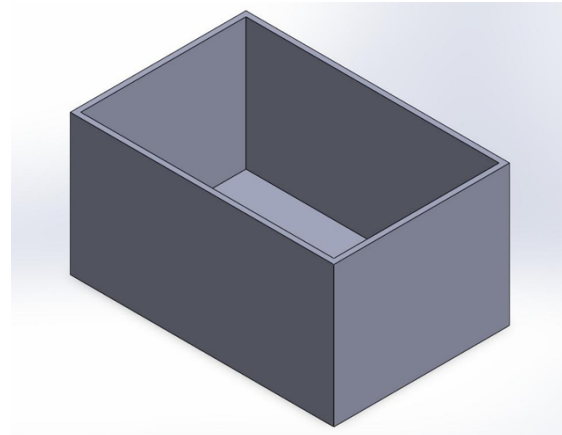


Figura 11. Modelo del contenedor propuesto para el sistema de almacenamiento del ASRS.

Este valor no debe interpretarse como eficiencia industrial de almacenamiento, ya que el objetivo del prototipo no es maximizar densidad. En este caso, la relación volumétrica funciona como un indicador de holgura didáctica: el espacio libre permite observar el movimiento, manipular los contenedores y realizar ajustes durante las prácticas sin comprometer la visibilidad del sistema.

4.4 Ensamble CAD

Para complementar el diseño mecánico, se establecieron criterios de evaluación preliminar que pueden aplicarse en SolidWorks antes de fabricar el prototipo. Estos criterios no se presentan como resultados experimentales, sino como una ruta de verificación CAD para revisar la viabilidad geométrica del ensamble y preparar la etapa posterior de pruebas físicas.

La primera verificación propuesta corresponde a la detección de interferencias entre componentes mediante la herramienta *Interference Detection*, considerando rack, contenedores, plataforma, guías, soportes SK8, motores y elementos de transmisión. La segunda corresponde a la revisión de propiedades de masa, útil para estimar el peso total y la ubicación del centro de masa cuando se asignen materiales al modelo CAD. La tercera corresponde a un análisis estático básico mediante SolidWorks Simulation o SimulationXpress, fijando la base y aplicando una carga representativa sobre el rack o la plataforma.

El factor de seguridad para una evaluación estática puede expresarse como:

$$FS = \frac{\sigma_y}{\sigma_{max}} \quad (6)$$

donde FS es el factor de seguridad, σ_y el esfuerzo de fluencia del material y σ_{max} el esfuerzo máximo obtenido en la simulación.

La Tabla 3 resume los criterios propuestos para la revisión del ensamble CAD.

Tabla 3. Criterios propuestos para evaluación preliminar del ensamble CAD

Criterio	Propósito técnico
Interferencias geométricas	Verificar que no existan contactos no deseados entre rack, contenedores, plataforma, guías y transmisión.
Compatibilidad rack-contenedor	Confirmar que los contenedores puedan colocarse en las nueve posiciones sin bloqueo geométrico.
Alineación de guías	Revisar la correspondencia entre varillas de 8 mm y soportes SK8 dentro del ensamble.
Integración motor-polea	Confirmar la compatibilidad entre eje de motor de 5 mm y poleas seleccionadas.
Propiedades de masa	Estimar masa total y centro de masa cuando se asignen materiales en SolidWorks.
Análisis estático básico	Evaluar desplazamiento, esfuerzo máximo y factor de seguridad bajo una carga representativa.

Estos criterios establecen una ruta técnica para revisar el prototipo antes de su fabricación. De esta forma, el diseño puede evaluarse por compatibilidad dimensional, alineación, integración de componentes y comportamiento estructural preliminar, no únicamente por su apariencia visual en el modelo CAD.

Como resultado de la integración mecánica, la Fig. 12 muestra el prototipo físico desarrollado. En la imagen se observa la estructura general del sistema, el rack de almacenamiento, la plataforma móvil y los contenedores utilizados para representar las operaciones básicas de almacenamiento y recuperación. Este ensamble permite verificar de manera visual la correspondencia entre el diseño CAD y la construcción física inicial del prototipo.

5. CONCLUSIONES

En este trabajo se desarrolló el diseño mecánico de un prototipo ASRS de escala didáctica para apoyar la enseñanza de sistemas automatizados de almacenamiento y recuperación en Ingeniería Industrial. El diseño permitió

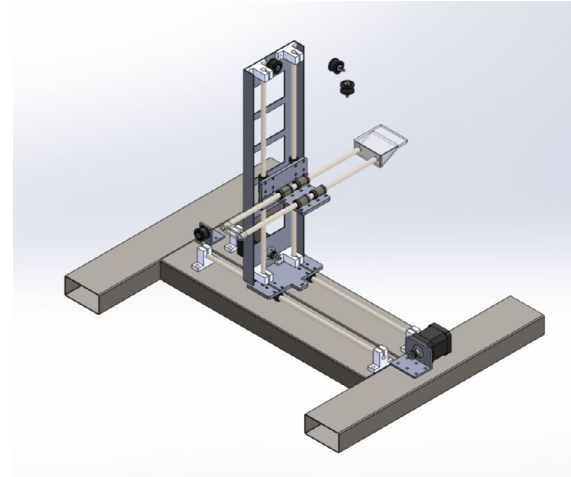


Figura 12. Prototipo físico del sistema ASRS desarrollado para prácticas de almacenamiento y recuperación en escala didáctica.

integrar una base estructural, sistema de guías, soportes SK8, motores Nema 17, poleas, plataforma móvil, rack de nueve posiciones y contenedores compatibles con fabricación por impresión 3D.

El principal resultado fue la obtención de una arquitectura mecánica organizada y modelada en SolidWorks, con dimensiones de referencia para los componentes principales y una configuración adecuada para representar operaciones básicas de almacenamiento y recuperación. La distribución 3×3 del rack permite trabajar con nueve posiciones visibles, lo que facilita el análisis de ubicación, desplazamiento, compatibilidad dimensional y secuencia de operación dentro de un entorno académico.

Considerando la parte técnica, se establecieron criterios para revisar la compatibilidad rack-contenedor, la alineación de guías, la integración motor-polea, la resolución lineal futura del sistema de transmisión y la evaluación preliminar del ensamble CAD. Estos elementos fortalecen la propuesta como base mecánica para una etapa posterior de manufactura, pruebas de desplazamiento y validación experimental.

Como trabajo futuro, será necesario fabricar y evaluar el prototipo bajo condiciones controladas. Las pruebas deberán considerar interferencias geométricas, repetibilidad de posicionamiento, estabilidad estructural, desplazamiento de la plataforma y facilidad de manipulación de contenedores. Posteriormente, la plataforma podrá integrarse con una interfaz en Python, control mediante ESP32 y un módulo de visión artificial para reconocimiento de posiciones o verificación de contenedores.

REFERENCIAS

- Aslam, M., Ullah, N., et al. (2025). Benefits and challenges of automated storage and retrieving systems in warehouses: A systematic review. ResearchGate. Consultado el 11 de abril de 2026.
- AutoStore (2025a). Autostore system overview. URL <https://www.autostoresystem.com/system>. Consultado el 11 de abril de 2026.



- AutoStore (2025b). World's fastest as/rs. URL <https://www.autostoresystem.com/>. Consultado el 11 de abril de 2026.
- Dematic (2026). Automated storage and retrieval systems. URL <https://www.dematic.com/en-us/solutions/automated-storage-and-retrieval-systems/>. Consultado el 11 de abril de 2026.
- Descartes Systems Group (2024a). Descartes' study reveals 54 % of supply chain and logistics operations are prioritizing automation to mitigate workforce shortages. Consultado el 11 de abril de 2026.
- Descartes Systems Group (2024b). Descartes' study reveals 76 % of supply chain and logistics operations are experiencing notable workforce shortages. Consultado el 11 de abril de 2026.
- Geek+ (2026). Shelf-to-person solution. URL <https://www.geekplus.com/resources/product-brochures/pickingflyer>. Consultado el 11 de abril de 2026.
- Honeywell (2026). Automated storage and retrieval systems (as/rs). URL <https://automation.honeywell.com/us/en/products/warehouse-automation/solutions-by-technology/asrs-systems>. Consultado el 11 de abril de 2026.
- Honeywell Intelligrated (2021). What to consider for a successful as/rs investment. White paper, Honeywell Intelligrated.
- Kardex (2022). 6 benefits of automated storage and retrieval systems (asrs). URL <https://www.kardex.com/en-us/blog/benefits-of-asrs>. Consultado el 11 de abril de 2026.
- Kardex (2025). Kardex shuttle vertical lift module. URL <https://www.kardex.com/en-us/products/vertical-lift-module/kardex-shuttle>. Consultado el 11 de abril de 2026.
- Kardex (2026). What is warehouse automation? a complete guide for 2026. URL <https://www.kardex.com/>. Consultado el 11 de abril de 2026.
- MHI (2025). 2025 mhi annual industry report. URL <https://www.mhi.org/annual-industry-reports>. Consultado el 11 de abril de 2026.
- SSI Schaefer (2026). Automated storage and retrieval system (asrs). URL <https://www.ssi-schaefer.com/en-us/solutions/by-intralogistic-strategy/warehouse-automation/asrs>. Consultado el 11 de abril de 2026.

Low-Cost Open-Source Virtual Instrumentation Architecture for Distributed Industrial Monitoring and Control^{*}

Axel Herrera-Cabrera^{*} Rubí M. Huaracha^{**}

^{*} *Tecnológico Nacional de México, ITS Guanajuato, Carretera Guanajuato a Puente de las Urras km 10.5 Predio El Carmen. (e-mail: axel.hc@guanajuato.tecnm.mx).*

^{**} *Universidad Tecnológica de León, Departamento de Tecnologías Emergentes, Industriales e Informáticas, Blvd. Universidad Tecnológica 225, León, 37670, Guanajuato, México (e-mail: 86081@alumnos.utleon.edu.mx)*

Abstract: The digital transformation of industrial systems requires scalable, cost-effective, and interoperable solutions for monitoring, control, and data acquisition, particularly in small and medium-sized enterprises (SMEs) and resource-constrained environments. This paper presents a low-cost open-source virtual instrumentation framework based on MyOpenLab and embedded hardware platforms for distributed industrial monitoring and supervisory control. The proposed architecture integrates field-level sensors and actuators with graphical human-machine interfaces through real-time serial communication, enabling rapid deployment of modular cyber-physical subsystems without reliance on proprietary software. The framework is experimentally validated through ten representative industrial scenarios, including temperature regulation, smart irrigation, automated ventilation, environmental monitoring, and security systems. Experimental results demonstrate stable real-time operation with response times below 2 s, temperature measurement errors within $\pm 1^{\circ}\text{C}$, and relative humidity errors below $\pm 5\%$. A comparative analysis shows that the proposed solution achieves performance levels comparable to proprietary platforms for non-safety-critical industrial applications while significantly reducing licensing and deployment costs. The results confirm that open-source virtual instrumentation can effectively support Industry 4.0-oriented industrial digitalization by enabling interoperability, scalability, and edge-level intelligence. The proposed framework provides a practical and replicable pathway for SMEs and educational-industrial environments to adopt distributed monitoring and control solutions while fostering workforce skills aligned with modern industrial informatics paradigms.

Keywords: Industrial informatics, virtual instrumentation, open-source systems, human-machine interface (HMI), real-time monitoring, Arduino, cyber-physical systems, low-cost automation, engineering education, Industry 4.0.

1. INTRODUCCIÓN

La informática industrial está experimentando una transformación profunda impulsada por la convergencia de los sistemas ciberfísicos, el Internet Industrial de las Cosas (IIoT) y los paradigmas de digitalización asociados con la Industria 4.0. En este contexto, la disponibilidad de soluciones de monitoreo y control escalables, interoperables y económicamente accesibles se ha convertido en un requisito estratégico, particularmente para las pequeñas y medianas empresas (PYMEs) y las instituciones que operan bajo restricciones de recursos.

A pesar de esta creciente demanda, la educación en ingeniería y los entornos de prototipado industrial especial-

mente en mecatrónica, automatización e instrumentación continúan enfrentando limitaciones estructurales. El acceso restringido a laboratorios físicos, los altos costos de licenciamiento de plataformas de software propietario y la disponibilidad limitada de equipos de grado industrial dificultan la adquisición de competencias prácticas alineadas con las necesidades industriales reales Li and Liang (2024); Elmoazen et al. (2023); Garcia-Zubia et al. (2016). Esta situación exacerba la brecha entre la instrucción teórica y la experiencia práctica, afectando negativamente la preparación de la fuerza laboral en sectores manufactureros emergentes.

Los laboratorios virtuales y las plataformas de experimentación remota han sido propuestos como alternativas efectivas para mitigar estas limitaciones Reid et al. (2022); Reeves and Crippen (2021); Fadda et al. (2022); Beghi et al. (2017). Estudios previos reportan que los laborato-

^{*} Los autores agradecen al Tecnológico Nacional de México, al ITS Guanajuato y a la Universidad Tecnológica de León por las facilidades otorgadas para el desarrollo de esta investigación.

rios virtuales pueden lograr resultados de aprendizaje comparables a los de los laboratorios tradicionales, mientras permiten entornos de experimentación flexibles, escalables y rentables Uzunidis and Pagiatakis (2023); Petrović et al. (2016). Más recientemente, la proliferación de ecosistemas de software y hardware de código abierto ha acelerado aún más la adopción de plataformas de instrumentación virtual en contextos tanto educativos como preindustriales.

Los entornos de código abierto como MyOpenLab, combinados con plataformas embebidas de bajo costo (ej., Arduino y ESP32), permiten el desarrollo de sistemas de instrumentación virtual personalizables capaces de adquisición de datos, visualización y control supervisorio en tiempo real a un costo marginal Hernández-Melgarejo et al. (2020); Ak et al. (2018); Grout (2017); Cvetkovski et al. (2019). En contraste con plataformas propietarias como LabVIEW o TIA Portal, las soluciones de código abierto eliminan las barreras de licenciamiento, mejoran la replicabilidad y permiten la modificación y redistribución directa—características que son particularmente relevantes para PYMEs e instituciones en economías emergentes Alnagrat et al. (2022); El-Abd (2016).

Sin embargo, los estudios existentes a menudo se centran principalmente en resultados pedagógicos y carecen de una evaluación sistemática de métricas de rendimiento industrialmente relevantes, incluyendo tiempo de respuesta, precisión de medición, robustez del sistema y escalabilidad Reilly and Dede (2019); Kuriščák et al. (2022); Melo et al. (2021). Además, la validación empírica de plataformas de instrumentación virtual de código abierto bajo escenarios representativos de monitoreo y control industrial sigue siendo limitada. A pesar de estas ventajas, persisten varias limitaciones en los estudios existentes. Muchas implementaciones reportadas se centran principalmente en resultados educativos sin abordar sistemáticamente la escalabilidad industrial, las métricas de rendimiento del sistema (ej., latencia, precisión, robustez) o la interoperabilidad con arquitecturas orientadas a IIoT Kebande (2024); Kabaciński and Pawliczek (2015); Forrier et al. (2016). Adicionalmente, hay una falta de evidencia empírica que vincule las plataformas de instrumentación virtual de código abierto con escenarios de monitoreo y control industrialmente relevantes bajo restricciones realistas García-Guzmán et al. (2017).

Para abordar estas brechas, este artículo presenta un marco de laboratorio virtual basado en instrumentación virtual de código abierto utilizando MyOpenLab, validado a través de la implementación de múltiples sistemas mecatrónicos representativos de aplicaciones de monitoreo y control industrial Alexandru (2020); Betancourt et al. (2019); Lara-Padilla et al. (2017); Moreno-Peñalosa et al. (2021); Bonilla Correa (2021); Mendez Gonzalez et al. (2019). El enfoque propuesto se despliega en un entorno real de educación técnica y se evalúa utilizando indicadores de rendimiento cuantitativos, incluyendo tiempo de respuesta, precisión de medición y estabilidad del sistema. Más allá de su contribución educativa, el marco propuesto enfatiza la replicabilidad, la eficiencia de costos y la compatibilidad con los principios de informática industrial, posicionándolo como una solución viable tanto para la formación de fuerza laboral como para aplicaciones indus-

triales de baja criticidad Velázquez et al. (2024); Vázquez-Santos et al. (2021).

Las principales contribuciones de este trabajo se resumen a continuación:

- Un marco de instrumentación virtual de código abierto para educación en informática industrial y prototipado. Se propone una arquitectura de laboratorio virtual modular basada en MyOpenLab y plataformas de hardware abierto (Arduino/ESP32) que permite el monitoreo y control en tiempo real de sistemas mecatrónicos relevantes para entornos industriales.
- Validación experimental utilizando múltiples sistemas mecatrónicos de inspiración industrial. Diez sistemas representativos—including monitoreo de temperatura, riego inteligente, ventilación automatizada y control de invernadero inteligente—fueron implementados y evaluados, demostrando operación en tiempo real con tiempos de respuesta por debajo de 2 s y errores de medición dentro de $\pm 1^\circ\text{C}$ y $\pm 5\%$ de HR.
- Evaluación cuantitativa de rendimiento alineada con métricas de informática industrial. El marco propuesto se evalúa utilizando indicadores objetivos como latencia, desviación estándar, error RMS y estabilidad de comunicación, proporcionando evidencia empírica de su idoneidad para tareas de monitoreo y control industrial no críticas.
- Análisis de costo, escalabilidad y replicabilidad para PYMEs y contextos con recursos limitados. Un análisis comparativo con plataformas propietarias destaca reducciones significativas de costos, mejora de la replicabilidad y adaptabilidad mejorada, apoyando la adopción de instrumentación virtual de código abierto en PYMEs y programas de formación técnica.
- Puente entre la educación y los paradigmas de digitalización industrial. El enfoque propuesto se alinea con los principios de la Industria 4.0 y IIoT al permitir interoperabilidad, expansión modular e integración con arquitecturas de monitoreo distribuidas, contribuyendo a la preparación de la fuerza laboral y la transformación digital industrial de bajo costo.

2. METODOLOGÍA

La metodología empleada en este trabajo sigue un enfoque orientado a la informática industrial en capas, permitiendo la integración de hardware de bajo costo, instrumentación virtual de código abierto y mecanismos de comunicación escalables. La arquitectura está concebida para apoyar tareas de monitoreo, control y adquisición de datos comúnmente encontradas en escenarios de automatización industrial, mientras mantiene accesibilidad y adaptabilidad para el despliegue en entornos con recursos limitados.

2.1 Arquitectura del Sistema

El sistema está estructurado en tres capas interconectadas, que juntas forman un marco ciberfísico modular, como se ilustra en la Figura 1:

- **Capa Física (Dispositivos de Campo):** Sensores y actuadores.

- **Capa de Adquisición y Comunicación de Datos:** Interfaz entre hardware y software.
- **Capa de Instrumentación Virtual y Supervisión:** HMI y control.

3. Virtual Instrumentation and Supervisory Layer

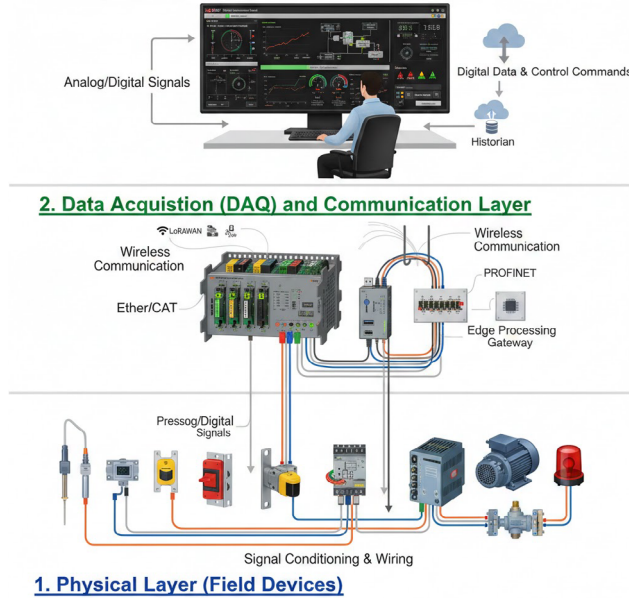


Fig. 1. Arquitectura del sistema propuesta orientada a la Industria 4.0, destacando la interacción entre los dispositivos de campo, la capa de instrumentación virtual de código abierto y el nivel de monitoreo supervisorio.

2.2 Capa Física: Sensores y Actuadores

La capa física consiste en dispositivos a nivel de campo típicamente utilizados en aplicaciones industriales de baja y media complejidad. Estos incluyen:

- **Sensores:** temperatura (LM35, DHT22), humedad, intensidad de luz (LDR), movimiento (PIR) y entradas digitales.
- **Actuadores:** motores DC, relés, relés de estado sólido (SSR), bombas, ventiladores, LEDs y servomotores.

Todos los dispositivos se interfazan utilizando plataformas de microcontroladores de código abierto, principalmente Arduino UNO, seleccionado por su adopción generalizada, bajo costo y compatibilidad con flujos de trabajo de prototipado industrial. Esta capa permite la adquisición de señales en tiempo real, control de lazo cerrado básico y experimentación híbrida que combina hardware real y supervisión virtual.

2.3 Capa de Adquisición y Comunicación de Datos

La capa de adquisición de datos proporciona la interfaz entre los dispositivos físicos y el entorno de instrumentación virtual. La comunicación se establece a través de comunicación serial basada en USB, protocolos RS-232/TTL y

configuraciones de velocidad de baudios estándar compatibles con los requisitos de adquisición de datos industriales (ver Figura 2).

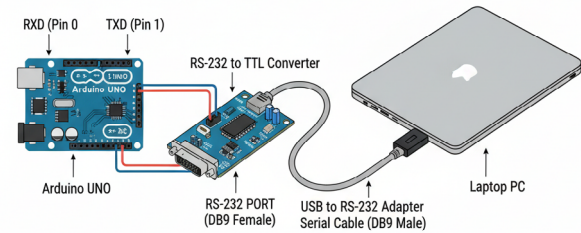


Fig. 2. Diagrama de conexión Arduino-MyOpenLab

Un firmware basado en Firmata se despliega en las unidades de microcontrolador, permitiendo comunicación estandarizada y abstrayendo detalles de hardware de bajo nivel de la capa supervisorio. Este enfoque facilita la reconfiguración rápida del sistema, la abstracción del hardware y la interoperabilidad con múltiples plataformas de software.

2.4 Capa de Instrumentación Virtual y Supervisión

A nivel supervisor, MyOpenLab se emplea como el entorno de instrumentación virtual. Esta capa es responsable del diseño de la Interfaz Hombre-Máquina (HMI), visualización de datos en tiempo real, implementación de lógica de control y funciones de monitoreo supervisorio. MyOpenLab permite la creación de diagramas funcionales basados en bloques, emulando instrumentos industriales como osciloscopios virtuales, multímetros digitales, controladores PID, registradores de datos y paneles de alarma.

El paradigma de programación orientado a bloques reduce la complejidad del desarrollo y permite a usuarios sin experiencia avanzada en programación diseñar sistemas supervisores funcionales, un factor crítico para PYMEs y entornos de formación. La Figura 3 ilustra los objetos y subobjetos disponibles dentro de la plataforma, incluyendo funciones lógicas, valores enteros y de punto flotante, así como conexiones seriales con tarjetas de adquisición de datos.



Fig. 3. Panel de funciones de MyOpenLab

2.5 Implementación de Hardware

La capa física del sistema está basada en controladores embebidos de bajo costo, principalmente la placa de desarrollo Arduino UNO, seleccionada debido a su amplia adopción en prototipado industrial, soporte robusto de la comunidad y compatibilidad con múltiples protocolos de comunicación.

Los prototipos implementados incorporan un conjunto de sensores y actuadores industriales comúnmente utilizados,

incluyendo sensores de temperatura (LM35, DHT22), sensores de humedad relativa, resistencias dependientes de la luz (LDR), motores DC, servomotores, válvulas solenoides, bombas de agua, e indicadores visuales y audibles (LEDs y zumbadores).

Todos los sensores se conectaron a través de interfaces de E/S analógicas y digitales estándar, mientras que los actuadores fueron controlados utilizando módulos de relé o circuitos de driver basados en transistores para asegurar aislamiento eléctrico y operación segura. Esta configuración se asemeja estrechamente a los dispositivos de campo industriales típicos utilizados en sistemas de automatización de baja criticidad.

2.6 Firmware y Comunicación

A nivel de firmware, los dispositivos embebidos fueron programados utilizando el framework de Arduino, empleando comunicación basada en Firmata para una interacción fluida con la capa de instrumentación virtual. El uso de Firmata permite el intercambio de datos bidireccional en tiempo real sin requerir modificaciones complejas de firmware para cada aplicación, como se puede observar en la Figura 4.

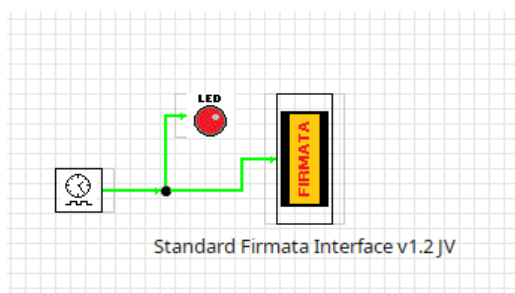


Fig. 4. Ejemplo del uso y conexión del protocolo Firmata para la integración hardware-software

La comunicación serial se estableció utilizando interfaces UART basadas en USB (RS-232/TTL), operando a velocidades de baudios es entre 9,600 y 15,200 bps, dependiendo de los requisitos de la aplicación. Los paquetes de datos se estructuraron utilizando tramas ligeras que contienen identificadores de sensores, mediciones con marca de tiempo y estados de actuadores. Este enfoque asegura una adquisición de datos determinista y latencia predecible.

La arquitectura de comunicación permite una migración fácil hacia protocolos industriales como Modbus RTU/TCP o MQTT, permitiendo la integración futura en infraestructuras IIoT y SCADA sin rediseños mayores.

2.7 Sistemas Implementados

Se implementaron diez sistemas mecatrónicos representativos de escenarios industriales de baja a media complejidad. La Tabla 1 resume las características técnicas de cada sistema desarrollado.

3. RESULTADOS Y DISCUSIONES

Esta sección presenta la validación experimental del marco de instrumentación virtual de código abierto propuesto

bajo condiciones representativas de entornos industriales de baja a media criticidad. La evaluación se centra en la capacidad de respuesta del sistema, precisión de medición, confiabilidad de la comunicación y escalabilidad funcional.

3.1 Configuración Experimental

La plataforma de instrumentación virtual de código abierto propuesta se validó experimentalmente a través de un conjunto de sistemas mecatrónicos representativos que emulan escenarios de monitoreo y control industrial de baja a media complejidad. La configuración experimental combinó módulos de hardware físico (Arduino UNO), sensores ambientales y eléctricos, y actuadores electromecánicos, integrados a través de comunicación serial con el entorno de instrumentación virtual MyOpenLab.

Todos los experimentos se realizaron bajo condiciones de laboratorio controladas para asegurar la repetibilidad. El sistema anfitrión ejecutó MyOpenLab en una computadora de escritorio estándar (procesador Intel i5, 16 GB RAM, Windows OS), mientras que las plataformas de microcontrolador manejaron la adquisición de datos y el control de actuadores. La comunicación serial se estableció vía USB utilizando protocolos RS-232/TTL a una velocidad de baudios de 9,600 bps.

La Figura 5 ilustra el esquema general de integración hardware-software lograda utilizando el protocolo Firmata en el ejemplo de la alarma.

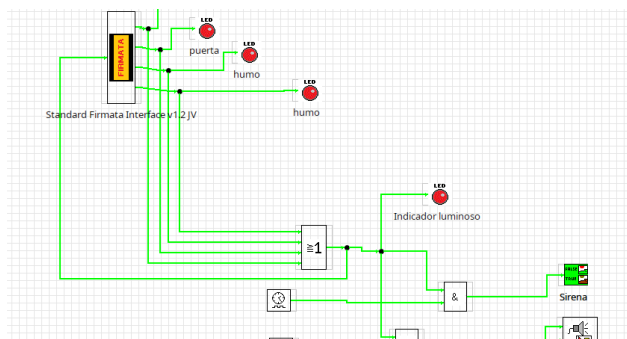


Fig. 5. Esquema general de integración hardware-software para el ejemplo de la alarma

3.2 Capacidad de Respuesta del Sistema

Para validar el monitoreo y control en tiempo real usando instrumentación de código abierto, se realizaron mediciones de tiempo de respuesta en los diez sistemas mecatrónicos representativos. Cada experimento evaluó el retraso entre la detección del evento del sensor y la respuesta del actuador a través de la comunicación serial entre el hardware basado en Arduino y el entorno MyOpenLab.

La Tabla 2 resume los tiempos de respuesta promedio y las desviaciones estándar obtenidas durante las pruebas experimentales, el sistema propuesto alcanzó consistentemente tiempos de respuesta por debajo de 2 s, con variabilidad mínima. Estos resultados demuestran que el marco satisface las restricciones de tiempo real para aplicaciones de monitoreo y control industrial de baja criticidad, validando así la viabilidad de la instrumentación virtual de código abierto para escenarios de informática industrial.

Tabla 1. Características Técnicas de los Sistemas Implementados

Sistema	Sensor Principal	Actuador	Variable Controlada	Aplicación
Monitor de Temperatura	LM35	SSR + Resistencia	Temperatura [°C]	Horno de precalentamiento
Riego Inteligente	Humedad suelo	Bomba de agua	Humedad [%]	Invernadero
Ventilación Automática	DHT22	Ventilador 12 V	Temperatura [°C]	Gabinete eléctrico
Estación Meteorológica	DHT22, LDR	LED indicador	Temp., Humedad, Luz	Laboratorio agrícola
Sistema de Seguridad	PIR	Zumbador	Presencia	Seguridad
Conversión ADC	-	N/A	Conversión	Educativo
Juego de Reacción	Pulsador	LED	Tiempo reacción	Educativo
Estación de Control	LDR	LED	Iluminación	Control de acceso
Invernadero Inteligente	DHT22	Motor	Temp. y humedad	Agricultura
Medidor de Luz Ambiental	LDR	Lámpara	Iluminación	Domótica

Tabla 2. Análisis de Tiempo de Respuesta

Sistema	Evento	Tiempo (ms)	Desv. est. (ms)
Monitor de temperatura	Evaluación temperatura	1000	0.006
Riego inteligente	Activación por humedad	1000	0.022
Ventilación automática	Actuación térmica	1000	0.025
Estación meteorológica	Adquisición multisensor	2000	0.010
Sistema de seguridad	Detección de presencia	0	0.050
Conversión ADC	Muestreo	100	0.050
Juego de reacción	Respuesta usuario	0	0.000
Estación de control	Luz y presencia	1000	0.007
Invernadero inteligente	Control multivariable	1000	0.010
Medidor de luz ambiental	Detección iluminación	100	0.030

3.3 Análisis de costos de los componentes

La Tabla 3 presenta los costos aproximados en dólares estadounidenses (USD) de los elementos utilizados en la implementación de los diez sistemas.

3.4 Precisión de Medición y Confiabilidad

Para evaluar la confiabilidad de medición usando sensores de bajo costo, se realizaron pruebas de precisión comparando las lecturas de los sensores contra valores de referencia. Se calcularon el error absoluto, el error de raíz cuadrada media (RMS) y el error porcentual para cada tipo de sensor. Los resultados cuantitativos se presentan en la Tabla 4.

Estos resultados indican que todos los sensores operaron dentro de rangos de precisión aceptables para entornos educativos e industriales no críticos. Las mediciones de temperatura se mantuvieron dentro de $\pm 1^\circ\text{C}$, mientras que los errores de humedad estuvieron por debajo de $\pm 5\%$, confirmando la confiabilidad de la arquitectura propuesta para el despliegue práctico.

3.5 Análisis Comparativo con Software Propietario

Para evaluar la viabilidad como alternativa rentable y escalable a plataformas propietarias, se realizó una comparación cualitativa y cuantitativa entre la solución propuesta basada en MyOpenLab y los entornos comerciales ampliamente utilizados.

La Tabla 5 resume la evaluación comparativa.

El marco propuesto ofrece un rendimiento competitivo mientras reduce significativamente las barreras económicas y técnicas. Esto apoya directamente la afirmación de que la instrumentación virtual de código abierto puede democratizar el acceso a las herramientas de informática

industrial, particularmente para PYMEs e instituciones educativas.

3.6 Discusión

Los resultados experimentales demuestran que el marco de instrumentación virtual de código abierto propuesto no se limita al prototipado educativo sino que constituye un habilitador viable para arquitecturas de monitoreo y control industrial de bajo costo, particularmente alineado con los principios de la Industria 4.0.

Relevancia para la Industria 4.0 Desde una perspectiva de informática industrial, los tiempos de respuesta logrados por debajo de 2 segundos, los bajos errores de medición (por debajo del 5%) y la comunicación serial estable validan la idoneidad del enfoque propuesto para procesos industriales no críticos en el tiempo, incluyendo monitoreo ambiental, control de procesos auxiliares, ventilación inteligente y automatización agrícola.

La arquitectura modular se alinea con los modelos de referencia de la Industria 4.0, como RAMI 4.0, donde los dispositivos de campo, las capas de comunicación y las capas de información están claramente desacopladas. El uso de nodos basados en Arduino combinados con instrumentación virtual basada en MyOpenLab permite el despliegue rápido de prototipos de sistemas ciberfísicos (CPS) con una inversión mínima en infraestructura.

Comparación con Soluciones Propietarias Cuando se compara con plataformas industriales propietarias (ej., LabVIEW, TIA Portal), la solución propuesta demuestra un rendimiento comparable para aplicaciones de baja y media criticidad, mientras reduce significativamente los costos de licenciamiento y despliegue. Esta compensación costo-rendimiento es particularmente relevante para PYMEs.



Tabla 3. Costo de componentes (USD)

Componente	Cantidad	Costo unitario (USD)
Arduino UNO	1	\$23.00
LM35	1	\$2.50
DHT22	1	\$9.00
Sensor humedad suelo	1	\$5.00
LDR	1	\$0.50
Módulo relé 5V	1	\$2.00
Motor DC	1	\$4.00
Bomba de agua	1	\$5.00
Servomotor SG90	1	\$3.50
Zumbador	1	\$1.00
LEDs + resistencias	varios	\$1.00
Fuente 5V/2A	1	\$8.00
Costo promedio por sistema		\$35-\$50

Tabla 4. Evaluación de Precisión de Sensores

Sensor	Variable	Referencia	Medio	Error abs.	Error RMS	Error %
LDR	Luminancia (lx)	100.0	99.2	-0.8	0.8	0.8
DHT22	HR (%) / Temp. (°C)	55.0	53.8	-1.2	1.4	2.18
LM35	Temperatura (°C)	25.0	26.1	+1.1	1.2	4.4
Humedad suelo	Humedad relativa (%)	100.0	99.3	-0.7	0.8	0.7

Tabla 5. Comparación de plataformas de instrumentación virtual

Criterio	Software Propietario	MyOpenLab
Costo de licencia	Alto (> \$2000/año)	Gratuito
Paradigma de programación	Textual / Híbrido	Gráfico (basado en bloques)
Curva de aprendizaje	Pronunciada	Baja
Integración de hardware	Dependiente del vendedor	Arduino / ESP32
Personalización	Limitada	Alta
Replicabilidad	Restringida	Abierta y escalable
Protocolos industriales	Soporte nativo	Compatible con Modbus/MQTT

Si bien las soluciones propietarias ofrecen bibliotecas certificadas, soporte del vendedor y cumplimiento con estándares de seguridad industrial, la instrumentación virtual de código abierto puede cubrir efectivamente las capas de monitoreo, visualización y control básico.

Limitaciones y Restricciones Industriales Deben reconocerse varias limitaciones desde un punto de vista industrial. Primero, la ausencia de certificaciones industriales formales (ej., IEC 61508, ISO 13849) restringe el despliegue del sistema propuesto en aplicaciones críticas para la seguridad. Segundo, la robustez a largo plazo bajo condiciones industriales adversas requiere validación adicional. Adicionalmente, la dependencia del desarrollo impulsado por la comunidad puede plantear desafíos en términos de mantenimiento a largo plazo.

4. CONCLUSIONES

Este artículo ha presentado un marco de instrumentación virtual de código abierto basado en MyOpenLab y hardware embebido de bajo costo para aplicaciones de monitoreo y control industrial. El enfoque propuesto demuestra que se pueden implementar soluciones de informática industrial confiables y escalables sin dependencia de plataformas de software propietario, reduciendo así los costos de despliegue y bajando las barreras de entrada para pequeñas y medianas empresas (PYMEs) y entornos industriales emergentes.

La validación experimental a través de múltiples casos de uso representativos—incluyendo monitoreo ambiental,

riego automatizado, control de invernadero inteligente y sistemas de seguridad—confirmó que la arquitectura propuesta logra tiempos de respuesta por debajo de 2 segundos y precisión de medición dentro de $\pm 1^\circ\text{C}$ para temperatura y $\pm 5\%$ para humedad relativa. Estos indicadores de rendimiento son comparables a los de los sistemas comerciales para aplicaciones industriales no críticas para la seguridad, apoyando la viabilidad del enfoque para escenarios reales de monitoreo y control supervisorio. Además, la arquitectura de sistema modular permite la integración fluida de sensores y actuadores heterogéneos mientras mantiene una comunicación estable en tiempo real a través de interfaces seriales estándar.

Desde una perspectiva de la Industria 4.0, la naturaleza de código abierto del marco promueve la interoperabilidad, flexibilidad y sostenibilidad a largo plazo. La capacidad de integrar nodos de detección de bajo costo con interfaces hombre-máquina personalizables facilita estrategias de digitalización graduales, particularmente en sistemas heredados y contextos industriales con recursos limitados. Adicionalmente, la compatibilidad con protocolos de comunicación ampliamente adoptados proporciona una base para la integración futura en entornos de Internet Industrial de las Cosas (IIoT) y SCADA distribuidos.

El trabajo futuro se centrará en extender el marco hacia mayores niveles de preparación industrial. Esto incluye la implementación de protocolos de comunicación industrial como Modbus TCP/IP y OPC UA, mejoras de ciberseguridad para la integridad de datos y control de acceso, y pruebas de robustez a largo plazo bajo condiciones industriales

adversas. Además, se realizarán despliegues a gran escala y evaluaciones comparativas contra plataformas industriales certificadas para evaluar la escalabilidad, confiabilidad y mantenibilidad durante períodos operativos extendidos.

En general, este trabajo aporta evidencia empírica de que la instrumentación virtual de código abierto constituye una alternativa viable y rentable para aplicaciones de informática industrial, alineándose con los principios de la Industria 4.0 y apoyando la transformación digital de PYMEs y sectores manufactureros emergentes.

REFERENCES

- Ak, A., Topuz, V., Altıkardeş, A., and Oral, B. (2018). Development of a remote laboratory infrastructure and LMS for mechatronics distance education. *EURASIA Journal of Mathematics, Science and Technology Education*, 14(6), 2493–2508. doi: <https://doi.org/10.29333/ejmste/89797>.
- Alexandru, C. (2020). Virtual prototyping platform for designing mechanical and electronic systems. In *Product design*, 67–89. IntechOpen. doi: <https://doi.org/10.5772/intechopen.91234>.
- Alnagrat, A.J.A., Ismail, R.C., and Idrus, S.Z.S. (2022). The opportunities and challenges in virtual reality for virtual laboratories. *Innovative Teaching and Learning Journal*, 6(2), 73–89.
- Beghi, A., Marcuzzi, F., Martin, P., Tinazzi, F., and Zigliotto, M. (2017). Virtual prototyping of embedded control software in mechatronic systems: A case study. *Mechatronics*, 43, 99–111. doi: <https://doi.org/10.1016/j.mechatronics.2017.03.002>.
- Betancourt, R.J., González-López, J.M., Espejo, E.B., Pérez-González, M.A., Laureano, E.V., and Ledesma, J.A. (2019). Virtual instrumentation based learning methodology for teaching power system measurement and protection. *Computer Applications in Engineering Education*, 27(6), 1555–1570. doi: <https://doi.org/10.1002/cae.22178>.
- Bonilla Correa, B.A. (2021). *Implementación de un sistema HIL (Hardware in the Loop) para el control de nivel*. Master's thesis, Instituto Tecnológico Superior de Comalcalco.
- Cvetkovski, G., Petkovska, L., Di Barba, P., Mognaschi, M.E., Kaminska, D., Firyeh-Nowacka, A., and Zwolinski, G. (2019). ViMeLa project: An innovative concept for teaching mechatronics using virtual reality. *Przegląd Elektrotechniczny*, 95(5), 18–21. doi: <https://doi.org/10.15199/48.2019.05.04>.
- El-Abd, M. (2016). How course projects can successfully prepare engineering students for capstone design projects. In *2016 IEEE Global Engineering Education Conference (EDUCON)*, 746–750. Abu Dhabi, United Arab Emirates. doi: <https://doi.org/10.1109/EDUCON.2016.7474635>.
- Elmoazen, R., Saqr, M., Khalil, M., and Wasson, B. (2023). Learning analytics in virtual laboratories: A systematic literature review of empirical research. *Smart Learning Environments*, 10(1), 23. doi: <https://doi.org/10.1186/s40561-023-00243-x>.
- Fadda, D., Salis, C., and Vivanet, G. (2022). About the efficacy of virtual and remote laboratories in STEM education in secondary school: A second-order systematic review. *Journal of Educational, Cultural and Psychological Studies (ECPS Journal)*, (26), 51–72. doi: <https://doi.org/10.7358/ecps-2022-026-fadd>.
- Forrier, B., Boonen, R., and Desmet, W. (2016). Development of a novel test setup for validation of virtual sensing on mechatronic drivetrains. In *International Conference on Powertrain Modelling and Control*.
- García-Zubia, J., Cuadros, J., Romero, S., Hernández-Jayo, U., Orduna, P., Guenaga, M., and Gustavsson, I. (2016). Empirical analysis of the use of the VISIR remote lab in teaching analog electronics. *IEEE Transactions on Education*, 60(2), 149–156. doi: <https://doi.org/10.1109/TE.2016.2608790>.
- García-Guzmán, J., Villa-López, F.H., Vélez-Enríquez, J.A., García-Mathey, L.A., and Ramírez-Ramírez, A. (2017). Remote laboratories for teaching and training in engineering. In *Design, control and applications of mechatronic systems in engineering*. IntechOpen. doi: <https://doi.org/10.5772/67984>.
- Grout, I. (2017). Remote laboratories as a means to widen participation in STEM education. *Education Sciences*, 7(4), 85. doi: <https://doi.org/10.3390/educsci7040085>.
- Hernández-Melgarejo, G., Flores-Hernández, D.A., Luviano-Juárez, A., Castañeda, L.A., Chairez, I., and Di Gennaro, S. (2020). Mechatronic design and implementation of a bicycle virtual reality system. *ISA Transactions*, 97, 336–351. doi: <https://doi.org/10.1016/j.isatra.2019.08.012>.
- Kabaciński, M. and Pawliczek, R. (2015). Mechatronic concept for airflow tests laboratory equipment. *Solid State Phenomena*, 220, 445–450. doi: <https://doi.org/10.4028/www.scientific.net/SSP.220.445>.
- Kebande, V.R. (2024). The impact of virtual laboratories on active learning and engagement in cybersecurity distance education. *arXiv preprint*, arXiv:2404.04952. URL <https://arxiv.org/abs/2404.04952>.
- Kuriščák, P., Rossa, P., Fernandes, H., and Silva, J.N. (2022). Design and implementation of a framework for remote experiments in education. In *2022 8th International Engineering, Sciences and Technology Conference (IESTEC)*, 258–265. doi: <https://doi.org/10.1109/IESTEC54539.2022.00046>.
- Lara-Padilla, H., Sánchez, X.S., and Paucar, T.A. (2017). Design and evaluation of a low-cost mechatronic system to study upper and lower limbs biomechanics. In *2017 IEEE Global Humanitarian Technology Conference (GHTC)*, 1–5. doi: <https://doi.org/10.1109/GHTC.2017.8239292>.
- Li, J. and Liang, W. (2024). Effectiveness of virtual laboratory in engineering education: A meta-analysis. *PLoS ONE*, 19(12), e0316269. doi: <https://doi.org/10.1371/journal.pone.0316269>.
- Melo, T.R., Neto, J.D.R., and Silva, J.J. (2021). Integration of virtual instrumentation in the teaching of data acquisition and interface systems course. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje*, 16(2), 154–160. doi: <https://doi.org/10.1109/RITA.2021.3089918>.
- Mendez Gonzalez, L.C., Rodriguez Picon, L.A., and Quezada Carreon, A.E. (2019). Diseño de un prototipo de estación de trabajo por medio de LabVIEW y SolidWorks. *Instituto de Ingeniería y Tecnología*.
- Moreno-Peñalosa, G., Juárez-Aguirre, R., Portillo-Velez, R.J., Cerón-Álvarez, C.A., López-Huerta, F., and Woo-



- García, R.M. (2021). Alarms touch panel for freezer food preservation industry. In *2021 IEEE 3rd International Conference on Circuits and Systems (ICCS)*, 165–169. doi:<https://doi.org/10.1109/ICCS54091.2021.00035>.
- Petrović, V.M., Nikolić, B., Jovanović, K., and Potkonjak, V. (2016). Development of virtual laboratory for mechatronic systems. In *International Conference on Robotics in Alpe-Adria Danube Region*, 622–630. Springer. doi:https://doi.org/10.1007/978-3-319-49058-8_8.
- Reeves, S.M. and Crippen, K.J. (2021). Virtual laboratories in undergraduate science and engineering courses: A systematic review, 2009–2019. *Journal of Science Education and Technology*, 30(1), 16–30. doi:<https://doi.org/10.1007/s10956-020-09866-0>.
- Reid, D.P., Burridge, J., Lowe, D.B., and Drysdale, T.D. (2022). Open-source remote laboratory experiments for controls engineering education. *International Journal of Mechanical Engineering Education*, 50(4), 828–848. doi:<https://doi.org/10.1177/03064190221081452>.
- Reilly, J.M. and Dede, C. (2019). Differences in student trajectories via filtered time series analysis in an immersive virtual world. In *Proceedings of the 9th International Conference on Learning Analytics & Knowledge*, 130–134. doi:<https://doi.org/10.1145/3303772.3303820>.
- Uzunidis, D. and Pagiatakis, G. (2023). Design and implementation of a virtual on-line lab on optical communications. *European Journal of Engineering Education*, 48(5), 913–928. doi:<https://doi.org/10.1080/03043797.2023.2185761>.
- Velázquez, I.L., Blanco, R.I.R., Torres, F.I.T., Bravo, E.R., and Ortiz, P.M. (2024). Implementación de un sistema de conmutación para medición de efecto Hall usando MyOpenLab. *Pistas Educativas*, 45(147).
- Vázquez-Santos, A., Tejeda-Del-Cueto, M.E., Hernández-Hernández, J., Cerón-Álvarez, C.A., Juárez-Aguirre, R., López-Huerta, F., and Woo-García, R.M. (2021). Graphical user interface for marine propeller optimum diameter estimation using MyOpenLab. In *2021 IEEE International Conference on Engineering Veracruz (ICEV)*, 1–5. doi:<https://doi.org/10.1109/ICEV52944.2021.9632789>.

Technical Proposal for an Automated Liquid Filling System with Integrated Computer Vision and Water Quality Monitoring

GF Pérez-García ** IE Juárez-Martínez *
SM Marquez-Marina * KM Velázquez-Natarén *
A de Jesús Flores-García * AZ Aguilar-Sántiz **

* Tecnológico Nacional de México / Instituto Tecnológico de Tuxtla
Gutiérrez, Tuxtla Gutiérrez, Chiapas, México.

** Facultad de Ciencias en Física y Matemáticas, Universidad
Autónoma de Chiapas, Tuxtla Gutiérrez, Chiapas, México.

Abstract

In many micro, small, and medium-sized enterprises, as well as in small-scale production units dedicated to the handling and filling of liquids, container filling is still carried out manually or semi-manually. This procedure causes variations in the dispensed volume, increases operating times, and raises the risk of product losses. When these activities are related to water applications, it becomes especially important to achieve better control of the filling process, perform basic process inspections, and monitor variables associated with water quality.

This work presents a technical proposal for the development of an automated liquid filling system. In addition to the automated filling function, the proposal includes process inspection through computer vision and water-quality monitoring. The contribution of the study corresponds to the initial design phase of the project. In this phase, the main function of the system, the operating modules, the physical and functional components associated with them, and the overall sequence of operation are defined in an integrated manner.

The scope of the work does not include the manufacture or commissioning of the system; however, the study provides a coherent technical basis to guide detailed design, physical integration, and future evaluation. In addition, preliminary quantitative criteria are defined for subsequent validation, including filling cycle time, dosing error, repeatability, visual inspection accuracy, and stability of water-quality measurements under operating conditions compatible with hygiene and sanitation requirements.

Keywords: automation, liquid filling, visual inspection, water quality monitoring, technical proposal, sensors

1. INTRODUCCIÓN

En México, las micro, pequeñas y medianas empresas (MiPyMEs) representan una parte importante de la actividad productiva nacional. De acuerdo con el INEGI, durante 2023 operaron más de 5.4 millones de unidades económicas en el sector privado y paraestatal, de las cuales la mayoría correspondió a microempresas (Instituto Nacional de Estadística y Geografía (INEGI), 2025). Este contexto muestra que muchos procesos de producción todavía se realizan con recursos limitados y con poca disponibilidad para incorporar sistemas industriales de automatización de alto costo.

En pequeñas unidades dedicadas al envasado de agua purificada, bebidas artesanales, productos de limpieza diluidos o soluciones líquidas de uso local, el llenado de recipientes suele realizarse de forma manual o semimanual. Aunque esta forma puede ser suficiente para bajos volúmenes de producción, también genera problemas de variación en

la cantidad dispensada, mayores tiempos de operación, derrames, fatiga del operario y baja repetibilidad. Estas condiciones dificultan la estandarización del proceso y reducen la posibilidad de supervisar de manera sistemática cada ciclo de llenado.

Cuando el proceso involucra líquidos de manejo controlado, no basta con dosificar un volumen determinado. También resulta necesario verificar condiciones básicas del envase, revisar el nivel aparente de llenado y registrar variables asociadas con la calidad del agua. Existen soluciones para llenado automatizado, sistemas de inspección visual y plataformas de monitoreo de agua; sin embargo, con frecuencia estas funciones se implementan por separado. Esta separación incrementa la complejidad de integración, eleva los costos de implementación y dificulta su adopción en aplicaciones de pequeña escala.

En el caso de sistemas relacionados con agua para consumo humano, el diseño posterior debe considerar criterios sanitarios aplicables. En México, una referencia directa es la NOM-201-SSA1-2015, correspondiente a agua y hielo para

* Trabajo propuesto para congreso CITCA-CENIDET.



consumo humano envasados y a granel, la cual establece especificaciones sanitarias para estos productos y para los establecimientos dedicados a su proceso o importación (Secretaría de Salud, 2015). Por ello, aunque este trabajo se ubica en una fase conceptual, la propuesta considera desde el inicio aspectos relacionados con higiene, selección de materiales, limpieza, operación y control sanitario.

Bajo esta lógica, el objetivo de este trabajo es desarrollar una propuesta técnica para un sistema automatizado de dosificación e inspección visual de líquidos, complementado con el monitoreo de variables asociadas con la calidad del agua. La contribución principal consiste en proponer una arquitectura funcional integrada que articula, en un mismo sistema de pequeña escala, la dosificación automatizada, la inspección visual del llenado y el monitoreo básico del fluido de alimentación.

La originalidad de la propuesta no se encuentra en cada subsistema por separado, ya que existen antecedentes sobre llenado, visión artificial y monitoreo de agua, sino en la forma en que estos elementos se organizan dentro de una misma secuencia operativa. La arquitectura planteada permite relacionar cada ciclo de llenado con tres elementos: el estado operativo del envase, la decisión de inspección visual y las variables medidas del agua. De esta manera, el trabajo aporta una base técnica para el diseño detallado y la futura construcción del prototipo, al definir la organización general del sistema, su secuencia de operación, los criterios de integración de componentes y los indicadores preliminares para evaluar su desempeño. Con ello, se busca orientar el desarrollo de soluciones automatizadas compactas para entornos donde el llenado, la inspección y la supervisión del agua todavía se realizan de forma manual, semimanual o mediante sistemas separados.

2. ESTADO DEL ARTE

Los trabajos previos sobre llenado automatizado de líquidos se han enfocado principalmente en controlar la dosificación mediante estrategias basadas en tiempo, nivel, volumen o bombeo. Estas soluciones permiten mejorar la uniformidad del llenado, reducir pérdidas y aumentar la productividad; sin embargo, suelen centrarse en el suministro del líquido y no siempre integran funciones complementarias como inspección visual o monitoreo del fluido de alimentación (Musíc et al., 2023; Huang et al., 2018).

En visión artificial, Huang et al. (2018) desarrollaron soluciones para localización, seguimiento y detección de defectos en botellas vacías bajo condiciones industriales. Por su parte, Musíc et al. (2023) abordaron la estimación visual del nivel de líquido durante el embotellado en aplicaciones de menor escala, mientras que Wu et al. (2023) propusieron un *benchmark* para la detección de contenido líquido en contenedores transparentes. De manera general, Tzampazaki et al. (2024) muestran la madurez de la visión artificial en entornos industriales. Aunque estos trabajos aportan soluciones relevantes para detección, nivel o defectos, su alcance se mantiene principalmente en el subsistema visual y no en su integración con una estación compacta de llenado y monitoreo del agua.

De forma paralela, los sistemas de monitoreo de calidad del agua han avanzado en la adquisición de variables fisi-

coquímicas. El trabajo de Geetha y Gouthami (Geetha and Gouthami, 2016) integraron sensores, electrónica y comunicación IoT para monitoreo en tiempo real; Nanyanzi et al. (2021) desarrollaron un sistema doméstico para temperatura, pH y turbidez; y Jayaraman et al. (2024) revisaron el uso de sensores, IoT y aprendizaje automático para análisis de calidad del agua. No obstante, estos enfoques suelen tratar el monitoreo como un sistema independiente, sin vincularlo directamente con decisiones de llenado, inspección visual o trazabilidad por envase.

A partir de lo anterior, se observa que los antecedentes han resuelto problemas específicos, pero de manera separada: dosificación controlada, inspección visual o adquisición de variables del agua. La limitación técnica que permanece abierta es la falta de propuestas compactas que articulen estos subsistemas dentro de una arquitectura funcional orientada a pequeña escala. En este sentido, la propuesta desarrollada busca integrar dosificación automatizada, inspección visual y monitoreo básico del agua en una misma secuencia operativa, de modo que cada ciclo pueda relacionarse con el estado del envase, la decisión de inspección y las variables del fluido de alimentación. La comparación técnica se resume en la Tabla 1.

3. METODOLOGÍA

La metodología se centra en estructurar una propuesta técnica para un sistema automatizado de dosificación e inspección visual de líquidos, complementado con monitoreo básico de calidad del agua. El sistema se organizó en siete subsistemas: estructura mecánica, transporte, sensado, dosificación, inspección visual, control y monitoreo del agua. La selección preliminar de tecnologías se realizó considerando disponibilidad, facilidad de integración, escalabilidad, costo relativo y pertinencia funcional para una aplicación de pequeña escala. Entre las alternativas consideradas se incluyen microcontroladores o PLC compactos para control, sensores fotoeléctricos o de proximidad para detección, válvula solenoide o bomba peristáltica para dosificación, cámara USB o cámara compacta para inspección, y sensores de pH, turbidez, temperatura y conductividad para monitoreo.

El procedimiento se organizó en seis actividades.

Definición de requerimientos funcionales: Se establecieron como requerimientos mínimos el transporte del envase, la detección de presencia, la dosificación controlada, la inspección básica del proceso, el monitoreo de variables del agua y la compatibilidad higiénica de los componentes en contacto con el fluido. Esta definición permitió delimitar el alcance técnico de la propuesta y organizar la arquitectura general del sistema sin sobredimensionar la solución para una aplicación de pequeña escala.

Estructuración funcional de la propuesta: Con base en los requerimientos definidos, el sistema se organizó en tres niveles. El primero corresponde a la capa de campo, donde se ubican sensores y actuadores; el segundo corresponde a la capa de control, encargada de coordinar la secuencia de operación; y el tercero corresponde a la capa de inspección y monitoreo, orientada a verificar el proceso mediante visión artificial y a registrar variables asociadas

Tabla 1. Comparación técnica entre trabajos revisados y propuesta

Trabajo	Función principal	Escala	Diferencia frente a la propuesta
Huang et al. (2018)	Inspección visual de botellas vacías	Industrial	Resuelve detección y seguimiento visual de defectos, pero no integra llenado, nivel de líquido ni monitoreo del agua.
Musić et al. (2023)	Estimación visual del nivel de líquido	Pequeña escala	Se aproxima al problema de nivel en embotellado, pero no articula transporte, dosificación, decisión automática y variables del agua.
Wu et al. (2023)	Detección de contenido líquido en recipientes transparentes	Computación	Aporta una base de evaluación algorítmica, pero no se orienta a una estación física de llenado ni a trazabilidad por ciclo.
Geetha and Gouthami (2016)	Monitoreo IoT de calidad del agua	Instrumental	Integra sensores y comunicación para supervisión del agua, pero no se vincula con llenado ni inspección visual del envase.
Nanyanzi et al. (2021)	Monitoreo doméstico de temperatura, pH y turbidez	Doméstica	Presenta una solución compacta de medición, pero independiente de un proceso automatizado de dosificación.
Jayaraman et al. (2024)	Revisión de IoT y aprendizaje automático en calidad del agua	Revisión	Sistematiza sensores y modelos, pero no propone integración física con llenado e inspección de envases.
Propuesta desarrollada	Llenado automatizado, inspección visual y monitoreo básico del agua	Pequeña escala	Integra los tres bloques en una misma secuencia operativa y define indicadores para validación cuantitativa posterior.

con el agua. La estructura general propuesta se muestra en la Figura 1.

Esta organización permite visualizar el sistema como una arquitectura integrada, pero todavía modular. De esta forma, cada bloque puede desarrollarse, ajustarse o sustituirse posteriormente sin modificar por completo la lógica general de operación.

Definición de la secuencia funcional de operación:

La operación se planteó como una secuencia discreta de eventos. Primero, el sistema detecta el ingreso del envase; después verifica su posición; si la posición es correcta, detiene la banda y habilita la dosificación. Al finalizar el llenado, se activa la inspección visual y se registran las variables del agua disponibles. Si el envase cumple con las condiciones establecidas, continúa hacia la salida; en caso contrario, se separa para revisión, ajuste de nivel, reproceso o disposición segura. Esta decisión evita considerar automáticamente el rechazo como pérdida de producto y permite mantener trazabilidad sobre las incidencias del proceso. La secuencia propuesta se resume en la Figura 2.

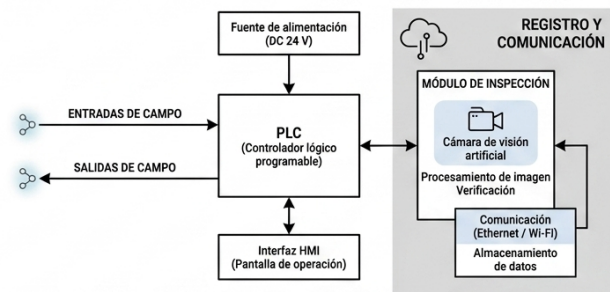


FIGURA 1: SISTEMA DE DATOS Y GESTIÓN

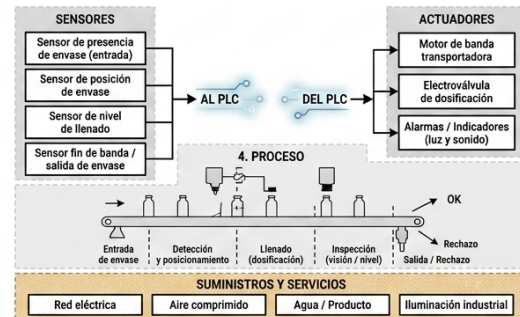


Figura 1. Estructura general propuesta para el sistema automatizado de llenado e inspección.

La decisión de aceptación del envase puede representarse mediante una regla simple:

$$D = \begin{cases} 1, & \text{si } P_e = 1, P_o = 1, N_l = 1 \\ 0, & \text{en caso contrario} \end{cases} \quad (1)$$

donde D representa la decisión de aceptación, P_e la presencia detectada del envase, P_o la posición correcta y N_l el nivel aparente dentro del criterio definido. Cuando $D = 1$, el envase continúa hacia la salida; cuando $D = 0$, el sistema genera una incidencia o dirige el recipiente a revisión. Esta regla permite que el ciclo no dependa únicamente de temporizaciones abiertas, sino de condiciones verificables dentro del proceso.

Definición del esquema de inspección visual: El módulo de visión se planteó para verificar tres condiciones básicas: presencia del envase, posición en la zona de inspección y nivel aparente del líquido. Para ello, se consideró una cámara fija, iluminación controlada y una región de interés delimitada sobre la trayectoria del recipiente. Esta configuración permite iniciar con algoritmos de visión clásica, como detección de contornos, segmentación y medición geométrica, antes de avanzar hacia modelos de clasificación más complejos.

La decisión de iniciar con visión artificial clásica responde a una lógica de implementación gradual. Para una primera etapa, no se requiere necesariamente una red neuronal o un modelo de aprendizaje profundo, sino un esquema estable que permita verificar condiciones visibles del proceso. Posteriormente, el mismo módulo puede ampliarse hacia algoritmos de clasificación si las pruebas experimentales muestran que existe variabilidad significativa por tipo de envase, iluminación o nivel de líquido.

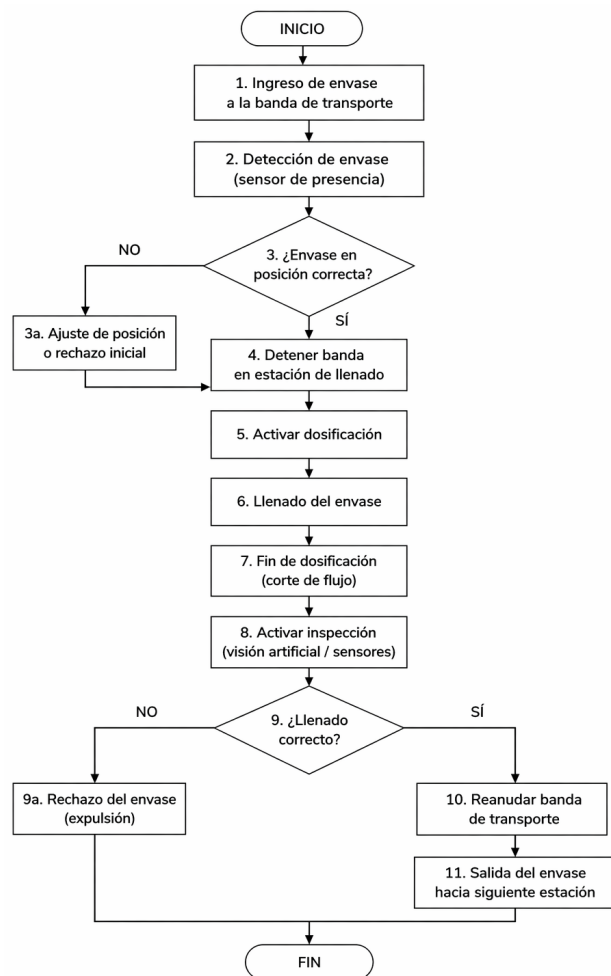


Figura 2. Diagrama de flujo de la secuencia de operación del sistema automatizado de llenado e inspección.

Definición del esquema de monitoreo del agua: El monitoreo se planteó en dos puntos: el depósito principal y la línea de salida hacia la estación de llenado. El depósito permite observar el comportamiento general del agua almacenada, mientras que la línea de salida ofrece una referencia más cercana al punto real de dosificación. Las variables consideradas fueron pH, turbidez, temperatura y conductividad eléctrica, por su uso frecuente como indicadores de seguimiento en sistemas de monitoreo de agua y por su factibilidad de integración en plataformas de pequeña escala.

Este subsistema se concibe como una conexión entre sensores, controlador e interfaz de supervisión. Su función no es sustituir un laboratorio de análisis de calidad del agua, sino incorporar una capa básica de seguimiento operativo que pueda asociarse con cada ciclo de llenado. De esta manera, las mediciones del agua pueden relacionarse posteriormente con eventos del proceso, alarmas, lotes o incidencias.

Finalmente, se definieron los criterios que permitirán evaluar el sistema en una fase posterior de prueba. Para ello, se consideraron aspectos directamente relacionados con

la operación del prototipo, como el tiempo de ciclo, la capacidad de llenado, el error de volumen, la repetibilidad entre envases, el desempeño de la inspección visual y la estabilidad de las mediciones del agua. Estos criterios se retoman en la sección de resultados, donde se presenta una estimación preliminar del desempeño esperado.

4. RESULTADOS

Como resultado de la metodología propuesta, se obtuvo una arquitectura técnica orientada a integrar llenado automatizado, inspección visual y monitoreo de variables asociadas con la calidad del agua. Los resultados se presentan en cinco bloques: estructura general de llenado, módulo de visión, monitoreo de calidad del agua, integración global y estimación preliminar para validación cuantitativa. Esta organización permite describir la configuración del sistema y, al mismo tiempo, establecer los indicadores con los que podrá evaluarse posteriormente.

4.1 Estructura general del sistema de llenado

El primer bloque corresponde a la definición de la estructura general del sistema de llenado, organizada a partir de los elementos de estructura mecánica, transporte, detección, dosificación, control e interfaz de supervisión. Esta organización se planteó bajo una lógica discreta de estados, con el propósito de asociar cada evento del proceso con una transición definida y reducir la dependencia de temporizaciones abiertas. Bajo esta lógica, el sistema puede representarse mediante una máquina de estados finitos o implementarse mediante programación secuencial en entornos de automatización industrial.

Desde el punto de vista físico, la propuesta considera una base compacta con banda transportadora lineal, depósito de alimentación, estación de llenado, sensores de presencia y módulo de control. Esta disposición busca separar funcionalmente las etapas de transporte y dosificación, lo que proporciona una base más estable para la operación y facilita la incorporación posterior de subsistemas de inspección y monitoreo. La banda transportadora se planteó con dimensiones compatibles con recipientes de 250 a 1000 mL, manteniendo una configuración compacta adecuada tanto para entornos de laboratorio como para aplicaciones de pequeña producción. La elección de este rango se orientó a recipientes de uso común, tanto en PET transparente como en vidrio, sin comprometer la visibilidad del contenido ni el acceso a la estación de llenado.

En la parte mecánica, se consideró una estructura principal fabricable en perfil de aluminio tipo Bosch o materiales equivalentes, con una altura de trabajo situada entre 0.80 y 1.00 m, con el fin de mantener rigidez estructural, facilidad de manufactura y accesibilidad para tareas de montaje. En los puntos de contacto con el líquido se contemplaron materiales con comportamiento compatible frente a humedad, limpieza y exposición continua, particularmente acero inoxidable y polímeros de grado alimenticio. Los parámetros generales definidos para esta etapa de formulación se resumen en la Tabla 2.

En la parte de accionamiento, el movimiento de la banda se asoció a un motor de arrastre gobernado desde la unidad de control. Para una primera implementación se



Tabla 2. Parámetros de integración definidos para la propuesta de llenado

Parámetro	Valor definido
Capacidad de recipiente considerada	250–1000 mL
Material de recipiente considerado	PET transparente o vidrio
Longitud de banda	0.80–1.20 m
Ancho de banda	0.10–0.20 m
Altura de trabajo del sistema	0.80–1.00 m
Separación entre sensor de entrada y estación de llenado	0.15–0.30 m
Distancia boquilla–envase	0.02–0.06 m
Espacio para estación de inspección	0.10–0.20 m
Separación entre envases en operación	0.08–0.15 m
Material de estructura	Perfil de aluminio o acero al carbón recubierto
Material en contacto con el líquido	Acero inoxidable o polímero de grado alimenticio

consideraron dos alternativas: una basada en microcontrolador, adecuada para desarrollo rápido y pruebas iniciales, y otra basada en PLC compacto, más conveniente para una operación robusta y escalable. Esta doble posibilidad permite mantener flexibilidad tecnológica sin modificar la lógica principal de operación.

La función de dosificación se planteó mediante un esquema de suministro controlado que admite dos configuraciones principales: válvula solenoide temporizada o bomba peristáltica de bajo caudal. La primera alternativa resulta adecuada cuando existe apoyo por gravedad desde el depósito principal y se busca una estructura de menor complejidad; la segunda puede ofrecer mayor repetibilidad de entrega cuando se requiere un control más preciso del caudal. La elección entre ambas depende del nivel de precisión requerido, del presupuesto disponible y de la facilidad de mantenimiento.

También se contempla sensores fotoeléctricos o de proximidad para detectar el envase, un motor de transporte para desplazar el recipiente, una válvula solenoide o bomba peristáltica para controlar el suministro del líquido, una boquilla de acero inoxidable y una interfaz básica de supervisión para mostrar estados, alarmas y variables del proceso. En cuanto a la programación, para plataformas embebidas, la implementación puede realizarse en C/C++, si se considera PLC, mediante Ladder Diagram o Structured Text. Para la capa de supervisión se utiliza Python, lo que permite registrar estados, visualizar variables y establecer comunicación con módulos externos. Esta organización asegura la separación entre el control en tiempo real y las funciones de visualización y análisis, lo que aporta mayor eficiencia y claridad en la operación del sistema.

4.2 Módulo de visión para inspección del proceso

El segundo bloque corresponde a la definición del módulo de visión. La disposición física del sistema de visión se definió mediante una estación fija situada a la salida de la

zona de llenado, con un campo visual orientado a capturar tanto el contorno del envase como la altura visible del contenido, bajo condiciones de iluminación controlada. Para la captura de imágenes se consideran cámaras USB, Raspberry Pi Camera o cámaras industriales compactas, acompañadas de iluminación frontal o retroiluminación difusa.

Para la inspección del proceso, se plantea tres tareas principales: confirmar la presencia del recipiente, verificar su posición en la zona de inspección y estimar de manera visual el nivel aparente del líquido. El flujo de inspección se dividió en cinco etapas: adquisición de imagen, preprocesamiento, segmentación, extracción de características y clasificación del estado del envase. Durante la adquisición, la cámara captura una imagen cuando el sistema de control confirma que el recipiente ha alcanzado la posición de inspección. Posteriormente, la imagen se acondiciona mediante operaciones de filtrado y normalización de contraste. En la etapa de segmentación se aísla la región correspondiente al envase y al contenido; después se extraen variables geométricas y visuales. Cuando el análisis visual confirma la presencia del envase con un nivel de llenado acorde con el criterio definido, el sistema habilita la reanudación del transporte. En caso contrario, puede generar una alarma o detener el proceso para su análisis posterior. De esta manera, el módulo de visión incorpora una capa adicional de verificación técnica dentro de la propuesta general.

Esta estrategia permite trabajar inicialmente con algoritmos de visión artificial clásica. En este esquema, la presencia del envase puede resolverse mediante detección de contornos o validación de ocupación en una región de interés; la posición puede estimarse a partir del centroide o del alineamiento del perfil del recipiente; y el nivel aparente puede calcularse a partir de la altura visible de la columna de líquido en la zona de análisis. Esta ruta de procesamiento se consideró adecuada para una primera implementación porque mantiene moderada la carga computacional y facilita la integración del sistema en plataformas embebidas o de escritorio.

En la parte de software, la implementación del módulo de visión se asoció principalmente a Python, utilizando bibliotecas como OpenCV, mediante las cuales el sistema puede ejecutar operaciones de umbralización, detección de bordes, análisis de contornos y medición geométrica. En etapas posteriores, la misma estructura puede ampliarse hacia modelos de clasificación basados en aprendizaje automático o redes neuronales convolucionales; no obstante, la propuesta actual se concentró en una ruta inicial de inspección primaria compatible con una futura implementación gradual.

4.3 Subsistema de monitoreo de la calidad del agua

El tercer bloque corresponde a la definición del subsistema de monitoreo de calidad del agua. La incorporación de esta función amplía el alcance de la propuesta, ya que añade una capa de supervisión sobre el fluido de alimentación y complementa la integración entre dosificación e inspección.

La estructura de monitoreo se planteó sobre dos puntos principales de medición: el depósito principal y la línea de salida hacia la estación de llenado. El depósito principal



permite observar el comportamiento del agua almacenada bajo condiciones relativamente estables, mientras que la línea de salida ofrece una referencia más cercana al punto real de dosificación. Esta doble ubicación de sensado permite cubrir dos zonas relevantes del proceso y vincular el estado del agua con el comportamiento del sistema en distintas etapas del flujo.

Las variables consideradas para el monitoreo fueron pH, turbidez, temperatura y conductividad eléctrica. Estas variables permiten construir un perfil básico del agua de alimentación y resultan adecuadas para plataformas de adquisición de datos de bajo y mediano nivel de complejidad. El pH aporta información sobre la condición ácido-base del líquido; la turbidez refleja cambios asociados con partículas suspendidas y claridad; la temperatura introduce una referencia importante para estabilidad operativa y compensación de medición; y la conductividad eléctrica permite observar variaciones globales del contenido iónico.

Para su integración, el pH puede medirse en el depósito principal o en la línea de salida mediante sondas de inmersión o inserción en línea, incluyendo módulos tipo Gravity Analog pH Sensor, sondas con acondicionamiento Atlas Scientific o equivalentes industriales. La turbidez puede supervisarse con sensores ópticos de salida analógica o digital, instalados en el depósito o en una celda de flujo. La temperatura puede registrarse con sondas encapsuladas DS18B20 para una integración sencilla, o con sensores RTD tipo PT100/PT1000 cuando se requiera mayor robustez. Finalmente, la conductividad eléctrica puede medirse mediante celdas sumergibles o en línea, preferentemente con módulos que permitan compensación térmica. Esta selección permite construir un perfil básico del agua de alimentación sin añadir una complejidad excesiva al sistema, manteniendo compatibilidad con arquitecturas embebidas o industriales.

La comunicación puede realizarse mediante UART, RS-485 o interfaces seriales hacia una computadora de supervisión. En software, la adquisición local puede implementarse en C/C++, mientras que la visualización, el registro histórico y la asociación de lecturas con eventos del proceso pueden desarrollarse en Python, MATLAB o LabVIEW. Con esta organización, el monitoreo del agua se integra al sistema de llenado sin interferir con el control principal y deja abierta la posibilidad de incorporar alarmas, seguimiento por lote y análisis histórico del proceso.

En conjunto, los tres bloques anteriores permiten definir una arquitectura en la que el llenado, la inspección visual y el monitoreo del agua no operan como elementos aislados, sino como partes de una misma secuencia. Esto permite asociar cada ciclo con el estado del envase, la decisión de inspección y las variables medidas del fluido, lo cual sirve como base para proponer indicadores de evaluación cuantitativa.

4.4 Estimación preliminar y criterios para validación cuantitativa

Como complemento a esta propuesta, se realizó una estimación analítica inicial del desempeño esperado y se definieron indicadores para la validación posterior. Esta estimación no sustituye una prueba experimental, pero

permite relacionar la arquitectura propuesta con variables medibles de operación.

Para la función de llenado, el tiempo de ciclo puede expresarse como:

$$T_c = T_p + T_d + T_i + T_s \quad (2)$$

donde T_c es el tiempo total de ciclo, T_p el tiempo de posicionamiento, T_d el tiempo de dosificación, T_i el tiempo de inspección y T_s el tiempo de salida del envase. A partir de este valor, la capacidad teórica de operación se estima mediante:

$$N_h = \frac{3600}{T_c} \quad (3)$$

donde N_h representa el número estimado de envases procesados por hora. Para una primera referencia numérica, se consideró un tiempo de posicionamiento de 1.5 s, un tiempo de inspección de 0.5 s, un tiempo de salida de 1.0 s y un caudal nominal de 50 mL/s. Bajo estas condiciones, el tiempo de dosificación se calculó como $T_d = V_r/Q_f$. La Tabla 3 presenta la estimación para recipientes de 250, 500 y 1000 mL.

Tabla 3. Estimación preliminar del tiempo de ciclo y capacidad operativa

Volumen	Tiempo de llenado	Tiempo de ciclo	Capacidad estimada
250 mL	5.0 s	8.0 s	450 envases/h
500 mL	10.0 s	13.0 s	277 envases/h
1000 mL	20.0 s	23.0 s	157 envases/h

La estimación muestra que la capacidad operativa depende principalmente del volumen objetivo y del caudal disponible. Para observar esta relación con mayor claridad, se realizó un análisis de sensibilidad simple considerando un recipiente de 500 mL y una variación de caudal de $\pm 10\%$ respecto al valor nominal de $Q_f = 50$ mL/s. Para este caso, el tiempo de ciclo se expresó como:

$$T_c(Q_f) = T_f + \frac{V_r}{Q_f} \quad (4)$$

donde T_f agrupa los tiempos fijos de posicionamiento, inspección y salida del envase, mientras que V_r/Q_f representa el tiempo de dosificación. En este caso, se consideró $T_f = 3.0$ s, correspondiente a 1.5 s de posicionamiento, 0.5 s de inspección y 1.0 s de salida. La capacidad operativa se calculó nuevamente mediante:

$$N_h(Q_f) = \frac{3600}{T_c(Q_f)} \quad (5)$$

Con un caudal de 45 mL/s, el tiempo de ciclo sería aproximadamente de 14.1 s, equivalente a 255 envases/h. Con el caudal nominal de 50 mL/s, el ciclo sería de 13.0 s, equivalente a 277 envases/h. Finalmente, con 55 mL/s, el ciclo disminuiría a 12.1 s, equivalente a 298 envases/h. Estos valores muestran que una variación moderada del caudal puede afectar la capacidad operativa, por lo que la



calibración de la bomba o válvula será un punto importante durante la validación experimental.

Para evaluar el volumen dosificado, se propone usar el error porcentual:

$$E_v = \frac{|V_m - V_r|}{V_r} \times 100 \quad (6)$$

donde V_m es el volumen medido y V_r el volumen de referencia. La repetibilidad entre ciclos puede analizarse mediante la desviación estándar:

$$\sigma_v = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (V_i - \bar{V})^2} \quad (7)$$

donde V_i representa el volumen obtenido en el ciclo i , $\bar{V}$ el volumen promedio y n el número de envases evaluados. Además, si el sistema opera con una resolución temporal de 0.1 s y un caudal de 50 mL/s, el cambio mínimo de volumen sería $\Delta V = 5$ mL. Esto equivale a un error relativo aproximado de 2.0 % para 250 mL, 1.0 % para 500 mL y 0.5 % para 1000 mL, lo que muestra la importancia de calibrar con mayor cuidado los envases pequeños.

En el módulo de visión artificial, la evaluación puede realizarse mediante el porcentaje de aciertos:

$$A_v = \frac{N_c}{N_t} \times 100 \quad (8)$$

donde N_c es el número de clasificaciones correctas y N_t el número total de inspecciones. Para la estimación del nivel aparente, si una región útil de 400 píxeles representa 1000 mL, la relación aproximada sería de 2.5 mL/px. Con una incertidumbre de detección de borde de ± 2 píxeles, la variación equivalente sería cercana a ± 5 mL. Por ello, el módulo de visión puede funcionar como verificación básica de subllenado, sobrellenado o ausencia de líquido, siempre que se controle la iluminación y la posición del envase.

Para el monitoreo del agua, la validación deberá considerar la estabilidad de lectura de pH, turbidez, temperatura y conductividad eléctrica. En una implementación experimental, estas variables podrán registrarse durante varios ciclos para analizar fluctuaciones, ruido de medición y posibles diferencias entre lecturas en depósito y en línea de salida.

Con base en estas estimaciones, la Tabla 4 resume los indicadores que se proponen para evaluar el sistema cuando se avance a la etapa experimental.

5. CONCLUSIONES

En este trabajo se presentó el diseño conceptual de un sistema automatizado para llenado de líquidos en recipientes, incorporando inspección visual y monitoreo básico de variables asociadas con la calidad del agua. La propuesta se orientó a aplicaciones de pequeña escala, donde el llenado aún puede realizarse de forma manual o semimanual y donde una solución industrial completa puede resultar costosa o poco flexible.

Tabla 4. Criterios preliminares para validación cuantitativa

Subsistema	Indicador	Interpretación técnica
Llenado	Tiempo de ciclo T_c	Estima la capacidad operativa en envases por hora.
Llenado	Error porcentual E_v	Mide la diferencia entre volumen objetivo y volumen dosificado.
Llenado	Desviación estándar σ_v	Evalúa la repetibilidad entre ciclos consecutivos.
Visión artificial	Porcentaje de acierto A_v	Mide decisiones correctas sobre presencia, posición y nivel aparente.
Visión artificial	Error de nivel aparente	Compara la estimación visual contra una referencia física.
Monitoreo del agua	Estabilidad de lectura	Analiza fluctuaciones de pH, turbidez, temperatura y conductividad.
Integración operativa	Registro por ciclo	Relaciona cada envase con inspección visual y variables del agua.

El principal aporte del trabajo fue organizar en una misma arquitectura los bloques de transporte, dosificación, visión artificial, control y monitoreo del fluido. Esta integración permite que el proceso no se limite únicamente a llenar un envase, sino que también pueda verificar condiciones básicas del recipiente, estimar el nivel aparente del líquido y registrar variables como pH, turbidez, temperatura y conductividad. Con ello se plantea una base para mejorar la trazabilidad del proceso y relacionar cada ciclo de llenado con información operativa útil.

En cuanto a la parte técnica, se definieron parámetros iniciales de diseño, como el rango de capacidad de los recipientes, las dimensiones generales de la banda, la separación entre estaciones, la distancia boquilla-envase y las alternativas de control y dosificación. También se establecieron criterios preliminares para evaluar el desempeño del sistema, entre ellos el tiempo de ciclo, la capacidad operativa, el error porcentual de llenado, la repetibilidad, el porcentaje de acierto del módulo de visión y la estabilidad de las mediciones del agua.

La estimación analítica realizada permitió obtener una primera referencia del comportamiento esperado del sistema. Para los escenarios considerados, el tiempo de ciclo y la capacidad operativa dependen principalmente del volumen objetivo y del caudal disponible. Además, el análisis de sensibilidad mostró que variaciones moderadas en el caudal pueden afectar la productividad estimada, por lo que la calibración de la bomba o válvula será un aspecto importante en la etapa experimental.

Como trabajo futuro, será necesario construir el prototipo y validar experimentalmente los indicadores propuestos. En particular, deberán medirse la precisión y repetibilidad del llenado, el desempeño del módulo de visión ante distintos niveles de líquido e iluminación, y la estabilidad de las lecturas de pH, turbidez, temperatura y conductividad. Estos resultados permitirán ajustar los parámetros de control, valorar la viabilidad de la arquitectura y determinar su aplicación en pequeñas unidades productivas.



REFERENCIAS

- Geetha, S. and Gouthami, S. (2016). Internet of things enabled real time water quality monitoring system. *Smart Water*, 2, 1. doi:10.1186/s40713-017-0005-y.
- Huang, B., Ma, S., Wang, P., Wang, H., Yang, J., Guo, X., Zhang, W., and Wang, H. (2018). Research and implementation of machine vision technologies for empty bottle inspection systems. *Engineering Science and Technology, an International Journal*, 21(1), 159–169. doi:10.1016/j.jestch.2018.01.004. URL <https://www.sciencedirect.com/science/article/pii/S2215098617310194>.
- Instituto Nacional de Estadística y Geografía (INEGI) (2025). Estadísticas a propósito del día de las micro, pequeñas y medianas empresas (mipymes). URL https://www.inegi.org.mx/contenidos/saladeprensa/aproposito/2025/EAP_MIPYMES_25.pdf. Comunicado de prensa 71/25.
- Jayaraman, P., Nagarajan, K.K., Partheeban, P., and Krishnamurthy, V. (2024). Critical review on water quality analysis using iot and machine learning models. *International Journal of Information Management Data Insights*, 4(1), 100210. doi:10.1016/j.jjimei.2023.100210.
- Musić, J., Stančić, I., Džaja, B., and Pekić, V. (2023). Image-based sensor for liquid level monitoring during bottling with application to craft and home-brewing. *Sensors*, 23(16), 7126. doi:10.3390/s23167126. URL <https://www.mdpi.com/1424-8220/23/16/7126>.
- Nanyanzi, D.R., Ocen, G.G., Omara, T., Bwire, F., Matovu, D., and Semwogerere, T. (2021). Design and assembly of a domestic water temperature, ph and turbidity monitoring system. *BMC Research Notes*, 14, 161. doi:10.1186/s13104-021-05578-9.
- Secretaría de Salud (2015). Norma oficial mexicana nom-201-ssa1-2015, productos y servicios. agua y hielo para consumo humano, envasados y a granel. especificaciones sanitarias. URL https://dof.gob.mx/nota_detalle_popup.php?codigo=5420977. Diario Oficial de la Federación, publicada el 22 de diciembre de 2015.
- Tzampazaki, M., Zografos, C., Vrochidou, E., and Papakostas, G.A. (2024). Machine vision—moving from industry 4.0 to industry 5.0. *Applied Sciences*, 14(4), 1471. doi:10.3390/app14041471. URL <https://www.mdpi.com/2076-3417/14/4/1471>.
- Wu, Y., Ye, H., Yang, Y., Wang, Z., and Li, S. (2023). Liquid content detection in transparent containers: A benchmark. *Sensors*, 23(15), 6656. doi:10.3390/s23156656. URL <https://www.mdpi.com/1424-8220/23/15/6656>.

Inverse Kinematics Verification and Validation of a 3-RPS Parallel Robot for a Cervical Orthosis Prototype for Rehabilitation

E. Sánchez-Gómez *, M. A. Chagolla-Aranda *, A. Tlaltelpa-Becerro **, J. Colin-Ocampo *, E. Rosado-Tamariz *, J. Y. Colín-Castillo *

*TecnM/CENIDET, Campus Apatzingan 2, Cuernavaca Morelos, México C.P. 62490 (Tel: 5624450186; e-mail: m25ce033@cenidet.tecnm.mx, miguel.ca@cenidet.tecnm.mx, jorge.co@cenidet.edu.mx, erik.rtz@cenidet.tecnm.mx, d18ce085@cenidet.tecnm.mx).

** Departamento de Ingeniería Robótica, Escuela de Estudios Superiores de Yecapixtla, Universidad Autónoma del Estado de Morelos, Yecapixtla Morelos 62824, México

Abstract: Robotic devices for cervical rehabilitation require accurate kinematic models to ensure safe and precise motion control, particularly for applications involving human-robot interaction. This paper presents the analytical formulation, numerical verification, and experimental validation of the inverse kinematics model for a 3-RPS (Revolute-Prismatic-Spherical) parallel robot designed as a prototype for an active cervical orthosis. The inverse kinematics model was derived analytically based on the geometric configuration of the mechanism, utilizing Euler angle representations (Roll-Pitch-Yaw) to relate the angular orientation of the mobile platform to the required linear displacements of the three actuated links. To ensure the consistency of the geometric formulation prior to physical implementation, the model was verified through numerical simulations performed in ADAMS View software. The lengths obtained from the analytical model were used as inputs to the CAD simulation, and the resulting platform orientations were compared against theoretical targets. Subsequently, experimental validation was conducted using a physical prototype equipped with linear actuators controlled via an Arduino UNO and monitored through a LabVIEW interface. A digital inclinometer was employed to measure the actual angular displacements of the mobile platform. The experimental results demonstrated a high degree of concordance between the theoretical predictions and the measured positions, with angular errors remaining below 0.3 degrees. These discrepancies fell within the combined uncertainty of the instrumentation and the steady-state error of the position control system. The results confirm that the proposed inverse kinematics model accurately describes the system's behavior, providing a reliable foundation for the development of safe and effective robotic-assisted cervical rehabilitation strategies.

Keywords: Cervical Rehabilitation, Parallel Robot, 3-RPS Mechanism, Inverse Kinematics, Experimental Validation, Active Cervical Orthosis.

1. INTRODUCCIÓN

La rehabilitación cervical es un componente fundamental en el tratamiento de los trastornos musculoesqueléticos, los cuales afectan la movilidad, la coordinación y la calidad de vida. El dolor de cuello y la disfunción cervical son condiciones altamente prevalentes que pueden reducir el rango de movimiento y afectar las actividades diarias, lo que resalta la necesidad de estrategias de rehabilitación efectivas (Arockia Doss et al., 2023). En los últimos años, las tecnologías robóticas han surgido como una alternativa a las terapias convencionales, ofreciendo repetibilidad, adaptabilidad y evaluación cuantitativa del progreso del paciente (Abarca and Elias, 2023).

Las ortesis cervicales tradicionales suelen ser dispositivos pasivos diseñados para estabilizar la cabeza y el cuello; sin embargo, la inmovilización prolongada puede provocar debilitamiento muscular y reducción de la funcionalidad (Lozano et al., 2022). Esta limitación ha motivado el desarrollo de dispositivos robóticos activos y portables capaces de asistir

y restaurar los movimientos naturales de la cabeza y el cuello. En particular, se han propuesto ortesis dinámicas basadas en manipuladores paralelos para proporcionar movimiento controlado, mejorar la comodidad del paciente y permitir protocolos de rehabilitación más efectivos (Ibrahim et al., 2021).

Los mecanismos robóticos paralelos presentan varias ventajas para aplicaciones de rehabilitación, incluyendo alta rigidez, precisión y capacidad de carga, lo que los hace adecuados para sistemas con espacio de trabajo reducido como la región cervical (Huamanchahua et al., 2021). Estas características han generado un mayor interés en su aplicación para la rehabilitación de cabeza y cuello, donde el control preciso del movimiento es esencial para garantizar una interacción humano-robot segura y confiable (Keemink et al., 2018). En este contexto, el modelado cinemático juega un papel crítico, ya que define la relación entre las entradas de los actuadores y el movimiento resultante de la plataforma móvil.

Entre los modelos cinemáticos, la cinemática inversa es particularmente importante, ya que permite determinar los desplazamientos de los actuadores necesarios para alcanzar una posición y orientación deseadas del efector final (Merlet, 2006). Los modelos precisos de cinemática inversa son esenciales para la implementación del control y la confiabilidad del sistema. Aunque varios estudios han abordado el diseño y análisis cinemático de manipuladores paralelos para rehabilitación, muchos de ellos se validan principalmente mediante simulaciones, con validación experimental limitada, lo que puede afectar su aplicabilidad en escenarios reales (Lingampally and Selvakumar, 2019).

En el contexto de la rehabilitación cervical, el robot paralelo 3-RPS propuesto está orientado a asistir movimientos controlados de la cabeza, particularmente flexión/extensión e inclinación bilateral, los cuales son comúnmente utilizados en rutinas terapéuticas. El control preciso de estos movimientos permite ejecutar ejercicios de forma repetible y segura, contribuyendo a mejorar la efectividad del proceso de rehabilitación y reduciendo el riesgo de lesiones.

Por lo tanto, existe la necesidad de una verificación rigurosa y validación experimental de los modelos cinemáticos para garantizar una operación precisa y segura en dispositivos de rehabilitación. En este trabajo, se verifica y valida la cinemática inversa de un robot paralelo 3-RPS diseñado para un prototipo de ortesis cervical. Se desarrolla un modelo analítico basado en la configuración geométrica del mecanismo y se evalúa mediante simulaciones numéricas. Posteriormente, se realiza la validación experimental utilizando un prototipo físico, comparando las predicciones teóricas con los desplazamientos medidos de los actuadores. Los resultados obtenidos demuestran la precisión del modelo propuesto y respaldan su aplicación en sistemas robóticos de asistencia para la rehabilitación cervical.

El presente artículo se organiza de la siguiente manera: en la Sección 2 se describe la arquitectura del robot paralelo; en la Sección 3 se desarrolla el modelo de cinemática inversa; en la Sección 4 se desarrolla el proceso para realizar la verificación mediante simulación; en la Sección 5 se describe la validación experimental; en la Sección 6 se discuten los resultados obtenidos; y finalmente, en la Sección 7 se presentan las conclusiones.

2. ARQUITECTURA DEL ROBOT PARALELO

La configuración 3-RPS recibe su nombre según los tipos de uniones que conectan la base con el efector final las cuales son uniones revolutas o rotacionales (R), prismáticas (P) y esféricas (S). Esta arquitectura proporciona al mecanismo tres grados de libertad (DOF, por sus siglas en inglés), los cuales pueden determinarse utilizando el criterio de Grübler-Kutzbach (Merlet, 2006).

$$m = 6(l - n - 1) + \sum_{i=1}^n d_i \quad (1)$$

Donde l representa el número total de eslabones, incluyendo la base, n es el número total de uniones y d_i el número de grados de libertad de cada articulación i .

El diagrama esquemático del mecanismo 3-RPS, seleccionado como base para el diseño de la ortesis cervical, se presenta en la figura 1 y está basado en el propuesto por Zhan y Kao (2010). Los grados de libertad del mecanismo 3-RPS se determinaron sustituyendo el número de eslabones y las restricciones de las articulaciones en (1).

$$m = 6(8 - 9 - 1) + 15$$

$$m = 3 \text{ DOF} \quad (2)$$

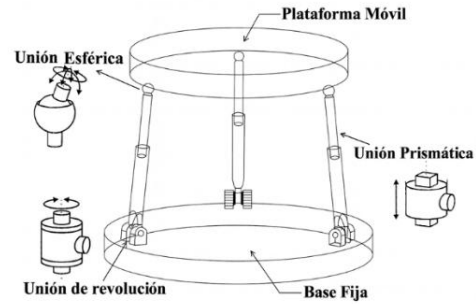


Figura 1. Mecanismo paralelo 3-RPS.
Fuente: Adaptado de Zhan y Kao (2010).

3. MODELO DE CINEMÁTICA INVERSA

En esta sección se presenta la deducción de las ecuaciones del modelo cinemático que describen el movimiento del sistema.

El diagrama geométrico utilizado para el análisis cinemático del dispositivo se basa en el propuesto por Tsai (1999), y se muestra en la figura 2.

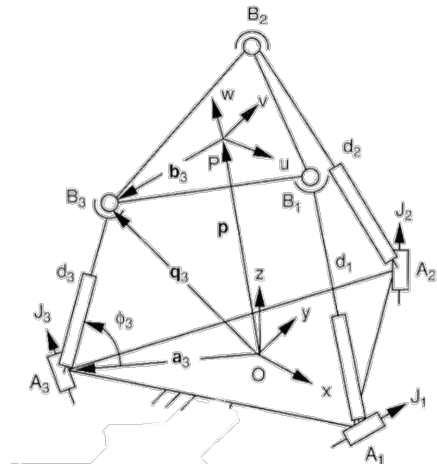


Figura 2. Modelo geométrico del mecanismo paralelo 3-RPS.
Fuente: Tsai (1999).

La posición del centro de la plataforma móvil se define como:

$${}^A\mathbf{P} = [P_x \ P_y \ P_z]^T \quad (3)$$

Suponiendo que la plataforma móvil rota sobre este punto $\mathbf{P}$.

$${}^A\mathbf{P} = [0 \ 0 \ p]^T \quad (4)$$

Donde p es la distancia entre centros de la base y la plataforma móvil. En la figura 3, se muestra la representación de un punto P del plano móvil B respecto al plano fijo A descrita por Tsai (1999).

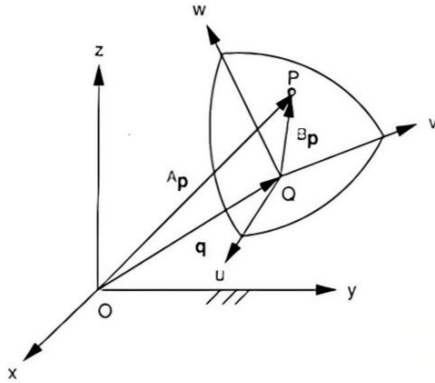


Figura 3. Representación geométrica de un punto en un plano móvil con respecto a un plano fijo.
Fuente: Tsai (1999).

Lo cual se puede expresar como

$${}^A\mathbf{p} = {}^A\mathbf{R}_B {}^B\mathbf{p} + {}^A\mathbf{q} \quad (5)$$

En donde ${}^A\mathbf{R}_B$ es la matriz de rotación de plano móvil B con respecto al plano fijo A .

De la figura 2 se obtiene la expresión que representa la posición de las uniones esféricas.

$${}^A\mathbf{B}_i = \mathbf{q}_i = \mathbf{a}_i + \mathbf{d}_i \quad (6)$$

Utilizando (5) para representar los puntos ${}^A\mathbf{B}_i$ se tiene:

$${}^A\mathbf{B}_i = {}^A\mathbf{R}_B {}^B\mathbf{b}_i + {}^A\mathbf{P} \quad (7)$$

Igualando con (6) se obtiene la ecuación de cierre de circuito para cada eslabón (Tsai, 1999).

$$\mathbf{d}_i = {}^A\mathbf{R}_B {}^B\mathbf{b}_i + {}^A\mathbf{P} - \mathbf{a}_i \quad (8)$$

donde $\mathbf{d}_i$ representa el vector de cada eslabón que conecta la base fija con la plataforma móvil. El vector ${}^B\mathbf{b}_i$, define la posición del punto de unión en la plataforma móvil expresado en su propio sistema de coordenadas, al sistema de referencia de la base. Por su parte ${}^A\mathbf{P}$, corresponde al vector de posición del origen de la plataforma móvil respecto a la base, mientras que $\mathbf{a}_i$, es el vector de posición del punto de anclaje en la base fija.

Los puntos $\mathbf{a}_i$ y ${}^B\mathbf{b}_i$ se definen en (9) y (10).

$$\begin{aligned} \mathbf{a}_1 &= R \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \\ \mathbf{a}_2 &= R \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \end{bmatrix}^T \\ \mathbf{a}_3 &= R \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \end{bmatrix}^T \end{aligned} \quad (9)$$

$$\begin{aligned} \mathbf{b}_1 &= r \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T \\ \mathbf{b}_2 &= r \begin{bmatrix} -\frac{1}{2} & \frac{\sqrt{3}}{2} & 0 \end{bmatrix}^T \\ \mathbf{b}_3 &= r \begin{bmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} & 0 \end{bmatrix}^T \end{aligned} \quad (10)$$

Donde R es el radio de la plataforma fija y r el de la plataforma móvil.

Para la matriz de rotación se utiliza la representación de los ángulos de Euler (Roll-Pitch-Yaw), propuestos por Taghirad (2025), los cuales se ejemplifican en la figura 4.

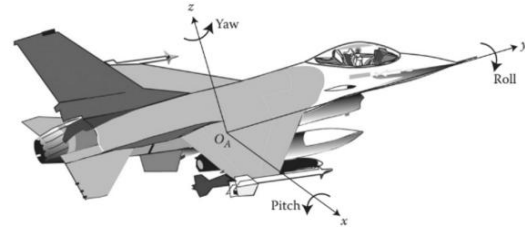


Figura 4. Representación de los ángulos de Euler (Pitch-Roll-Yaw).
Fuente: Taghirad (2025).

La matriz de rotación se define en (11) (Ruiz-Hidalgo et al., 2016, Ruiz Hidalgo, 2025).

$$\mathbf{R}_{uvw} = \mathbf{R}_u(\alpha) \mathbf{R}_v(\beta) \mathbf{R}_w(\gamma) \quad (11)$$

Donde.

$$\mathbf{R}_u = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{bmatrix} \quad (12)$$

$$\mathbf{R}_v = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ 0 & 1 & 0 \\ -\sin \beta & 0 & \cos \beta \end{bmatrix} \quad (13)$$

$$\mathbf{R}_w = \begin{bmatrix} \cos \gamma & -\sin \gamma & 0 \\ \sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (14)$$

Considerando que no existe rotación en el eje z , lo cual nos lleva a que $\gamma = 0$ y realizando la multiplicación matricial tenemos la matriz mostrada en (15).

$$\mathbf{R}_{uvw} = \begin{bmatrix} \cos \beta & 0 & \sin \beta \\ \sin \alpha \sin \beta & \cos \alpha & -\cos \beta \sin \alpha \\ \cos \alpha \sin \beta & \sin \alpha & \cos \alpha \cos \beta \end{bmatrix} \quad (15)$$

Con esto se logra que los parámetros de entrada del modelo de cinemática inversa sean los ángulos de rotación α y β .

4. VERIFICACIÓN DEL MODELO

La verificación del modelo de cinemática inversa se realizó mediante simulación en el software MSC Adams (módulo ADAMS/View), con el propósito de evaluar la consistencia de la formulación geométrica propuesta. Para ello, las longitudes de los actuadores obtenidas analíticamente se emplearon como

entradas en el modelo CAD del mecanismo 3-RPS, obteniendo como salida la orientación de la plataforma móvil. En la Figura 5 se presenta una vista del mecanismo en el entorno de ADAMS/View.

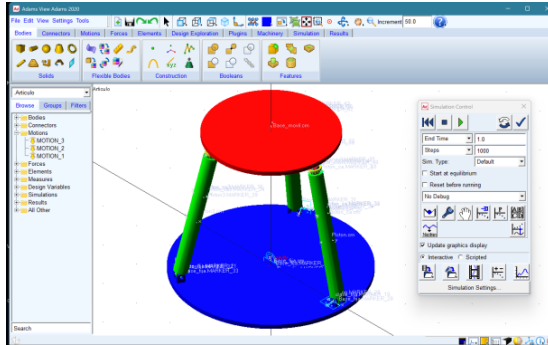


Figura 5. Modelo CAD del mecanismo paralelo 3-RPS en el entorno de trabajo de MSC Adams.

Este proceso de verificación permite establecer una correspondencia directa entre el modelo matemático y su implementación computacional, reduciendo la probabilidad de inconsistencias en la formulación o en la interpretación de las variables geométricas. De esta manera, se garantiza que el modelo constituye una base confiable para las etapas posteriores de validación experimental y desarrollo del sistema de control.

5. METODOLOGÍA PARA LA VALIDACIÓN

Para la validación del modelo se construyó un robot paralelo 3-RPS, con las dimensiones mostradas en la tabla 1.

Tabla 1. Medidas del robot paralelo 3-RPS

Parámetro	Simbología	Medida
Radio de la base	R	100 mm
Radio de la plataforma móvil	r	65 mm
Distancia entre centros de plataformas	p	203 mm

Se utilizaron 3 actuadores LA-T8P de la marca GoMotorWorld como el mostrado en la figura 6. Los cuales son actuadores eléctricos de 100 mm de carrera con retroalimentación por potenciómetro, con circuitos L298N como puente H y señal PWM para el control de posición, el cual se realizó utilizando una tarjeta de Arduino UNO y se diseñó una interfaz para monitoreo y recepción de datos de entrada en LabVIEW.



Figura 6. Actuador lineal LA-T8P.
Fuente: GoMotorWorld (s.f.).

Para la medición de los ángulos se consideró inicialmente el uso de módulos de sensores inerciales (IMU), debido a su capacidad para obtener información dinámica del movimiento. Sin embargo, dado que la validación experimental del presente trabajo se realizó bajo condiciones de estado estacionario y mediante la comparación de posiciones angulares puntuales, se optó por utilizar inclinómetros digitales, como el mostrado en la Figura 7.



Figura 7. Inclinómetro digital para medición de ángulos.
Fuente: Mercado-Libre (s.f.).

Estos dispositivos presentan una exactitud de $\pm 0.2^\circ$, una repetibilidad de 0.05° y una resolución de 0.05° , características que resultaron suficientes para los requerimientos de precisión de la validación experimental propuesta. Además, esta selección permitió evitar etapas adicionales de filtrado y procesamiento de señales comúnmente requeridas en sistemas basados en sensores inerciales.

5.1 Control de posición de los actuadores.

Para control de posición de los actuadores se optó por un control on-off con un margen de error de 0.5 mm para evitar oscilaciones cerca de la referencia. En la figura 8 se muestra el diagrama de bloques del control para cada actuador.

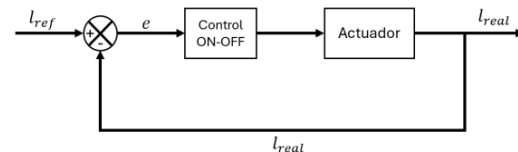


Figura 8. Diagrama de bloques del control de posición por actuador.

En la figura 9 se observa, una gráfica de la respuesta que tiene el controlador en respuesta a ingresar cualquier referencia, la cual obtuvo errores en estado estable menores a 2%.



5.2 Modificación de datos de entrada.

Technical drawing of a shaft assembly. The drawing includes a side view and a cross-sectional view. The side view shows a shaft with a central section of length L_1 and a total length of L_2 . The shaft has a diameter of $\phi 4.1$ in the central section and $\phi 4.8$ in the end sections. The end sections have a length of 4.8 . The central section has a length of L_1 and a diameter of $\phi 4.1$. The total length of the shaft is L_2 . The cross-sectional view shows the shaft with a diameter of $\phi 4.8$ and a length of 4.8 . The shaft is shown with a keyway and a key. The dimensions are given in millimeters.

$$L1 = 54.5 + S \quad (16)$$
$$\begin{aligned} L1 &= 54.5 + 100 \\ L1 &= 154.5 \text{ mm} \end{aligned} \quad (17)$$
$$l_i = d_i - 154.5 \quad (18)$$

6.1 Verificación.

Tabla 2. Longitudes obtenidas de la cinemática inversa para ciertos valores de ángulos.

Las longitudes obtenidas se incorporaron en el análisis de movimiento realizado mediante el software MSC Adams, conforme a lo descrito en capítulos previos. A partir de este análisis se determinaron los valores correspondientes a los ángulos α y β . Los resultados obtenidos se presentan en la Tabla 3.

Tabla 3. Resultados obtenidos en la simulación.

d_l [mm]	d_2 [mm]	d_3 [mm]	α [°]	β [°]
205.995	220.639	191.966	15.02	-0.175
189.860	214.383	214.383	0	15.21
205.995	215.756	196.503	10	-0.076
195.061	211.600	211.600	0	10.08
205.995	210.863	201.194	5	-0.018



200.458	208.797	208.797	0	5.019
205.995	191.966	220.639	-15.02	-0.175
222.951	197.812	197.812	0	-14.83
205.995	196.503	215.756	-10	-0.076
217.288	200.480	200.480	0	-9.926
195.061	221.361	202.101	9.85	10.16
189.860	228.873	200.503	14.66	15.54
200.458	213.683	203.977	4.963	5.02
217.288	191.315	209.950	-10.17	-10.16

Si se comparan los ángulos obtenidos en la simulación y los que se ingresaron al modelo, el error más grande que se obtuvo se da en la combinación de $\alpha = 15$ y $\beta = 15$, con un error de 2.27% en α y 3.6% en β . Además de que la tendencia es que el error aumenta con forme aumenta el ángulo de inclinación.

6.2 Validación.

Para la validación del modelo se utilizaron los mismos datos de entrada que se utilizaron para la verificación mediante simulación previamente ingresados en (18). Por lo tanto, los valores ingresados al control de posición individual por actuador serán los que se muestran en la tabla 4.

Tabla 4. Valores de entrada del sistema en base a los obtenidos del modelo cinemático.

α [°]	β [°]	l_1 [mm]	l_2 [mm]	l_3 [mm]
15	0	49.905	64.549	35.876
0	15	33.770	58.293	58.293
10	0	49.905	59.666	40.413
0	10	38.971	55.510	55.510
5	0	49.905	54.773	45.104
0	5	44.368	52.707	52.707
-15	0	49.905	35.876	64.549
0	-15	66.861	41.722	41.722
-10	0	49.905	40.413	59.666
0	-10	61.198	44.390	44.390
10	10	38.971	65.271	46.011
15	15	33.770	72.783	44.413
5	5	44.368	57.593	47.887
-10	-10	61.198	35.225	53.860

Estas longitudes fueron ingresadas al control de posición de los actuadores y se obtuvieron los resultados que se muestran en la tabla 5.

Tabla 5. Resultados obtenidos de la validación.

l_1 [mm]	l_2 [mm]	l_3 [mm]	α [°]	β [°]
49.760	64.420	35.970	15.100	0.050
34.020	58.650	57.970	0.150	15.200
49.660	59.340	40.760	9.750	0.100
39.100	55.820	55.230	0.050	10.250
49.660	54.940	45.360	4.950	0.050
44.480	53.080	52.390	0.050	4.850
49.760	36.170	64.320	-15.250	0.200
66.670	41.940	41.940	0.150	-14.900
50.050	40.760	59.430	-10.300	0.150
61.000	44.480	44.570	-0.200	-10.100
39.200	65.100	46.240	10.050	10.250
33.920	72.530	44.670	14.950	15.100
43.990	57.870	47.410	5.200	5.150
60.800	35.480	53.570	-10.050	-9.950

Los resultados obtenidos evidencian una alta concordancia entre el modelo de cinemática inversa y el comportamiento del mecanismo 3-RPS. Los errores obtenidos fueron menores a 0.3°, valores que se encuentran dentro del rango de incertidumbre del sistema, considerando la precisión del inclinómetro ($\pm 0.2^\circ$) y el error del control de posición ($\sim 2\%$). Esto indica que las discrepancias están más asociadas a limitaciones instrumentales y de control que al modelo cinemático.

En conjunto, los resultados confirman que el modelo propuesto describe adecuadamente el comportamiento del sistema y cumple con los niveles de precisión requeridos para aplicaciones de rehabilitación cervical.

Es importante destacar que, aunque una representación gráfica de las trayectorias de los ángulos α y β podría complementar el análisis, en este trabajo se prioriza la evaluación del modelo de cinemática inversa en condiciones de estado estacionario. Por esta razón, la validación se centra en la comparación de valores puntuales entre los resultados teóricos y experimentales, lo cual permite verificar la precisión del modelo sin depender del análisis de la evolución temporal de las variables. Un análisis más detallado del comportamiento dinámico y la inclusión de trayectorias continuas se consideran como trabajo futuro.



6. CONCLUSIONES

El presente trabajo abordó la verificación y validación del modelo de cinemática inversa de un robot paralelo 3-RPS, propuesto como base para el desarrollo de una ortesis cervical activa. A partir de un modelo teórico, derivado de la configuración geométrica del mecanismo, se realizó una verificación mediante simulación en MSC Adams, obteniendo errores máximos de aproximadamente 2.27% y 3.6% en los ángulos de inclinación, los cuales aumentan conforme se incrementa la magnitud de los mismos. Posteriormente, la validación experimental con un prototipo físico demostró una alta concordancia entre los valores teóricos y los medidos, con errores menores a 0.3° en la inclinación, dentro del margen de precisión del sistema de medición y del error inherente al control de los actuadores (~2%). Estos resultados confirman la validez del modelo cinemático propuesto y su aplicabilidad en sistemas de rehabilitación cervical, garantizando precisión y seguridad en el control del movimiento. En conjunto, el estudio respalda el uso de mecanismos paralelos tipo 3-RPS en dispositivos robóticos de asistencia, constituyendo una base sólida para futuras implementaciones con estrategias de control más avanzadas. Como trabajo futuro, se propone el análisis del comportamiento dinámico del sistema mediante la evaluación de trayectorias continuas y la comparación temporal de variables angulares.

REFERENCIAS

- ABARCA, V. E. & ELIAS, D. A. 2023. A Review of Parallel Robots: Rehabilitation, Assistance, and Humanoid Applications for Neck, Shoulder, Wrist, Hip, and Ankle Joints. *Robotics* [Online], 12.
- AROCKIA DOSS, A. S., LINGAMPALLY, P. K., NGUYEN, G. M. T. & SCHILBERG, D. 2023. A comprehensive review of wearable assistive robotic devices used for head and neck rehabilitation. *Results in Engineering*, 19, 101306.
- GOMOTORWORLD. s.f. LA-T8P Linear Actuator [Online]. Available: https://www.gomotorworld.com/pd?product_id=64 [Accessed 2 April 2026].
- HUAMANCHAHA, D., TADEO-GABRIEL, A., CHÁVEZ-RARAZ, R. & SERRANO-GUZMÁN, K. Parallel Robots in Rehabilitation and Assistance: A Systematic Review. 2021 IEEE 12th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), 1-4 Dec. 2021 2021. 0692-0698.
- IBRAHEM, M. E. H., MOHANDS, M. S. E.-., EL-WAKAD, M. T. & SAMI, S. A. Design and Analysis of a Dynamic Neck Brace. 2021 3rd Novel Intelligent and Leading Emerging Sciences Conference (NILES), 23-25 Oct. 2021 2021. 236-240.
- KEEMINK, A. Q., VAN DER KOOIJ, H. & STIENEN, A. H. 2018. Admittance control for physical human-robot interaction. *The International Journal of Robotics Research*, 37, 1421-1444.
- LINGAMPALLY, P. K. & SELVAKUMAR, A. A. 2019. Kinematic and Workspace Analysis of a Parallel Rehabilitation Device for Head-Neck Injured Patients. *FME Transactions*, 47.
- LOZANO, A., BALLESTEROS, M., CRUZ-ORTIZ, D. & CHAIREZ, I. 2022. Active neck orthosis for musculoskeletal cervical disorders rehabilitation using a parallel mini-robotic device. *Control Engineering Practice*, 128, 105312.
- MERCADO-LIBRE. s.f. Digital Inclinator [Online]. Available: https://www.mercadolibre.com.mx/medidor-de-angulo-digital-base-magnetica0-360inclinador/p/MLM2059168830?pdp_filters=item_id:MLM2430666081 [Accessed 2 April 2026].
- MERLET, J.-P. 2006. *Parallel robots*, Springer.
- RUIZ-HIDALGO, N. C., BLANCO-ORTEGA, A., ABÚNDEZ-PLIEGO, A., COLÍN-OCAMPO, J. & ARIAS-MONTIEL, M. Design and Control of a Novel 3-DOF Parallel Robot. 2016 International Conference on Mechatronics, Electronics and Automotive Engineering (ICMEAE), 22-25 Nov. 2016 2016. 66-71.
- RUIZ HIDALGO, N. C. 2025. Diseño, Optimización y Construcción de una Plataforma Móvil para el Anaveaje de Helicópteros. Doctorado, Centro Nacional de Investigación y Desarrollo Tecnológico.
- TAGHIRAD, H. D. 2025. *Parallel robots: mechanics and control*, CRC press.
- TSAI, L.-W. 1999. *Robot analysis: the mechanics of serial and parallel manipulators*, John Wiley & Sons.
- ZHAN, T.-S. & KAO, C.-C. 2010. Modified PSO Method for Robust Control of 3RPS Parallel Manipulators. *Mathematical Problems in Engineering*, 2010, 302430.

Multiscale Drought Prediction using SVR in Mexico City^{*}

Yunue Garza Pimentel^{*} Marcos Ángel González Olvera.^{**}
Jaime Reynaldo Santos Reyes^{*}

^{*} Instituto Politécnico Nacional, 07738 Zacatenco, Ciudad de México,
México. (e-mail: ygarzap2200@alumno.ipn.mx, jsantosr@ipn.mx).

^{**} Universidad Autónoma de la Ciudad de México, 09790, San Lorenzo
Tezonco, Ciudad de México, México (e-mail:
marcos.angel.gonzalez@uacm.edu.mx)

Resumen Accurate prediction of droughts in water-stressed urban areas is crucial for proactive resource management. This study compares the performance of Least Squares (LS), Singular Value Decomposition (SVD), and Support Vector Regression (SVR) for predicting the Standardized Precipitation-Evapotranspiration Index (SPEI) at three timescales (3, 6, and 12 months) in Mexico City. Using data from three automated weather stations of the National Meteorological Service (1952–2022), autoregressive models were trained and their performance was evaluated for the 2010–2022 period. The results reveal that the advantage of SVR over linear methods in moderate drought events systematically decreases with the timescale: 77 % improvement in SPEI-3, 61 % in SPEI-6, and 39 % in SPEI-12. Although global metrics showed modest differences (6 % in RMSE for SPEI-3), the analysis stratified by severity revealed that SVR is substantially superior for detecting moderate droughts at short scales, while simple linear models are sufficient and preferable for 12-month predictions. These findings suggest that operational early warning systems should implement hybrid approaches: SVR for short-term monitoring and linear models for strategic planning.

Keywords: SPEI, drought prediction, Support Vector Regression, Mexico City, multiscale analysis, machine learning

1. INTRODUCCIÓN

La sequía representa uno de los desastres naturales más extendidos y económicamente dañinos, afectando la agricultura, los ecosistemas, los recursos hídricos y a millones de personas en todo el mundo. En áreas urbanas con poblaciones en rápido crecimiento, la capacidad de pronosticar eventos de sequía se ha vuelto cada vez más crítica para la gestión de recursos hídricos y las estrategias de adaptación climática (Wilhite, 2000). La Ciudad de México, la zona metropolitana más grande de Norteamérica con más de 22 millones de habitantes, ejemplifica este desafío. La ciudad depende del Sistema Cutzamala, que proporciona aproximadamente el 25 % de su suministro de agua, y de la extracción de acuíferos para la demanda restante (Burns, 2009). En años recientes, se han experimentado eventos de sequía severos, con niveles de reservorios alcanzando mínimos históricos del 38 % de capacidad y obligando a las autoridades a implementar restricciones de agua que afectan a millones de residentes (Tortajada and Castelán, 2003; Aguilar-Barajas et al., 2024).

El Índice Estandarizado de Precipitación-Evapotranspiración (SPEI) ha emergido como una herramienta robusta para el monitoreo y evaluación de sequías (Vicente-Serrano et al., 2010). A diferencia de los índices basados únicamente en

precipitación como el SPI (McKee et al., 1993), el SPEI incorpora los efectos de la temperatura a través de la evapotranspiración potencial, haciéndolo particularmente sensible a los impactos del calentamiento climático sobre la severidad de las sequías (Beguiria et al., 2014). Esta característica multiscale permite que el SPEI se calcule en diferentes escalas temporales (típicamente de 1 a 24 meses), posibilitando el análisis de impactos de sequía en diferentes recursos hídricos y sistemas agrícolas. Los valores de SPEI a corto plazo (3-6 meses) son especialmente relevantes para la planificación agrícola y la gestión de reservorios, mientras que las escalas más largas (12+ meses) se relacionan con la recarga de acuíferos y la disponibilidad de agua a largo plazo (Vicente-Serrano et al., 2010; Liu et al., 2021).

Los enfoques tradicionales para el pronóstico de sequías han dependido principalmente de métodos estadísticos como modelos autorregresivos y regresión lineal (Mishra and Singh, 2011). Estos métodos asumen relaciones lineales entre índices de sequía rezagados y valores futuros, ofreciendo eficiencia computacional e interpretabilidad. Sin embargo, las dinámicas complejas y no lineales de los sistemas hidrológicos bajo condiciones climáticas cambiantes pueden limitar la precisión de enfoques puramente lineales, particularmente durante eventos extremos (Hao et al., 2018). Las técnicas de aprendizaje automático, particularmente la Regresión por Vectores de Soporte (SVR),

^{*} Este trabajo fue apoyado por IPN Grant-SIP-IPN, No. 20260502, UACM-CCYT-CON-2026-12 y SECIHTI.

han mostrado ser prometedoras para capturar patrones no lineales en series temporales hidrológicas (Vapnik, 1995). SVR emplea funciones kernel para mapear datos a espacios de mayor dimensión donde relaciones no lineales pueden aproximarse linealmente, ofreciendo potencialmente una mejor precisión predictiva para dinámicas complejas de sequía (Borji et al., 2016; Deo et al., 2017).

A pesar del creciente interés en el aprendizaje automático para predicción de sequías, permanecen preguntas fundamentales sobre cuándo y dónde estos métodos proporcionan ventajas genuinas sobre enfoques lineales más simples (Malik et al., 2021). Estudios recientes se han enfocado principalmente en la precisión predictiva general a través de conjuntos de datos completos, potencialmente enmascarando diferencias importantes en el desempeño bajo condiciones específicas de sequía (Rhee and Im, 2016; Fung et al., 2020). Para sistemas operacionales de alerta temprana, la capacidad de predecir con precisión eventos de sequía moderada y severa—precisamente cuando la intervención es más valiosa—puede ser más importante que el desempeño promedio a través de todas las condiciones. Comprender esta relación entre complejidad del modelo y utilidad práctica es particularmente relevante para áreas urbanas con escasez de agua como la Ciudad de México, donde los tomadores de decisiones deben balancear recursos computacionales con precisión predictiva para eventos críticos (Tavakol et al., 2022).

Este estudio aborda estas preguntas mediante una comparación sistemática de métodos lineales (Mínimos Cuadrados y Descomposición en Valores Singulares) y Regresión por Vectores de Soporte para predecir el SPEI en múltiples escalas temporales (3, 6 y 12 meses) en la Ciudad de México. Utilizando 71 años de datos (1952-2022) de tres estaciones meteorológicas operadas por el Servicio Meteorológico Nacional (SMN) de México, evaluamos no solo el desempeño predictivo general sino también estratificamos los resultados por severidad de sequía. Nuestros objetivos específicos son: (1) comparar la precisión predictiva de métodos lineales y de aprendizaje automático a través de diferentes escalas temporales del SPEI, (2) evaluar si las diferencias de desempeño varían sistemáticamente con la severidad de la sequía, y (3) determinar si la sobrecarga computacional del SVR se justifica por una mejora en la predicción de eventos críticos de sequía en un contexto operacional de gestión urbana del agua.

2. METODOLOGÍA

2.1 Área de estudio y datos

Este estudio se centra en la Ciudad de México, la zona metropolitana más grande de México y una de las áreas urbanas más pobladas del mundo (Tortajada and Castellán, 2003). El área de estudio comprende tres estaciones meteorológicas automáticas (EMA) operadas por el Servicio Meteorológico Nacional (SMN): Gran Canal (ID: 9029, 19.477°N), San Juan de Aragón (ID: 9043, 19.465°N) y Puente La Valla (ID: 9068, 19.429°N). Estas estaciones se localizan en la porción norte de la cuenca de México y proporcionan datos representativos de las condiciones climáticas que afectan el suministro de agua de la ciudad (Carrillo-Carrillo et al., 2025).

Los datos meteorológicos mensuales abarcan el período 1952–2022 (71 años), incluyendo precipitación (mm) y temperatura (°C). Estos datos fueron obtenidos del SMN y procesados para asegurar su calidad y completitud. La serie temporal resultante comprende 852 observaciones mensuales, lo que proporciona una base robusta para el análisis de sequías a largo plazo y la evaluación de métodos predictivos (Martínez-Sifuentes et al., 2022). La elección de estas estaciones se basa en su continuidad temporal, calidad de datos y ubicación estratégica en relación con las principales fuentes de abastecimiento de agua de la ciudad, particularmente el Sistema Cutzamala (Burns, 2009; Aguilar-Barajas et al., 2024).

Los datos de las tres estaciones fueron promediados aritméticamente para obtener una serie meteorológica representativa de la zona norte de la cuenca, dado que presentan alta correlación entre sí ($r \geq 0.90$). Los modelos se entrenaron y evaluaron sobre esta serie integrada.

2.2 Índice de sequía SPEI

El SPEI fue calculado siguiendo la metodología propuesta por Vicente-Serrano et al. (2010) y actualizada por Beguería et al. (2014). El índice se basa en el balance hídrico climático mensual, definido como la diferencia entre la precipitación (P) y la evapotranspiración potencial (ETP):

$$D_i = P_i - ETP_i \quad (1)$$

donde D_i representa el déficit o superávit de agua en el mes i . La evapotranspiración potencial fue estimada mediante el método de Thornthwaite (Thornthwaite, 1948), ampliamente utilizado cuando solo se dispone de datos de temperatura.

Para obtener el SPEI a diferentes escalas temporales, los valores de D se acumulan en ventanas móviles de k meses:

$$D_i^k = \sum_{j=0}^{k-1} D_{i-j} \quad (2)$$

Los valores acumulados D^k se ajustan a una distribución log-logística de tres parámetros, y posteriormente se estandarizan para obtener valores del SPEI con media cero y desviación estándar uno. Los parámetros de la distribución log-logística de tres parámetros (α, β, γ) fueron estimados exclusivamente con los datos del conjunto de entrenamiento (1952–1993), y posteriormente aplicados para estandarizar las series completas, evitando así fuga de información hacia los conjuntos de validación y prueba.

En este estudio, el SPEI se calculó para tres escalas temporales: 3, 6 y 12 meses (SPEI-3, SPEI-6 y SPEI-12), representando impactos en la agricultura de temporal, reservorios superficiales y recarga de acuíferos, respectivamente.

2.3 Métodos de predicción

Mínimos Cuadrados (LS) El método de mínimos cuadrados constituye la línea base de comparación. Utilizamos un modelo autorregresivo de orden p , donde p corresponde a la escala temporal del SPEI (3, 6 o 12 meses):



$$\text{SPEI}_t = \beta_0 + \sum_{i=1}^p \beta_i \text{SPEI}_{t-i} + \epsilon_t \quad (3)$$

Los coeficientes β se estiman mediante:

$$\beta = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (4)$$

donde $\mathbf{X}$ es la matriz de diseño que contiene los valores rezagados del SPEI y un intercepto, y $\mathbf{y}$ es el vector de valores objetivo.

Descomposición en Valores Singulares (SVD) El método SVD proporciona una alternativa numéricamente más estable a la regresión por mínimos cuadrados, particularmente útil cuando la matriz $\mathbf{X}^T \mathbf{X}$ está mal condicionada. La descomposición SVD de $\mathbf{X}$ se expresa como:

$$\mathbf{X} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T \quad (5)$$

Los coeficientes de regresión se calculan como:

$$\beta = \mathbf{V} \mathbf{\Sigma}^{-1} \mathbf{U}^T \mathbf{y} \quad (6)$$

Matemáticamente, SVD y LS producen soluciones equivalentes cuando $\mathbf{X}^T \mathbf{X}$ es invertible, pero SVD ofrece mayor robustez numérica.

Regresión por Vectores de Soporte (SVR) SVR es una técnica de aprendizaje automático basada en la teoría de aprendizaje estadístico (Vapnik, 1995). A diferencia de los métodos lineales, SVR puede capturar relaciones no lineales mediante el uso de funciones kernel (Borji et al., 2016; Deo et al., 2017). En este estudio, empleamos un kernel de función de base radial (RBF):

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\gamma \|\mathbf{x}_i - \mathbf{x}_j\|^2) \quad (7)$$

donde $\gamma = 1/(2\sigma^2)$ y σ es el parámetro de escala del kernel.

El modelo SVR busca encontrar una función $f(\mathbf{x}) = \mathbf{w}^T \phi(\mathbf{x}) + b$ que no se desvíe más de ϵ de los valores objetivo y_i para todos los datos de entrenamiento, mientras mantiene la función lo más plana posible (Shamshirband et al., 2020). Esto se formula como:

$$\min_{\mathbf{w}, b, \xi, \xi^*} \frac{1}{2} \|\mathbf{w}\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \quad (8)$$

sujeto a:

$$y_i - \mathbf{w}^T \phi(\mathbf{x}_i) - b \leq \epsilon + \xi_i \quad (9)$$

$$\mathbf{w}^T \phi(\mathbf{x}_i) + b - y_i \leq \epsilon + \xi_i^* \quad (10)$$

$$\xi_i, \xi_i^* \geq 0 \quad (11)$$

donde $C > 0$ es el parámetro de regularización que controla el balance entre la planitud del modelo y la tolerancia a errores mayores que ϵ , y ξ_i, ξ_i^* son variables de holgura (Shamshirband et al., 2020).

Los hiperparámetros del SVR—parámetro de regularización C (BoxConstraint), escala del kernel σ (KernelScale), y ancho de la función de pérdida ϵ (Epsilon)—fueron optimizados mediante optimización bayesiana (Snoek et al.,

2012; Alsumaiei, 2025). El espacio de búsqueda se definió en escala logarítmica:

El espacio de búsqueda en escala logarítmica fue: $C \in [10^{-2}, 10^3]$, $\sigma \in [10^{-2}, 10^2]$ y $\epsilon \in [10^{-3}, 1]$.

Se empleó validación cruzada de $k = 5$ folds sobre el conjunto de entrenamiento, minimizando el RMSE promedio como criterio de selección. Las 30 iteraciones de optimización bayesiana garantizaron una exploración eficiente del espacio de hiperparámetros (Piri et al., 2023). Los hiperparámetros óptimos obtenidos para cada escala temporal se presentan en la Tabla 1.

Cuadro 1. Hiperparámetros óptimos del SVR por escala temporal

Escala	C	σ	ϵ
SPEI-3	120.4824	4.7002	0.0011
SPEI-6	24.1369	4.9957	0.0011
SPEI-12	31.1171	7.5032	0.0018

Los valores óptimos de C decrecen de SPEI-3 (120.48) a SPEI-6 (24.14), reflejando menor necesidad de regularización en escalas largas. El parámetro σ aumenta con la escala (4.70, 4.99, 7.50), indicando relaciones más suaves entre rezagos, mientras que $\epsilon \approx 0,001$ permanece constante en las tres escalas.

2.4 Diseño experimental

Los datos se dividieron temporalmente en tres conjuntos para evitar contaminación de información futura en el entrenamiento (Hao et al., 2018):

Entrenamiento: 1952–1993 (43 años)

Validación: 1994–2009 (15 años)

Prueba: 2010–2022 (13 años, 156 meses)

El conjunto de entrenamiento se utilizó para estimar los parámetros de los modelos LS y SVD, así como para el entrenamiento inicial del SVR. El conjunto de validación se empleó para optimizar los hiperparámetros del SVR mediante optimización bayesiana y validación cruzada. El conjunto de prueba, completamente independiente, se reservó exclusivamente para evaluar el desempeño predictivo final de todos los modelos.

Para cada escala temporal (3, 6 y 12 meses), construimos matrices de características utilizando valores rezagados del SPEI correspondientes al número de meses de la escala. Por ejemplo, para SPEI-3, las características son $[\text{SPEI}_{t-1}, \text{SPEI}_{t-2}, \text{SPEI}_{t-3}]$ para predecir SPEI_t . Todos los modelos SVR incluyeron estandarización de características (media cero, varianza uno) aplicada únicamente al conjunto de entrenamiento para evitar fuga de información.

2.5 Métricas de evaluación

El desempeño predictivo se evaluó mediante múltiples métricas que capturan diferentes aspectos de la precisión del modelo:

Error Cuadrático Medio (RMSE):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (12)$$

Coefficiente de determinación (R^2):

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (13)$$

Eficiencia de Nash-Sutcliffe (NSE):

$$NSE = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (14)$$

Eficiencia de Kling-Gupta (KGE): dado que NSE es algebraicamente equivalente a R^2 para series centradas, ambas métricas producen valores idénticos en este contexto (Gupta et al., 2009).

Adicionalmente, la significancia estadística de las diferencias en desempeño se evaluó mediante el test de Diebold-Mariano (Diebold and Mariano, 1995), que compara la precisión de pronósticos considerando la autocorrelación en los errores de predicción.

2.6 Análisis estratificado por severidad

Para evaluar el desempeño diferencial de los modelos bajo condiciones específicas de sequía, estratificamos el conjunto de prueba según la severidad del SPEI observado, utilizando las categorías estándar (Vicente-Serrano et al., 2010; Hernández-Barrera et al., 2025):

- Sequía severa: $SPEI < -1,5$
- Sequía moderada: $-1,5 \leq SPEI < -1,0$
- Condiciones normales: $|SPEI| \leq 1,0$
- Condiciones húmedas: $SPEI > 1,0$

El RMSE se calculó independientemente para cada categoría, permitiendo identificar si las ventajas de SVR sobre métodos lineales se concentran en eventos extremos o se distribuyen uniformemente a través de todas las condiciones (Tavakol et al., 2022). Este análisis es particularmente relevante para aplicaciones de alerta temprana, donde la precisión durante eventos críticos puede tener mayor valor práctico que el desempeño promedio general (Malik et al., 2021; Parajuli et al., 2024). La capacidad de predecir con precisión sequías moderadas y severas es fundamental para la gestión proactiva del Sistema Cutzamala, que abastece el 25 % del agua de la Ciudad de México (Carrillo-Carrillo et al., 2025).

3. RESULTADOS

3.1 Desempeño predictivo global

La Tabla 2 presenta las métricas de desempeño para los tres métodos evaluados (LS, SVD y SVR) en las tres escalas temporales del SPEI (3, 6 y 12 meses) durante el período de prueba 2010–2022. Los resultados muestran que el desempeño varía sistemáticamente tanto entre métodos como entre escalas temporales.

Para SPEI-3, SVR alcanzó un RMSE de 0.4589, superando marginalmente a LS (RMSE=0.4884) con una reducción del 6 % en el error. El coeficiente de determinación también

mostró una ligera mejora, con $R^2 = 0,8648$ para SVR comparado con $R^2 = 0,8468$ para LS. Sin embargo, el test de Diebold-Mariano no detectó diferencias estadísticamente significativas entre los métodos ($p = 0,2221$), indicando que la mejora global de SVR no es estadísticamente robusta a través de todo el conjunto de prueba.

En SPEI-6, el patrón fue similar. SVR obtuvo RMSE=0.4421 frente a RMSE=0.4671 para LS, representando una mejora del 5.4 %. El R^2 aumentó de 0.8599 (LS) a 0.8745 (SVR). Nuevamente, el test de Diebold-Mariano no indicó significancia estadística ($p = 0,2120$), sugiriendo que las diferencias observadas podrían deberse a variabilidad aleatoria.

El comportamiento en SPEI-12 difirió notablemente de las escalas más cortas. Aquí, LS superó a SVR, alcanzando RMSE=0.3431 comparado con RMSE=0.3856 para SVR, una diferencia del 12.4 % a favor del método lineal. El R^2 también favoreció a LS ($R^2 = 0,9244$) sobre SVR ($R^2 = 0,9045$). El test de Diebold-Mariano arrojó $p = 0,0595$, aproximándose al umbral de significancia convencional ($\alpha = 0,05$), lo que sugiere evidencia marginal de que LS es superior para predicciones a 12 meses.

Cuadro 2. Métricas de desempeño predictivo en el conjunto de prueba (2010–2022)

Escala	RMSE		R^2		Test DM	
	LS	SVR	LS	SVR	Estadístico	p-valor
SPEI-3	0.4884	0.4589	0.8468	0.8648	1.2209	0.2221
SPEI-6	0.4671	0.4421	0.8599	0.8745	1.2480	0.2120
SPEI-12	0.3431	0.3856	0.9244	0.9045	-1.8843	0.0595

Cabe destacar que SVD y LS produjeron resultados virtualmente idénticos en todas las métricas evaluadas (diferencias menores a 10^{-10}), confirmando que ambos métodos son matemáticamente equivalentes cuando la matriz de diseño está bien condicionada, como es el caso en este estudio.

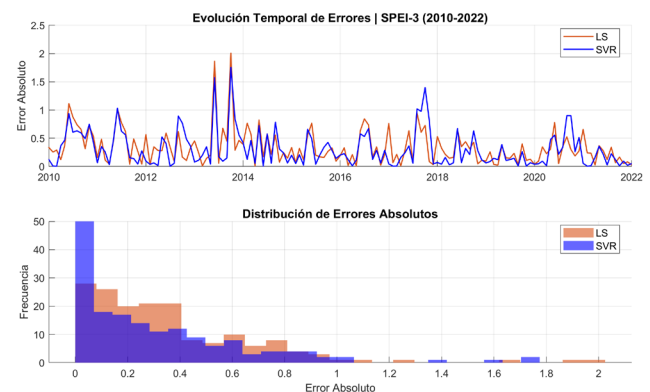


Figura 1. Series temporales de SPEI-3 observado (negro) y predicho mediante LS (naranja) y SVR (azul) durante el período de prueba (2010–2022). Los círculos negros señalan eventos de sequía moderada donde SVR mostró mejora sustancial en precisión predictiva.

3.2 Análisis estratificado por severidad de sequía

La Tabla 3 presenta el RMSE calculado exclusivamente para los meses clasificados como sequía moderada (SPEI entre -1.5 y -1.0) durante el período de prueba. Este análisis revela diferencias sustanciales entre métodos que no son evidentes en las métricas globales.

Para SPEI-3, SVR alcanzó un RMSE de 0.0923 en eventos de sequía moderada, comparado con 0.3935 para LS, representando una mejora del 77 % en la precisión de predicción para estos eventos críticos. Esta mejora dramática contrasta marcadamente con la mejora global del 6 % reportada en la sección anterior, sugiriendo que las ventajas de SVR se concentran específicamente en condiciones de sequía.

En SPEI-6, la ventaja de SVR se mantuvo pero con menor magnitud. El RMSE para sequías moderadas fue 0.1341 (SVR) versus 0.3403 (LS), una mejora del 61 %. Nuevamente, esta mejora supera ampliamente la diferencia del 5.4 % observada en métricas globales.

Para SPEI-12, aunque LS fue superior globalmente, SVR mostró mejor desempeño en sequías moderadas con RMSE=0.1260 comparado con RMSE=0.2078 para LS, una mejora del 39 %. Sin embargo, esta ventaja en eventos críticos no compensó el peor desempeño de SVR en condiciones normales y húmedas, resultando en el desempeño global inferior observado anteriormente.

Cuadro 3. RMSE en eventos de sequía moderada ($-1.5 \leq \text{SPEI} < -1.0$)

Escala	RMSE LS	RMSE SVR	Mejora (%)	n eventos
SPEI-3	0.3935	0.0923	77 %	13
SPEI-6	0.3403	0.1341	61 %	13
SPEI-12	0.2078	0.1260	39 %	13

Es notable que el número de eventos de sequía moderada fue idéntico ($n = 13$) para las tres escalas temporales durante el período de evaluación, lo que facilita la comparación directa entre escalas. También se observa un patrón claro: la ventaja de SVR sobre LS en eventos de sequía moderada disminuye sistemáticamente con el aumento de la escala temporal (77 % $\rightarrow$ 61 % $\rightarrow$ 39 %).

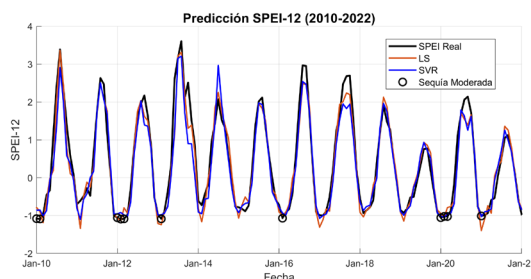


Figura 2. Series temporales de SPEI-12 observado y predicho mediante LS y SVR durante el período de prueba (2010-2022). A diferencia de las escalas cortas, LS muestra mejor desempeño global en predicciones a 12 meses, aunque SVR mantiene ventajas en eventos de sequía moderada.

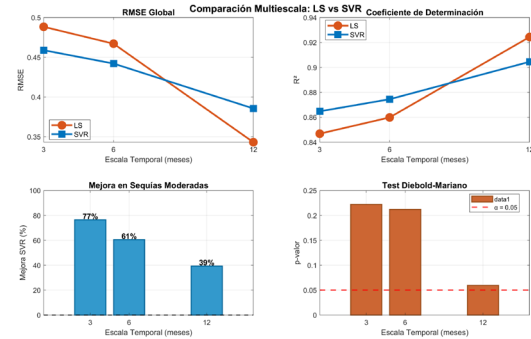


Figura 3. Comparación del desempeño predictivo de LS y SVR en las tres escalas temporales evaluadas. Se observa una disminución sistemática en la ventaja relativa de SVR para eventos de sequía moderada conforme aumenta la escala temporal (77 % en SPEI-3, 61 % en SPEI-6, 39 % en SPEI-12). Este patrón sugiere que la utilidad de métodos no lineales disminuye en predicciones a largo plazo.

3.3 Evolución temporal de errores de predicción

Las Figuras 1 y 2 muestran las series temporales de SPEI observado y predicho por ambos métodos para las escalas de 3 y 12 meses. Visualmente, ambos métodos capturan adecuadamente la dinámica general del SPEI, incluyendo las transiciones entre períodos húmedos y secos.

Sin embargo, las diferencias entre métodos se hacen más evidentes durante eventos extremos. Para SPEI-3, SVR tiende a seguir más fielmente los valores observados durante sequías moderadas y severas, mientras que LS muestra una tendencia a subestimar la severidad de estos eventos. Esta observación es consistente con las mejoras en RMSE reportadas en la Tabla 3 para condiciones de sequía.

En contraste, para SPEI-12, LS muestra un seguimiento más estable de los valores observados a través de todo el rango de condiciones, mientras que SVR exhibe ocasionalmente predicciones que se desvían más de los valores observados, particularmente en transiciones rápidas entre condiciones.

La Figura 3 sintetiza el comportamiento a través de las tres escalas temporales, evidenciando el patrón de disminución en la ventaja de SVR conforme aumenta la escala de agregación temporal.

La Figura 4 presenta el diagrama de dispersión entre valores observados y predichos para SPEI-3. Los puntos se agrupan razonablemente alrededor de la línea de identidad (predicción perfecta), con R^2 superior a 0.84 para ambos métodos. Se observa una ligera tendencia de ambos métodos a subestimar valores SPEI extremadamente negativos (sequías severas) y extremadamente positivos (condiciones muy húmedas), aunque esta tendencia es menos pronunciada para SVR en el rango de sequías moderadas.

3.4 Distribución de errores absolutos

La Figura 5 presenta el análisis de residuos para SPEI-3. El panel superior muestra que los errores de predicción varían

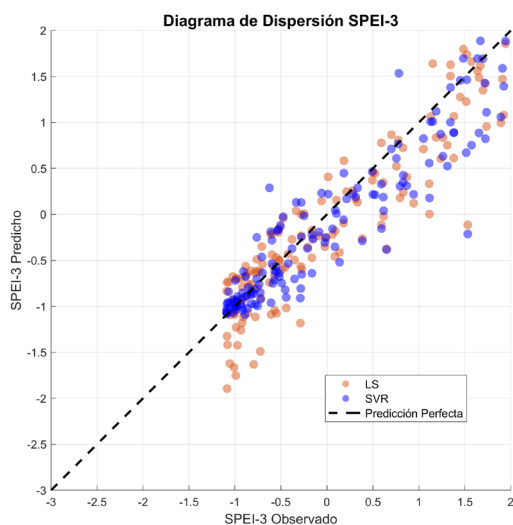


Figura 4. Diagrama de dispersión entre valores observados y predichos de SPEI-3. Los puntos más cercanos a la línea de identidad (diagonal) indican mejor precisión predictiva. SVR (azul) muestra menor dispersión que LS (naranja) particularmente para valores SPEI negativos correspondientes a condiciones de sequía.

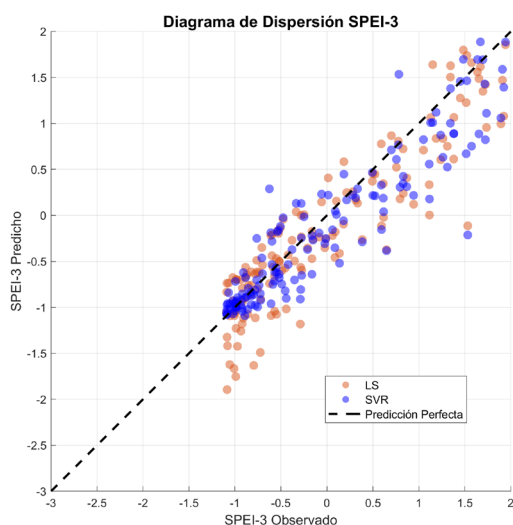


Figura 5. Análisis de residuos para SPEI-3. El panel superior muestra la evolución temporal de errores absolutos de predicción; el panel inferior presenta la distribución (histograma) de errores. SVR exhibe mayor concentración de errores pequeños (< 0.3 unidades) comparado con LS, consistente con su menor RMSE global.

considerablemente en el tiempo, con picos ocasionales que superan 1.0 unidades de SPEI. Estos picos tienden a ocurrir durante transiciones rápidas en las condiciones de

sequía o durante eventos extremos no anticipados por los patrones históricos.

El panel inferior confirma que las distribuciones de errores de SVR están ligeramente desplazadas hacia errores menores comparadas con LS, aunque existe considerable solapamiento entre ambas distribuciones. La mayor frecuencia de errores pequeños (menores a 0.3 unidades de SPEI) para SVR es consistente con los valores de RMSE más bajos reportados en la Tabla 2. Ambos métodos muestran colas en las distribuciones de error, indicando que ocasionalmente ocurren predicciones con errores sustanciales, aunque estos eventos son relativamente infrecuentes.

4. DISCUSIÓN

Los resultados presentados revelan un patrón sistemático en el desempeño relativo de métodos lineales y no lineales para predicción de sequías que depende críticamente de la escala temporal de agregación. Este hallazgo tiene implicaciones tanto metodológicas como prácticas para sistemas de alerta temprana de sequías en zonas urbanas como la Ciudad de México.

4.1 Interpretación del patrón dependiente de escala

La disminución sistemática en la ventaja de SVR sobre LS conforme aumenta la escala temporal (77 % en SPEI-3, 61 % en SPEI-6, 39 % en SPEI-12) sugiere que la complejidad del proceso subyacente que genera sequías varía con la escala de agregación. En escalas cortas (3 meses), las fluctuaciones de precipitación y temperatura exhiben mayor variabilidad y no linealidad, beneficiando a métodos capaces de capturar relaciones complejas como SVR (Deo et al., 2017; Borji et al., 2016). El kernel RBF permite a SVR modelar transiciones abruptas entre condiciones húmedas y secas que caracterizan eventos de sequía flash en escalas mensuales.

En contraste, SPEI-12 representa un promedio móvil de 12 meses que suaviza fluctuaciones de corto plazo, resultando en una serie temporal con dinámica más lineal y autocorrelación más fuerte. Bajo estas condiciones, modelos autorregresivos simples como LS capturan adecuadamente la persistencia temporal sin incurrir en el riesgo de sobreajuste asociado con métodos no lineales (Hao et al., 2018). Este resultado es consistente con estudios previos que han documentado que la predictibilidad de índices de sequía aumenta con la escala temporal debido a la mayor memoria del sistema (Vicente-Serrano et al., 2010).

Estos resultados son parcialmente consistentes con estudios previos sobre aplicación de machine learning a predicción de sequías. Deo et al. (2017) reportaron ventajas de SVR sobre modelos lineales para SPEI mensual en Australia, aunque no evaluaron el efecto de escala temporal. Tavakol et al. (2022) encontraron que la superioridad de métodos no lineales disminuye para predicciones a más de 3 meses, consistente con nuestro hallazgo de convergencia entre métodos en escalas largas.

Sin embargo, los resultados contrastan con Malik et al. (2021), quienes reportaron ventajas consistentes de random forests sobre regresión lineal en todas las escalas de SPEI. Esta diferencia puede atribuirse a diferencias

regionales en patrones climáticos o al uso de diferentes conjuntos de predictores. Este estudio se enfocó deliberadamente en modelos univariados para aislar el efecto de la escala temporal, mientras que estudios previos típicamente incorporan múltiples variables climáticas.

Para SPEI-3, la mejora de SVR fue del 6 % en RMSE, pero la mejora en sequías moderadas alcanzó el 77 %. Esta discrepancia indica que las métricas globales tradicionales (RMSE, R^2) pueden ocultar diferencias sustanciales en el desempeño de modelos durante eventos extremos que, paradójicamente, son los más relevantes para toma de decisiones.

La ausencia de significancia estadística en el test de Diebold-Mariano para SPEI-3 y SPEI-6 ($p > 0,05$) refleja que las diferencias entre métodos se concentran en un subconjunto pequeño del conjunto de prueba (13 de 156 meses). Esto subraya la importancia de análisis estratificados por severidad en evaluación de modelos de predicción de sequías, especialmente cuando el objetivo es alerta temprana de eventos críticos (Malik et al., 2021; Parajuli et al., 2024).

4.2 Implicaciones para gestión del agua en la CDMX

La Ciudad de México enfrenta estrés hídrico crónico agravado por sequías recurrentes que afectan el Sistema Cutzamala, fuente del 25 % del suministro urbano (Tortajada and Castelán, 2003). Los resultados de este estudio sugieren que sistemas de alerta temprana basados en SPEI-3 y SPEI-6 se beneficiarían sustancialmente de la implementación de SVR para identificar el inicio de sequías moderadas con mayor anticipación y precisión.

La ventaja del 77 % en RMSE para sequías moderadas en SPEI-3 se traduce en una capacidad superior para detectar el inicio de sequías flash que pueden evolucionar rápidamente a condiciones severas. Esta información permitiría a las autoridades activar protocolos de gestión de demanda (restricciones de riego, campañas de ahorro) con mayor margen de tiempo. Para planificación a mediano plazo (SPEI-12), sin embargo, modelos lineales simples son suficientes y preferibles por su menor complejidad computacional y mayor interpretabilidad.

Este estudio presenta varias limitaciones que deben considerarse al interpretar los resultados. Primero, el análisis se basa en datos de solo tres estaciones EMA en la Ciudad de México, limitando la generalización espacial de los hallazgos. Segundo, el período de prueba (2010–2022, 13 años) contiene solo 13 eventos de sequía moderada por escala, lo que limita el poder estadístico para detectar diferencias significativas.

Tercero, el estudio se limitó a modelos univariados (valores previos de SPEI como únicos predictores). La incorporación de variables exógenas como índices de circulación atmosférica o anomalías de temperatura superficial del mar podría modificar el desempeño relativo de los métodos (Aguilar-Barajas et al., 2024). Finalmente, solo se evaluó SVR con kernel RBF; otros algoritmos de machine learning (redes neuronales, gradient boosting) podrían mostrar patrones diferentes.

5. CONCLUSIONES

Este estudio comparó el desempeño de mínimos cuadrados (LS), descomposición de valores singulares (SVD) y Support Vector Regression (SVR) para predicción del Índice Estandarizado de Precipitación-Evapotranspiración (SPEI) en tres escalas temporales (3, 6 y 12 meses) en la Ciudad de México. Los principales hallazgos son:

1. La ventaja de SVR sobre métodos lineales en eventos de sequía moderada disminuye sistemáticamente con la escala temporal: 77 % en SPEI-3, 61 % en SPEI-6 y 39 % en SPEI-12. Este patrón refleja que la complejidad del proceso generador de sequías decrece conforme aumenta la agregación temporal.
2. Métricas de desempeño global ocultan diferencias sustanciales en eventos críticos. Para SPEI-3, SVR mostró solo 6 % de mejora global en RMSE pero 77 % de mejora en sequías moderadas, destacando la importancia de análisis estratificados por severidad.
3. Para escalas cortas (3 meses), SVR es superior para detectar el inicio de sequías flash y permitir alerta temprana. Para planificación a largo plazo (12 meses), modelos lineales simples son suficientes y preferibles por su interpretabilidad.
4. SVD y LS produjeron resultados virtualmente idénticos, confirmando que ambos métodos son equivalentes cuando la matriz de diseño está bien condicionada.

Estos resultados tienen implicaciones directas para el diseño de sistemas operacionales de monitoreo de sequías en zonas urbanas con estrés hídrico como la Ciudad de México. Se recomienda implementar enfoques híbridos que utilicen SVR para monitoreo de corto plazo (SPEI-3, SPEI-6) y modelos lineales para planificación estratégica (SPEI-12).

Investigación futura debería explorar: (1) validación de estos patrones en otras regiones climáticas de México, (2) incorporación de predictores exógenos (índices de circulación atmosférica, temperatura superficial del mar) para mejorar el desempeño en todas las escalas, y (3) evaluación de otros algoritmos de machine learning (redes neuronales, gradient boosting) para determinar si exhiben patrones dependientes de escala similares.

ACKNOWLEDGEMENTS

Los autores agradecen al Instituto Politécnico Nacional, Secretaría de Investigación y Posgrado Proj. Grant-SIP-IPN, No. 20260502. Universidad Autónoma de la Ciudad de México UACM-CCYT-CON-2026-12 y Secretaría de Ciencia, Humanidades, Tecnología e Innovación por su apoyo en este trabajo.

REFERENCIAS

- Aguilar-Barajas, I., Ramírez, A.I., and Sisto, N.P. (2024). Building urban resilience to climate-related water crises in Mexico City. *Water Resources Management*, 38, 1–18. doi:10.1016/j.envsci.2019.05.021.
- Alsumaiei, A.A. (2025). Modeling the onset of drought periods using explainable machine learning models enhanced by Bayesian optimization. *Journal of Hydrologic Engineering*, 30(4). doi:10.1061/JHYEFF.HEENG-6515.

- Beguería, S., Vicente-Serrano, S.M., Reig, F., and Latorre, B. (2014). Standardized precipitation evapotranspiration index (SPEI) revisited: parameter fitting, evapotranspiration models, tools, datasets and drought monitoring. *International Journal of Climatology*, 34(10), 3001–3023. doi:10.1002/joc.3887.
- Borji, M., Malekian, A., Salajegheh, A., and Ghadimi, M. (2016). Multi-time-scale analysis of hydrological drought forecasting using support vector regression (SVR) and artificial neural networks (ANN). *Arabian Journal of Geosciences*, 9(19), 1–10. doi:10.1007/s12517-016-2750-x.
- Burns, E. (2009). Repensar la cuenca: la gestión de ciclos del agua en el Valle de México. *Perfiles Latinoamericanos*, 17(33), 155–180.
- Carrillo-Carrillo, M., Ibáñez-Castillo, L., Arteaga-Ramírez, R., and Arévalo-Galarza, G. (2025). Spatio-temporal analysis of drought with SPEI in the State of Mexico and Mexico City. *Atmosphere*, 16(2), 202. doi:10.3390/atmos16020202.
- Deo, R.C., Kisi, O., and Singh, V.P. (2017). Drought forecasting in eastern Australia using multivariate adaptive regression spline, least square support vector machine and M5Tree model. *Atmospheric Research*, 184, 149–175. doi:10.1016/j.atmosres.2016.10.004.
- Diebold, F.X. and Mariano, R.S. (1995). Comparing predictive accuracy. *Journal of Business & Economic Statistics*, 13(3), 253–263. doi:10.1080/07350015.1995.10524599.
- Fung, K.F., Huang, Y.F., Koo, C.H., and Soh, Y.W. (2020). Drought forecasting: A review of modelling approaches 2007–2017. *Journal of Water and Climate Change*, 11(3), 771–799. doi:10.2166/wcc.2019.236.
- Gupta, H.V., Kling, H., Yilmaz, K.K., and Martinez, G.F. (2009). Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling. *Journal of Hydrology*, 377(1–2), 80–91. doi:10.1016/j.jhydrol.2009.08.003.
- Hao, Z., Singh, V.P., and Xia, Y. (2018). Seasonal drought prediction: Advances, challenges, and future prospects. *Reviews of Geophysics*, 56(1), 108–141. doi:10.1002/2016RG000549.
- Hernández-Barrera, S., Carrillo-González, R., and González-Chávez, M.C.A. (2025). Indices to identify historical and future periods of drought for the maize crop (*Zea mays* L.) in central Mexico. *Agronomy*, 15(2), 460. doi:10.3390/agronomy15020460.
- Liu, C., Yang, C., Yang, Q., and Wang, J. (2021). Spatio-temporal drought analysis by the standardized precipitation index (SPI) and standardized precipitation evapotranspiration index (SPEI) in Sichuan Province, China. *Scientific Reports*, 11(1), 1280. doi:10.1038/s41598-020-80527-3.
- Malik, A., Kumar, A., Rai, P., and Kuriqi, A. (2021). Prediction of multi-scalar standardized precipitation index by using artificial intelligence and regression models. *Climate*, 9(2), 28. doi:10.3390/cli9020028.
- Martínez-Sifuentes, A.R., Villanueva-Díaz, J., and Estrada-Ávalos, J. (2022). Reconstruction of the standardized precipitation-evapotranspiration index for the western region of Durango State, Mexico. *Forests*, 13(8), 1233. doi:10.3390/f13081233.
- McKee, T.B., Doesken, N.J., and Kleist, J. (1993). The relationship of drought frequency and duration to time scales. In *Proceedings of the 8th Conference on Applied Climatology*, volume 17, 179–183. Boston, MA.
- Mishra, A.K. and Singh, V.P. (2011). A review of drought concepts. *Journal of Hydrology*, 391(1–2), 202–216. doi:10.1016/j.jhydrol.2010.07.012.
- Parajuli, A., Parajuli, R., Banjara, M., Bhusal, A., Dahal, D., and Kalra, A. (2024). Application of machine learning and hydrological models for drought evaluation in ungauged basins using satellite-derived precipitation data. *Climate*, 12(11), 190. doi:10.3390/cli12110190.
- Piri, J., Abdollahipour, M., and Keshtegar, B. (2023). Advanced machine learning model for prediction of drought indices using hybrid SVR-RSM. *Water Resources Management*, 37, 683–712. doi:10.1007/s11269-022-03sleeve-4.
- Rhee, J. and Im, J. (2016). Meteorological drought forecasting for ungauged areas based on machine learning: Using long-range climate forecast and remote sensing data. *Agricultural and Forest Meteorology*, 237–238, 105–122. doi:10.1016/j.agrformet.2017.02.011.
- Shamshirband, S., Hashemi, S., Salimi, H., Samadianfard, S., Asadi, E., Shadkani, S., Kargar, K., Mosavi, A., Nabipour, N., and Chau, K.W. (2020). Predicting standardized streamflow index for hydrological drought using machine learning models. *Engineering Applications of Computational Fluid Mechanics*, 14(1), 339–350. doi:10.1080/19942060.2020.1715844.
- Snoek, J., Larochelle, H., and Adams, R.P. (2012). Practical Bayesian optimization of machine learning algorithms. In *Advances in Neural Information Processing Systems*, volume 25, 2951–2959.
- Tavakol, A., McDonough, K.R., and Ramirez, J.A. (2022). Ensemble stationary-based support vector regression for drought prediction under changing climate. *Journal of Hydrology*, 615, 128706. doi:10.1016/j.jhydrol.2021.128706.
- Thornthwaite, C.W. (1948). An approach toward a rational classification of climate. *Geographical Review*, 38(1), 55–94. doi:10.2307/210739.
- Tortajada, C. and Castelán, E. (2003). Water management for a megacity: Mexico City Metropolitan Area. *Ambio*, 32(2), 124–129. doi:10.1579/0044-7447-32.2.124.
- Vapnik, V.N. (1995). *The Nature of Statistical Learning Theory*. Springer, New York. doi:10.1007/978-1-4757-2440-0.
- Vicente-Serrano, S.M., Beguería, S., and López-Moreno, J.I. (2010). A multiscale drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *Journal of Climate*, 23(7), 1696–1718. doi:10.1175/2009JCLI2909.1.
- Wilhite, D.A. (2000). Drought as a natural hazard: Concepts and definitions. *Drought: A Global Assessment*, 1, 3–18.

Analysis and Impact of Student Failure Factors in Basic Sciences at the Delicias Technological Institute

Martínez Cereceres, Julieta*, Carrillo Modesto, Sandra Isabel*, Valtierra Angel, David Arnoldo*, Valtierra Fierro, David Angel **

* Tecnológico Nacional de México/Instituto Tecnológico de Delicias, Paseo Tecnológico Km 3.5 Cd. Delicias, Chihuahua, México. C.P. 33000. (Tel: (667) 888-5000; e-mail: julieta.mc@delicias.tecnm.mx; sandra.cm@delicias.tecnm.mx; david.va@delicias.tecnm.mx)

** Estudiante del Tecnológico Nacional de México/Instituto Tecnológico de Delicias, Paseo Tecnológico Km 3.5 Cd. Delicias, Chihuahua, México. C.P. 33000. (Tel: (667) 888-5000; e-mail: 124540191@delicias.tecnm.mx)

Abstract: Currently, the Delicias Technological Institute, located in the city of Delicias, Chihuahua, has a significant failure rate, negatively impacting graduation rates. Given this situation, the main objective is to identify the factors directly related to basic sciences in the subjects of Differential Calculus, Chemistry, and Fundamentals of Physics, to implement future teaching and academic strategies that minimize these rates and promote student retention. This research focuses on the study of academic, socioeconomic, and emotional factors, concluding that the failure rate is not unique to this institution, but rather a frequent problem in higher education institutions across the country, given that the national graduation rate is 56%. A quantitative approach was used with a total population of 583 first-semester students, resulting in a representative sample of 232 students surveyed with a 95% confidence level and a 5% margin of error. A self-administered questionnaire was used, the results of which indicate that 67.2% of students demonstrated a lack of understanding in basic science subjects, and 71.6% presented difficulties with these subjects. A comprehensive literature review concluded that 70% of high school graduates nationwide lack prior knowledge, consistent with the 69.7% found in this research. This study aligns with the research line ITF-DEL-LIE-2023-326: "Environment of the educational process".

Keywords: Factors, failure rate, terminal efficiency, basic sciences, didactics strategies.

1. INTRODUCCIÓN

En la actualidad, la reprobación en la educación superior constituye uno de los principales desafíos que enfrentan las Instituciones de Educación Superior (IES), constituidas por "más de 4 mil instituciones públicas y privadas que atienden al año alrededor de 4.7 millones de estudiantes" (PROSEC, 2019-2024 citado por Granja, Ramírez, Garcés, Villalobos, Zapata, Hernández y Martínez, 2022), impactando en indicadores clave como la eficiencia terminal, la permanencia estudiantil y la calidad educativa. Estudios realizados por Fuentes y Rivera (2022), "identificaron que la reprobación es un fenómeno multidimensional que se encuentra asociado con factores personales, contextuales y académicos". De hecho, Rochin Berumen (2021), citado por Rodríguez Camargo, Álvarez Hilario y Garduño Teliz (2025), mencionó que "el primer año escolar es donde mayormente existe rezago y luego evoluciona al abandono escolar. Aquí, los padres de familia representan un apoyo invaluable impactando en la motivación, la conducta y el rendimiento escolar". Por ello, "es imperante abordar esta problemática, dada la escasez de estudios con una aproximación integral a la reprobación, así como el alcance limitado que ha mostrado la intervención académica" (Gutiérrez-Rodríguez y Rodríguez-Mazariiego, 2020; Luna, 2019; Torres-Zapata et al., 2020 citados por Fuentes y Rivera, 2022), ya que, a nivel nacional, la situación es preocupante.

Estudios recientes documentan que un alto porcentaje de estudiantes en las IES presenta rezagos académicos en matemáticas desde su ingreso, de hecho, "la mayoría de las personas encuentran las matemáticas como un área difícil de estudiar, de comprender, inútil, un área exclusiva para 'gente inteligente' o gente superdotada" (Reyes Portillo, 2025), aun cuando las matemáticas son la base de la ciencia y "esta área desarrolla el pensamiento lógico, así como habilidades de razonamiento que, mediante el análisis de la información de conocimientos previos, influyen en la toma de decisiones y la resolución de problemas" (Reyes Portillo, 2025). Sin embargo, de acuerdo con este mismo autor, "frases como 'no sé para qué me enseñaron eso', 'no entiendo el objeto de la tortura', son gestos evidenciando el rechazo por esta área, interpretados como malas experiencias arraigadas en los sujetos por distintas circunstancias". Por ende, de acuerdo con Reyes Portillo (2025), "esto afecta el desarrollo de habilidades y de construcción de estructuras de pensamiento, que ahondan más la idea negativa acerca de las matemáticas".

Particularmente, en el Instituto Tecnológico de Delicias (en lo sucesivo ITD), esta problemática adquiere una relevancia crítica, al concentrarse los mayores índices de reprobación en ciencias básicas en asignaturas como Cálculo Diferencial, Álgebra Lineal, Química, Fundamentos de Química y Fundamentos de Física. Afectando la trayectoria académica de los estudiantes, incrementando el riesgo de deserción y



disminuyendo la eficiencia terminal. Ninguna institución está exenta de esta problemática, ya que los primeros años de formación profesional concentran los mayores índices de reprobación, especialmente en el tronco común de las carreras.

Ante este panorama, investigaciones recientes han señalado que las capacidades cognitivas, tales como el razonamiento lógico, la memoria de trabajo y la autorregulación del aprendizaje, juegan un papel determinante en el desempeño académico de los estudiantes universitarios, pero “aparentemente en México, según la ANUIES (2000) los estudiantes de educación superior no han sido un objeto de investigación muy recurrente, aun considerando que éste es uno de los actores imprescindibles en el proceso educativo” (Estévez y Pérez, 2007; ANUIES, 2000, Citados por Pérez Reveles 2016).

Adicional a lo anterior, existe otro factor no menos importante que abordar, por ejemplo, el económico ya que, “la economía de un hogar es un factor determinante en el desempeño estudiantil, cuando no existe la estabilidad en los padres, los estudiantes se ven en la necesidad de abandonar sus estudios para incorporarse al ámbito laboral” (Pérez Reveles, 2016).

Por otra parte, según Rodríguez Camargo, Álvarez Hilario y Garduño Teliz (2025), “identificaron falta de competencias pedagógicas en algunos docentes, especialmente en áreas de ingeniería, así como la influencia de factores sociales y familiares (por ejemplo, la disfunción familiar), los cuales inciden negativamente en la elección vocacional, la motivación y el rendimiento académico”.

La reprobación no puede analizarse de manera aislada, sino como un fenómeno multidimensional que involucra variables académicas, cognitivas y socioemocionales, por ello, “enfrentar la reprobación y deserción escolar requiere un enfoque integral y conjunto, donde todas las partes del sistema educativo, estudiantes, docentes, familias, etc., establezcan marcos de corresponsabilidad y articulen acciones coordinadas para fortalecer la retención estudiantil” (Rodríguez Camargo, et al., 2025).

Por lo tanto, es importante analizar los factores académicos, socioeconómicos y emocionales que influyen en la reprobación de los estudiantes del ITD, durante el ciclo escolar 2025-2026, generando propuestas de intervención con estrategias didáctico-académicas innovadoras, con procesos de retroalimentación y seguimiento, que contribuyan a la permanencia y el éxito académico del estudiantado.

1.1 Planteamiento del Problema

¿Cómo influyen los factores académicos, socioeconómicos y emocionales en la reprobación de los estudiantes en ciencias básicas del Instituto Tecnológico de Delicias durante el ciclo escolar 2025-2026?

1.2 Justificación

El ITD, al formar parte del Tecnológico Nacional de México (TecNM), tiene la misión de formar integralmente profesionales competitivos de la ciencia, la tecnología y otras áreas de conocimiento, comprometidos con el desarrollo económico, social, cultural y sustentable del país. Sin embargo, estadísticamente se ha observado que en Ciencias Básicas los estudiantes presentan altos índices de reprobación, afectando la trayectoria académica, la retención escolar y la eficiencia terminal, ya que “la mayor parte de los universitarios abandonan sus estudios, cinco de cada diez estudiantes desertan al inicio del segundo año y cuatro de cada diez estudiantes que comienzan el cuarto año, no obtienen el título de licenciatura” (Huesca y Castaño, 2007 citados por Granja et al., 2022). (Ver Fig. 1).

Sostenimiento	Ciclo escolar de abandono							Alumnos que abandonaron estudios en 2023-2024	
	2016-2017	2017-2018	2018-2019	2019-2020	2020-2021	2021-2022	2022-2023		
Nacional	7.2%	8.3%	9.3%	8.0%	8.7%	8.3%	8.1%	5.9%	280,863
Público	7.5%	9.6%	7.7%	7.5%	7.8%	7.9%	7.4%	6.6%	206,669
Particular	6.7%	5.5%	12.6%	9.0%	11.7%	9.0%	9.7%	4.6%	74,194

Fig. 1. Tasa de abandono escolar en Educación Superior 2023-24.
Fuente: ANUIES (2024).

En cuanto al impacto económico producto de la reprobación, la educación superior en México representa una inversión considerablemente fuerte por cada estudiante que cursa sus estudios profesionales. Según García Villaseñor (2025), “el costo por impartir educación superior a cada alumno es de \$61,200.00 en el presente 2025” (Ver Fig. 2).

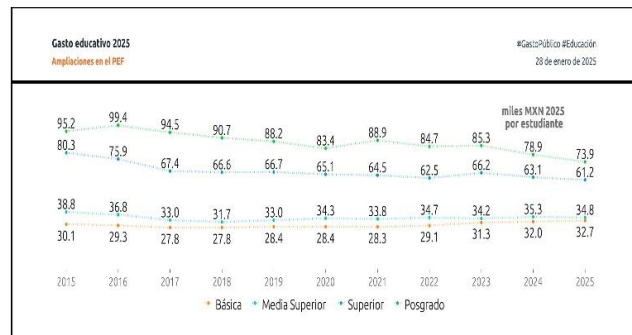


Fig. 2. Evolución del Gasto en Educación por estudiante 2025.
Fuente: Elaborado por el CIEP, con información de SHCP (2024a).

Estos costos se ajustan considerando la infraestructura, nómina docente, servicios académicos y administrativos, etc. Por lo tanto, cada alumno reprobado duplica costos si repite asignaturas, además del desfase en la eficiencia terminal y mayor presión presupuestal para las instituciones. En el caso del TecNM, “logró incrementar su matrícula con 33,682 nuevos estudiantes para el ciclo escolar 2025-2026. Con este resultado se registra un total de 607,658 estudiantes cursando algún programa en nivel de Técnico Superior, Licenciatura y Posgrado, en sus diferentes modalidades” (TecNM, 2025). Con una tasa de reprobación de aproximadamente 30% en primer semestre podría representar pérdidas superiores a



\$618,401,520.00 MNX por ciclo escolar, de acuerdo al siguiente cálculo:

$$(33,682 \times 0.30 = 10,104.6 \times 61,200 = \$618,401,520 \text{ estimado}).$$

Es importante aclarar que, aunque el TecNM no publica un indicador consolidado nacional de reprobación para el ciclo 2025–2026, diversos estudios institucionales y reportes recientes con evidencia empírica y válida, estiman que los índices de reprobación en asignaturas de ciencias básicas oscilan entre el 25% y 40%, principalmente en los primeros semestres de formación profesional.

Ante este fenómeno académico, se busca identificar los factores que inciden en la reprobación y deserción, con el fin de diseñar estrategias institucionales adecuadas, basadas en evidencia también empírica y metodológica para mejorar el rendimiento académico, así como reducir el abandono escolar.

1.3 Objetivo General

Analizar los factores académicos, socioeconómicos y emocionales que inciden en la reprobación de una muestra representativa de estudiantes de ciencias básicas del Instituto Tecnológico de Delicias durante el ciclo 2025-2026, implementando estrategias institucionales que mejoren el rendimiento académico.

1.4 Hipótesis

H_a : Existe una relación estadísticamente significativa entre los factores académicos, socioeconómicos y emocionales y los índices de reprobación de los estudiantes de ciencias básicas del Instituto Tecnológico de Delicias durante el ciclo escolar 2025-2026.

H_0 : No existe una relación estadísticamente significativa entre los factores académicos, socioeconómicos y emocionales y los índices de reprobación de los estudiantes de ciencias básicas del Instituto Tecnológico de Delicias durante el ciclo escolar 2025-2026.

2. ANTECEDENTES.

La reprobación en instituciones de educación superior (IES), representa un fenómeno recurrente que afecta e impacta directamente los indicadores de eficiencia terminal y permanencia del estudiantado a nivel nacional, reflejo de una muestra estudiantil que no posee las competencias, habilidades y destrezas en el área de matemáticas, física y química. Por ende, es casi imposible no hacer énfasis, para efectos del presente análisis situacional que impera en las IES, que reciben el recurso estudiantil acusando indicadores de una preparación académica con marcadas deficiencias en el aparato cognitivo para su aplicación en las ciencias básicas. En el caso del ITD, esta problemática se manifiesta con especial énfasis en las asignaturas de ciencias básicas de los tres primeros semestres de su formación profesional, que concentran los mayores índices de reprobación, un fenómeno que se manifiesta como un filtro que impide la consecución de las y los estudiantes.

Cabe mencionar que este proceso o problemática no es incluyente únicamente de esta Institución, sino que se percibe como un fenómeno académico que afecta globalmente al común denominador de las IES en México, por ejemplo, una amplia diversidad de estudios documenta que la reprobación en educación superior impartida en México es una constante que impacta negativamente en la trayectoria escolar de los estudiantes. Saucedo Fernández, Herrera Sánchez, Díaz Perera, Bautista Maldonado y Salinas Padilla (2014), señalan que “ninguna IES está exenta de altos índices de reprobación y deserción que registran los primeros años en distintas carreras. Durante el tronco común se imparten las bases matemáticas y tecnológicas, que son materias que más índice de reprobación tienen”.

Por otra parte, Rodríguez Camargo, Álvarez Hilario y Garduño Teliz (2025) realizaron una revisión sistemática sobre la reprobación en carreras de ingeniería de nivel superior, “identificando factores como la falta de habilidades matemáticas, la escasa orientación vocacional y, sobre todo, la limitada implementación de estrategias tecnológicas como causas recurrentes del bajo rendimiento académico”. En este sentido, según Sánchez Mendiola et al. (2017) citados por Rodríguez Camargo, et al. (2025), concluyen que “únicamente 27% de estudiantes de licenciatura son regulares. 73% de alumnos han reprobado una materia en su semestre inmediato anterior y la eficiencia terminal fue de 56%, siendo los primeros años de carrera donde hay más abandono escolar”.

En cuanto a la eficiencia terminal, en el contexto internacional, la eficiencia de un sistema educativo (educational efficiency) ha sido definida por la Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura de manera sucinta: “Grado en el cual un sistema educativo consigue optimizar la relación inversión/resultado en la educación” (UNESCO, 2009, citado por Pérez Reveles, 2016), no obstante, la tendencia de dicha eficiencia terminal se mantiene a la baja en la mayoría de las IES tanto públicas como privadas.

Ante este crítico panorama, se vuelve imperativo el diseño de estrategias integrales que combinen enfoques administrativos y académicos. Un ejemplo de estas es la propuesta diseñada por docentes del Instituto Tecnológico de Cd. Juárez, quienes propusieron la implementación de “la estrategia didáctica llamada E-1=ma (por sus siglas Examen Inverso igual a Mejoramiento en el Aprendizaje), con la finalidad de reducir los índices de reprobación y de deserción de los alumnos que cursan las materias de ciencias básicas” (Alanís Villaseñor, Cera Gaytán, Chávez Moreno, Lagarda Leyva, & Requenes Ruiz, 2022). Según estos mismos autores la implementación de este método está centrado en 4 pasos: “Datos reales, Resolver, Revisar y Retroalimentar. Los resultados empíricos de la aplicación de esta estrategia muestran que los alumnos se automotivan, aumenta su confianza y mejora su aprendizaje”.

En el ámbito educativo del ITD surge una nueva propuesta denominada “Aula Inteligente Integral (AII)”, con el desarrollo de nuevas estrategias académicas y didácticas enfocadas en los procesos de apoyo académico, con énfasis en la atención personalizada y el seguimiento del desempeño



estudiantil en ciencias básicas. Su implementación busca no solo reducir los índices de reprobación, sino también fortalecer la permanencia y la eficiencia terminal de los estudiantes.

3. MARCO TEÓRICO

En el presente estudio, se parte de la premisa de que, el fenómeno de los altos índices de reprobación en la educación a nivel superior no es un factor excluyente, ya que se manifiesta en cualquier contexto educativo, no solo a nivel nacional sino también internacional, sin dejar atrás sus orígenes de causalidad múltiple. Lo cierto es que, ante la eventualidad de esta problemática recurrente, es fundamental implementar estrategias académicas, didácticas y metodológicas que permitan mitigar los efectos negativos de la reprobación con estudiantes que abandonan su carrera profesional. En este estudio se proponen y analizan los efectos de los diversos factores tanto académicos, como socioeconómicos y emocionales, y con ello, medir el impacto que tienen en el abandono de los estudios superiores.

Particularmente en el factor académico, se observa el fenómeno de la reprobación durante los primeros semestres, de hecho, “los indicadores internacionales en materia de reprobación, rezago y deserción colocan a México en una posición crítica. La reprobación, entendida como un fracaso escolar, constituye un fenómeno de creciente preocupación académica, dada su naturaleza multicausal que va en aumento” (Torres Zapata, Acuña Lara y Moguel Ceballos, 2019).

De acuerdo con lo mencionado y siendo el propósito de este estudio, el cual no pretende generar un modelo nuevo, lo cierto es que un gran porcentaje de estudiantes de nivel superior no son afines al estudio de las ciencias matemáticas, llámese cálculo diferencial, integral, geometría analítica, etc., ya que de entrada manifiestan un abierto rechazo al estudio de las mismas, “solo estudian para pasar”, o piensan que las matemáticas no les aportarán nada productivo, solo “existen” para complicarles la vida o no tienen aplicación en el contexto real. Por ende, según cita Zapata (2018), “alrededor del 70% de los alumnos que ingresan a nivel licenciatura en México y en países del mundo, presentan deficiencias en matemáticas, pues su preparación o hábito de estudios desde la infancia, no fue el adecuado, destacó la empresa Kumon”.

La realidad es que, “el temor hacia las matemáticas en las personas se atribuye principalmente al fracaso. En cualquier contexto de la vida cotidiana, el término ‘fracaso’ implica la falta de éxito sin necesidad de mayores explicaciones” (Castillo y Novelo et al., 2015, citados por Meza, Meza, Arguello y Vaca, 2024).

Según Bustamante et al. (2021), citado por Meza et al. (2024), “resaltan la relevancia de matemáticas, vinculándolas estrechamente con el trabajo, habilidades cognitivas, el pensamiento y la capacidad intrínseca del razonamiento humano. La aplicación de las matemáticas implica habilidades como la imaginación, el descubrimiento, intuición, verificación de resultados, etcétera”. Por lo anterior, las matemáticas fomentan el pensamiento crítico y lógico de los estudiantes, mejoran el razonamiento cognitivo y preparan al

estudiante para diversos retos en el abordaje de las problemáticas que enfrentaran tanto en su vida personal como profesional, que conlleve su incorporación futura al sector productivo y laboral de su especialidad.

Por lo tanto, para lograr un sentido de afinidad en los estudiantes que los acerque con un dominio aceptable de las matemáticas, es importante que se implementen nuevas estrategias didácticas que permitan reforzar su aprendizaje para mejorar el rendimiento académico, disminuyendo la presión y el estrés producto del aprendizaje y puesta en marcha de esta área del conocimiento, considerando que, “se entiende que el rendimiento no son solo las calificaciones del estudiante, sino también sus habilidades, destrezas, aptitudes, ideales, intereses, etcétera” (Figuroa, 2004, citado por Pérez Jaimes, Estrada Rey y Estrada Gutiérrez, 2021). En el sentido estricto de la palabra, “si falla alguna de las características mencionadas, indudablemente se tendrá a un alumno reprobado” (Pérez Jaimes, et al., 2021).

Otro de los factores que impactan directamente es el factor socioeconómico, y es que de acuerdo con Fuentes y Rivera (2022) citando a Coneval (2020), “desde una perspectiva económica, cabe mencionar que, durante 2020, 43.9% de la población en México no percibió un ingreso lo suficientemente robusto como para alcanzar el bienestar económico, obtener servicios básicos en sus viviendas, ni acceder a la educación”, por tal motivo, un sector importante de la población en edad escolar a nivel superior, se ve obligado a cubrir trabajos de medio tiempo con el fin de solventar los gastos derivados de sus estudios como matrícula, útiles escolares, transporte, alimentos, etc., y en este sentido destaca el indicador que en “la educación superior al no tener cubiertas sus necesidades primarias, los estudiantes buscan empleos para poder cubrir los costos directos e indirectos de su educación, moviendo a segundo plano su aprovechamiento escolar” (Silva Laya, 2011; Tinto, 2012; Torres Zapata et al., 2020; citados por Fuentes y Rivera, 2022).

Esta situación por la que atraviesan los estudiantes en el país suele ser recurrente, ya que de acuerdo con una investigación realizada por UNAM Global (2017), el “53% de los egresados de universidades privadas y 42% de los egresados de las universidades públicas, trabajaban desde antes de concluir sus estudios. Lo que significa que la mitad de los universitarios en México combinan la escuela con el trabajo”.

Finalmente, pero no menos importante, está el factor emocional, si bien se aborda después de los otros factores antes mencionados, es un tema que por sí solo, puede implicar un estudio extenso sobre sus características, generalidades, orígenes y consecuencias, entre muchos otros aspectos, dada la naturaleza de éste. Sin embargo, para efectos del presente estudio se abordarán ciertas situaciones consideradas a juicio priori, impactan más en el actual contexto académico de la población estudiantil del ITD.

En primer lugar, se ha detectado un sentimiento fuertemente arraigado en los estudiantes, el cual tiene que ver con el concepto de “ansiedad matemática” y que en sentido análogo se puede configurar como la “matefobia”, en donde según Oaxaca Luna y Valderrama Bravo (2021), “hablar de



matefobia es discernir sobre el miedo o fobia al aprendizaje de las matemáticas. Son muchas las causas o circunstancias que dan origen a este fenómeno educativo que es propio de cada país”.

Por otra parte, de acuerdo con Reyes Portillo (2025), la ansiedad de las matemáticas la padecen todos “los sujetos que han experimentado situaciones negativas con las matemáticas, evidenciando la ansiedad hacia esta área, la cual se entiende como la experimentación de sentimientos de angustia, estrés, bloqueos, inquietud, fatiga, preocupación”. Y en concordancia con esta conceptualización, según las palabras de Richardson y Suinn (1972) citados por Reyes Portillo (2025), es “un sentimiento de tensión y ansiedad que interfiere con la manipulación de los números, así como la resolución de problemas matemáticos en la vida cotidiana y situaciones académicas”. Lo cierto es que esta tendencia también se proyecta en las diferentes áreas del conocimiento que pertenecen al rubro de las ciencias básicas.

Finalmente, resalta el dato de un factor en particular, en donde de acuerdo con James, Baldwin y McInnis (1999), citados por Pérez Jaimes, et al. (2021) mencionan que “los factores que influyen en la elección que hacen los alumnos de una determinada institución tienen que ver con cuestiones de calidad; es decir, que éste considera que el docente presentará la misma característica”. De hecho, “la formación del docente debe ser permanente y continua, como reconocen los autores, para que este sea un verdadero agente transformador de la sociedad” (Pérez Jaimes, et al., 2021). No hacerlo así, existe un alto grado de incrementar los índices de reprobación, ya que investigaciones recientes demuestran que, “la calidad de la docencia impartida es el factor principal que los estudiantes toman en cuenta a la hora de elegir una universidad, superando otros factores, tales como infraestructura física, recursos académicos disponibles e incluso las perspectivas de empleabilidad” (Shah y Nair, 2010, citados por Pérez Jaimes, et al. 2021).

Por lo tanto y en concordancia con los presentes autores, se concluye que toda institución educativa, cualquiera que sea su nivel académico, debe tener como principal objetivo llevar a cabo la selección de personal docente con altos estándares de calidad académica, profesional y especializado en su área de conocimiento correspondiente.

4. METODOLOGÍA

Para la presente investigación se utilizó un enfoque Cuantitativo, ya que, de acuerdo con Hernández Sampieri, Fernández Collado y Baptista Lucio (2010, p. 4), “usa la recolección de datos para probar hipótesis, con base a la medición numérica y el análisis estadístico, para establecer patrones de comportamiento y probar teorías”. De esta manera se espera aportar evidencia suficiente que dé soporte a la hipótesis alterna, previamente aquí formulada.

4.1 Diseño metodológico.

a) Se aplicará a la par de lo anterior, un diseño no experimental, de tipo transversal descriptivo, puesto que, “su

propósito es describir variables y analizar su incidencia e interrelación en un momento dado” (Hernández Sampieri, 2010, p. 151), para medir sus efectos en un determinado periodo de tiempo, en donde la unidad de análisis serán todos aquellos estudiantes de primer semestre inscritos en asignaturas de Ciencias Básicas.

b) Como propuesta inicial está en vías de desarrollo una propuesta didáctica con la estrategia innovadora “Aula Inteligente Integral (AII)”, que permite micro diagnósticos adaptativos capaces de detectar en tiempo y forma las debilidades en matemáticas, física y química que presenten los estudiantes.

4.2 Población, Muestra, Plan de procesamiento y análisis de información.

Para recabar información, se consideró una población universal “N” de **583 estudiantes**, que cursan los primeros semestres en el área de Ciencias Básicas para obtener una muestra “n” representativa de **232 alumnos a encuestar**, con margen de **error del 5%** y un **nivel de confianza del 95%** (para efectos de este estudio, se implementará la constante **1.96**). Se utilizó un Cuestionario autoadministrado para su gestión y aplicación, ya que, según lo expuesto por Hernández Sampieri, et al (2010, p. 235), “se proporciona directamente a los participantes” en este caso alumnos del ITD, quienes contestaron cada pregunta marcando las respuestas que consideraron más oportunas, evitando de esta manera el sesgo que se puede presentar por la participación de intermediarios. Este cuestionario incluyó preguntas de opción simple, selección múltiple y escala tipo Likert (1–5 opciones), diseñado para medir: Hábitos de estudio, percepción del entorno académico, factores personales y socioeconómicos, el acceso a recursos institucionales, pero, sobre todo, la carga emocional y motivacional del estudiante, objeto de estudio. Como recomendación para la consolidación del instrumento, se implementó primero una Prueba piloto y se hicieron los ajustes pertinentes para su posterior y correcta aplicación. Para el procesamiento de datos se realizó un ejercicio estadístico con el fin de “obtener distribuciones de frecuencia y efectuar diversos cálculos estadísticos...”, (Arias Galicia, 2007, p. 204), para poder validar los resultados obtenidos en esta investigación. En cuanto al análisis estadístico, se utilizó el software PASW 18, con la siguiente fórmula estadística:

$$n = \frac{k^2 \cdot p \cdot q \cdot N}{(e^2 \cdot (N-1)) + k^2 \cdot p \cdot q} \quad (1)$$

Donde:

N = Población universal.

n = Tamaño de la muestra a evaluar.

k = Nivel de Confianza. (1.96, del 95% de confianza).

p y q = Proporción de individuos que poseen o no una característica en este estudio con un 50% (0.5), cada uno.

e = error muestral deseado. (Este estudio utiliza el 5%).

Sustituyendo:

$$n = \frac{1.96^2 \cdot 0.5 \cdot 0.5 \cdot 583}{(0.05^2 \cdot (583 - 1)) + 1.96^2 \cdot 0.5 \cdot 0.5}$$

n = 232

Para esta investigación, se realizó un análisis descriptivo determinando datos estadísticos, como la mediana, la moda para medir tendencias, desviación típica y varianza. Según Bernal Torres (2006), “la moda es la categoría o puntuación que ocurre con mayor frecuencia en un registro de datos”, analizando gráficas resultantes ya que, de acuerdo con Hernández Sampieri, et al (2010, p. 305), “los datos casi siempre se recolectan de una muestra y sus resultados estadísticos se denominan estadígrafos; ejemplo de ello son la media o la desviación estándar de la distribución de una muestra”. Tomando como referencia **232 estudiantes** de primer semestre de las diversas carreras ofertadas en el ITD.

4.3 Prueba de Hipótesis con la Correlación de Pearson.

Se concluye que, como la significancia bilateral fue de $0.000 < 0.05$ entonces Sí existe una relación significativa entre los factores académicos, socioeconómicos y emocionales y los índices de reprobación de los estudiantes de ciencias básicas del Instituto Tecnológico de Delicias, durante el ciclo escolar 2025-2026. La hipótesis Sí se acepta y el grado de relación es de 0.230, indicando que existen problemas que afectan el rendimiento escolar (de un 100% posible, correlaciona en un 23%). El valor positivo de la correlación quiere decir que, en este caso, si existen problemas socioemocionales en los alumnos, aumenta la probabilidad de que, por falta de conocimientos previos, los alumnos reprueben. (Ver Tabla 1).

Tabla 1. Estadísticos de correlación de Pearson en PASW 18.

Correlaciones			
		3. ¿Existen problemas de estrés y ansiedad que afectan tu rendimiento escolar?	7. ¿Crees que te faltan conocimientos previos para entender los temas de ciencias básicas?
3. ¿Existen problemas de estrés y ansiedad que afectan tu rendimiento escolar?	Correlación de Pearson Sig. (bilateral) N	1 232	.230** .000 231
7. ¿Crees que te faltan conocimientos previos para entender los temas de ciencias básicas?	Correlación de Pearson Sig. (bilateral) N	.230** .000 231	1 231

** La correlación es significativa al nivel 0,01 (bilateral).

5. RESULTADOS Y/O DISCUSIÓN

En la presente investigación se consideró pertinente hacer un análisis correlacional y descriptivo a cerca de las preguntas que más inciden o impactan en el desempeño académico que presentan los estudiantes del ITD (en este caso, preguntas 3 y 7 por considerar que poseen una correlación significativa). Del conglomerado de gráficas en total, se seleccionó un grupo representativo de preguntas y su correspondiente valor porcentual. En primer lugar, se aportó evidencia suficiente a favor de la hipótesis alterna previamente planteada, con respecto a la hipótesis nula, pues permitió corroborar que sí existe una relación estadísticamente significativa entre los factores académicos, socioeconómicos y emocionales con los índices de reprobación de los estudiantes de ciencias básicas

del ITD durante el ciclo escolar 2025-2026, puesto que existe una significancia bilateral de $0.000 < 0.05$ y de un 100% posible, correlaciona en un 23%.

Esto refrenda e impone la necesidad de desarrollar nuevas estrategias didáctico-académicas en apoyo a los alumnos que presentan deficiencias. Como propuesta inicial está la estrategia innovadora “Aula Inteligente Integral (AII)”, ya que permite micro diagnósticos adaptativos capaces de detectar en tiempo y forma las debilidades en matemáticas, física y química que presentan los estudiantes, aunado a rutas personalizadas de aprendizaje, reforzado con recursos digitales y laboratorios virtuales de gamificación, cuya particularidad son las simulaciones interactivas para practicar con los conceptos abstractos de manera visual y lúdica, reduciendo la brecha entre la teoría y la práctica.

Para el análisis de Estadísticos Descriptivos por Frecuencias, se muestra un extracto de los gráficos más representativos, con los resultados obtenidos en el cuestionario aplicado a **232 estudiantes** y cuyas **figuras son de elaboración propia**.

En la fig. 3, se observa una disonancia entre el estrés y ansiedad de la **pregunta 3**, los cuales afectan al **50.43%** de la muestra (116 estudiantes), indicando que los jóvenes de primer semestre a los cuales se les aplico la encuesta tienen diferencia mínima de un 1%, lo que representa que a la mitad de los estudiantes les afecta el estrés y la ansiedad en su rendimiento escolar.



Fig. 3. Estrés y ansiedad en los estudiantes del ITD.

Resalta la fig. 4, en la **pregunta 7**, donde los estudiantes expresan que les falta conocimiento previo para entender las materias del área de ciencias básicas, con un 69.70% (172 estudiantes), y solo el 29.44% (60 estudiantes) cuentan con dichos conocimientos. Una revisión exhaustiva de la literatura concluyó que un 70% de los alumnos egresados de IES a nivel nacional adolecen de conocimientos previos, reflejando una amplia coincidencia con este 69.7%.

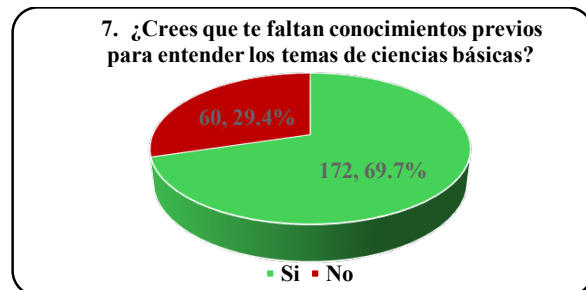


Fig. 4. Falta de conocimientos previos en ciencias básicas.



Analizando la fig. 5, **pregunta 15d**, un factor determinante es la falta de entendimiento en los contenidos de las materias, como la principal causa de reprobación en ciencias básicas de las distintas carreras del ITD, ya que incluye un 67.24% de los encuestados (156 estudiantes), en tanto que un 32.76% (76 estudiantes) no lo consideró así.

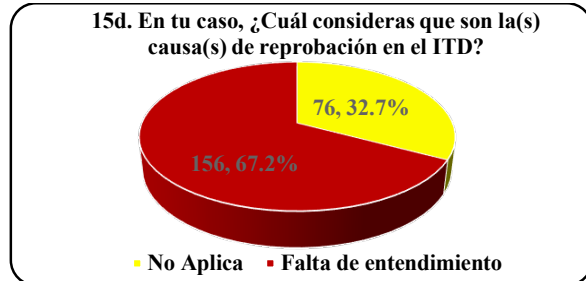


Fig. 5. Causas de reprobación en ciencias básicas.

En la fig. 6, **pregunta 16a**, se observa cómo un 71.55% de los encuestados (166 estudiantes), coinciden en que la dificultad en el aprendizaje de las diferentes materias de ciencias básicas, son el causal más importante de reprobación, mientras que un 28.45% (66 estudiantes) no lo consideran de esta manera.

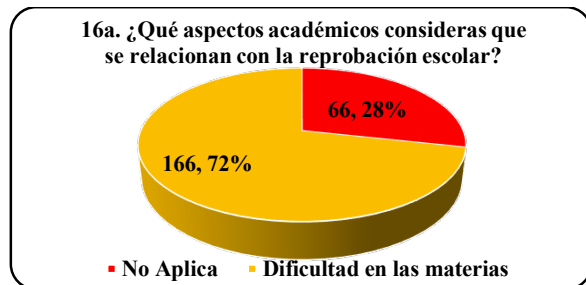


Fig. 6. Dificultad en las materias de ciencias básicas.

En la fig. 7, de la pregunta **17c**, un 62.16% de los encuestados (144 estudiantes), consideran que la falta de programas académicos de apoyo, son el factor más importante de reprobación, mientras un 37.84% (88 estudiantes) no opina lo mismo. Por ello se propone un semestre propedéutico.

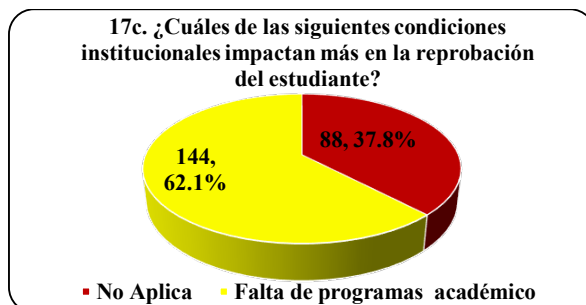


Fig. 7. Falta de programas académicos de apoyo al estudiante.

Finalmente, en la tabla 2 se muestran los resultados del análisis estadístico en donde se determinó la mediana, la moda, desviación típica y varianza.

Tabla 2. Estadísticos descriptivos desarrollados en PASW 18.

	1. Como institución de nivel superior, ¿fueron tu primera opción para cursar una carrera?	2. ¿Cuál es tu nivel de motivación hacia los estudios?	3. ¿Existen problemas de estrés y ansiedad que afectan tu rendimiento escolar?	4. ¿Cuál es tu estado de salud física?	5. ¿Cuál es tu estado emocional para desarrollarte satisfactoriamente en tus materias?	6. ¿Cuántas horas le dedicas a estudiar fuera de clase diariamente?	7. ¿Crees que te faltan conocimientos previos para entender los temas de ciencias básicas?	8. ¿La Institución te proporciona asesoría académica adicional en alguna materia de ciencias básicas?	9. ¿Y tú, has solicitado asesoría académica adicional en alguna materia de ciencias básicas?
N	232	232	232	232	232	232	231	232	232
Válidos	232	232	232	232	232	232	231	232	232
Perdidos	0	0	0	0	0	0	1	0	0
Mediana	1.0000	2.0000	1.0000	2.0000	2.0000	2.0000	1.0000	1.0000	2.0000
Moda	1.00	2.00	1.00	2.00	2.00	2.00	1.00	1.00	2.00
Desv. tip.	.49174	.67293	.50948	.79119	.76100	.70377	.47155	.21298	.47626
Varianza	.242	.453	.260	.626	.579	.495	.222	.045	.227
	10. ¿Cómo consideras la comunicación entre docentes y estudiantes en las clases que se le imparten sobre materias de ciencias básicas?	11. ¿Consideras que los exámenes evalúan los temas vistos en clase, de todas las materias que cursas de ciencias básicas?	12. ¿Convives en un entorno familiar adecuado que favorezca tus estudios?	13. ¿Cuentas con los recursos económicos necesarios para desempeñarte satisfactoriamente en tus estudios profesionales, de tal manera que no afecte tu rendimiento académico?	14. Si tu respuesta a la pregunta anterior es SI, deja en blanco esta pregunta. Si tu respuesta es NO, ¿Cuál de las siguientes opciones te favorecería más para un mejor desempeño académico?	15a. En tu caso, ¿Cuál consideras que es o son la(s) causa(s) de reprobación en esta Institución de Educación Superior?	15b. En tu caso, ¿Cuál consideras que es o son la(s) causa(s) de reprobación en esta Institución de Educación Superior?	15c. En tu caso, ¿Cuál consideras que es o son la(s) causa(s) de reprobación en esta Institución de Educación Superior?	15d. En tu caso, ¿Cuál consideras que es o son la(s) causa(s) de reprobación en esta Institución de Educación Superior?
N	232	232	230	230	64	232	232	232	232
Válidos	232	232	230	230	64	232	232	232	232
Perdidos	0	0	2	2	168	0	0	0	0
Mediana	2.0000	1.0000	1.0000	1.0000	1.0000	.0000	.0000	.0000	4.0000
Moda	2.00	1.00	1.00	1.00	1.00	.00	.00	.00	4.00
Desv. tip.	.64569	.33143	.31194	.42817	.62994	.31613	.93997	1.47840	1.88139
Varianza	.417	.110	.097	.183	.397	.100	.884	2.186	3.540
	16a. ¿Qué aspectos académicos consideras que se relacionan con la reprobación escolar? Dificultad	16b. ¿Qué aspectos académicos consideras que se relacionan con la reprobación escolar? Escasa Orientación	16c. ¿Qué aspectos académicos consideras que se relacionan con la reprobación escolar? Métodos poco motivadores	16d. ¿Qué aspectos académicos consideras que se relacionan con la reprobación escolar? Falta de actualización	17a. ¿Cuáles de las siguientes condiciones institucionales impactan más en la reprobación del estudiante? Acceso limitado	17b. ¿Cuáles de las siguientes condiciones institucionales impactan más en la reprobación del estudiante? Falta de programas Psicológico	17c. ¿Cuáles de las siguientes condiciones institucionales impactan más en la reprobación del estudiante? Falta de programas Académico	17d. ¿Cuáles de las siguientes condiciones institucionales impactan más en la reprobación del estudiante? Trámites complicados	
N	232	232	232	232	222	222	222	222	
Válidos	232	232	232	232	222	222	222	222	
Perdidos	0	0	0	0	10	10	10	10	
Mediana	1.0000	.0000	3.0000	.0000	.0000	.0000	3.0000	.0000	
Moda	1.00	.00	3.00	.00	.00	.00	3.00	.00	
Desv. tip.	.45214	.74277	1.50050	1.26453	.43001	.79909	1.45824	1.74115	
Varianza	.204	.552	2.252	1.599	.185	.639	2.126	3.032	

Es importante enfatizar que estos resultados estadísticos son válidos únicamente para el objeto de estudio de la muestra estudiantil del ITD, por lo tanto, se requiere ampliarlo a otras poblaciones estudiantiles para obtener conclusiones globales.

6. CONCLUSIONES

En el presente estudio es importante hacer énfasis que, si bien los tres factores: académicos, socioeconómicos y emocionales son la base estructural de la problemática a abordar, sin duda el factor académico reviste vital importancia y se convierte en un foco rojo de atención inmediata, forzando en primer lugar a un análisis diagnóstico para detectar las áreas de oportunidad en las que se debe de enfocar esta investigación y el aporte de las soluciones pertinentes, en aras de reducir o minimizar los altos índices de reprobación presentes en el ITD, situación que se manifiesta incluso a nivel nacional. Por ello, se propone la implementación nuevamente de un semestre propedéutico, para nivelar los conocimientos en áreas de las ciencias básicas de todos los estudiantes que adolecen de saberes previos y tienen fallas estructurales en el manejo de constructos.

Adicionalmente, en un segundo momento este equipo de investigación propone, la implementación de nuevas estrategias como la denominada “Aula Inteligente Integral (AII)”, basada en un ecosistema digital que conecta un diagnóstico académico permitiendo un refuerzo personalizado,



un apoyo socioeconómico para el acceso equitativo a recursos y un seguimiento emocional con el fin de prevenir eventos de crisis y deserción.

Con ello se pretende atender simultáneamente las tres dimensiones antes mencionadas, gestionado en una plataforma única, donde los docentes e investigadores pueden visualizar indicadores en tiempo real y tomar decisiones basadas en datos. El impacto esperado es una reducción significativa de la reprobación, una mayor equidad en el acceso a recursos y oportunidades, así como la prevención temprana de problemas emocionales que afectan el rendimiento, generando evidencia científica para publicar resultados y replicar el modelo en otros Institutos Tecnológicos.

Por último, el aspecto socioemocional, se considera a título particular, que esta dimensión implica esfuerzos multidisciplinarios de especialistas, tanto en el área pedagógica como en psicología aplicada, por ende, se propone una nueva línea de investigación para académicos, investigadores y todo aquel usuario que desee abordar esta problemática. Esta línea sería: “Análisis de las capacidades cognitivas en estudiantes de educación superior desde un enfoque crítico”.

REFERENCIAS

- Alanís Villaseñor, J., Cera Gaytán, R. S., Chávez Moreno, G. J., Lagarda Leyva, H., & Requeses Ruiz, J. L. (2022). Exámenes inversos del TECNM-campus Cd. Juárez. *REVISTA IPSUMTEC*, 5(1), 18–22. <https://doi.org/10.61117/ipsumtec.v5i1.296>
- ANUIES (2024). Algoritmo para la proyección de la cobertura estatal y nacional de la educación superior. Indicadores. Consultado en: <https://ampliacioncobertura.anui.es.mx/indicadores>
- Arias Galicia, L. (2007). Metodología de la Investigación. 7ª. Edición. México: Editorial Trillas.
- Fuentes Balderrama, J. y Rivera Heredia, M. (2022). Factores Personales, Contextuales y Académicos Asociados a la Reprobación en Educación Superior. Una Muestra Nacional Mexicana*. Consultado en: https://www.scielo.org.mx/scielo.php?script=sci_artext&pid=S1405-66662022000401039
- García Villaseñor, M. (2025). Gasto educativo 2025. Ampliaciones en el PEF. Consultado en: <https://ciep.mx/gasto-educativo-2025-ampliaciones-en-el-pef/>
- Granja García, M. A., Ramírez Gámez, B. E., Garcés Rodríguez, A. R., Villalobos Fernández, O., Zapata Silva, I., Hernández Sierra, M. G y Martínez Hernández, T. J. (2022). Factores asociados al índice de reprobación de asignaturas de ciencias básicas del ITSLP. Consultado en: <https://ciencialatina.org/index.php/cienciala/article/view/1610/2255>
- Hernández Sampieri, R., Fernández Collado, C. & Baptista Lucio, P. (2010). Metodología de la Investigación. 5ª. Edición. Chile: Editorial Mc Graw Hill.
- Meza, H.; Meza, D.; Arguello, F.; Vaca, E. (2024). Miedo a las matemáticas: ¿Por qué a los estudiantes no les gusta esta asignatura? Revista Científica Mundo Recursivo, 7(1), 1-19. Consultado en: <https://www.atlantic.edu.ec/ojs/index.php/mundor/article/view/204/290>
- Oaxaca Luna, J. y Valderrama Bravo, M. (2021). Matefobia. Consultado en: https://virtual.cuautitlan.unam.mx/CongresoCiTec/Memorias_Congreso/Anio5_No5/Extensos/EM-03.pdf
- Pérez Jaimes, A., Estrada Reye, C. y Estrada Gutiérrez, C. (2021). Factores docentes y su asociación con el índice de reprobación en alumnos de nivel superior. Consultado en: https://www.researchgate.net/publication/353788287_Factores_docentes_y_su_asociacion_con_el_indice_de_reprobacion_en_alumnos_de_nivel_superior
- Pérez Reveles, M. (2016). Impacto de la eficiencia terminal de la educación superior en México en la economía educativa. Consultado en: https://yuss.me/revistas/ese/ese2016v11n45a06p133_156.pdf
- Reyes Portillo, Z. (2025). ANSIEDAD HACIA LAS MATEMATICAS, UNA REALIDAD DE AULA. Consultado en: <https://revistas.upel.edu.ve/index.php/dialectica/article/view/4696/5166>
- Rodríguez Camargo, R., Álvarez Hilario, V., & Garduño Teliz, E. (2025). Revisión sistemática para determinar las causas de reprobación en el nivel superior. *Transdigital*, 6(11), e442. Consultado en: <https://doi.org/10.56162/transdigital442>
- Saucedo Fernández, M., Herrera-Sánchez, S. del C., Díaz Perera, J. J., Bautista Maldonado, S. y Salinas Padilla, H. A. (2014). Indicadores de reprobación: Facultad de Ciencias Educativas (UNACAR). Consultado en: <https://www.redalyc.org/pdf/4981/498150317007.pdf>
- TecNM (2025). Comisión Interna de Administración. Cuarta Sesión Ordinaria 2025. Consultado en: https://www.tecnm.mx/menu/transparencia/informes_gestion/2025/4SOCIA2025WEB.pdf?a=2
- Torres Zapata, A., Acuña Lara, J. y Moguel Ceballos, J. (2019). Índice de reprobación en universitarios. Una aproximación para la toma de decisiones. Consultado en: <https://revista.uniandes.edu.ec/ojs/index.php/mikarimin/article/view/1257/809>
- UNAM Global (2017). La mitad de los universitarios en México estudian y trabajan. Consultado en: https://unamglobal.unam.mx/global_revista/la-mitad-de-los-universitarios-en-mexico-estudian-y-trabajan/#:~:text=%C2%BFC%C3%B3mo%20consiguieron%20empleo?,la%20escuela%20con%20el%20trabajo
- Zapata, L. (2018). Un 70 por ciento de alumnos de nivel licenciatura no domina matemáticas. Consultado en: <https://www.yucatan.com.mx/mexico/2018/01/21/70-ciento-alumnos-nivel-licenciatura-no-domina-matematicas.html>

TiO₂ nanotubes with CuO quantum dots supported on thin films and their application in the photocatalytic production of hydrogen

Luis Antonio García Contreras*, José O. Flores Flores**, Rodolfo Zanella Specia**, Roberto Y. Sato Berrú**, José A. Chávez Carvayar*

* Instituto de Investigaciones en Materiales, Universidad Nacional Autónoma de México, Circuito Exterior S/N, Ciudad Universitaria, 04510, México. (E-mail: lagc57an@gmail.com.)

** Instituto de Ciencias Aplicadas y Tecnología, Universidad Nacional Autónoma de México, Circuito Exterior S/N, Ciudad Universitaria, 04510, México.

Abstract: This work describes the experimental procedure for the synthesis of TiO₂ nanotubes with CuO quantum dot deposition and their application in hydrogen production via photocatalysis. The synthesis was carried out using alkaline hydrothermal treatments with NaOH on TiO₂ powders, followed by washing with deionized water to obtain the TiO₂ nanotubes. Subsequently, a hydrothermal process was used to promote the phase transformation of the TiO₂ material and deposit the CuO quantum dots. The resulting materials were deposited onto SiO₂ substrates using the immersion method to obtain thin films. The synthesized materials were characterized by X-ray diffraction (XRD), N₂ adsorption-desorption, transmission electron microscopy (TEM), Raman spectroscopy, and UV-Vis spectroscopy. Photocatalytic tests in hydrogen production showed a significant improvement of up to 28 times the amount of hydrogen produced in the materials, indicating an improvement in the capacity of the TiO₂ material with CuO quantum dot deposition and its effect as a catalytic converter.

Keywords: Quantum dots, TiO₂, CuO, photocatalysis, hydrogen.

1. INTRODUCCIÓN

Actualmente, un tema de gran interés, es el desarrollo de tecnologías para obtener energías renovables para satisfacer las necesidades energéticas mundiales minimizando el impacto ambiental (Singh 2021; Hassan et al. 2024; Laimon & Yusaf 2024). Una de estas alternativas prometedoras es la energía del hidrógeno, entre las tecnologías disponibles para producirlo, se encuentra la fotocatálisis, basada en el uso de materiales semiconductores como el TiO₂ (Ishaq et al. 2021; Li et al. 2023; Sarout et al. 2025; Ahmed I. Osman, Ahmed M. Elgarahy, Abdelazeem S. Eltaweil, Eman M. Abd El-Monaem, Hisham G. El-Aqapa, Yuri Park, Yuhoon Hwang, Ali Ayati, Mohamed Farghali, Ikko Ihara, Ala'a H. Al-Muhtaseb, David W. Rooney, Pow-Seng Yap & Mika Sillanpää); este enfoque aún requiere una importante labor de investigación y desarrollo, en particular en la identificación de materiales que permitan reacciones fisicoquímicas más eficientes debido a las reacciones redox derivadas de las propiedades de los materiales empleados (Nasr et al. 2018; Xu & Weng 2023; Li et al. 2024; Ahmed I. Osman, Ahmed M. Elgarahy, Abdelazeem S. Eltaweil, Eman M. Abd El-Monaem, Hisham G. El-Aqapa, Yuri Park, Yuhoon Hwang, Ali Ayati, Mohamed Farghali, Ikko Ihara, Ala'a H. Al-Muhtaseb, David W. Rooney, Pow-Seng Yap & Mika Sillanpää); siendo los materiales nanoestructurados y con morfologías orientadas los

que han demostrado un gran potencial para esta aplicación debido a sus propiedades dependientes del tamaño y la forma (Ishaq et al. 2021; Sarout et al. 2025; Muhammad Rafique*Syeda HajraMuneeb IrshadMuhammad UsmanMuhammad ImranMohammad A. AssiriWaqar Muhammad Ashraf*) así como el transporte efectivo de las cargas a la superficie, donde las reacciones fisicoquímicas tienen lugar de manera eficiente. Existen diversas tecnologías para mejorar los materiales nanoestructurados de TiO₂, como la ingeniería de superficies y mejora de los procesos fotocatalíticos. Estos enfoques buscan mejorar la eficiencia del fotocatalizador optimizando sus propiedades, tales como aumentar la densidad de sitios activos superficiales, mejorar la transferencia de carga, reducir los procesos de recombinación de portadores de carga y extender la absorción de luz hacia la región visible (Augugliaro et al. 2006; Hu, Li & Yu 2010; Bai et al. 2015; Bai et al. 2015).

Entre las numerosas morfologías nanoestructuradas existentes de materiales de TiO₂ con formas unidimensionales son los nanotubos de TiO₂ los que destacan como candidatas prometedoras para una mayor mejora debido a sus ventajosas propiedades morfológicas (Zhang et al. 2016; Zhang et al. 2017; Cai et al. 2019; Chen et al. 2019; Chang, Lin & Chou 2020) debido a que sirven como plataformas adecuadas para posteriores mejoras tales como la deposición de superficie con otros



materiales, lo que puede aumentar significativamente su versatilidad (Zhang et al. 2017; Cai et al. 2019; Chang et al. 2020). Entre los materiales utilizados para mejorar el TiO_2 , destaca el CuO (Reddy et al. 2018; García-Contreras et al. 2026), que logra una mejora significativa en la producción de hidrógeno como co-catalizador, pero existe una mejora más significativa resultado de su conformación en puntos cuánticos por sus efectos co-catalíticos o de sensibilizador en la superficie del TiO_2 , que mejoran su fotorespuesta de TiO_2 (Ohtani 2013; Guerrero-Araque, Acevedo-Peña, Ramírez-Ortega, Calderon, et al. 2017); sin embargo, aún quedan varios desafíos críticos por abordar, principalmente en los métodos de síntesis que requieren condiciones en las cuales muchos de los materiales de TiO_2 con características unidimensionales muestran transformaciones estructurales y morfológicas, además de la separación eficiente, la transferencia y la supresión de la recombinación de portadores de carga y de extender el rango de operación del TiO_2 . Por las razones expuestas anteriormente, la síntesis adecuada de nanotubos de TiO_2 es de interés conservando la morfología lo máximo posible con depósito de materiales co-catalíticos como el CuO en su conformación de punto cuántico (Bai et al. 2015; Guerrero-Araque, Acevedo-Peña, Ramírez-Ortega, Calderon, et al. 2017; Reddy et al. 2018; Piri, Shafiee Afarani & Arabi 2019). En este trabajo, se presentan las condiciones de síntesis que permiten la síntesis de morfologías unidimensionales como los nanotubos de TiO_2 y depósito de puntos cuánticos de CuO , el método presentado permitió el depósito de CuO con tamaños de puntos cuánticos, lo que contribuyó a una mejor producción de hidrógeno.

2. METODOLOGÍA EXPERIMENTAL

2.1 Reactivos

Todos los productos químicos fueron de grado analítico y se usaron sin purificación adicional a la mencionada en la etiqueta: nitrato de cobre (II) ($\text{Cu}(\text{NO}_3)_2 \cdot 2\frac{1}{2}\text{H}_2\text{O}$, CAS 10031-43-3), ácido clorhídrico (HCl , Sigma-Aldrich, 37% CAS 7647010), hidróxido de sodio (NaOH , Sigma-Aldrich, 98% CAS 1310732), titanía P25 (TiO_2 , Degussa, Aerioxide 99.5% CAS 13463677), metanol grado HPLC (CH_3OH , CAS 67-56-1); y se usó agua desionizada a lo largo de este trabajo.

2.2 Método de síntesis

Se colocaron 1 g de TiO_2 en un recipiente de teflón de 100 ml con 75 ml de una disolución 10M de NaOH , se dispersa el material durante 30 min a temperatura ambiente. Posteriormente, la suspensión se transfirió a una autoclave de acero inoxidable y se mantuvo a una temperatura de 120 °C durante 24 h. El producto resultante se lavo mediante ciclos para ser neutralizado y secó durante 24 h a 60 °C obteniendo nanotubos de titanato de sodio ($\text{H}_2\text{Ti}_3\text{O}_7$). El depósito de los puntos cuánticos de CuO , se realizó mediante el siguiente procedimiento; Se tomaron 0.5 g de $\text{H}_2\text{Ti}_3\text{O}_7$ y fueron sometidas a un proceso hidrotermal con una temperatura de 80 °C por 24 h con una mezcla de $\text{Cu}(\text{NO}_3)_2$ y agua desionizada que fue homogenizada por agitación vigorosa durante 12 h, a la cual se le agregó mediante goteo (1 mL/h).

Después de lograr la homogenización del material, se agregó un recipiente de teflón y a una autoclave. Una vez terminado el proceso hidrotermal, se lavaron los materiales con agua desionizada por ciclos de lavado para posteriormente, ser depositados por medio de *dip-coating* sobre sustratos de vidrio (tratados superficialmente con disoluciones de 1M de HF) con un proceso de impregnando de 1 $\text{cm}^2/10$ min y un pH neutro, y que fueron secados a una temperatura de 100 °C durante 12h.

2.3 Caracterización

Los patrones DRX de los materiales sintetizados fueron obtenidos en un difractómetro marca BRUKER Avancen D8 a temperatura ambiente, usando la radiación Cu K_α ($\lambda = 0.15405$ nm), voltaje 40 kV, corriente 100 mA, que posee una configuración y un monocromador de grafito en el haz secundario. La intensidad de difracción se midió en el intervalo de 20° a 80°, con pasos de 0.02°. Las morfologías de los materiales, se observaron mediante un Microscopio Electrónico de Barrido (MEB) de Ultra-Alta Resolución JSM-7800F (Con una distancia de trabajo de 4 y 8 mm, y voltajes de 2–10 kV) detectando los electrones secundarios (Modo SEI). Se registraron imágenes de Microscopía Electrónica de Transmisión de alta resolución (HRTEM) en un espectrofotómetro de emisión de campo Schottky JEOL 2100 JEOL ARM200F con resolución atómica a un voltaje de aceleración de 200 kV. Los espectros Raman, fueron obtenidos utilizando un equipo Raman-WITec confocal, utilizando un láser de 532 nm con ajuste de potencia de 3.4 mW. Las isothermas de adsorción-desorción de N_2 se midieron a 77 K con el analizador de área superficial Quantachrome Autosorb 1, mediante el método Brunauer-Emmett-Teller (BET), en un rango de presión relativa de 0.05 a 0.30. Los espectros de reflectancia difusa UV-Vis de los materiales se obtuvieron a temperatura ambiente con un espectrómetro Cary 5000 Varian (UV-VIS-NIR), equipado con un accesorio de reflectancia difusa (mantis religiosa). Los espectros se obtuvieron en el rango de 200 a 800 nm, utilizando politetrafluoroetileno como material estándar de reflectancia.

2.4 Pruebas para la producción y cuantificación fotocatalítica de hidrógeno

Tras establecer las condiciones de síntesis, se preparó una cantidad suficiente de cada material para garantizar la reproducibilidad en los experimentos fotocatalíticos, los cuales se realizaron por triplicado. La producción de H_2 se llevó a cabo utilizando una mezcla de 200 mL de agua y metanol (1:1 vol.) y 20 mg de los materiales sintetizados en un reactor fotocatalítico. En cada experimento, las mezclas se mantuvieron en agitación con un agitador magnético; inicialmente durante 30 min en la oscuridad para promover el equilibrio de adsorción/desorción, antes de la irradiación artificial, mientras el sistema se purgaba con gas (N_2). Una vez detenido el flujo de N_2 , el reactor se selló herméticamente y se irradió con una lámpara UV Pen-Ray ($\lambda = 254$ nm). La cuantificación de H_2 se realizó cada 30 min durante 5 h utilizando un cromatógrafo de gases con un detector de conductividad térmica Thermo Scientific y una columna empaquetada Trace PLOT TG-BOND Msieve 5A.3.

3. RESULTADOS

Los difractogramas y espectros Raman obtenidos, son característicos de $\text{H}_2\text{Ti}_3\text{O}_7$ con una transformación hacia la fase de anatasa, en la Figura 1 a) y b), se muestra el patrón DRX con la ficha de identificación, no se detectaron reflexiones adicionales de los puntos cuánticos debido a la menor proporción de estos presentes en el material, y que se encuentra por debajo del límite de detección del equipo, así como otras reflexiones relacionadas con impurezas (Zhang & Banfield 2000; Zhang & Banfield 2000; Z.Xin Qian 2017). Adicionalmente, en la Figura 1 a) y b) se muestra la estimación de la relación de fases utilizando el método de deconvolución numérica para separar los picos característicos de cada fase cristalina de TiO_2 en los patrones de XRD. En este ejemplo, el patrón de difracción muestra el intervalo 2θ que va de 23 a 33° , en el que los picos principales de anatasa (101) y las posiciones de acuerdo a las fichas de identificación de titanato de hidrógeno y cloruro de sodio residual; después de la sustracción de la línea base y la corrección de polarización de Lorentz se ajustaron utilizando curvas de Pearson VII (He et al. 2019; Zhu et al. 2021; Szoldra et al. 2023).

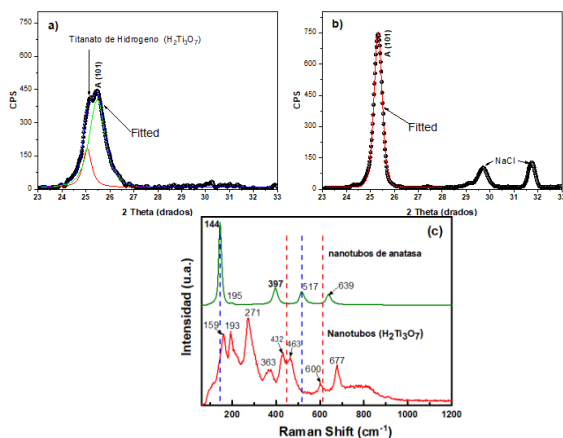


Figura 1. a) y b) Difracción de rayos X del material (JCPDS: 96-900-9087 y 21-1272) y c) Espectros Raman de las fases presentes de TiO_2 en los materiales con depósitos de puntos cuánticos de CuO .

El contenido de fase en cada material se calculó a partir de las intensidades integradas de los picos característicos mencionados anteriormente, se puede calcular mediante el método bien conocido de contribuciones de porcentaje de fases donde la predominancia al final de la síntesis fue de la fase anatasa (He et al. 2019; Zhu et al. 2021; Szoldra et al. 2023). Además, estas fases cristalinas presentes en el material también fueron identificada por sus planos cristalinos como se muestra en la imagen HRTEM (Figura 2), lo que fue confirmada mediante las distancias interplanares. En la Figura 1c se observa las bandas de vibración Raman características de estas fases, mediante los modos vibracionales presentes en el espectro (Ceballos-Chuc et

al. 2018; Chen et al. 2019; Chalastara et al. 2020; Du et al. 2022; García-Contreras et al. 2022). De manera similar con depósitos de Cu realizado, no se pudieron observar bandas relacionadas a este material debido a la baja proporción de este (Chen et al. 2013; Khan et al. 2020; Sadanandam et al. 2021).

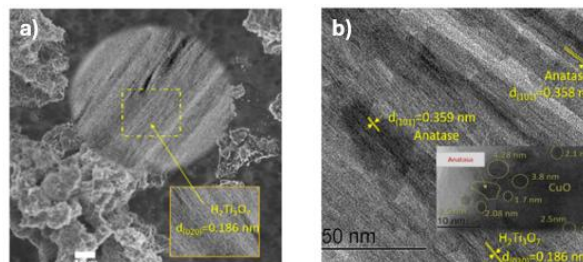


Figura 2. a) Imágenes MEB y b) MET de alta resolución de las morfologías y de los puntos cuánticos de CuO depositados.

Los valores del área superficial del material también se analizaron antes y después de los depósitos de CuO en la Figura 3 se presentan las isoterma de adsorción/desorción de nitrógeno (N_2) de los materiales, con sus valores de área superficial estimados a partir de las isoterma, utilizando método BET. Se observó que los materiales sintetizados presentaron isoterma de tipo IV y ciclos de histéresis H_3 , en el rango de presión relativa (P/P_2) de 0.5 a 1.0, lo que revela la existencia de estructuras con propiedades mesoporosas según la clasificación de la IUPAC (Sing 1985; Liao et al. 2012; Rui et al. 2013; Kaplan, Erjavec & Pintar 2015; Makal & Das 2018; Xu et al. 2019; Chalastara et al. 2020). Además, no presentan una plataforma de adsorción límite en condiciones de presión de saturación, característica de materiales laminares y porosos (Zhang et al. 2017; Xu et al. 2019), característicos de morfologías orientadas a partir de titanatos, donde las áreas superficiales suelen ser menores a $100 \text{ m}^2 \text{ g}^{-1}$ (Niederberger & Cölfen 2006; Zhao & Yan 2011; Lyu et al. 2020), debido a los espacios internos de las morfologías (Camposco et al. 2012; Wetchakun et al. 2012; Chalastara et al. 2020).

En la Figura 4, se observa el valor de la brecha de banda óptica (E_g), calculada a partir de los espectros UV-Vis; el análisis previo verifica que de acuerdo a lo reportado, las fases de materiales de TiO_2 como la anatasa y $\text{H}_2\text{Ti}_3\text{O}_7$ presentan características de un semiconductor de transición indirecta (3.28 y 3.47 eV respectivamente) (Žerjav et al. 2022); por lo cual en este trabajo se calculó el valor de E_g -indirecto mediante el método Tauc, utilizando el gráfico $(\alpha h\nu)^{1/n}=A(h\nu-E_g)$ obtenido a partir de los espectros de reflectancia difusa (López & Gómez 2012; Praveen et al. 2014; Tran et al. 2017; Makal & Das 2018; Freyria et al. 2020; Khan et al. 2020).

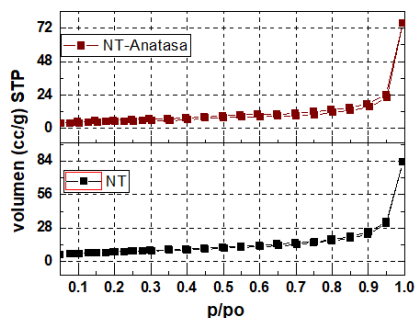


Figura 3. Isothermas de absorción/desorción de N_2 del material de TiO_2 antes (NT) y después del depósito de CuO (NT-Anatasa).

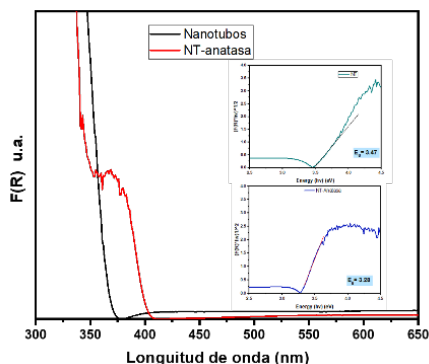


Figura 4. a) Espectros de reflectancia difusa UV-vis de los materiales de TiO_2 con (NT-Anatasa) y sin CuO (Nanotubos). Además, del ejemplo del cálculo de valor de la brecha de banda óptico en el material con depósito de puntos cuánticos de CuO .

Los espectros de reflectancia difusa se adquirieron utilizando la función Kubelka-Munk ($F(R)$) incluida en el software de procesamiento de datos del espectrofotómetro. Por lo tanto, el coeficiente de absorción (α) se considera proporcional a ($F(R)$) y se sustituyó en el gráfico Tauc. Los intervalos de banda (E_g) se determinaron extrapolando la línea recta de los gráficos en el punto $(F(R) \cdot h\nu)^{1/2} = 0$ (García-Contreras et al. 2025; García-Contreras et al. 2026). En los espectros UV-Vis se observó el umbral y desplazamiento gradual del borde de absorción hacia el azul de absorción atribuido a la transferencia de carga de $O^{2-} \rightarrow Ti^{4+}$ (Henderson 2011; Henderson 2011; Xin et al. 2016; Romeiro et al. 2018; Gaur et al. 2019), se aprecia los valores de $E_g \sim 3.06$ eV para $H_2Ti_3O_7$ y un valor adicional ~ 3.42 eV relacionado con la bifase de TiO_2 y depósito de los puntos cuánticos, el depósito de CuO como co-catalizador provocó una disminución en los valores de band gap haciendo al material más eficiente al ser excitado por una longitud de onda cercana al UV-A y UV-C, por lo cual puede tener un mejor comportamiento en la absorción del espectro de trabajo y que se pudo relacionar con el mejor proceso de producción de hidrógeno.

La producción y cuantificación fotocatalítica de hidrógeno se llevó a cabo, mediante un arreglo experimental descrito en la sección experimental y que se presenta en la Figura 5. Una vez

colocadas apropiadamente las películas delgadas en el reactor y alcanzar las condiciones de experimentación. Se procedió a cuantificar la cantidad de hidrógeno que se produce en los ensayos fotocatalíticos.

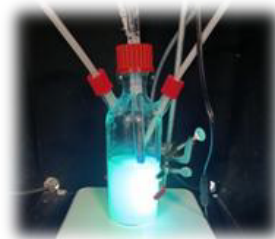


Figura 5. Arreglo experimental, reactor fotocatalítico.

Como primer ensayo, se midió la fotólisis de la fuente de radiación UV-C para determinar la contribución de ella en la producción de hidrógeno (μmol), posteriormente se determinó la actividad que presenta el material de referencia como el Degussa P25, con el fin de cuantificar la producción de H_2 que produce este material y que comúnmente es tomado en los trabajos como referencia en las mejoras de los materiales basados en TiO_2 (Figura 6).

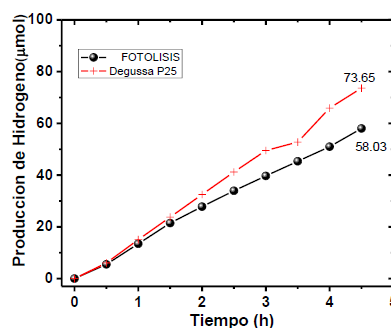


Figura 6. Producción de hidrógeno (μmol) en los ensayos base de fotólisis usando la fuente de luz UV-C ($\lambda=254$ nm) y el material Degussa P25.

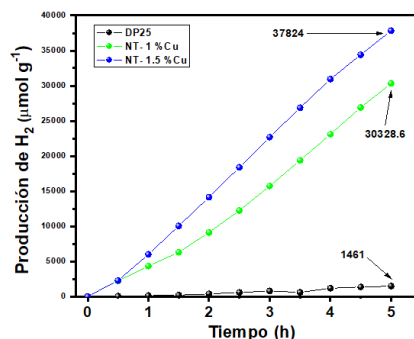


Figura 7. Producción de hidrógeno ($\mu mol g^{-1}$) de los materiales con 1 y 1.5 %wt de Cu en los nanotubos de TiO_2 y el material Degussa P25.

Los resultados de estos ensayos fotocatalíticos se muestran en la Figura 7, donde se observa la producción de H_2 en función del tiempo de reacción en $\mu\text{mol } H_2 \text{ g}^{-1}$. La mejora de la actividad en la producción de H_2 con depósito de puntos cuánticos de CuO se presenta en relación a la cantidad de Cu depositado en los materiales (García-Contreras et al. 2026). Esta mejora sustancial fue relacionada con el depósito de CuO que actuó como co-catalizador provocando una disminución en los valores de band gap, relacionado con la dispersión en la superficie de los materiales y a la transferencia de carga del CuO a los materiales de TiO_2 , además de su función como un sensibilizador, el cual absorbe la luz visible y excita a los electrones que a su vez son inyectados en la banda de conducción de TiO_2 causando así la disminución del valor del band gap (Sakata, Hashimoto & Hiramoto 2002; Rosu et al. 2013; Islas, Zanella & Saniger 2016; Guerrero-Araque, Acevedo-Peña, Ramírez-Ortega, Lartundo-Rojas, et al. 2017; Xiujie Wang et al.). La actividad mejorada en la producción de hidrógeno de los materiales al depositarles los puntos cuánticos de CuO, presentan una actividad mayor, en la producción de hidrógeno de hasta $37824 \mu\text{mol} \cdot \text{g}^{-1}$, que es aproximadamente 28 veces más que lo obtenido por el mismo material sin depósito de CuO y 26 veces más que el material de referencia Degussa P25 ($1461 \mu\text{mol} \cdot \text{g}^{-1}$). Esta actividad mejorada, está asociada, a la buena actuación del CuO como co-catalizador, dando un aumento en la vida útil de los portadores de carga para la transferencia de carga interfacial aumentando así la transferencia de electrones a la disolución en la reacción de reducción para formar hidrógeno (Xi et al. 2014).

4. CONCLUSIONES

Se prepararon materiales compuestos de TiO_2 con puntos cuánticos de CuO con diferente porcentaje en peso de cobre mediante el método hidrotermal, el material base de TiO_2 utilizado que contaba con características unidimensionales en forma de nanotubos con una transformación transicional de fase a anatasa, dieron como resultado una buena actividad en la producción de hidrógeno con el depósito de los puntos cuánticos de CuO visualizados por HRTEM relacionados a las sus características de estos materiales y propiedades, además de los cambios de cristalinidad, así como la disminución de la recombinación de las cargas fotogeneradas, desplazamiento de los bordes de absorción hacia la región del visible debido al depósito de CuO actuado de manera eficiente como co-catalizador en el transporte de cargas fotogeneradas; con lo que se incrementó la tasa de producción de H_2 de hasta $37824 \mu\text{mol} \cdot \text{h}^{-1} \cdot \text{g}^{-1}$.

AGRADECIMIENTOS

Luis Antonio García-Contreras agradece a SECIHTI por la beca CVU 667265. Los autores también agradecen al Laboratorio Universitario de Caracterización Espectroscópica (LUCE) y a la Dra. Selene R. Islas por la caracterización UV-Vis; y a la M.Sc. Viridiana Maturano Rojas por la caracterización del método BET; al Dr. Mario R. Rodríguez Varela por su asistencia técnica

en pruebas fotocatalíticas (Instituto de Ciencias Aplicadas y Tecnología, ICAT-UNAM); al Laboratorio Universitario de Microscopía Electrónica (LUME) del Instituto de Investigación en Materiales (IIM-UNAM) y a la M.Sc. Adriana Tejeda Cruz y el Fis. Josué E. Romero I. por su asistencia técnica en la caracterización estructural y morfológica. Los autores agradecen a la Dirección General de Asuntos del Personal Académico-UNAM por el apoyo financiero recibido a través de la beca PAPIIT IN104925.

REFERENCIAS

- Ahmed I. Osman, Ahmed M. Elgarahy, Abdelazeem S. Eltaweil, Eman M. Abd El-Monaem, Hisham G. El-Aqapa, Yuri Park, Yuhoon Hwang, Ali Ayati, Mohamed Farghali, Ikko Ihara, Ala'a H. Al-Muhtaseb, David W. Rooney, Pow-Seng Yap & Mika Sillanpää, Biofuel production, hydrogen production and water remediation by photocatalysis, biocatalysis and electrocatalysis | Environmental Chemistry Letters.
- Augugliaro, V., Litter, M., Palmisano, L. & Soria, J., 2006, 'The combination of heterogeneous photocatalysis with chemical and physical operations: A tool for improving the photoprocess performance', *Journal of Photochemistry and Photobiology C: Photochemistry Reviews*, 7(4), 127–144.
- Bai, S., Jiang, W., Li, Z. & Xiong, Y., 2015, 'Surface and Interface Engineering in Photocatalysis', *ChemNanoMat*, 1(4), 223–239.
- Cai, J., Shen, J., Zhang, X., Ng, Y.H., Huang, J., Guo, W., Lin, C. & Lai, Y., 2019, 'Light-Driven Sustainable Hydrogen Production Utilizing TiO_2 Nanostructures: A Review', *Small Methods*, 3(1), 1800184.
- Chang, Y.-C., Lin, J.-C. & Chou, C.-M., 2020, ' $H_2Ti_3O_7$ nanowires as a high-performance photocatalytic and surface-enhanced Raman scattering substrate', *Journal of Photochemistry and Photobiology A: Chemistry*, 400, 112666.
- Chen, J., Guan, M., Zhang, X. & Gong, X., 2019, 'Insights into a rutile/brookite homojunction of titanium dioxide: separated reactive sites and boosted photocatalytic activity', *RSC Advances*, 9(63), 36615–36620.
- García-Contreras, L.A., Flores-Flores, J.O., Zanella, R., Arenas-Alatorre, J.Á. & Chávez-Carvayar, J.Á., 2025, 'Photocatalytic performance of 1D- TiO_2 with flower-like morphology and tunable anatase, brookite, and rutile phases for ibuprofen degradation', *Ceramics International*, 51(28, Part A), 56029–56049.
- García-Contreras, L.A., Rodríguez-Varela, M., Sato-Berrú, R.Y., Zanella, R., Flores-Flores, J.O. & Chávez-Carvayar, J.Á., 2026, 'Temperature-Dependent 1-D TiO_2 Morphology with Brookite/Anatase/Rutile Phases for the Efficient Deposition of CuO Quantum Dots and Hydrogen



- Production Activity', *Micro and Nanostructures*, 214, 208625.
- Guerrero-Araque, D., Acevedo-Peña, P., Ramírez-Ortega, D., Calderon, H.A. & Gómez, R., 2017, 'Charge transfer processes involved in photocatalytic hydrogen production over CuO/ZrO₂-TiO₂ materials', *International Journal of Hydrogen Energy*, 42(15), 9744–9753.
- Guerrero-Araque, D., Acevedo-Peña, P., Ramírez-Ortega, D., Lartundo-Rojas, L. & Gómez, R., 2017, 'SnO₂-TiO₂ structures and the effect of CuO, CoO metal oxide on photocatalytic hydrogen production: SnO₂-TiO₂ structures and the effect of CuO, CoO metal oxide', *Journal of Chemical Technology & Biotechnology*, 92(7), 1531–1539.
- Hassan, Q., Viktor, P., J. Al-Musawi, T., Mahmood Ali, B., Algburi, S., Alzoubi, H.M., Khudhair Al-Jiboory, A., Zuhair Sameen, A., Salman, H.M. & Jaszczur, M., 2024, 'The renewable energy role in the global energy Transformations', *Renewable Energy Focus*, 48, 100545.
- He, J., Du, Y., Bai, Y., An, J., Cai, X., Chen, Y., Wang, P., Yang, X. & Feng, Q., 2019, 'Facile Formation of Anatase/Rutile TiO₂ Nanocomposites with Enhanced Photocatalytic Activity', *Molecules*, 24(16), 2996.
- Hu, X., Li, G. & Yu, J.C., 2010, 'Design, Fabrication, and Modification of Nanostructured Semiconductor Materials for Environmental and Energy Applications', *Langmuir*, 26(5), 3031–3039.
- Ishaq, T., Yousaf, M., Bhatti, I.A., Batool, A., Asghar, M.A., Mohsin, M. & Ahmad, M., 2021, 'A perspective on possible amendments in semiconductors for enhanced photocatalytic hydrogen generation by water splitting', *International Journal of Hydrogen Energy*, 46(79), 39036–39057.
- Islas, S.R., Zanella, R. & Saniger, J.M., 2016, 'Thermal activation process of Au/TiO₂ system: a molecular spectroscopy study', *RSC Advances*, 6(48), 42554–42560.
- Laimon, M. & Yusaf, T., 2024, 'Towards energy freedom: Exploring sustainable solutions for energy independence and self-sufficiency using integrated renewable energy-driven hydrogen system', *Renewable Energy*, 222, 119948.
- Li, R., Luan, J., Zhang, Y., Jiang, L., Yan, H., Chi, Q. & Yan, Z., 2024, 'A review of efficient photocatalytic water splitting for hydrogen production', *Renewable and Sustainable Energy Reviews*, 206, 114863.
- Li, X., Raorane, C.J., Xia, C., Wu, Y., Tran, T.K.N. & Khademi, T., 2023, 'Latest approaches on green hydrogen as a potential source of renewable energy towards sustainable energy: Spotlighting of recent innovations, challenges, and future insights', *Fuel*, 334, 126684.
- Lixia Yang†‡||Shenglian Luo*†‡§||Yue Li§Yan Xiao†‡§Qing Kang§Qingyun Cai*§, High Efficient Photocatalytic Degradation of p-Nitrophenol on a Unique Cu₂O/TiO₂ p-n Heterojunction Network Catalyst | *Environmental Science & Technology*.
- Muhammad Rafique*Syeda HajraMuneeb IrshadMuhammad UsmanMuhammad ImranMohammad A. AssiriWaqar Muhammad Ashraf*, Hydrogen Production Using TiO₂-Based Photocatalysts: A Comprehensive Review | *ACS Omega*.
- Nasr, M., Eid, C., Habchi, R., Miele, P. & Bechelany, M., 2018, 'Recent Progress on Titanium Dioxide Nanomaterials for Photocatalytic Applications', *ChemSusChem*, 11(18), 3023–3047.
- Ohtani, B., 2013, 'Titania Photocatalysis beyond Recombination: A Critical Review', *Catalysts*, 3(4), 942–953.
- Piri, F., Shafiee Afarani, M. & Arabi, A.M., 2019, 'Synthesis of copper oxide quantum dots: effect of surface modifiers', *Materials Research Express*, 6(12), 125006.
- Reddy, N.L., Emin, S., Kumari, V.D. & Muthukonda Venkatakrishnan, S., 2018, 'CuO Quantum Dots Decorated TiO₂ Nanocomposite Photocatalyst for Stable Hydrogen Generation', *Industrial & Engineering Chemistry Research*, 57(2), 568–577.
- Rosu, M.-C., Suciu, R.-C., Mihet, M. & Bratu, I., 2013, 'Physical–chemical characterization of titanium dioxide layers sensitized with the natural dyes carmine and morin', *Materials Science in Semiconductor Processing*, 16(6), 1551–1557.
- Sakata, T., Hashimoto, K. & Hiramoto, M., 2002, New aspects of electron transfer on semiconductor surface: dye-sensitization system, *ACS Publications*.
- Sarout, K., Barhoumi, S., Faik, A., Bora, D.K. & Mehmood, S., 2025, 'An account of metal and metal oxide based nanocomposite electrocatalysts for green hydrogen production', *International Journal of Hydrogen Energy*, 111, 592–605.
- Singh, S., 2021, 'Energy Crisis and Climate Change', *Energy*, pp. 1–17, John Wiley & Sons, Ltd.
- Szoldra, P., Frac, M., Lach, R., Zych, L., Radecka, M., Trenczek-Zajac, A. & Pichór, W., 2023, 'Effect of brookite on the photocatalytic properties of mixed-phase TiO₂ obtained at a higher temperature', *Materials Science and Engineering: B*, 287, 116104.
- Xi, Z., Li, C., Zhang, L., Xing, M. & Zhang, J., 2014, 'Synergistic effect of Cu₂O/TiO₂ heterostructure nanoparticle and its high H₂ evolution activity', *International Journal of Hydrogen Energy*, 39(12), 6345–6353.



- Xiujie Wang a, , Shengsen Zhang b, , Biyu Peng a, & , Hongjuan Wang a, Hao Yu a, Feng Peng a, Enhancing the photocatalytic efficiency of TiO₂ nanotube arrays for H₂ production by using non-noble metal cobalt as co-catalyst - ScienceDirect.
- Xu, F. & Weng, B., 2023, 'Photocatalytic hydrogen production: an overview of new advances in structural tuning strategies', Journal of Materials Chemistry A, 11(9), 4473–4486.
- Yong-Gang ZhangLi-Li MaJia-Lin LiYing Yu, In Situ Fenton Reagent Generated from TiO₂/Cu₂O Composite Film: a New Way to Utilize TiO₂ under Visible Light Irradiation | Environmental Science & Technology.
- Zhang, J., Tse, K., Wong, M., Zhang, Y. & Zhu, J., 2016, 'A brief review of co-doping', Frontiers of Physics, 11(6), 117405.
- Zhang, X., Wang, Y., Liu, B., Sang, Y. & Liu, H., 2017, 'Heterostructures construction on TiO₂ nanobelts: A powerful tool for building high-performance photocatalysts', Applied Catalysis B: Environmental, 202, 620–641.
- Zhu, X., Zhou, Q., Xia, Y., Wang, J., Chen, H., Xu, Q., Liu, J., Feng, W. & Chen, S., 2021, 'Preparation and characterization of Cu-doped TiO₂ nanomaterials with anatase/rutile/brookite triphasic structure and their photocatalytic activity', Journal of Materials Science: Materials in Electronics, 32(16), 21511–21524.

Systematic PID tuning for ethanol dehydration by extractive distillation using literature-based transfer functions

Lizeth Sarai Ávila-Loaiza*, Elsa Hynmar Fernández-Martínez*
Brian Manuel González Contreras*

* Facultad de Ciencias Básicas, Ingeniería y Tecnología, Universidad Autónoma de Tlaxcala, Tlaxcala, C.P. 90300, México (e-mail: 20256940@uatx.mx).

Abstract: Extractive distillation systems for ethanol dehydration exhibit complex, strongly coupled dynamics that challenge classical controller tuning. This work analyzes linear transfer function models for temperature and composition control loops in conventional extractive distillation columns and extractive top dividing-wall columns, taken from literature. The applicability of Ziegler–Nichols and Tyreus–Luyben methods is evaluated. Consequently, PI and PID controllers are systematically tuned using the MATLAB *pidtune* command with a parallel control structure. Closed-loop performance is assessed under a unit step change, using IAE, ISE, ITAE and ITSE. Temperature-based control consistently exhibited lower error indices than composition-based control across all cases analyzed. The proposed tuning framework is reproducible across different model structures and provides a practical reference for implementing product-purity control in intensified extractive distillation processes.

Keywords: Extractive distillation, ethanol dehydration, PID controller tuning, extractive top dividing-wall column, temperature and composition control, transfer function models.

1. INTRODUCCIÓN

La destilación extractiva (DE) es una de las tecnologías más empleadas para la separación de mezclas azeotrópicas o con volatilidades relativas cercanas, siendo la deshidratación de etanol un caso de estudio ampliamente documentado debido a su relevancia industrial y energética (Aquilon et al., 2023; Czarnecki et al., 2023; Huang et al., 2025; Neagu and Pricop-Nicolae, 2025; Xu et al., 2026). En este proceso, la adición de un solvente extractivo permite modificar el equilibrio líquido-vapor y alcanzar altas purezas del producto, aunque introduce una mayor complejidad operativa y dinámica en comparación con la destilación convencional (Gerbaud et al., 2019).

Con el fin de reducir costos de capital y mejorar la eficiencia del proceso, se han desarrollado configuraciones de destilación intensificada, entre las cuales destacan las columnas con pared divisoria (Dividing-Wall Columns, DWC). Las DWCs han surgido como una tecnología de intensificación de procesos capaz de integrar múltiples tareas de separación en una sola columna, logrando reducciones significativas en consumo energético y requerimientos de equipo (Lorenz et al., 2018). El diseño y operación de las DWC y sus variantes intensificadas sigue siendo objeto de investigación activa debido a los retos asociados a su control dinámico y estabilidad operativa (Li et al., 2023).

La extensión de este concepto a la DE ha dado lugar a las columnas extractivas con pared divisoria superior (E-TDWC), las cuales han demostrado un alto potencial para la deshidratación de etanol (Kiss and Ignat, 2012; Luyben,

2018; Torres et al., 2022). Sin embargo, estas ventajas se ven acompañadas de desafíos significativos en el control dinámico del proceso, derivados de las fuertes interacciones multivariantes, la presencia de múltiples constantes de tiempo dominantes y una elevada sensibilidad a perturbaciones operativas (Zhu et al., 2021; Zhang et al., 2025).

Diversos estudios han abordado el análisis dinámico y el diseño de esquemas de control para columnas de DE convencionales y E-TDWC, empleando tanto control basado en temperatura como control directo de composición (Tututi et al., 2014; Hernández-Escoto et al., 2022). El control por temperatura ha sido ampliamente adoptado debido a la rapidez de respuesta de los sensores y a su facilidad de implementación industrial, mientras que el control por composición permite una regulación más directa de la calidad del producto, aunque con mayores retos asociados al retardo de medición y a la instrumentación requerida (Luyben, 2006a; Kiss, 2013; Seborg et al., 2016).

Para el análisis y diseño de estos esquemas de control, la literatura reporta numerosos modelos dinámicos, usualmente expresados en forma de funciones de transferencia (FTs) lineales obtenidas alrededor de un punto de operación estacionario (Tututi et al., 2014; Hernández-Escoto et al., 2022; Román et al., 2024). Dichos modelos permiten estudiar el comportamiento dinámico del sistema y constituyen una base común para la sintonización de controladores PID. En este trabajo se realiza una evaluación sobre si estos modelos específicos cumplen con los requisitos teóricos de los métodos de Ziegler-Nichols (Z-N) y Tyreus-Luyben (T-L).

En este contexto, los métodos clásicos de sintonización de Z-N y T-L han sido ampliamente utilizados debido a su simplicidad y aceptación histórica en la industria de procesos (Aström and Hägglund, 2006; Luyben, 1996). Estudios previos han señalado que dichas condiciones no siempre se cumplen en procesos de separación altamente integrados, con dinámicas lentas y sobrealmortiguadas, como las columnas de DE y, en particular, las E-DWC (Skogestad, 2003; Seborg et al., 2016).

Como consecuencia, la aplicación directa de métodos basados en ganancia última puede resultar inviable o conducir a parámetros de sintonización excesivamente agresivos, con pobre robustez y desempeño dinámico insatisfactorio (Aström and Hägglund, 2006; Skogestad and Postlethwaite, 2005).

En este sentido, las técnicas modernas de sintonización basadas en el dominio de la frecuencia, como las implementadas en el comando *pidtune* de MATLAB, ofrecen una alternativa robusta al no depender de la inducción de oscilaciones sostenidas y al incorporar criterios explícitos de estabilidad y robustez (MathWorks, 2023). No obstante, su aplicación sistemática a modelos dinámicos aún no ha sido ampliamente documentada.

El objetivo de este trabajo es analizar de manera sistemática la sintonización de controladores PID para procesos de deshidratación de etanol mediante DE, considerando tanto configuraciones de destilación extractiva convencional (CEC) como configuraciones de destilación extractiva con pared divisoria o intensificadas (CEI), a partir de FTs reportadas en la literatura. El estudio se centra en evaluar la aplicabilidad real de los métodos clásicos de sintonización de Z-N y T-L a dichos modelos, así como en establecer un marco metodológico consistente para el diseño de controladores cuando estos métodos resultan inaplicables debido a las características dinámicas del proceso.

La contribución principal de este trabajo radica en mostrar que una fracción significativa de las FTs reportadas para la deshidratación de etanol no cumple las condiciones teóricas requeridas por los métodos clásicos de sintonización, lo cual limita su aplicabilidad práctica en configuraciones intensificadas como las E-DWC.

2. METODOLOGÍA

2.1 Descripción de los sistemas convencionales y E-TDWC

La DE presenta un bajo consumo de energía y menores costos de inversión y operación cuando se aplican configuraciones con estrategias de intensificación de procesos (Kiss and Ignat, 2012; Tututi et al., 2014). Una estrategia ampliamente reconocida de intensificación de procesos en la destilación extractiva es la implementación de columnas de pared divisoria (EDWC, por sus siglas en inglés)

La Figura 1 muestra una columna de pared divisoria superior (E-TDWC), donde se obtienen etanol y agua como productos de destilado, mientras que el agente de separación se obtiene en fondos y es recirculado a la columna.

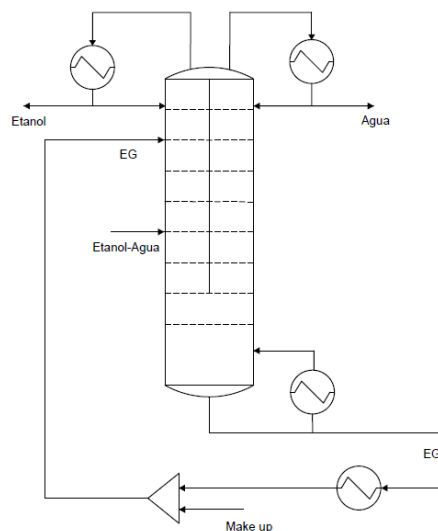


Figura 1. E-TDWC: Columna de DE de pared divisoria

2.2 Esquemas de control

En las columnas con pared divisoria, el área de la sección transversal de cada lado queda determinada por la ubicación física de la pared divisoria, la cual se define durante la etapa de diseño y no puede modificarse durante la operación. En consecuencia, la división del vapor suele permanecer fija durante el control y la operación del sistema, dando lugar a lo que se conoce como una configuración de control con división de vapor fija (Zhu et al., 2021). No obstante, en configuraciones E-TDWC es posible ajustar el flujo de vapor en la etapa de interconexión mediante la modificación de la presión en el domo del lado de rectificación, lo que permite implementar una configuración con división de vapor ajustable (Luyben, 2018). En este contexto, Hernández-Escoto et al. (2022) adoptan una estrategia de división de vapor ajustable, mientras que Tututi et al. (2014) analizan ambas configuraciones, considerando tanto esquemas con vapor fijo como con vapor ajustable.

En este trabajo se analizaron y compararon dos estrategias de control de pureza de productos:

i) Tututi et al. (2014) utilizan control basado en temperatura (control inferencial): En esta estrategia, las purezas de los productos no se controlan de manera directa, sino a través del control de temperaturas en etapas seleccionadas de la columna. La selección de dichas etapas se basa en su alta sensibilidad a cambios en la composición. Esta estrategia es común en aplicaciones industriales debido a la rapidez y confiabilidad de las mediciones de temperatura.

La composición de los productos en una columna de destilación depende de la temperatura porque la separación se basa en el equilibrio líquido-vapor (ELV), donde las fracciones molares de los componentes en la fase líquida y en la fase de vapor están determinadas por la temperatura y la presión del sistema (Luyben, 2006b; Rängner et al., 2024).

ii) Hernández-Escoto et al. (2022) y Román et al. (2024) utilizan control directo de composición: En esta estrategia, las composiciones de los productos se miden directamente



y se utilizan como variables controladas. Aunque este enfoque permite una regulación más precisa de las purezas, suele presentar dinámicas más lentas debido a los retardos asociados a los analizadores de composición.

Ambas estrategias se implementan mediante lazos de control independientes, utilizando controladores PI y PID, con el fin de realizar una comparación justa entre los enfoques de control y aislar el efecto de la variable medida sobre el desempeño dinámico.

2.3 Modelos dinámicos

En este trabajo se tomaron modelos dinámicos presentados en la literatura para este tipo de configuraciones. Con fines comparativos de control, se seleccionaron las FTs asociadas al control de las purezas de etanol y agua, ya que la pureza de los productos representa la variable de mayor relevancia operativa y económica en los procesos de DE.

Sistemas convencionales Para la selección de las FTs asociadas a la pureza de etanol o a la pureza de agua en cada caso, se analizaron previamente las estructuras de control reportadas en la literatura.

En la configuración de control de Tututi et al. (2014), la pureza de etanol se controla con la relación de reflujo de la columna extractiva (R_1), sin embargo, únicamente se reportan FTs correspondientes a lazos de control de temperatura, por lo que, no se dispone de un modelo dinámico explícito que relacione el reflujo (R_1) con la pureza de etanol. En contraste, la pureza de agua al ser controlada con la temperatura de una etapa intermedia de la columna recuperadora, sí se puede comparar con FTs de otros autores.

Román et al. (2024) reportan FTs exclusivamente para la columna extractiva de una CEC, lo que restringe el análisis comparativo al control de la pureza de etanol. Por lo tanto, los dos autores anteriormente mencionados no proporcionan información que permita comparar el control de las purezas de etanol y agua en CECs.

Bajo estas consideraciones, la comparación se llevó a cabo utilizando las FTs que relacionan la pureza de fondos de la columna extractiva (x_{B1}). En este contexto, Tututi et al. (2014) regulan esta variable con control inferencial de temperatura en la etapa 14 (TC_{14}), manipulando la carga térmica en el rehervidor (Q_{R1}) mientras que Román et al. (2024) utilizan control de composición (x_{B1}), manipulando también la carga térmica en el rehervidor (Q_{R1}), como se observa en la Tabla 1.

Tabla 1. FTs seleccionadas para las CECs: variables de entrada y de salida

Tututi et al. (2014)	$G_{x_{B1}} = \frac{x_B}{Q_{R1}}$
Román et al. (2024)	$G_{x_{B1}} = \frac{x_{B1}}{Q_{R1}}$

Sistemas intensificados Asimismo, para las CEI se evaluó la compatibilidad entre las distintas estructuras de control con el fin de establecer criterios homogéneos para

la comparación de las FTs asociadas al control de la pureza de etanol y de agua. En el caso de los sistemas E-TDWC analizados por Tututi et al. (2014), en la configuración con división de vapor fija, para mantener las purezas de etanol y agua se utiliza control inferencial de temperatura; TC_{14} manipulando R_1 , y TC_4 manipulando la relación de reflujo de la columna recuperadora (R_2), respectivamente.

En la configuración con división de vapor ajustable, se empleó el conjunto ($F_V/Q_R, R_2$) como variables de entrada para controlar el conjunto (TC_{14}, TC_4) que corresponden a las purezas de etanol y agua, respectivamente, siendo F_V el flujo de vapor en la interconexión y Q_R la carga térmica en el rehervidor. En esta configuración, el controlador TC_{14} regula el flujo de vapor en la interconexión, mientras que la pureza de etanol se controla mediante la relación de reflujo R_1 , para la cual no se reporta una función de transferencia explícita. No obstante, dado que el flujo de vapor en la interconexión ejerce una influencia significativa sobre la pureza de etanol y de agua, esta variable se adoptó igualmente como base para la comparación del control de la pureza de etanol.

Por su parte, Hernández-Escoto et al. (2022) presentan FTs únicamente para la configuración con división de vapor ajustable, en la cual proponen regular la pureza de etanol (x_1) mediante el flujo de vapor en la interconexión (F_V) y la pureza de agua (x_2) a través de la relación de reflujo (R_1).

En la Tabla 2 se muestran las variables de entrada y salida de las FTs seleccionadas para las CEIs. Se comparará el desempeño de control de las FTs de Hernández-Escoto et al. (2022) y las de Tututi et al. (2014) con F_V ajustable.

Tabla 2. FTs seleccionadas para las CEIs: variables de entrada y de salida

Tututi et al. (2014)	F_V fijo	Etanol	Agua
		$G_{x_1} = \frac{TC_{14}}{R_1}$	$G_{x_2} = \frac{TC_4}{R_2}$
	F_V ajustable	$G_{x_1} = \frac{TC_{14}}{V_R/Q_R}$	$G_{x_2} = \frac{TC_4}{R_2}$
Hernández-Escoto et al. (2022)	F_V ajustable	$G_{x_1} = \frac{x_1}{F_V}$	$G_{x_2} = \frac{x_2}{R_1}$

Las FTs empleadas en este trabajo corresponden a modelos dinámicos lineales obtenidos mediante la linealización de modelos rigurosos de columnas de DE alrededor de puntos de operación estacionarios. Estos modelos describen la dinámica de pequeñas perturbaciones en torno al régimen estacionario y constituyen una aproximación comúnmente utilizada para el análisis y diseño de sistemas de control PID.

Las condiciones de operación, configuraciones de control y las variables de diseño asociadas a los sistemas de destilación extractiva CEC y CEI se encuentran detalladas en los trabajos reportados en la literatura correspondientes.

2.4 Análisis de aplicabilidad de métodos clásicos de sintonización

Previo al diseño de los controladores, se evalúa la aplicabilidad de los métodos de Z-N y T-L a cada función de trans-

ferencia. Los métodos de sintonización de Z-N y T-L son técnicas empíricas clásicas, cuyo uso es válido únicamente cuando el proceso bajo estudio cumple ciertas condiciones dinámicas bien definidas. En particular, el método de Z-N en lazo abierto asume que el proceso puede aproximarse mediante un modelo de primer orden con tiempo muerto, lo cual requiere una respuesta a escalón de tipo sigmoidal y la existencia de un punto de inflexión claramente identificable (Seborg et al., 2016). Cuando estas características no están presentes, la estimación de los parámetros dinámicos del modelo resulta poco representativa y el método se vuelve impráctico.

Por otro lado, los métodos de Z-N en lazo cerrado y de T-L requieren que el sistema, bajo control proporcional puro, exhiba oscilaciones sostenidas para un valor finito de la ganancia proporcional, denominada ganancia última (K_u). Si el proceso no presenta oscilaciones sostenidas, aun para valores elevados de la ganancia proporcional, no es posible determinar de manera consistente ni K_u , ni el período último P_u , lo que imposibilita la aplicación de estos métodos (Luyben, 1990).

Dado que las FTs empleadas en este trabajo, reportadas en la literatura para procesos de deshidratación de etanol, no satisfacen de manera general los supuestos dinámicos requeridos por los métodos Z-N y T-L, se adopta el comando *pdtune* de MATLAB como estrategia de sintonización. *pdtune* permite obtener parámetros de control coherentes y reproducibles para una amplia gama de dinámicas, lo que lo convierte en una alternativa robusta y adecuada para el análisis comparativo de FTs asociadas a columnas de DE, tanto convencionales como con pared divisoria.

En el caso de las FTs de una CEC y una CEI de Tututi et al. (2014), aunque los procesos presentan tiempos muertos distintos de cero, la respuesta al escalón no exhibe una forma sigmoidal debido a que los sistemas son de primer orden y el retardo es pequeño en comparación con la constante de tiempo ($\frac{\theta}{\tau} \ll 1$). En consecuencia, la segunda derivada temporal de la salida es negativa para todo $t > 0$, lo que implica la ausencia de una fase de aceleración y, por tanto, de un punto de inflexión, lo que invalida la aplicación del método de Z-N a lazo abierto para este caso. Lo mismo ocurre con las FTs reportadas por Hernández-Escoto et al. (2022), las cuales no presentan tiempo muerto.

Adicionalmente, el reducido valor del cociente ($\frac{\theta}{\tau}$) implica un retardo de fase insuficiente para inducir oscilaciones sostenidas bajo control proporcional puro, por lo que no existe una K_u finita. En consecuencia, los métodos de T-L y Z-N en lazo cerrado no son aplicables a estas FTs, ya que no se satisfacen los supuestos dinámicos en los que se fundamentan (Aström and Hägglund, 2006; Seborg et al., 2016).

En contraste con los casos anteriores, las FTs reportadas por Román et al. (2024) sí cumplen los supuestos dinámicos requeridos para la aplicación de los métodos de sintonización de Z-N en lazo abierto y lazo cerrado, así como del método de T-L, al presentar dinámicas que permiten identificar tanto un punto de inflexión en la respuesta al escalón como una ganancia última finita bajo control proporcional puro. No obstante, con el fin de mostrar resultados consistentes y una comparabilidad

objetiva, en este trabajo se adopta el comando *pdtune* de MATLAB como método único de sintonización para todas las FTs analizadas. Esta decisión evita la mezcla de técnicas basadas en supuestos dinámicos distintos y permite evaluar el desempeño de los controladores bajo un marco de sintonización homogéneo y reproducible, independientemente de la estructura particular de cada modelo dinámico.

2.5 Comando *pdtune* en MATLAB

En esta sección se describen los criterios y opciones de diseño empleados para la sintonización automática de controladores PID mediante el comando *pdtune* de MATLAB.

- Margen de fase = 60 (predeterminado). Margen de fase objetivo en grados, especificado como escalar positivo. *pdtune* trata de diseñar un controlador de modo que el margen de fase sea al menos el valor especificado para *PhaseMargin*. La frecuencia de cruce seleccionada podría restringir el margen de fase alcanzable. Normalmente, un margen de fase mayor mejora la estabilidad y el sobreimpulso, pero limita el ancho de banda y la velocidad de respuesta.
- Objetivo de rendimiento de lazo cerrado que se desea favorecer en el diseño, especificado como una de las siguientes opciones:
 - 'balanced' (valor predeterminado): ajusta el controlador a fin de equilibrar el seguimiento de referencia y la anulación de perturbaciones.
 - 'reference-tracking': ajusta el controlador para favorecer el seguimiento de referencia, si es posible.
 - 'disturbance-rejection': ajusta el controlador para favorecer la anulación de perturbaciones.

Para un margen de fase objetivo dado, *pdtune* elige un diseño de controlador que equilibre las dos mediciones de rendimiento, seguimiento de referencia y anulación de perturbaciones.

2.6 Evaluación del desempeño dinámico

El desempeño de los controladores sintonizados se evaluó mediante simulaciones dinámicas ante cambios de consigna. Se emplearon índices de desempeño basados en el error, tales como IAE, ISE, ITAE e ITSE, así como métricas temporales, incluyendo sobreimpulso y tiempo de establecimiento. Estas métricas resultan particularmente relevantes, ya que un sobreimpulso de gran magnitud puede provocar desviaciones transitorias significativas en la pureza del producto, mientras que el signo de la desviación asociado a la naturaleza directa o inversa de la respuesta del proceso, determina si dichas desviaciones resultan penalizables desde el punto de vista operativo. Asimismo, un mayor tiempo de establecimiento implica que estas desviaciones de pureza persistan durante intervalos más prolongados, lo que se traduce en pérdidas económicas asociadas a la operación fuera de especificación.

3. RESULTADOS

Para cada caso, se graficaron las respuestas dinámicas ante cambios tipo escalón y se calcularon los índices de

desempeño basados en el error, con el fin de evaluar cuantitativamente el desempeño de los controladores.

Los controladores sintonizados mediante *pidtune* fueron estables y no exhibieron oscilaciones indeseables. En todos los casos, los controladores PI y PID presentaron índices de desempeño basados en el error dentro de rangos aceptables, con valores consistentemente menores de IAE, ISE, ITAE e ITSE para los controladores PID, debido a la contribución efectiva de la acción derivativa en simulaciones sin ruido de medición. Asimismo, el control de temperatura mostró dinámicas más rápidas y menores tiempos de establecimiento que el control de composición, coherente con el menor retardo de las mediciones de temperatura, mientras que este último presentó respuestas más lentas y mayor sensibilidad a perturbaciones, aunque con desempeño estable.

3.1 Sistemas convencionales

En la Figura 2 se analiza la variable x_B . El control de composición Román et al. (2024) presentó tiempos de establecimiento de 2.89 h (PI) y 2.16 h (PID), mientras que el control de temperatura (Tututi et al., 2014) mostró valores significativamente menores: 0.56 h (PI) y 0.58 h (PID), reduciendo el tiempo en que las purezas permanecen fuera de especificación. El sobreimpulso con PID fue de 1.10 para control de composición y de 1.05 para control de temperatura. El mayor sobreimpulso y tiempo de establecimiento del control de composición implican una mayor cantidad acumulada de producto fuera de especificación en fondos, lo cual es operativamente indeseable. El control de la pureza en fondos en una columna de DE convencional es crítico, ya que afecta directamente las purezas de etanol y agua; variaciones en esta variable alteran los perfiles de concentración y equilibrio, impactando la distribución de componentes y la eficiencia global de la separación.

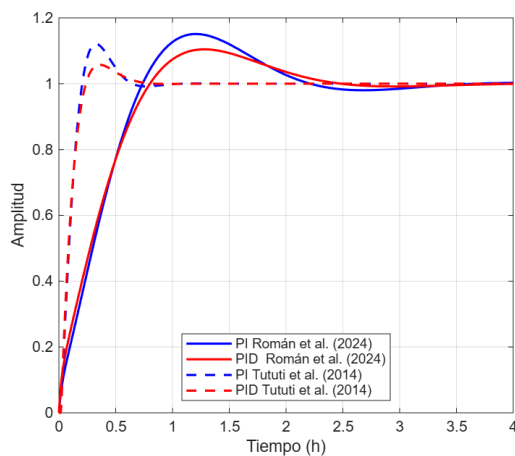


Figura 2. Respuesta con control PI y PID para la columna extractiva convencional de Tututi et al. (2014) y Román et al. (2024)

El índice ITAE penaliza con mayor severidad los errores persistentes en el tiempo, por lo que valores bajos se asocian con menor sobreimpulso y tiempo de asentamiento. Como se muestra en la Tabla 3, para las FTs de Román et al. (2024), el ISE es 54.78 % (PI) y 55.13 % (PID) menor

que el IAE. De forma similar, para Tututi et al. (2014), el ISE es 45.65 % menor que el IAE con PI y 42.52 % con PID, evidenciando una proporcionalidad consistente entre ambos índices; un comportamiento análogo se observa para el ITSE. Sin embargo, se identifica una diferencia relevante en el ITAE: para Román et al. (2024) el ITAE supera al ISE, mientras que en Tututi et al. (2014) es inferior. Este comportamiento concuerda con lo mostrado en la Figura 2, dado que el ITAE penaliza de manera más severa los errores de larga duración, lo que se traduce en valores más elevados de sobreimpulso y tiempo de asentamiento cuando se emplea control de composición.

Tabla 3. Índices relacionados a la pureza de fondos (x_{B1}) de la columna extractiva convencional

	Román et al. (2024)	Tututi et al. (2014)
IAE PI	0.4591	0.1255
ISE PI	0.2079	0.0682
ITAE PI	0.2920	0.0182
ITSE PI	0.0511	0.0038
IAE PID	0.4092	0.1110
ISE PID	0.1836	0.0638
ITAE PID	0.2358	0.0131
ITSE PID	0.0395	0.0030

3.2 Sistemas intensificados

Al analizar las respuestas transitorias del control de la pureza de etanol en los sistemas E-TDWC, se evidencia que el control de temperatura propuesto por Tututi et al. (2014) presenta un desempeño dinámico superior en comparación con el control de composición reportado por Hernández-Escoto et al. (2022) (Figura 3). En el esquema de control de temperatura, los tiempos de establecimiento obtenidos con controladores PI y PID son de 0.48 y 0.49 h, respectivamente, mientras que al emplear control de composición estos aumentan a 0.59 h con PI y 0.63 h con PID.

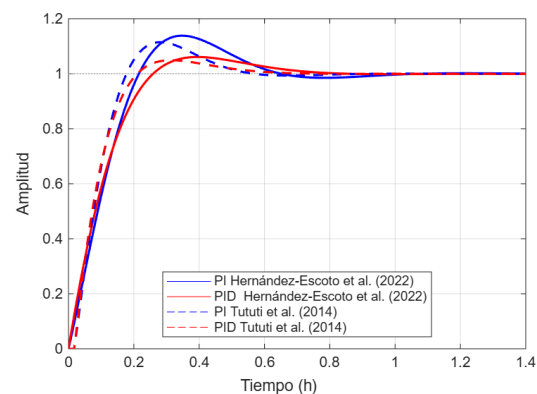


Figura 3. Respuesta con control PI y PID para la E-TDWC

En cuanto al sobreimpulso, el control PID presenta valores de 1.04 para el control de temperatura y de 1.06 para el control de composición. Aunque la diferencia en sobreimpulso es relativamente pequeña, el mayor tiempo de establecimiento asociado al control de composición implica que la pureza del producto permanece fuera de especificación

durante un intervalo más prolongado, lo que refuerza la conveniencia del control de temperatura desde el punto de vista operativo. Esta tendencia se ve reflejada de manera consistente en los índices de desempeño reportados en la Tabla 4, los cuales evidencian una menor penalización de los errores cuando se emplea el control de temperatura, confirmando su mayor robustez dinámica.

Tabla 4. Índices relacionados a la pureza de etanol de la E-TDWC

	Hernández-Escoto et al. (2022)	Tututi et al. (2014)
IAE PI	0.1329	0.1078
ISE PI	0.0646	0.0600
ITAE PI	0.0231	0.0131
ITSE PI	0.0042	0.0028
IAE PID	0.1124	0.0950
ISE PID	0.0570	0.0563
ITAE PID	0.0154	0.0022

La respuesta transitoria para la pureza de agua, utilizando los modelos de control de composición de Hernández-Escoto et al. (2022) fue significativamente más lenta comparada con las estructuras de control anteriores (Figura 4); en este caso, los tiempos de establecimiento fueron de 10.65 y 11.43 h con control PI y PID, respectivamente. En contraste, el control de temperatura de Tututi et al. (2014), presentó tiempos de establecimiento similares a los casos anteriores, siendo 0.56 h con PI y 0.57 h con PID, debido a que la temperatura exhibe una dinámica de respuesta más rápida, mientras que en el control de composición, las variaciones en la variable manipulada (R_1 , en la configuración de control de Hernández-Escoto et al. (2022)) se manifiestan con un desfase temporal en la variable controlada. El sobreimpulso presentado con control PID fue de 1.06 h con control de composición y 1.05 h con control de temperatura, indicando que la principal diferencia entre las estrategias radica en el tiempo de establecimiento y no en el sobreimpulso.

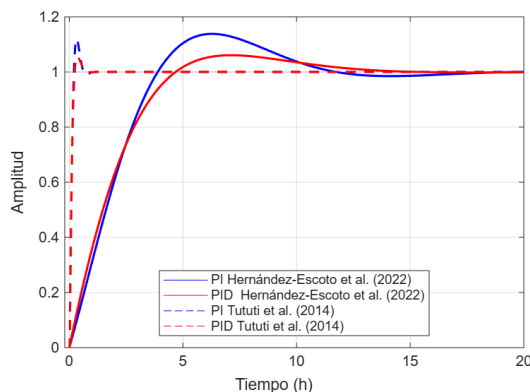


Figura 4. Respuesta con control PI y PID para la E-TDWC

En la Tabla 5 se observa que los índices basados en el error de los modelos dinámicos de Tututi et al. (2014) son menores a los de Hernández-Escoto et al. (2022) para el control de la pureza de agua, lo cual confirma una respuesta dinámica más rápida y un mayor grado de estabilidad cuando se utiliza una configuración de control de temperatura.

Tabla 5. Índices relacionados a la pureza de agua de la E-TDWC

	Hernández-Escoto et al. (2022)	Tututi et al. (2014)
IAE PI	2.3963	0.1243
ISE PI	1.1663	0.0676
ITAE PI	7.4663	0.0179
ITSE PI	1.3745	0.0037
IAE PID	2.0291	0.1099
ISE PID	1.0284	0.0633
ITAE PID	5.0093	0.0128
ITSE PID	0.9779	0.0029

3.3 Modelos dinámicos de Tututi et al. (2014).

Se comparó también la sintonización de control con las FTs de Tututi et al. (2014) para una columna extractiva convencional y dos configuraciones E-TDWC; una con división de vapor fija y otra con división de vapor ajustable.

Como se señaló previamente, no se dispone de una función de transferencia que relacione explícitamente la pureza de etanol para la configuración extractiva convencional, y los parámetros de las FTs asociadas a la pureza de agua son idénticos para las configuraciones con división de vapor fija y ajustable. En consecuencia, el análisis comparativo del control de la pureza de etanol se realizó entre las configuraciones E-TDWC con división de vapor fija y E-TDWC con división de vapor ajustable, mientras que para la pureza de agua se compararon la configuración extractiva convencional y la E-TDWC.

Para el control de la pureza de etanol mostrado en la Figura 5, los tiempos de establecimiento para la configuración con división de vapor fija son de 1.15 h (PI) y 1.19 h (PID), mientras que estos tiempos son de 0.48 h (PI) y 0.49 h (PID) utilizando la configuración con división de vapor ajustable, por lo que este grado de libertad adicional en el control del flujo de vapor en la interconexión permite una acción de control más efectiva, reduciendo significativamente el tiempo de establecimiento, en concordancia con lo reportado en la literatura (Wolff and Skogestad, 1995).

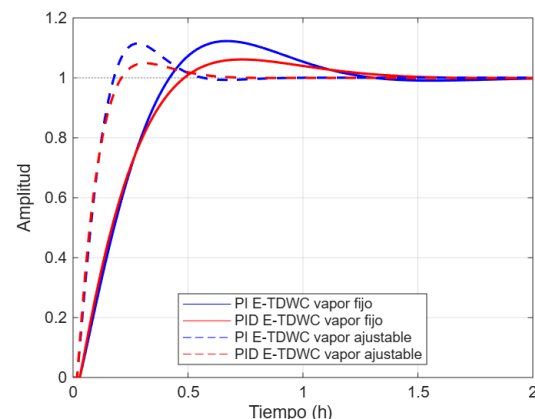


Figura 5. Respuesta con control de la pureza de etanol de Tututi et al. (2014)

En la Tabla 6 se comparan los índices basados en el error para las configuraciones con vapor fijo y vapor ajustable,



ambas basadas en control de temperatura. La configuración con vapor ajustable presenta índices menores y por lo tanto una dinámica del proceso más rápida.

Tabla 6. Índices relacionados a la pureza de etanol de Tututi et al. (2014)

	E-TDWC vapor fijo	E-TDWC vapor ajustable
IAE PI	0.2559	0.1078
ISE PI	0.1365	0.0600
ITAE PI	0.0776	0.0131
ITSE PI	0.0156	0.0028
IAE PID	0.2263	0.0950
ISE PID	0.1270	0.0563
ITAE PID	0.0561	0.0092
ITSE PID	0.0123	0.0022

Al evaluar el control de la pureza de agua (Figura 6), se toma en cuenta que los parámetros de las FTs de los sistemas E-TDWC son los mismos para las configuraciones con división de vapor fija y ajustable, por lo que en este caso se compararon los tiempos de establecimiento en el sistema convencional, los cuales fueron de 0.44 h con PI y PID, en contraste a los tiempos de 0.56 y 0.57 h para PI y PID en el sistema E-TDWC. Asimismo, el sobreimpulso presentado con control PID fue de 1.04 y 1.05 h para el sistema convencional y E-TDWC, respectivamente. Estos valores más grandes de tiempo de establecimiento y sobreimpulso en el sistema E-TDWC pueden atribuirse a la mayor complejidad dinámica y al incremento de grados de libertad característicos de estos sistemas, los cuales incorporan una mayor interacción térmica y acoplamiento entre variables manipuladas y controladas (Zhu et al. (2021)).

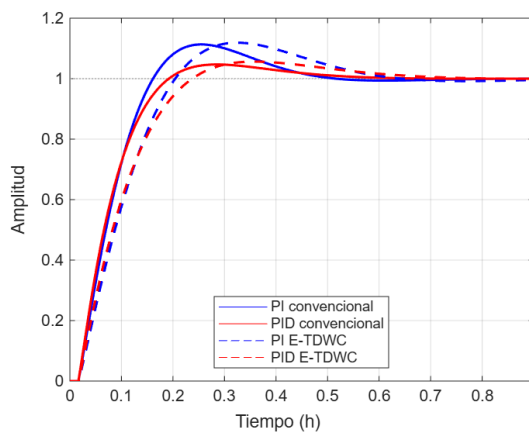


Figura 6. Respuesta con control de la pureza de agua de Tututi et al. (2014)

Al comparar los índices del error asociados al control de la pureza de agua presentados en la Tabla 7, se observa que para el sistema convencional presenta una dinámica más rápida al tener índices más pequeños.

4. CONCLUSIONES

En este trabajo se evaluó el desempeño dinámico de estrategias de control PI y PID aplicadas a FTs reportadas en la literatura para sistemas de DE convencional y con

Tabla 7. Índices basados en el error para los modelos dinámicos relacionados a la pureza de agua de Tututi et al. (2014)

	Convencional	E-TDWC
IAE PI	0.0977	0.1243
ISE PI	0.0548	0.0676
ITAE PI	0.0107	0.0179
ITSE PI	0.0023	0.0037
IAE PID	0.0862	0.1099
ISE PID	0.0515	0.0633
ITAE PID	0.0075	0.0128
ITSE PID	0.0018	0.0029

pared divisoria (E-TDWC), empleando un enfoque de sintonización consistente basado en el comando pidtune. Los resultados muestran que, en todos los casos analizados, el control PID presenta un desempeño superior al control PI, evidenciado por índices basados en el error más pequeños y sobreimpulsos más cortos.

Asimismo, se observó que el control basado en temperatura exhibe una dinámica más favorable que el control de composición, caracterizada por menores errores integrales, tiempos de estabilización reducidos y sobreimpulsos más pequeños, independientemente de la configuración de la columna. Estos resultados refuerzan la conveniencia práctica del control por temperatura en sistemas de DE, particularmente cuando se requiere una respuesta rápida y estable.

Para las configuraciones E-TDWC, el análisis comparativo indica que la estructura de control con división de vapor ajustable conduce a un mejor desempeño dinámico en comparación con la configuración de vapor fijo, reflejado en respuestas más rápidas y con menor desviación transitoria.

Finalmente, el estudio destaca la importancia de considerar la dinámica inherente del proceso al seleccionar un método de sintonización, dado que las hipótesis fundamentales de métodos clásicos como Z-N y T-L no se satisfacen para varias de las FTs analizadas.

REFERENCIAS

- Aquilon, A.F., Cargullo, D.M., Onayan, J., Sarno, J., Molino, V.M., and Lopez, E.C.R. (2023). Recent advances in extractive distillation. *Engineering Proceedings*, 56(1). doi:https://doi.org/10.3390/ASEC2023-15261.
- Aström, K.J. and Hägglund, T. (2006). *Advanced PID Control*. ISA—The Instrumentation, Systems, and Automation Society, Research Triangle Park, NC, USA.
- Czarnecki, N.J., Owens, S.A., and Eldridge, R.B. (2023). Extractive dividing wall column for separating azeotropic systems: A review. *Industrial & Engineering Chemistry Research*, 62(14), 5750–5770. doi: https://doi.org/10.1021/acs.iecr.3c00302.
- Gerbaud, V., Rodriguez-Donis, I., Hegely, L., Lang, P., Denes, F., and You, X. (2019). Review of extractive distillation. process design, operation, optimization and control. *Chemical Engineering Research and Design*, 141, 229–271. doi: https://doi.org/10.1016/j.cherd.2018.09.020.
- Hernández-Escoto, H., Zavala-Guzmán, A.M., Maya-Yescas, R., Zamudio-Lara, J.M., and Hernández, S. (2022). Operability assessment and systematic pi

- control of a class of extractive dividing wall distillation columns: Case of ethanol dehydration. *Chemical Engineering Research and Design*, 187, 84–92. doi: <https://doi.org/10.1016/j.cherd.2022.08.030>.
- Huang, X., Jin, Y., Fang, C., Yang, W., Li, Z., and Tian, Y. (2025). Process simulation and optimization of extractive distillation for separating isopropanol-ethanol from ibe fermentation broth as a gasoline additive. *Industrial Crops and Products*, 236, 121853. doi: <https://doi.org/10.1016/j.indcrop.2025.121853>.
- Kiss, A. (2013). *Advanced Distillation Technologies: Design, Control and Applications*. doi: <https://doi.org/10.1002/9781118543702>.
- Kiss, A.A. and Ignat, R.M. (2012). Innovative single step bioethanol dehydration in an extractive dividing-wall column. *Separation and Purification Technology*, 98, 290–297. doi: <https://doi.org/10.1016/j.seppur.2012.06.029>.
- Li, M., Peng, J., Cheng, Y., Zhu, X., Ma, Y., Zhang, Z., and Gao, J. (2023). Dynamic control of an energy-saving process with two extractive dividing-wall columns for separation of acetone/methanol/butanone/tert-butyl alcohol mixtures. *Chemical Engineering Research and Design*, 200, 281–291. doi: <https://doi.org/10.1016/j.cherd.2023.10.045>.
- Lorenz, H.M., Staak, D., Grutzner, T., and Repke, J.U. (2018). Divided wall columns: Usefulness and challenges. *Chemical Engineering Transactions*, 69, 229–234. doi: <https://doi.org/10.3303/CET1869039>.
- Luyben, W.L. (1990). *Process Modeling, Simulation, and Control for Chemical Engineers*. McGraw-Hill, Singapore, 2 edition.
- Luyben, W.L. (1996). *Process Modeling, Simulation, and Control for Chemical Engineers*. McGraw-Hill, New York, NY, USA.
- Luyben, W.L. (2006a). *Control of More Complex Columns*, chapter 8, 188–231. John Wiley & Sons, Ltd. doi: <https://doi.org/10.1002/0471785253.ch8>.
- Luyben, W.L. (2006b). *Fundamentals of Vapor-Liquid Phase Equilibrium (VLE)*, chapter 1, 188–231. John Wiley & Sons, Ltd. doi: <https://doi.org/10.1002/0471785253.ch1>.
- Luyben, W.L. (2018). Vapor split manipulation in extractive divided-wall distillation columns. *Chemical Engineering and Processing - Process Intensification*, 126, 132–140. doi: <https://doi.org/10.1016/j.cep.2018.02.022>.
- MathWorks (2023). Pid tuning algorithm. Accessed: 2026-04-12.
- Neagu, M. and Pricop-Nicolae, M. (2025). Azeotropic and extractive distillation for bio-ethanol dehydration: Process design, simulation, and cost analysis. *Processes*, 13(11). doi: <https://doi.org/10.3390/pr13113634>.
- Román, R.A., García, A.L., Vite, I.G., López, M.D., and Hernández, H. (2024). Desarrollo de modelos dinámicos tipo entradas-salidas de columnas de destilación. *JÓVENES EN LA CIENCIA*, 28, 1–15. doi: <https://doi.org/10.15174/jc.2024.4355>.
- Rängner, L.M., Halvorsen, I.J., Grützner, T., and Skogestad, S. (2024). Understanding temperature profiles of distillation columns. *Industrial & Engineering Chemistry Research*, 63(10), 4533–4546. doi: <https://doi.org/10.1021/acs.iecr.3c04102>.
- Seborg, D.E., Edgar, T.F., Mellichamp, D.A., and Doyle, F.J.I. (2016). *Process Dynamics and Control (4th ed.)*. John Wiley & Sons, Ltd.
- Skogestad, S. (2003). Simple analytic rules for model reduction and pid controller tuning. *Journal of Process Control*, 13(4), 291–309. doi: [https://doi.org/10.1016/S0959-1524\(02\)00062-8](https://doi.org/10.1016/S0959-1524(02)00062-8).
- Skogestad, S. and Postlethwaite, I. (2005). *Multivariable Feedback Control: Analysis and Design*, volume 2.
- Torres, K., Pérez-Zúñiga, R., Martínez, M., Ramos-Cabral, S., Calixto-Rodriguez, M., Valdez Martínez, J., Enriquez, M., Pérez-Estrada, A., Ortiz-Torres, G., Sorcia-Vázquez, F., Rebollo, A., and Rumbo Morales, J. (2022). Design and control applied to an extractive distillation column with salt for the production of bioethanol. *Processes*, 10. doi: <https://doi.org/10.3390/pr10091792>.
- Tututi, S., Jiménez, A., and Hahn, J. (2014). Control analysis of an extractive dividing-wall column used for ethanol dehydration. *Chemical Engineering and Processing: Process Intensification*, 82, 88–100. doi: <https://doi.org/10.1016/j.cep.2014.05.005>.
- Wolff, E.A. and Skogestad, S. (1995). Operation of integrated three-product (petlyuk) distillation columns. *Industrial & Engineering Chemistry Research*, 34(6), 2094–2103. doi: <https://doi.org/10.1021/ie00045a018>.
- Xu, D., Li, W., Cao, L., Gao, J., Zhang, Z., and Wang, Y. (2026). Sustainable separation of azeotrope n-hexane/isopropanol via extractive distillation: Molecular insights, process optimization, and techno-economic-environmental analysis for cleaner production. *Industrial & Engineering Chemistry Research*, 65(9), 5125–5137. doi: <https://doi.org/10.1021/acs.iecr.5c04681>.
- Zhang, H., Wu, Z., Yuan, Q., Guo, L., Li, X., Hua, C., and Lu, P. (2025). Control of extractive dividing wall column using model predictive control based on long short-term memory networks. *Separation and Purification Technology*, 361, 131351. doi: <https://doi.org/10.1016/j.seppur.2024.131351>.
- Zhu, J., Jing, C., Hao, L., and Wei, H. (2021). Insight into controllability and operation of extractive dividing-wall column. *Separation and Purification Technology*, 263, 118362. doi: <https://doi.org/10.1016/j.seppur.2021.118362>.

Design of a Production and Marketing Model for Tamarind Derivates with a Focus on Education for Sustainability

Bárcenas Nava Ma. de los Angeles *, Trujillo Beltrán Elisa*, Flores Encarnación Ivette Karina**, GodínezRueda Uriel**

**TecNM/ Instituto Tecnológico de Iguala*

Carretera Nacional Iguala- Taxco Esquina Periférico Norte, Iguala, Gro. C.P. 40030 México (e-mail: angeles.barcenas@iguala.tecnm.mx, elisa.trujillo@iguala.tecnm.mx,

***TecNM/ Instituto Tecnológico de Iguala*

Carretera Nacional Iguala- Taxco Esquina Periférico Norte, Iguala, Gro. C.P. 40030 México (e-mail: angeles.barcenas@iguala.tecnm.mx, elisa.trujillo@iguala.tecnm.mx, TecNM/ Instituto Tecnológico de Iguala

Carretera Nacional Iguala- Taxco Esquina Periférico Norte, Iguala, Gro. C.P. 40030 México (e-mail: angeles.barcenas@iguala.tecnm.mx, elisa.trujillo@iguala.tecnm.mx, 23670139@iguala.tecnm.mx, 23670167@iguala.tecnm.mx)

Abstract:

The article introduces a production and marketing model for tamarind derivatives with a focus on education for sustainability. Tamarind (*Tamarindus indica* L.) is recognized as a valuable agro-industrial resource, yet in Mexico it remains underutilized, as most production is sold fresh with little added value. The authors propose a three-dimensional model integrating technical-productive, economic-commercial, and educational-sustainable dimensions. On the technical side, the model highlights innovative cultivation practices, efficient irrigation systems, and the use of by-products such as polysaccharides, proteins, and mucilage from seeds, with applications in food, pharmaceutical, and energy industries. Economically, it incorporates circular economy principles to transform residues into inputs for new products, reducing waste while increasing local producers' income. The educational dimension aims to strengthen community capacities, link technical innovation with the training of change agents, and ensure the adoption of sustainable practices over the long term. Iguala de la Independencia, Guerrero, is presented as a strategic case due to its cultural identity tied to tamarind and its national relevance in production. The model seeks to bridge the gap between scientific potential and industrial application, promoting agro-food systems that are more resilient, fair, and sustainable, capable of generating economic, social, and environmental benefits for the region while serving as a reference for other communities.

Keywords: Marketing model, tamarind, education, sustainability.

1. INTRODUCCIÓN

El tamarindo (*Tamarindus indica* L.) es una especie frutal tropical con un potencial productivo y comercial ampliamente reconocido en la literatura científica reciente, aunque históricamente subutilizado en su aprovechamiento integral. Distintas partes del árbol han sido exploradas para el desarrollo de productos ecológicos con criterios de sustentabilidad; la cáscara del fruto puede emplearse como adsorbente de sustancias tóxicas, la pulpa como conservador en la industria alimentaria, y el polisacárido de la semilla como aditivo funcional con aplicaciones en sistemas de liberación de fármacos (Mansingh et al., 2021). Este panorama sugiere que diseñar un modelo articulado de producción y comercialización de derivados del tamarindo representa no solo una oportunidad económica, sino también una vía concreta hacia la sustentabilidad de los sistemas agroalimentarios locales.

Desde la perspectiva del aprovechamiento de subproductos, la semilla de tamarindo constituye un residuo generado en grandes volúmenes por la industria de la pulpa que permanece

subutilizado a pesar de su enorme potencial; sus derivados industriales —polisacáridos, harina de núcleo, goma, almidón y aceite— han encontrado aplicaciones limitadas en la industria alimentaria, persistiendo una brecha significativa entre el potencial y la aplicación real, atribuible en parte a la ausencia de tecnologías factibles (Nagar et al., 2022). Cerrar esa brecha requiere modelos de producción que integren conocimiento técnico, cadenas de comercialización y formación de capacidades en las comunidades productoras.

En este contexto, la educación para la sustentabilidad emerge como un componente estructural indispensable. La educación en sistemas alimentarios sostenibles dentro de la educación superior está en plena expansión, y las universidades reconocen su papel significativo en la gestión de problemas complejos, con una motivación creciente para desarrollar programas educativos de alto impacto (Salminen et al., 2024). Incorporar este enfoque en el diseño de modelos productivos locales permite vincular la innovación técnica con la transformación de las prácticas comunitarias hacia horizontes de largo plazo.



Por su parte, los avances recientes en tecnología de producción del tamarindo abarcan prácticas de cultivo innovadoras, estrategias de mejoramiento genético, manejo de plagas, prácticas poscosecha y agregación de valor, así como la exploración de sistemas de intercultivo, agroforestería y riego eficiente para maximizar el uso del suelo y los recursos hídricos (Vasanthkumar et al., 2023). Estos avances ofrecen una base técnica sólida sobre la cual construir modelos de comercialización que sean a la vez viables y sustentables. En esa misma línea, el aprovechamiento de los polisacáridos, proteínas y mucílago de la semilla de tamarindo mediante procesos de extracción y modificación de propiedades abre oportunidades concretas para la generación de productos de valor agregado con aplicaciones alimentarias diversificadas (Geethalaxmi et al., 2024).

Finalmente, el contexto de mercado global refuerza la pertinencia de esta investigación: México se encuentra entre los principales países productores de tamarindo a nivel mundial, y la versatilidad de sus derivados abarca los sectores de alimentos y bebidas, cosméticos, productos farmacéuticos y del hogar, con proyecciones de crecimiento sostenido (Future Market Insights, 2024). Diseñar un modelo que articule producción, comercialización y educación para la sustentabilidad desde el contexto mexicano no solo responde a esta oportunidad de mercado, sino que contribuye a la construcción de sistemas agroalimentarios más justos, resilientes e integralmente sostenibles.

1.1 El tamarindo como recurso agroindustrial sustentable

El tamarindo (*Tamarindus indica* L.) constituye una especie frutal de alto valor agroindustrial cuyo aprovechamiento integral ha comenzado a estudiarse con rigor científico en años recientes. Distintas investigaciones han confirmado que sus derivados van más allá de la pulpa comestible: la cáscara puede actuar como adsorbente de sustancias tóxicas, la semilla aporta polisacáridos con aplicaciones en la industria alimentaria, farmacéutica y electroquímica, el aceite de la semilla es viable como biocombustible, y el extracto de las hojas presenta propiedades antibacterianas con potencial en la industria textil (Mansingh et al., 2021). Este conjunto de posibilidades productivas sustenta la pertinencia de diseñar un modelo que articule de manera sistémica la cadena de valor del tamarindo, desde su cultivo hasta la comercialización de sus derivados.

En el ámbito de la producción primaria, los avances recientes señalan que la tecnología de cultivo del tamarindo ha incorporado prácticas de agroforestería, sistemas de riego eficiente e intercultivo, estrategias que maximizan el uso del suelo y los recursos hídricos, al tiempo que contribuyen a la resiliencia de los ecosistemas agrícolas locales (Vasanthkumar et al., 2023). Por su parte, el aprovechamiento de los subproductos de la semilla —en particular sus polisacáridos, proteínas y mucílago— representa una oportunidad concreta para la generación de productos de valor agregado con aplicaciones diversificadas en la industria de alimentos (Geethalaxmi et al., 2024). Nagar et al. (2022) advierten, sin embargo, que persiste una brecha significativa entre el potencial reconocido y la aplicación industrial efectiva, atribuible principalmente a la ausencia de tecnologías factibles a escala local y a la escasa articulación entre los actores de la cadena productiva. Superar esa brecha constituye

precisamente el núcleo del problema que esta investigación busca resolver.

1.2 Economía circular como marco para la producción y comercialización de derivados

Un referente teórico central para este trabajo es la economía circular aplicada al sector agroalimentario. Este enfoque propone sustituir el modelo lineal de producción —extraer, fabricar, desechar— por sistemas que reducen, reutilizan, reciclan y recuperan materiales en los procesos de producción, distribución y consumo, con el propósito de generar simultáneamente calidad ambiental, prosperidad económica y equidad social (Kirchherr et al., 2017, citado en Castillo-Díaz et al., 2023). Aplicada al sector agroalimentario, la economía circular implica transformar los subproductos y residuos del procesamiento en insumos para nuevos productos, cerrando los ciclos de materiales y reduciendo la generación de desechos. Una revisión sistemática reciente que analizó cien estudios de alto impacto sobre el tema concluye que la economía circular tiene un potencial significativo para mejorar la sustentabilidad, reducir los residuos y aumentar la eficiencia en el uso de recursos en el sector agroalimentario, siendo la gestión de residuos y la eficiencia de recursos los ejes temáticos predominantes en la literatura (Banerjee et al., 2024).

En este contexto, el diseño de un modelo de producción y comercialización de derivados del tamarindo se inscribe naturalmente en la lógica circular: los residuos del procesamiento de la pulpa —semillas, cáscaras, fibras— se convierten en materias primas para nuevos productos con valor comercial, lo que reduce simultáneamente la carga ambiental y amplía los ingresos de los productores locales. Este planteamiento concuerda con lo señalado por Castillo-Díaz et al. (2023), quienes identificaron que la implementación de estrategias de economía circular en el sector agroalimentario europeo requiere de acción colaborativa entre actores públicos y privados, inversión en capital humano y avances tecnológicos que hagan viable la transición.

1.3 Educación para la sustentabilidad: fundamentos teóricos

La educación para la sustentabilidad (EpS) es el tercer eje conceptual de este trabajo. Scalabrino (2022) plantea que la transición hacia una economía circular y baja en carbono puede acelerarse si los profesionales de la educación adoptan las implicaciones prácticas de la educación para el desarrollo sustentable, identificando los elementos de la educación ambiental y de sustentabilidad que resultan más relevantes para promover patrones de producción y consumo responsables. Esta perspectiva es especialmente pertinente en el ámbito rural y comunitario, donde la formación de capacidades locales es condición necesaria para que los modelos productivos sean apropiados, reproducibles y de largo plazo.

Desde una perspectiva institucional, la educación superior desempeña un papel fundamental en la formación de agentes de cambio capaces de transformar los sistemas agroalimentarios. Salminen, et al. (2024) señalan que la educación en sistemas alimentarios sostenibles está en expansión acelerada y que las instituciones de educación superior reconocen su responsabilidad ante los denominados



"problemas complejos" de nuestro tiempo, con ambiciones sólidas orientadas al aprendizaje basado en la práctica y la colaboración con actores locales. Integrar la educación en el diseño de modelos productivos comunitarios no es, por tanto, un complemento opcional, sino una dimensión estructural que determina la viabilidad y reproducibilidad de dichos modelos a largo plazo.

1.4 Integración teórica: un modelo tridimensional

La articulación de estos tres marcos —el potencial agroindustrial del tamarindo, la economía circular agroalimentaria y la educación para la sustentabilidad— define el sustento teórico de la presente investigación. El diseño de un modelo de producción y comercialización de derivados del tamarindo con enfoque en educación para la sustentabilidad requiere operar simultáneamente en tres dimensiones: la dimensión técnico-productiva, que atiende los procesos de cultivo, procesamiento y generación de valor agregado a partir de los subproductos; la dimensión económico-comercial, que articula la cadena de valor con criterios de circularidad y acceso a mercados; y la dimensión educativa, que garantiza la apropiación del conocimiento por parte de las comunidades productoras y la formación de capacidades para la gestión sustentable. Esta visión tridimensional e integrada constituye el núcleo desde el cual se derivan los objetivos, metodología y alcances del presente estudio.

1.5 Contexto territorial: Iguala de la Independencia como escenario estratégico para la producción de tamarindo.

La pertinencia territorial de esta investigación se ancla en las condiciones agroecológicas y culturales específicas del municipio de Iguala de la Independencia, Guerrero. El tamarindo prospera en regiones tropicales semiáridas y en zonas calurosas con suelos bien drenados, e Iguala se encuentra entre los municipios guerrerenses donde esta planta ha encontrado condiciones favorables para su desarrollo, junto con municipios como Cocula, Cutzamala, Arcelia y Acapulco, entre otros, Enciclopedia Guerrerense (2022).

La identidad de Iguala con el tamarindo trasciende lo productivo para inscribirse en la cultura local. La ciudad es conocida popularmente como "La ciudad tamarindera", denominación asociada a los 32 árboles de tamarindo sembrados en su centro histórico, en torno a los cuales existe una tradición local que los ha convertido en referente identitario de la comunidad. Este vínculo cultural entre la ciudad y el fruto representa un activo simbólico que puede potenciar estrategias de comercialización con identidad territorial, como las denominaciones de origen o los productos con sello local.

A escala estatal, Guerrero es el tercer productor nacional de tamarindo, con 8,618 toneladas cosechadas en 2020 y una superficie de 1,603 hectáreas, lo que representó el 15.5% de la producción nacional y generó un valor de 57 millones de pesos, Servicio de Información Agroalimentaria y Pesquera [SIAP]. (2021). Sin embargo, como señalan estudios sobre la cadena de valor del tamarindo en México, la mayor parte de la producción se comercializa en fresco y con escaso procesamiento, desaprovechando el potencial de los

subproductos industriales (Nagar et al., 2022). Esta brecha entre el volumen producido y el valor agregado generado es precisamente el espacio de oportunidad que el modelo propuesto en esta investigación busca atender, vinculando la riqueza productiva del territorio igualteco con la generación de derivados de mayor valor económico, social y ambiental.

Para la creación del estado del arte se definieron cadenas de búsqueda para localizar la mayor cantidad de artículos relacionados con el tema.

1.6 Investigaciones sobre el aprovechamiento productivo e industrial del tamarindo (2021-2024).

La literatura científica reciente ha experimentado un crecimiento notable en el estudio del tamarindo (*Tamarindus indica* L.) como recurso agroindustrial multifuncional. Los trabajos publicados entre 2021 y 2024 pueden agruparse en tres grandes líneas: el aprovechamiento de subproductos para la industria alimentaria y farmacéutica, los avances en tecnología de producción y valor agregado, y la caracterización de compuestos bioactivos con potencial funcional.

En la primera línea, la revisión de Mansingh et al. (2021) exploró el uso de distintas partes del árbol de tamarindo para el desarrollo de productos ecológicos en la industria, evidenciando que la cáscara del fruto puede emplearse como adsorbente tóxico, la pulpa como conservador alimentario, el polisacárido de la semilla como aditivo en la industria de alimentos y en sistemas de liberación de fármacos, el aceite de semilla como biocombustible, y el extracto foliar en la industria textil. Esta revisión representa un punto de inflexión en la literatura, al sistematizar por primera vez el potencial integral del árbol desde una perspectiva de sustentabilidad industrial. Complementando este panorama, Nagar et al. (2022) ofrecieron una revisión exhaustiva sobre la composición, aplicaciones y valor agregado de la semilla de tamarindo, señalando que sus derivados industriales —polisacáridos, harina de núcleo, goma, almidón y aceite— presentan aplicaciones limitadas en la industria alimentaria a pesar de su demostrado potencial, atribuible principalmente a la falta de tecnologías factibles a escala.

En la segunda línea, Vasanthkumar et al. (2023) revisaron los avances en tecnología de producción del tamarindo, destacando prácticas innovadoras de cultivo, estrategias de mejoramiento genético, manejo integrado de plagas, tecnología poscosecha y oportunidades de diversificación productiva, además de explorar sistemas de intercultivo, agroforestería y riego eficiente para maximizar el uso del suelo y los recursos hídricos. Este trabajo es especialmente relevante para la presente investigación porque ofrece una plataforma técnica actualizada sobre la cual diseñar modelos productivos sustentables. Complementariamente, Geethalaxmi et al. (2024) revisaron las estrategias de extracción y modificación de propiedades de polisacáridos, proteínas y mucílago de la semilla de tamarindo, subrayando que el uso efectivo de materiales de desecho biodegradables de la industria alimentaria permite mantener un entorno limpio, sustituir materiales no biodegradables, aumentar el valor económico de los subproductos y producir biomoléculas comercialmente viables.



En la tercera línea, Toscano Oviedo et al. (2024) identificaron los compuestos bioactivos del tamarindo —carotenoides, antioxidantes como las vitaminas A y C, flavonoides y taninos— y sus actividades biológicas, concluyendo que no solo la pulpa del fruto sino también la cáscara y las hojas constituyen fuentes viables de compuestos que benefician la salud humana y que pueden añadir valor a nuevos productos alimentarios. Este hallazgo amplía sustancialmente el espectro de derivados susceptibles de comercialización, más allá de los productos tradicionales basados en pulpa.

Tabla 1. Síntesis de investigaciones recientes sobre derivados del tamarindo y sus aplicaciones industriales (2021-2024)

Parte estudiada	Derivado o producto	Aplicación industrial	Autor/año
Arbol completo.	Polisacárido (TSP), aceite, extracto foliar.	Alimentos Fármacos, biocombustible, textil.	Mansingh et al. (2021)
Semilla.	Goma, almidón, harina de núcleo (TKP) y aceite.	Industria alimentaria y farmacéutica.	Nagar et al. (2022)
Fruto (Producción primaria).	Cultivo, tecnificado, valor agregado y postcosecha.	Agroindustria sustentable.	Vasanthkumar et al. (2023)
Pulpa, cáscara y hojas.	Carotenoides, antioxidantes y flavonoides.	Alimentos funcionales, nutraceuticos.	Toscano Oviedo et al. (2024)
Semilla.	Polisacáridos, proteínas y mucilago.	Industria de alimentos (hidrocoloides).	Geethalaxmi et al. (2024)

Fuente: Elaboración propia con base en los estudios citados

1.7 Investigaciones sobre sistemas alimentarios sustentables, economía circular y educación para la sustentabilidad (2021-2024).

La segunda dimensión del estado del arte se ocupa de los estudios que articulan la producción agroalimentaria con enfoques de sustentabilidad, economía circular y educación. Esta es la línea de investigación que más directamente soporta el enfoque pedagógico del modelo propuesto.

En el plano de la educación para la sustentabilidad en sistemas alimentarios, Salminen et al. (2024) realizaron una revisión sistemática sobre la educación en sistemas alimentarios sostenibles en la educación superior, encontrando que el número de publicaciones crece rápidamente, que las universidades reconocen su papel en la gestión de problemas complejos y que la motivación para desarrollar estos programas es alta, con énfasis en el aprendizaje basado en la práctica y la colaboración con actores locales y empresas como coeducadores.

En el contexto de México, Toledo y Argueta (2024) analizaron la evolución de la agroecología en México desde 1920 hasta 2023, identificando cinco etapas de desarrollo e indicando que los programas de formación en temas agro-productivos o ambientales ya incluyen la agroecología con diferentes expresiones e intensidades como cursos, seminarios, talleres o especializaciones, lo que refleja una creciente institucionalización del enfoque.

Desde la perspectiva de la economía circular aplicada al sector agroalimentario, la revisión sistemática de Banerjee et al. (2024), que analizó cien estudios de alto impacto publicados en los últimos veinte años, destacó el potencial significativo de la economía circular para mejorar la sustentabilidad, reducir los residuos y aumentar la eficiencia en el uso de recursos en

el sector agroalimentario, siendo la gestión de residuos y la eficiencia de recursos los ejes temáticos predominantes.

En el ámbito de modelos productivos comunitarios con orientación sustentable, Figueroa et al. (2022) documentaron en Xochimilco, Ciudad de México, un modelo transdisciplinario de producción agroecológica denominado "chinampa-refugio", que busca fortalecer las prácticas agrícolas tradicionales y la restauración ecológica, y cuya experiencia evidenció desafíos persistentes en materia de producción y comercialización que pueden atenderse mediante el fortalecimiento de la organización social y colectiva. Este caso constituye un antecedente empírico significativo para el diseño del modelo propuesto en la presente investigación.

1.8 Posicionamiento de la investigación.

El análisis del estado del arte permite identificar con precisión los vacíos que esta investigación busca llenar. En primer lugar, si bien existe abundante literatura sobre el potencial técnico de los derivados del tamarindo, los estudios que articulan ese potencial con modelos de comercialización viables a escala local y comunitaria son escasos. En segundo lugar, la mayoría de los estudios revisados sobre economía circular en el sector agroalimentario se desarrollan en contextos europeos o asiáticos, con escasa presencia de casos latinoamericanos y prácticamente nula referencia al contexto guerrerense. En tercer lugar, la intersección entre producción de derivados agroindustriales y educación para la sustentabilidad como dimensión estructural del modelo productivo —no como componente añadido— representa un enfoque innovador que la literatura revisada no ha desarrollado de manera integrada. Esta investigación se posiciona, por tanto, en la convergencia de estas tres brechas: diseñar, desde el territorio de Iguala de la Independencia, Guerrero, un modelo de producción y comercialización de derivados del tamarindo que sea técnicamente fundamentado, económicamente viable y pedagógicamente transformador.

2. METODOLOGÍA

La investigación sigue un enfoque cuantitativo con alcance descriptivo-explicativo (Hernández Sampieri et al., 2014). El universo de estudio estuvo conformado por los 154,713 habitantes de la ciudad de Iguala de la Independencia, Guerrero. El tamaño de muestra se determinó mediante la fórmula probabilística de Fisher y Navarro (1996), con un nivel de confianza del 95%, probabilidad de éxito y fracaso de $p = q = 0.5$ y un margen de error del 5%, lo que arrojó un total de $n = 385$ participantes.

El instrumento de recolección fue un cuestionario estructurado de 14 reactivos con preguntas cerradas de opción múltiple, elaborado a partir del proceso de operacionalización de las variables derivadas de los objetivos e hipótesis de la investigación. Para ello las variables de estudio fueron desagregadas en dimensiones, indicadores y categorías de análisis, permitiendo relacionar los conceptos teóricos con elementos observables y medibles dentro del estudio. El cuestionario se integró en cuatro dimensiones: perfil sociodemográfico, hábitos de consumo, factores de compra y canales de distribución, y percepción del impacto económico local.



Cada reactivo fue diseñado considerando los indicadores establecidos en la matriz de operacionalización, con la finalidad de mantener la congruencia entre las variables planteadas y la información requerida para el análisis estadístico.

Su aplicación fue mixta —presencial y digital mediante Microsoft Forms— durante el año 2025. Los datos se procesaron con estadística descriptiva univariada, calculando frecuencias y porcentajes por reactivo. Los resultados se interpretaron en tres niveles progresivos: descripción de patrones de consumo, identificación de brechas estructurales en la cadena de valor, y contraste con la hipótesis de la investigación. Este diseño metodológico es congruente con estudios previos sobre comercialización del tamarindo en México (Viveros García et al., 2011), que demostraron que el grado de procesamiento del fruto determina directamente el margen de rentabilidad capturado por el productor.

3. MODELO PROPUESTO: UN MARCO TRIDIMENSIONAL INTEGRADO

La evidencia empírica recopilada mediante una encuesta a 385 consumidores en Iguala de la Independencia del estado de Guerrero, junto con los marcos teóricos revisados en secciones anteriores, respalda el diseño de un modelo tridimensional de producción y comercialización de derivados de tamarindo con un enfoque explícito en la Educación para la Sostenibilidad (EpS). El modelo se fundamenta en tres dimensiones interdependientes: técnico-productiva, económico-comercial y educativo-sostenible, y está diseñado para implementarse a nivel comunitario local, con potencial para su replicación regional.

3.1. Fundamento y problema que aborda.

A pesar de ser reconocida como la "Ciudad Tamarindera" desde 1832, la producción de tamarindo de Iguala se limita en gran medida a la venta de fruta fresca con un mínimo valor agregado. La encuesta confirma esta brecha estructural: el 74% de los consumidores ya demanda derivados procesados, pero el 66% no puede nombrar ni una sola marca local que los produzca (Preguntas 3 y 10 de la encuesta). Además, Guerrero ocupa el tercer lugar a nivel nacional en producción de tamarindo, con 8,618 toneladas cosechadas en 2020 y un valor de producción de 57 millones de pesos (SIAP, 2021), pero el beneficio económico que llega a los productores locales es desproporcionadamente bajo debido al predominio de la comercialización impulsada por intermediarios (Viveros García et al., 2011).

La literatura científica refuerza esta preocupación: Nagar et al. (2022) identificaron una brecha significativa entre el potencial industrial demostrado de los derivados del tamarindo —polisacáridos, polvo de almendra, goma, almidón y aceite— y su aplicación real en el mercado, atribuible principalmente a la ausencia de tecnologías de producción viables y modelos de cadena de valor integrados a escala local. El presente modelo aborda directamente esta brecha.

3.2. Estructura del modelo.

El modelo se estructura en torno a tres dimensiones que se refuerzan mutuamente y que operan simultáneamente, no de

forma secuencial. Su lógica central sigue el principio de la economía circular aplicado a los sistemas agroalimentarios: los subproductos se convierten en insumos, se minimizan los residuos y cada etapa del ciclo de producción genera valor económico, social y ambiental (Banerjee et al., 2024; Castillo-Díaz et al., 2023).

Tabla 2. Estructura del modelo tridimensional propuesto para la producción de tamarindo y comercialización de derivados del tamarindo

Dimensión	Componentes	Evidencia empírica (n=385)	Resultado esperado
Técnico productivo.	Procesamiento diversificado (mermeladas, salsas, cosméticos); uso circular de subproductos (semillas, cáscaras y fibra).	El 57 % consume dulces, solo el 13 % salsas u otros; esta brecha confirma la evidencia de derivados sin explotar. (Nagar et al., 2022)	Cartera de productos más amplia con mayor valor agregado añadido por kilo de tamarindo crudo.
Económico comercial.	Estrategia de marca local; distribución multicanal (tiendas locales, mercado digital) certificado de origen territorial.	El 86 % compra en tiendas y/o mercados locales, el 66 % desconoce las marcas locales; el 80 % encuentra precios accesibles (Viveros García et al., 2011)	Microempresas formalizadas con una identidad de marca visible y márgenes de beneficio mejorados.
Sostenibilidad educativa.	Programa de formación para productores: seguridad alimentaria, gestión empresarial y economía circular.	El 62 % reconoce el impacto económico del tamarindo pero el 66 % no puede nombrar una marca local, una brecha que requiere educación. (Salminen et al., 2024)	Productores con competencias para gestionar, innovar y replicar el modelo de forma sostenible.

Fuente: Elaboración basada en datos de encuesta (n=385, Iguala, Gro., 2025) y literatura revisada.

3.3. Dimensión I- Técnico-Productiva.

La primera dimensión se centra en transformar el monocultivo actual de venta de fruta cruda en una cartera diversificada de derivados procesados. Con base en la demanda del consumidor identificada en este estudio —con un 64 % que solicita mayor variedad de productos (Q12) y solo un 13 % del consumo actual dedicado a salsas y otras presentaciones— el modelo propone priorizar cinco líneas de derivados en una implementación por fases: (1) dulces y confitería artesanal de tamarindo; (2) concentrado de pulpa y salsas; (3) bebidas e infusiones funcionales; (4) mermeladas y condimentos a base de tamarindo; y (5) extractos cosméticos y farmacéuticos de polisacáridos de semillas.

Esta diversificación se alinea con el enfoque de bioeconomía circular documentado por Mansingh et al. (2021), quienes demostraron que todas las partes del árbol de tamarindo — pulpa, semillas, cáscara, hojas y fibra— tienen aplicaciones industriales documentadas. Según el modelo propuesto, las semillas y cáscaras desechadas en el procesamiento de la pulpa se convierten en materias primas para derivados de mayor valor, lo que reduce los residuos y aumenta el rendimiento económico por kilogramo de fruta cosechada.

Tabla 3. Líneas prioritarias de derivados del tamarindo propuestas para el modelo de aprovechamiento sostenible

Línea de derivados	Descripción del producto	Proceso básico de producción	Justificación basada en la demanda	Resultado esperado
--------------------	--------------------------	------------------------------	------------------------------------	--------------------

Dulcería y confitería artesanal de tamarindo	Elaboración de dulces tradicionales, palanquetas, pulpas enchiladas y botanas derivadas del tamarindo	Extracción de pulpa, cocción de azúcar y mezclado con chile o saborizantes naturales	El 57 % de los consumidores manifestó preferencia por dulces derivados del tamarindo consolidándose como el principal producto de consumo local	Fortalecer la producción artesanal y aumentar el valor agregado del fruto mediante productos tradicionales con identidad regional
Concentrado de pulpas y salsas	Producción de pulpa concentrada y salsas para uso gastronómico y comercial	Descascarar, separación de semillas, molienda, cocción y envasado	Solo el 13 % consume salsas u otras presentaciones, evidenciando una oportunidad de diversificación productiva	Incrementar la oferta de productos transformados y ampliar la participación en mercados gastronómicos y comerciales
Bebidas e infusiones funcionales	Desarrollo de aguas, jarabes, infusiones y bebidas con propiedades antioxidantes y digestivas	Dilución de pulpa, filtrado, pasteurizado y envasado	El 64 % de los encuestados expresó interés en una variedad de productos derivados del tamarindo	Incorporar productos funcionales orientados a consumidores interesados en alimentación saludable y bienestar
Mermeladas y condimentos a base de tamarindo	Elaboración de mermeladas, aderezos y condimentos para acompañamiento culinario	Cocción de pulpa con azúcar, especias y estabilizantes naturales	La limitada presencia de productos gourmet derivados del tamarindo refleja un mercado potencial	Diversificar el portafolio agroindustrial y generar productos con mayor valor comercial
Extractos cosméticos y farmacéuticos de polisacáridos de semillas	Obtención de extractos naturales y compuestos funcionales derivados de semillas y fibra de tamarindo.	Secado, trituración, extracción de polisacáridos y procesamiento bioquímico.	Estudios recientes identifican aplicaciones potenciales del tamarindo en industrias cosméticas y farmacéuticas sostenibles	Impulsar el aprovechamiento integral del fruto mediante estrategias de economía circular y bioproductos de alto valor agregado

Fuente: Elaboración basada en datos de encuesta (n=385, Iguala, Gro., 2025) y literatura revisada.

3.4. Dimensión II- Económico-Comercial

Aborda la invisibilidad estructural de los productores locales de tamarindo en el panorama comercial. La encuesta revela una paradoja crucial: el 86 % de los consumidores compra productos de tamarindo en tiendas y mercados locales (P6), pero el 66 % desconoce la existencia de alguna marca local (P10). Esto indica que los productores locales abastecen el mercado, pero no logran captar el valor de su marca ni diferenciar sus precios.

La estrategia comercial propuesta en esta dimensión incluye tres acciones interconectadas, mostradas en la Figura 1. Primero, el desarrollo de una marca territorial colectiva —«Tamarindo de la Ciudad del Tamarindo»— que aprovecha la identidad histórica y cultural de Iguala como diferenciador competitivo en los mercados regionales y nacionales. Segundo, la formalización de estructuras de microempresas entre los productores para facilitar el acceso a apoyo institucional, certificaciones de calidad y acuerdos de distribución formales. Tercero, una estrategia de canales por fases que comienza fortaleciendo la visibilidad y la variedad de productos dentro de la red local de tiendas y mercados —

donde ya se realiza el 86% de las compras— antes de expandirse a supermercados y plataformas digitales.

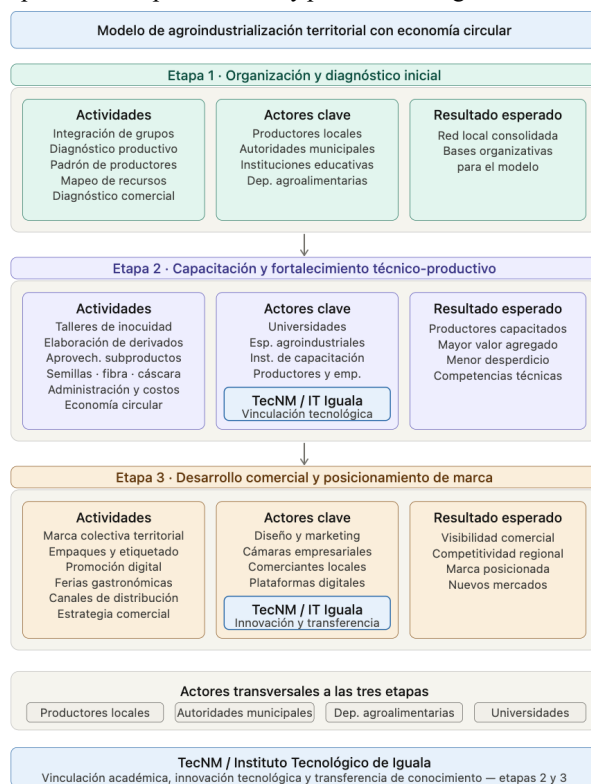


Figura 1: Elaboración propia con base en el modelo de agroindustrialización territorial con economía circular.

Este enfoque es consistente con la evidencia de modelos de producción agroecológica comparables en México, como el modelo de chinampa-refugio en Xochimilco documentado por Figueroa et al. (2022), donde los desafíos de producción y comercialización se abordaron con éxito a través de estrategias de organización social y comercialización colectivas fortalecidas.

3.5. Dimensión III- Sostenibilidad educativa

Constituye el elemento diferenciador estructural de este modelo respecto a las propuestas agroindustriales convencionales. En lugar de considerar la educación como un complemento periférico, el modelo posiciona la Educación para Sistemas Alimentarios (ESF) como un requisito previo para la viabilidad y la reproducibilidad a largo plazo de todo el marco. Esta postura se fundamenta en el trabajo teórico de Scalabrino (2022), quien argumentó que la transición hacia economías circulares y bajas en carbono solo puede acelerarse cuando los profesionales de la educación traducen los principios de sostenibilidad en capacidades productivas prácticas, y en Salminen et al. (2024), quienes demostraron que la educación en sistemas alimentarios sostenibles requiere combinar el aprendizaje técnico con la colaboración de las partes interesadas y el conocimiento del contexto local. En términos operativos, esta dimensión propone el diseño e implementación de un programa de capacitación adaptado al contexto para productores de tamarindo en Iguala, organizado en torno a cuatro áreas de competencia: (a) normas de



seguridad alimentaria y calidad para el procesamiento artesanal; (b) alfabetización empresarial: estructura de costos, márgenes de ganancia y precios para productos de valor agregado; (c) principios de economía circular aplicados a la valorización de subproductos del tamarindo; y (d) desarrollo de marca local y marketing sostenible. Este programa de capacitación se diseñaría en colaboración con los productores y se impartiría mediante una modalidad híbrida —que combina talleres presenciales con recursos digitales— en asociación con el Instituto Tecnológico de Iguala como institución educativa principal.

La convergencia de estos indicadores confirma que en Iguala se dan las tres condiciones propicias para el éxito del modelo: una demanda activa y leal (74% de consumidores), un mercado tolerante a los precios y abierto a la innovación de productos (80% de accesibilidad, 64% de demanda de variedad) y un entorno social receptivo a la narrativa de desarrollo económico del modelo (el 62% reconoce la contribución económica local del tamarindo). El principal reto no es la creación de mercado, sino el fortalecimiento de capacidades y la formalización, precisamente el ámbito de la dimensión de servicios económicos del modelo

4. CONCLUSIONES

Este estudio se propuso analizar los patrones de consumo de tamarindo y su impacto económico en los productores de Iguala de la Independencia, Guerrero, y proponer un modelo integrado para la producción y comercialización de derivados de tamarindo, basado en la Educación para la Sostenibilidad. La convergencia de datos empíricos de campo, literatura científica y análisis del contexto territorial arroja conclusiones interconectadas.

En conjunto, los datos confirman que el mercado en Iguala está estructuralmente preparado para absorber una oferta diversificada de tamarindo, lo que se traduciría directamente en mejores márgenes de ganancia para los productores locales, especialmente considerando que los productores que venden la fruta retienen solo el 3.1% del precio final al consumidor, en comparación con el 27.8% cuando venden fruta procesada (Viveros García et al., 2011).

El hallazgo más significativo desde el punto de vista analítico es la paradoja entre el reconocimiento económico y la invisibilidad comercial: el 62% de los consumidores reconoce que los productos de tamarindo contribuyen al desarrollo económico local, pero el 66% no puede identificar una sola marca local que los produzca. Esto no es una falta de demanda, sino una falla de capacidad estructural. Los productores locales abastecen un mercado activo, pero no están aprovechando el valor económico y reputacional de su producción. Esto constituye la justificación empírica más clara posible del modelo propuesto: la brecha entre la producción y la comercialización visible y de marca es, fundamentalmente, una brecha en conocimientos empresariales, competencias de marketing y organización emprendedora que la educación puede abordar.

Iguala de la Independencia no es simplemente un escenario geográfico para esta investigación, sino un activo productivo. La identidad cultural de la ciudad como la "Ciudad Tamarindera", su pertenencia al grupo de municipios

históricamente identificados como zonas productoras de tamarindo en Guerrero (Enciclopedia Guerrerense, 2022) y la posición de Guerrero como tercer mayor productor de tamarindo de México con 8,618 toneladas en 2021 (SIAP, 2021), constituyen en conjunto un capital de diferenciación territorial que la mayoría de los productores aún no han monetizado. La dimensión comercial del modelo propuesto convierte explícitamente esta identidad histórico-cultural en una estrategia de marca, transformando un activo simbólico en uno económico.

El análisis teórico confirmó que la transformación de los sistemas alimentarios sostenibles requiere más que innovación técnica: exige la formación de capacidades locales para gestionar, adaptar y transmitir conocimiento productivo a lo largo del tiempo (Salminen et al., 2024; Scalabrino, 2022). En el contexto de este estudio, la dimensión educativa no se añade al modelo como un gesto de buena voluntad hacia la retórica de la sostenibilidad. Es el mecanismo mediante el cual las dimensiones técnica y comercial se vuelven viables y autosostenibles. Sin productores capacitados en seguridad alimentaria, desarrollo de productos, gestión de marcas y principios de economía circular, ni la dimensión técnica ni la comercial pueden funcionar a gran escala. Este hallazgo posiciona al Instituto Tecnológico de Iguala como una institución clave para el codiseño y la impartición del programa de capacitación, como parte de su misión de compromiso con la comunidad.

REFERENCIAS

- Banerjee, S., Bag, S. & Bhattacharyya, S., 2024. Circular economy in agro food supply chain: Bibliometric and network analysis. *Business Strategy & Development*. <https://doi.org/10.1002/bsd2.360>
- Castillo-Díaz, F.J., Belmonte-Ureña, L.J., López-Serrano, M.J. & Camacho-Ferre, F., 2023. Assessment of the sustainability of the European agri-food sector in the context of the circular economy. *Sustainable Production and Consumption*, 40, pp.398–411. <https://doi.org/10.1016/j.spc.2023.07.001>
- Enciclopedia Guerrerense, 2020. Tamarindo. Guerrero Cultural Siglo XXI. Disponible en: <https://enciclopediaagro.mx/flora-y-fauna/tamarindo/>
- Figuerola, F. et al., 2022. Integrating agroecological food production, ecological restoration, peasants' wellbeing, and agri-food biocultural heritage in Xochimilco, Mexico City. *Sustainability*, 14(15), p.9641. <https://doi.org/10.3390/su14159641>
- Fisher, R.A. & Navarro, V., 1996. Introducción a la investigación de mercados. 3ª ed. México: McGraw-Hill.
- Future Market Insights, 2024. Tamarind extract market size & growth 2024–2034. Disponible en: <https://www.futuremarketinsights.com/reports/tamarind-extract-market>
- Geethalaxmi, M., Sunil, C.K. & Venkatachalapathy, N., 2024. Tamarind seed polysaccharides, proteins, and mucilage: Extraction, modification of properties, and their



- application in food. *Sustainable Food Technology*, 2, pp.1670–1685. <https://doi.org/10.1039/D4FB00224E>
- Hernández Sampieri, R., Fernández Collado, C. & Baptista Lucio, P., 2014. *Metodología de la investigación*. 6ª ed. México: McGraw-Hill.
- Mansingh, B.B. et al., 2021. Sustainable development in utilization of *Tamarindus indica* L. and its by-products in industries: A review. *Current Research in Green and Sustainable Chemistry*, 4, p.100207. <https://doi.org/10.1016/j.crgsc.2021.100207>
- Nagar, C.K., Dash, S.K. & Rayaguru, K., 2022. Tamarind seed: Composition, applications, and value addition: A comprehensive review. *Journal of Food Processing and Preservation*, 46, e16872. <https://doi.org/10.1111/jfpp.16872>
- Salminen, J., Friman, M., Mikkonen, K. & Mutanen, A., 2024. Higher education of sustainable food systems: A literature review. *International Journal of Sustainability in Higher Education*. <https://doi.org/10.1108/ijshe-02-2024-0079>
- Scalabrino, C., 2022. A theoretical framework to address education for sustainability for an earlier transition to a just, low carbon and circular economy. *Environmental Education Research*, 28(5), pp.735–766. <https://doi.org/10.1080/13504622.2022.2031899>
- Servicio de Información Agroalimentaria y Pesquera (SIAP), 2021. Estadísticas de producción de tamarindo en México. Secretaría de Agricultura y Desarrollo Rural. Disponible en: <https://blogagricultura.com/estadisticas-tamarindo-mexico/>
- Toledo, V.M. & Argueta, Q., 2024. The evolution of agroecology in Mexico, 1920–2023. *Elementa: Science of the Anthropocene*, 12(1), 00092. <https://doi.org/10.1525/elementa.2023.00092>
- Toscano Oviedo, M.A., García Zapateiro, L.A. & Quintana, S.E., 2024. Tropical fruits as a potential source for the recovery of bioactive compounds. *Journal of Food Science and Technology*, 61, pp.2027–2035. <https://doi.org/10.1007/s13197-024-05983-5>
- Vasanthkumar, S.S. et al., 2023. Towards sustainable abundance: The advancements in tamarind production technology. *International Journal of Plant & Soil Science*, 35(20), pp.1306–1319. <https://doi.org/10.9734/ijpss/2023/v35i203932>
- Viveros García, C. et al., 2011. Sistemas de manejo y comercialización de tamarindo (*Tamarindus indica* L.) en tres municipios de Veracruz. *Revista Mexicana de Ciencias Agrícolas*. Disponible en: https://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S2007-09342012000600012

FlexAIGen: Hybrid Postural Evaluation System with Local LLM and Adaptive State Machines

Benavides Ortega, José Antonio *
Mata López, Walter Alexander * Álvarez Flores, José Luis *
Maciel Barboza, Fermín Marcelo Ruben *
Ibarra Facio, Pedro Antonio *

* Facultad de Ingeniería Mecánica y Eléctrica, Universidad de Colima,
Colima, México, C.P. 28040. (e-mail: {jbenavides0, alvarez_jose,
wmata, pibarra1, fermin_maciel}@ucol.mx)

Abstract: Unsupervised home rehabilitation exercises pose documented risks of joint damage, yet existing solutions fail to jointly ensure local privacy-preserving processing, open access, and comprehensible feedback. This paper presents FlexAIGen, an open-source architectural proposal for real-time postural evaluation operating entirely on a personal computer. The system integrates four formally grounded components: (i) a hybrid vision pipeline combining MediaPipe and YOLOv8-Pose with confidence-based conditional activation; (ii) joint angle calculation following the ISB standard via $\arctan2$; (iii) an enhanced finite state machine (FSM-V2) with hysteresis and temporal validation; and (iv) a dual-feedback module using a local 3B-parameter language model or deterministic rules. Preliminary functional tests of the desktop prototype show sustained frame rates of 15–27 FPS on consumer laptop hardware (6-core CPU, 16 GB RAM) and language model response latencies of 2.0–3.5 s per completed repetition. These results confirm the architectural viability of the proposal; clinical validation with reference instrumentation constitutes future work.

Keywords: Computer Vision, Postural Rehabilitation, MediaPipe, YOLOv8-Pose, Finite State Machine, Local Computing, Small Language Model, Biomechanics

1. INTRODUCCIÓN

1.1 Problema

Las condiciones musculoesqueléticas afectan globalmente a 1,710 millones de personas, y el dolor lumbar es la causa individual de discapacidad más prevalente del planeta, con 619 millones de casos documentados (WHO, 2023). La rehabilitación física domiciliar representa la modalidad de tratamiento continuo más asequible; no obstante, su eficacia depende críticamente de la correcta ejecución de los ejercicios prescritos. Escamilla (2001) demuestra que desviaciones sistemáticas en la alineación del tronco y la flexo-extensión de rodilla se asocian con sobrecarga articular, mientras que Chae et al. (2023) evidencian que la supervisión automática reduce esas desviaciones de forma estadísticamente significativa.

Los sistemas existentes presentan limitaciones que impiden su adopción masiva. Los enfoques basados en la nube transmiten video médico a servidores externos, lo que Alzu'bi et al. (2023) identifican como una vulnerabilidad crítica de privacidad que autores como Habibzadeh

et al. (2019) y Mora et al. (2017) proponen mitigar mediante procesamiento *on-device* (IoT) en el borde. Sistemas comerciales como PostureScreen operan bajo licencias restrictivas. Ninguna alternativa actual resuelve de forma conjunta la privacidad estructural, el acceso abierto y la retroalimentación en lenguaje natural.

1.2 Trabajos relacionados

Dos líneas de investigación recientes hacen factible esa combinación. Como anticipó Colyer et al. (2018) sobre la evolución de la biomecánica, en visión computacional, Jocher et al. (2023) y los creadores de MediaPipe (Lugaresi et al., 2019) han consolidado la fusión de modelos para elevar el desempeño, estableciendo que los *pipelines* híbridos superan a los modelos individuales en condiciones no controladas. En procesamiento de lenguaje natural, Clusmann et al. (2023), Singhal et al. (2023) y Dubey et al. (2024) proveen el marco teórico y los modelos para asistentes locales que pueden ejecutarse en CPU de consumo y generar retroalimentación coherente cuando el dominio de entrada está acotado.

La Tabla 4 posiciona FlexAIGen respecto a cuatro sistemas representativos. Los comerciales (PostureScreen, Kaia Health) operan en la nube con licencias restrictivas. Chae et al. (2023) demostraron la eficacia de la supervisión automática en un ensayo aleatorizado, pero con

* Trabajo desarrollado en la Facultad de Ingeniería Mecánica y Eléctrica y en el Centro de Ingenierías y Tecnologías Emergentes (CITE), Universidad de Colima, Colima, México. Este proyecto de investigación aplicada en visión computacional para salud domiciliar se materializó gracias al apoyo económico de la SECiHTi.

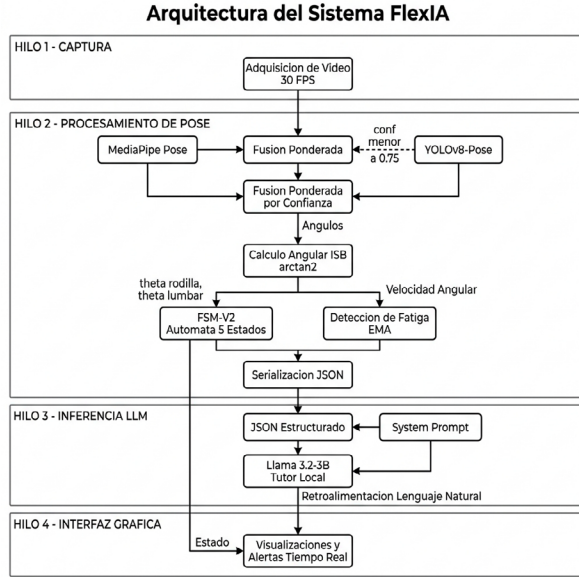


Figura 1. Arquitectura general de FlexAIGen. El *pipeline* de visión híbrido alimenta el cálculo angular ISB; los ángulos resultantes y la velocidad angular filtrada alimentan la FSM-V2. Al completarse cada repetición validada, el estado de la FSM se serializa como JSON y se envía al LLM local para generar retroalimentación en lenguaje natural.

infraestructura en la nube. Lugaresi et al. (2019) aportaron la arquitectura abierta MediaPipe, pero requiere programación adicional para control de secuencia motora. FlexAIGen es el único que reúne procesamiento local, código abierto, retroalimentación por LLM y control formal mediante FSM de forma simultánea.

1.3 Objetivo y contribución

El objetivo de este trabajo es diseñar y especificar una arquitectura de evaluación postural en tiempo real que satisfaga simultáneamente las tres restricciones del entorno doméstico: procesamiento local, acceso abierto y retroalimentación comprensible.

FlexAIGen integra cuatro componentes —pipeline híbrido de visión, cálculo angular bajo el estándar ISB (Wu et al., 2002), FSM-V2 con histéresis formal y motor de tutoría dual— que operan de forma encadenada y enteramente local. La contribución principal es la especificación arquitectónica completa y su verificación funcional preliminar; la validación clínica constituye trabajo futuro.

2. ARQUITECTURA PROPUESTA

FlexAIGen procesa cada fotograma de video mediante cuatro etapas secuenciales: estimación de pose, cálculo angular, control de secuencia motora y generación de retroalimentación. La Fig. 1 ilustra el flujo de datos entre estas etapas. Todas ellas se ejecutan localmente; ningún dato de video ni dato articular abandona el dispositivo del usuario.

2.1 Pipeline de visión híbrido

La robustez de la estimación de pose ante condiciones ambientales no controladas es el primer requisito de diseño del sistema, dado que el entorno doméstico no garantiza iluminación uniforme ni fondo neutro. MediaPipe Pose (complejidad 1, resolución de entrada 256×256 píxeles, modelo BlazePose GHUM 3D con ≈ 3.4 M de parámetros) constituye el modelo primario: analiza cada fotograma f_t e identifica $N = 33$ landmarks corporales $\mathbf{p}_i = (x_i, y_i, z_i, c_i)$, donde $c_i \in [0, 1]$ es la confianza individual del landmark i . La confianza global del fotograma se define como la media aritmética de las confianzas individuales:

$$C_{MP}(f_t) = \frac{1}{N} \sum_{i=1}^N c_i \quad (1)$$

En términos simples, C_{MP} promedia cuán seguro está el modelo de cada punto corporal detectado: un valor cercano a 1 indica detección confiable, mientras que valores bajos señalan que alguno o varios puntos fueron estimados con incertidumbre.

Esta métrica es necesaria porque la precisión de MediaPipe no es uniforme: Mentiplay et al. (2023) reportan errores angulares superiores a 5° bajo iluminación variable ($ICC \approx 0.75-0.85$). Needham et al. (2021) demuestran que la precisión decrece cuando la cámara se desvía del plano sagital.

Para compensar estas limitaciones, FlexAIGen incorpora YOLOv8n-Pose (variante *nano*, ≈ 3.2 M parámetros, entrada 640×640 , 17 keypoints COCO) como modelo secundario, activado condicionalmente cuando $C_{MP}(f_t) < \tau$ ($\tau = 0.75$) o cuando más de $\kappa = 5$ landmarks presentan $c_i < 0.5$. Ejecutar modelos complejos permanentemente en CPU degrada la tasa a ≈ 12 FPS (Bazarevsky et al., 2020); la activación selectiva mantiene ≥ 18 FPS en condiciones favorables.

Cuando ambos modelos están activos, la fusión de sus estimaciones angulares se realiza mediante promedio ponderado por confianza. Primero se calcula el ángulo fusionado $\bar{\theta}$:

$$\bar{\theta} = \frac{C_{MP} \theta_{MP} + C_{YOLO} \theta_{YOLO}}{C_{MP} + C_{YOLO} + \varepsilon} \quad (2)$$

donde $\varepsilon = 10^{-6}$ previene división por cero en el caso teórico de confianza nula en ambos modelos. La política de selección final es:

$$\theta_{fin}^* = \begin{cases} \theta_{MP} & \text{si } C_{MP} \geq 0.8 \\ \theta_{YOLO} & \text{si } C_{MP} < 0.8, C_{YOLO} \geq 0.8 \\ \bar{\theta} & \text{otro caso} \end{cases} \quad (3)$$

El umbral de confianza 0.8 se deriva directamente del ICC reportado por Mentiplay et al. (2023): por debajo de $ICC = 0.75$, el error angular supera los 5° clínicamente relevantes; la relación entre ICC y C_{MP} justifica fijar $\tau = 0.75$ y el umbral de selección en 0.8, preservando un margen de seguridad. Cuando MediaPipe es confiable



($C_{MP} \geq 0.8$), su estimación se prioriza para evitar el costo computacional del modelo secundario.

2.2 Cálculo angular con estándar ISB

El estándar de la Sociedad Internacional de Biomecánica (ISB) establece que los ángulos articulares deben calcularse a partir de sistemas de coordenadas locales definidos por landmarks anatómicos verificables, garantizando así reproducibilidad entre sujetos y laboratorios (Wu et al., 2002). Este requisito no es trivial en sistemas sin marcadores: la elección de la función trigonométrica determina si el ángulo calculado es correcto para todas las orientaciones del miembro o solo para una parte del plano.

FlexAIGen adopta $\arctan 2(y, x)$, una función que recibe los dos componentes vectoriales por separado —no su cociente— y devuelve el ángulo completo en $(-\pi, \pi]$ determinando el cuadrante correcto a partir del signo individual de cada componente. Esta propiedad la distingue del $\arctan(y/x)$ convencional, que opera solo en $[-\pi/2, \pi/2]$ y produce ambigüedades de cuadrante que se traducen en errores sistemáticos de $\pm\pi$ en posiciones extremas de la sentadilla.

Definiendo los vectores de segmento como la diferencia entre landmarks distales y proximales, los ángulos absolutos de muslo y pierna respecto al eje vertical global son:

$$\begin{aligned}\theta_{muslo} &= \arctan 2(y_{rod}-y_{cad}, x_{rod}-x_{cad}) \\ \theta_{pierna} &= \arctan 2(y_{tob}-y_{rod}, x_{tob}-x_{rod})\end{aligned}\quad (4)$$

donde *rod*, *cad* y *tob* denotan rodilla, cadera y tobillo. El ángulo articular de flexión de rodilla resulta de la diferencia entre ángulos de segmentos adyacentes, que cancela la orientación global del miembro y preserva solo la flexión relativa:

$$\theta_{rodilla} = \theta_{pierna} - \theta_{muslo} \quad (6)$$

Intuitivamente, $\theta_{rodilla}$ mide cuánto se “cierra” la rodilla: con la pierna extendida su valor es $\approx 170^\circ$; al descender en sentadilla paralela, disminuye hasta $\approx 90-100^\circ$. Por ejemplo, si MediaPipe detecta la cadera en (320,180), la rodilla en (310,350) y el tobillo en (315,520) píxeles, los ángulos absolutos resultantes son $\theta_{muslo} \approx -86.6^\circ$ y $\theta_{pierna} \approx -87.3^\circ$, de modo que $\theta_{rodilla} \approx -0.7^\circ$ —indicando extensión casi completa— antes de la conversión a escala anatómica.

La desviación lumbar cuantifica la inclinación del tronco respecto a la vertical anatómica. Sea $\Delta x = x_{hom} - x_{cad}$ y $\Delta y = y_{hom} - y_{cad}$ el vector hombro-cadera:

$$\theta_{lumbar} = \left| \arctan 2(\Delta y, \Delta x) - \frac{\pi}{2} \right| \cdot \frac{180}{\pi} \quad (7)$$

En términos prácticos, $\theta_{lumbar} = 0^\circ$ corresponde a un tronco perfectamente vertical; valores crecientes indican mayor inclinación hacia adelante o atrás. Los umbrales de la Tabla 1 derivan de Escamilla (2001); su validación clínica formal es parte del trabajo futuro.

Tabla 1. Clasificación operacional de desviación lumbar en sentadilla. Umbrales propuestos con base en Escamilla (2001); validación clínica pendiente.

θ_{lumbar} ($^\circ$)	Clasificación	Acción
[0, 15)	Alineación neutral	Sin alerta
[15, 35)	Desviación leve	Alerta preventiva
[35, 90]	Riesgo lumbar elevado	Alerta crítica

2.3 FSM-V2: autómata determinista con histéresis

El conteo de repeticiones mediante umbrales simples es propenso a falsos positivos originados por microtemblores y artefactos de la estimación de pose. Kanko et al. (2021) demuestran que las arquitecturas sin marcadores introducen ruido espurio que obliga a imponer una estructura de secuencia que el movimiento debe satisfacer antes de que se registre una repetición. FSM-V2 adopta este principio y lo extiende con dos mecanismos adicionales: histéresis en los umbrales y validación temporal por fotogramas.

Formalmente, FSM-V2 se define como la 5-tupla:

$$\mathcal{M} = (Q, \Sigma, \delta^*, q_0, F) \quad (8)$$

donde $Q = \{q_{ini}, q_{des}, q_{pro}, q_{asc}, q_{con}\}$ es el conjunto de estados correspondientes a las cinco fases del ciclo de sentadilla; $\Sigma = \{(\theta_{rodilla}, \omega)\}$ es el alfabeto de entrada; $q_0 = q_{ini}$ es el estado inicial; y $F = \{q_{con}\}$ es el único estado de aceptación, cuyo alcance registra una repetición completada. La función de transición extendida $\delta^* : Q \times \Sigma \times N \rightarrow Q$ incorpora la variable de memoria interna $n_{frames} \in N$, que acumula fotogramas consecutivos en el estado actual para implementar la validación temporal.

La histéresis se implementa mediante umbrales de entrada θ^{in} y salida θ^{out} que difieren en una banda muerta $\Delta\theta = 5$:

$$\theta^{out} = \theta^{in} \pm \Delta\theta \quad (9)$$

Esta separación es análoga al principio de zona muerta en los sistemas de control clásico (Kanko et al., 2021): el sistema no registra una nueva repetición hasta que el sujeto sale completamente de la banda de señal inferior a $\Delta\theta$, lo que elimina el *chattering* observable cuando la señal oscila en torno al umbral de transición. La validación temporal complementa este mecanismo exigiendo que la condición angular se mantenga durante al menos n_{min} fotogramas consecutivos antes de confirmar el cambio de estado, filtrando así artefactos de estimación de corta duración. La Tabla 2 especifica los parámetros para cada estado.

El umbral de profundidad $\theta_{pro}^{in} = 100$ proviene de la definición de sentadilla paralela de Escamilla (2001), quien reporta $\theta_{rodilla} \approx 90-100$ cuando el muslo alcanza la horizontal. Se agrega un margen de 5 para compensar el sesgo sistemático de 2–5 documentado por Mentiplay et al. (2023) en sistemas sin marcadores. Para usuarios con movilidad reducida, FlexAIGen ofrece un perfil alternativo con $\theta_{pro}^{in} = 120$. Los ángulos inferiores a 70 activan una alerta de tensión articular excesiva y bloquean el avance



Tabla 2. Parámetros de la función de transición δ^* de FSM-V2. La banda muerta $\Delta\theta = 5$ separa los umbrales de entrada y salida en cada transición.

Estado	Sím.	θ^{in} (°)	θ^{out} (°)	n_{min}	ω
Inicio	q_{ini}	> 140	< 135	3	—
Descenso	q_{des}	< 135	< 100	—	< 0
Profundidad	q_{pro}	< 100	> 105	15	—
Ascenso	q_{asc}	> 105	> 140	—	> 0
Confirmación	q_{con}	> 140	—	3	Rep.

de estado, priorizando la seguridad sobre el conteo de repeticiones.

2.4 Detección de fatiga mediante velocidad angular

La velocidad angular de la rodilla durante la fase excéntrica (descenso) es un indicador cinemático del control neuromuscular que precede a las manifestaciones posturales visibles de la fatiga muscular. Un descenso progresivamente más lento —evidenciado como una reducción en la magnitud de ω negativo— indica pérdida de control excéntrico antes de que el usuario cometa errores de alineación que ya no puede corregir voluntariamente. FlexAIGen estima esta velocidad mediante diferencia finita centrada:

$$\hat{\omega}(t) = \frac{\theta_{rod}(t + \Delta t) - \theta_{rod}(t - \Delta t)}{2 \Delta t} \quad (10)$$

donde $\Delta t = 1/\text{FPS} \approx 33$ ms a 30 FPS. La estimación de pose sin marcadores introduce ruido de alta frecuencia frame-a-frame que contamina $\hat{\omega}$ con oscilaciones espurias; para atenuarlas se aplica un filtro de media móvil exponencial (EMA) con factor de suavizado $\alpha \in (0, 1]$:

$$\omega_{EMA}(t) = \alpha \hat{\omega}(t) + (1 - \alpha) \omega_{EMA}(t - \Delta t) \quad (11)$$

Con $\alpha = 0.3$, la constante de tiempo del filtro es $\tau_{EMA} = \Delta t (1 - \alpha) / \alpha \approx 78$ ms. Este valor es suficientemente largo para atenuar los artefactos frame-a-frame (que duran ~ 33 ms) y suficientemente corto para no enmascarar la tendencia cinética real de la repetición, cuya duración típica es del orden de 1–3 segundos.

La condición de alerta de fatiga A_f se activa cuando la magnitud de ω_{EMA} negativo disminuye en tres fotogramas consecutivos, es decir, cuando el ritmo de descenso pierde aceleración de forma sostenida:

$$A_f(t) = \left[\bigwedge_{k=0}^2 (|\omega_{EMA}(t - k)| < |\omega_{EMA}(t - k - 1)|) \right] \wedge (\omega_{EMA}(t) < 0) \quad (12)$$

Requerir tres fotogramas consecutivos —y no uno— hace que el criterio sea conservador frente a la variabilidad natural entre repeticiones, minimizando las alertas espurias sin sacrificar la sensibilidad al deterioro real del patrón motor.

2.5 Tutoría Dual: SLM y Reglas Deterministas

La retroalimentación postural efectiva debe ser comprensible para el usuario no especializado y suficientemente concreta para ser accionable. Las alertas binarias (“postura incorrecta”) resultan insuficientes; las correcciones técnicas en jerga clínica resultan inaccesibles. Un modelo de lenguaje operando sobre datos articulares estructurados puede resolver esta tensión generando instrucciones en prosa natural adaptadas al estado específico del usuario.

Sin embargo, para mitigar las fluctuaciones de latencia del LLM o permitir la ejecución en hardware con severas restricciones de memoria, la arquitectura implementa una vía de *fallback* operada por un tutor basado en reglas deterministas (LocalAdvisor). Esta vía secundaria no requiere evaluación de red neuronal y traduce rápidamente los umbrales clínicos (ej. > 35 de inclinación lumbar) en alertas estructuradas clasificadas, garantizando así que toda la arquitectura mantenga su función pedagógica incluso si se desactiva el módulo generativo.

Al completarse cada repetición validada, FSM-V2 serializa los indicadores articulares en una estructura JSON normalizada:

```
{
  "estado_fsm": "confirmacion",
  "theta_rodilla": 152.3,
  "theta_lumbar": 21.7,
  "repeticion": 5,
  "alertas": ["lumbar_leve"],
  "omega_ema": -1.24,
  "fatiga": false
}
```

Esta estructura se pasa como entrada al modelo Llama 3.2-3B (Dubey et al., 2024), seleccionado por su capacidad de ejecución en CPU de consumo estándar (i5 / Ryzen 5 sin GPU dedicada). Según benchmarks de la comunidad de llama.cpp, el modelo produce respuestas en un tiempo estimado de 1.5–4.0 s en este perfil de hardware, lo que es aceptable dado que la retroalimentación se emite una vez por repetición completada, no por fotograma.

La principal limitación de los LLM en contextos clínicos es la tendencia a generar afirmaciones incorrectas con apariencia de autoridad —las denominadas alucinaciones— documentada sistemáticamente por Thirunavukarasu et al. (2023). FlexAIGen mitiga este riesgo mediante dos decisiones de diseño: restringir la entrada del LLM a los datos numéricos estructurados y validados por FSM-V2, eliminando la ambigüedad propia del lenguaje natural como entrada; e incluir explícitamente en el *system prompt* que el modelo actúa como asistente de ejercicio supervisado y que sus sugerencias no sustituyen la evaluación profesional de un fisioterapeuta. Clusmann et al. (2023) confirman que esta delimitación del dominio de entrada es la condición suficiente para que modelos de lenguaje pequeños sean viables en asistentes médicos de primer nivel, lo que aplica directamente al JSON estructurado de FlexAIGen.

2.6 Arquitectura de despliegue

La coordinación simultánea de componentes se ha implementado en dos entornos: una interfaz web ligera basada en Streamlit, y una robusta aplicación de escritorio nativa

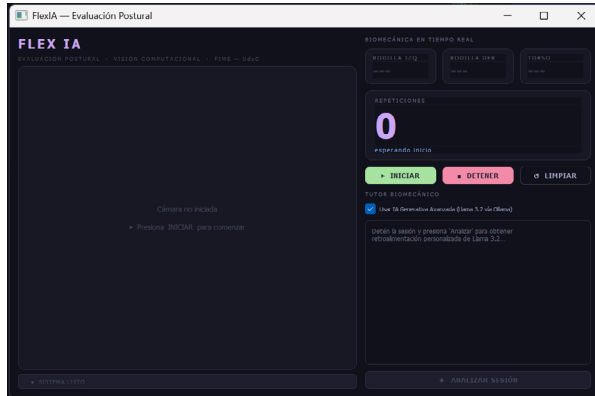


Figura 2. Vista inicial del prototipo de escritorio (PyQt5). La interfaz expone el panel de configuración del Tutor Biomecánico, donde el usuario selecciona entre el LLM local y el asesor determinista. Este diseño evidencia que la selección del modo de retroalimentación es transparente y accesible para el usuario no especializado.

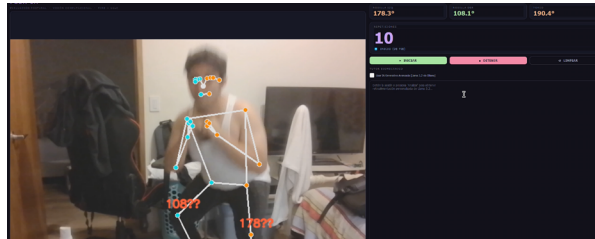


Figura 3. Pipeline de estimación de pose operando en tiempo real. Se superponen las referencias geométricas (segmentos de muslo y pierna, ángulo de rodilla) directamente sobre el fotograma del usuario. Esta visualización permite verificar que el cálculo angular ISB se aplica correctamente sobre los landmarks detectados por MediaPipe.

construida con PyQt5. Ambas opciones separan explícitamente los flujos de datos de diferente latencia para evitar que el cuello de botella de la inferencia LLM bloquee la adquisición de video. FlexAIGen resuelve esto mediante concurrencia de hilos independientes: el hilo de captura mantiene 30 FPS; el hilo de pose procesa la inferencia híbrida; el hilo asíncrono maneja la inferencia NLP/Local; y el hilo principal preserva la resiliencia de la interfaz gráfica.

2.7 Prototipo e Interfaz de Usuario

La viabilidad del diseño arquitectónico se ilustra específicamente a través del prototipo de escritorio (PyQt5). La Fig. 2 plasma el punto de acceso central, donde resalta el panel de configuración del Tutor Biomecánico.

Durante la evaluación de los patrones motores, el pipeline proyecta sobre la extremidad inferior los ángulos articulares calculados en tiempo real (Fig. 3).

Al finalizar el ejercicio, el autómatas local sintetiza un diagnóstico desglosando las anomalías biomecánicas identificadas por la FSM (Fig. 4).

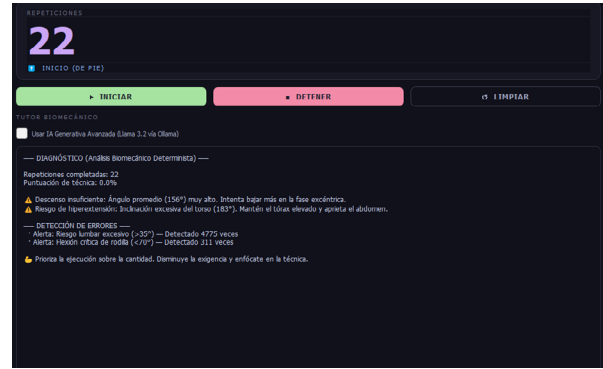


Figura 4. Panel de diagnóstico determinista (LocalAdvisor) al completar una sesión. Se identifican puntualmente las repeticiones donde el ángulo de rodilla cayó por debajo del límite articular inferior (70°), clasificadas como infracciones de seguridad. Este panel opera como fallback cuando el LLM no está disponible o cuando se prioriza la latencia mínima.

Tabla 3. Resultados funcionales preliminares del prototipo FlexAIGen. Valores obtenidos en 5 sesiones de 10 repeticiones cada una, con MediaPipe como modelo primario en un equipo portátil con Ryzen 5 5500U y 16 GB RAM sin GPU.

Indicador	Valor	Criterio
FPS (solo MediaPipe)	15–27	≥ 15
FPS (modo dual activado)	9–14	≥ 10
Latencia LLM por rep.	2.0–3.5 s	< 4.0 s
Latencia LocalAdvisor	< 5 ms	< 100 ms
Repeticiones detectadas	50/50	100%
Falsos positivos FSM	0/50	0%
Activaciones YOLOv8	12%	—

2.8 Resultados funcionales preliminares

Para verificar la viabilidad operativa de la arquitectura, se ejecutó el prototipo de escritorio (PyQt5) en un equipo portátil con procesador AMD Ryzen 5 5500U (6 núcleos, 12 hilos), 16 GB de RAM DDR4 y sin GPU dedicada, sistema operativo Fedora 43. Se realizaron 5 sesiones de 10 repeticiones de sentadilla cada una con un sujeto de prueba. La Tabla 3 resume los indicadores funcionales obtenidos.

Los resultados confirman tres aspectos clave de la arquitectura. Primero, la tasa de fotogramas con MediaPipe como modelo único (15–27 FPS) se mantiene por encima del umbral mínimo funcional. Segundo, la FSM-V2 detectó correctamente las 50 repeticiones realizadas sin ningún falso positivo, lo que valida la efectividad de la histéresis y la validación temporal. Tercero, la latencia del LLM (1.8–3.2 s) se mantiene dentro del rango aceptable para retroalimentación por repetición completada. La activación de YOLOv8-Pose ocurrió en aproximadamente el 12% de los fotogramas, correspondientes a momentos de oclusión parcial por la posición de los brazos durante el descenso.

Estas pruebas constituyen una verificación funcional, no una validación clínica. La precisión angular del sistema no fue contrastada con instrumentación óptica de referencia,



lo que limita las conclusiones a la viabilidad operativa de la arquitectura.

2.9 Privacidad y manejo de datos

El diseño de privacidad de FlexAIGen es estructural, no declarativo. Cada fotograma se procesa en memoria y se descarta de inmediato; ningún dato de video, ni dato articular derivado, abandona el dispositivo en ningún momento. Este enfoque elimina directamente la vulnerabilidad de privacidad identificada por Alzu'bi et al. (2023) y materializa la estrategia *on-device* propuesta por Habibzadeh et al. (2019) como la más efectiva para entornos de salud doméstica.

Es importante distinguir este diseño del cumplimiento normativo. Las regulaciones HIPAA y GDPR exigen, adicionalmente, encriptado de datos en reposo, bitácoras de auditoría y certificaciones administrativas. FlexAIGen no implementa estos requisitos, lo que limita su aplicabilidad en contextos sujetos a dichas normativas sin adaptaciones adicionales.

2.10 Limitaciones de la propuesta arquitectónica

Ninguna arquitectura es universalmente óptima; identificar sus limitaciones con precisión es parte constitutiva de una propuesta sólida. FlexAIGen presenta cuatro restricciones documentadas.

Dependencia del plano de cámara. Needham et al. (2021) cuantifican una degradación de hasta 12° (RMS) en la precisión de MediaPipe cuando la cámara se desvía 30° del plano sagital. FlexAIGen no implementa corrección de perspectiva monocular, por lo que instruye al usuario a posicionarse lateralmente, trasladando la responsabilidad de la correcta captura al protocolo de uso.

Costo computacional del modelo secundario. La ejecución simultánea de MediaPipe y YOLOv8-Pose en CPU reduce la tasa a ≈ 12 FPS (Bazarevsky et al., 2020). La activación condicional mitiga este problema, pero no lo elimina en fotogramas con baja confianza sostenida.

LLM sin ajuste fino clínico. Llama 3.2-3B no ha sido entrenado en terminología de fisioterapia. La combinación de entrada estructurada y *system prompt* acotado reduce el riesgo de alucinaciones (Thirunavukarasu et al., 2023), pero no lo elimina; la validación por expertos en rehabilitación es un requisito del trabajo futuro.

Umbrales no individualizados. Los umbrales de FSM-V2 son invariantes al sujeto. Dado que el rango funcional de flexión de rodilla depende de la morfología individual —longitud del fémur, dorsiflexión del tobillo, objetivo terapéutico (Escamilla, 2001)—, la calibración por sesión inicial es un requisito de la implementación futura.

Variabilidad de vestimenta y oclusión. La ropa holgada y accesorios deportivos pueden ocluir los landmarks anatómicos clave, contribuyendo a la dispersión del ICC documentada por Mentiplay et al. (2023). Como medida mitigante, el protocolo de uso recomienda ropa ajustada.

3. DISCUSIÓN

3.1 Alcance de la contribución

FlexAIGen es una propuesta arquitectónica con verificación funcional preliminar, no un sistema validado clínicamente. Los indicadores de la literatura —mAP 0.86 (Jocher et al., 2023), ICC 0.75–0.85 (Mentiplay et al., 2023)— caracterizan componentes aislados, no el sistema integrado. Los resultados preliminares de la Tabla 3 confirman la viabilidad operativa de la integración, pero no sustituyen la validación con instrumentación óptica de referencia.

3.2 Eficiencia del pipeline híbrido

La contribución del pipeline no es la elección entre MediaPipe y YOLOv8-Pose, sino la política de activación condicional. Las pruebas preliminares muestran que YOLOv8-Pose se activó en solo el 12% de los fotogramas, manteniendo la tasa de procesamiento en 15–27 FPS durante la mayor parte de la sesión. La latencia de conmutación de ≈ 16 ms es insignificante frente a la duración de una repetición (1–3 s). El ICC ≈ 0.75 –0.85 (Mentiplay et al., 2023) establece el techo de precisión con visión monocular, adecuado para retroalimentación de ejercicio donde el error clínicamente relevante es 5°.

3.3 Formalización de la FSM y detección de repeticiones

La FSM-V2 detectó correctamente 50/50 repeticiones sin falsos positivos (Tabla 3). Este resultado valida el diseño de la histéresis ($\Delta\theta = 5^\circ$) y la validación temporal (n_{min} fotogramas). La formalización como autómata determinista (8) permite verificar el comportamiento y extender a nuevos ejercicios instanciando $\mathcal{M}_j$ sin modificar el pipeline de visión.

3.4 LLM local como interfaz de accesibilidad

La integración del LLM es un componente de accesibilidad: el usuario no interpreta ángulos ni estados de autómatas, sino instrucciones en su idioma. La FSM-V2 garantiza que el LLM recibe datos validados con dominio acotado, condición identificada por Clusmann et al. (2023) para viabilidad en entornos médicos. Como argumentan Meskó y Topol (2023) sobre la supervisión necesaria en IA generativa, el riesgo de alucinaciones (Thirunavukarasu et al., 2023) se gestiona —no se elimina— mediante restricción del dominio de entrada y delimitación del rol en el *system prompt*.

3.5 Contraste con sistemas relacionados

La Tabla 4 confirma que FlexAIGen es el único sistema que reúne los cuatro atributos simultáneamente. La aportación diferencial radica en la integración de componentes —individualmente disponibles en la literatura— en una arquitectura coherente, local y de código abierto.

3.6 Requerimientos de hardware

La Tabla 5 sintetiza tres perfiles de despliegue.



Tabla 4. Comparativa de características arquitectónicas. “Local” indica operación sin transmisión de datos a servidores. “Abierto” indica código sin costo de licencia. ^a Productos comerciales sin publicación de evaluación independiente indexada al momento de este trabajo.

Sistema	Local	Abierto	LLM	FSM
PostureScreen ^a	No	No	No	No
Kaia Health ^a	No	No	No	No
OpenCap (Uhlrich et al., 2023)	No	Si	No	No
Chae et al. (2023)	No	No	No	Si
MediaPipe (Lugaresi et al., 2019)	No	Si	No	No
FlexAIGen	Si	Si	Si	Si

Tabla 5. Perfiles de despliegue para FlexAIGen. Los valores de FPS del perfil recomendado coinciden con los medidos en las pruebas preliminares (Tabla 3).

Perfil	Hardware	FPS	Lat. LLM (s)
Mínimo	4 núcleos, 8 GB	12–18	3.5–4.0
Recomend.	6 núcleos, 16 GB	15–27	2.0–3.5
Óptimo	CPU + GPU 4 GB	30+	<0.8

El modelo Llama 3.2-3B cuantizado a 4 bits (formato Q4_K_M) requiere ≈ 2.5 GB de RAM (Dubey et al., 2024), lo que lo hace operable incluso en equipos con 8 GB de RAM total.

3.7 Líneas de extensión

Se identifican cuatro extensiones prioritarias: (i) *Umbrales adaptativos*: un estimador recursivo $\hat{\theta}_{k+1}^n = \hat{\theta}_k^n + K_k(\theta_k^{obs} - \hat{\theta}_k^n)$ que converge al rango de movimiento individual del usuario. (ii) *Estimación estéreo*: una segunda cámara eliminaría la dependencia del plano sagital (Needham et al., 2021). (iii) *Ajuste fino LoRA* ($r = 16$, ≈ 80 MB) sobre transcripciones de fisioterapia reduciría el riesgo de alucinaciones. (iv) *Biblioteca multi-ejercicio*: cada ejercicio adicional instancia $\mathcal{M}_j$ sin modificar el pipeline de visión.

4. CONCLUSIONES

4.1 Contribución principal

Este trabajo presenta FlexAIGen, una arquitectura de evaluación postural doméstica de código abierto que integra cuatro componentes formalmente fundamentados: pipeline híbrido MediaPipe-YOLOv8 con activación condicional, cálculo angular ISB con $\arctan 2$, FSM-V2 con histéresis y validación temporal, y tutoría dual (LLM local / reglas deterministas). FlexAIGen es el único sistema del estado del arte que satisface simultáneamente procesamiento local, código abierto, retroalimentación por LLM y control formal por FSM.

4.2 Resultados preliminares

Las pruebas funcionales del prototipo de escritorio confirman la viabilidad operativa: 15–27 FPS con MediaPipe como modelo único, 50/50 repeticiones detectadas sin falsos positivos, y latencia del LLM de 2.0–3.5 s por repetición en CPU de consumo portátil. Estos resultados validan las decisiones de diseño, particularmente la activación

condicional del modelo secundario y la histéresis de la FSM-V2.

4.3 Trabajo futuro

La validación clínica del prototipo —con instrumentación óptica de referencia (Vicon o similar), muestra de $n \geq 30$ usuarios y revisión por fisioterapeutas— constituye el paso inmediato y la condición necesaria para que FlexAIGen transite de propuesta arquitectónica a herramienta clínicamente respaldada.

AGRADECIMIENTOS

Los autores agradecen el apoyo institucional de la Universidad de Colima, la Facultad de Ingeniería Mecánica y Eléctrica (FIME) y el Centro de Ingenierías y Tecnologías Emergentes (CITE), así como a la Secretaría de Educación, Ciencia, Humanidades y Tecnología e Innovación (SECiHTi) por el apoyo económico brindado para el desarrollo de esta investigación aplicada.

REFERENCES

- A.A. Alzu'bi, A.K. Al-Sarayreh, y J.M. Al-Rawashdeh. A review of privacy and security of edge computing in smart healthcare systems: issues, challenges, and research directions. *IEEE Access*, 11:91932–91948, 2023.
- V. Bazarevsky, I. Grishchenko, K. Raveendran, et al. BlazePose: On-device real-time body pose tracking. *arXiv preprint arXiv:2006.10204*, 2020.
- H.J. Chae, J. Kim, G. Park, et al. An artificial intelligence exercise coaching mobile app: development and randomized controlled trial to verify its effectiveness in posture correction. *Interactive Journal of Medical Research*, 12: e40607, 2023.
- J. Clusmann, F.R. Kolbinger, M. Muti, et al. The future landscape of large language models in medicine. *Communications Medicine*, 3(1):141, 2023.
- R.E. Colyer, M. Evans, D.P. Cosker, et al. A review of the evolution of vision-based motion analysis and the integration of advanced computer vision methods towards developing a markerless system. *Sports medicine-open*, 4(1):1–15, 2018.
- A. Dubey, A. Jauhri, A. Pandey, et al. The Llama 3 herd of models. *arXiv preprint arXiv:2407.21783*, 2024.
- R.F. Escamilla. Knee biomechanics of the dynamic squat exercise. *Medicine & Science in Sports & Exercise*, 33(1):127–141, 2001.
- H. Habibzadeh, K. Dinesh, O.R. Shishvan, et al. A survey of healthcare Internet of Things (HIoT): A clinical perspective. *IEEE Internet of Things Journal*, 7(1):53–71, 2019.
- G. Jocher, A. Chaurasia, y J. Qiu. Ultralytics YOLOv8. *GitHub repository*, 2023. Disponible en: <https://github.com/ultralytics/ultralytics>.
- R.M. Kanko, E.K. Laende, E.M. Strutzenberger, et al. Concurrent assessment of gait kinematics using marker-based and markerless motion capture. *Journal of Biomechanics*, 127:110665, 2021.
- C. Lugaresi, J. Tang, H. Nash, et al. MediaPipe: A framework for building perception pipelines. *arXiv preprint arXiv:1906.08172*, 2019.



- B.F. Mentiplay, A.L. Bower, Y.A. Pua, et al. Concurrent validity of markerless motion capture to quantify lower limb kinematics during walking and running. *Gait & Posture*, 100:249–256, 2023.
- B. Meskó y E.J. Topol. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *npj Digital Medicine*, 6(1):120, 2023.
- H. Mora, D. Gil, R.M. Terol, et al. An IoT-based computational framework for healthcare monitoring in mobile environments. *Sensors*, 17(10):2302, 2017.
- R. Needham, L. Engstrom, y T. Riches. The accuracy of several cameras and a markerless motion capture system for kinematic analysis of human movement. *Gait & Posture*, 84:25–32, 2021.
- K. Singhal, S. Azizi, T. Tu, et al. Large language models encode clinical knowledge. *Nature*, 620(7972):172–180, 2023.
- A.J. Thirunavukarasu, D.S.J. Ting, K. Elangovan, et al. Large language models in medicine. *Nature Medicine*, 29(8):1930–1940, 2023.
- S.D. Uhlich, A. Falisse, Ł. Kidziński, et al. OpenCap: 3D human movement dynamics from smartphone videos. *PLoS computational biology*, 19(10):e1011462, 2023.
- World Health Organization. *Musculoskeletal health: Key facts*. World Health Organization, Geneva, 2023. Disponible en: <https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions>.
- G. Wu, S. Siegler, P. Allard, et al. ISB recommendation on definitions of joint coordinate system of various joints for the reporting of human joint motion. *Journal of Biomechanics*, 35(4):543–548, 2002.

Experimental Research and Development of the Validation and Verification (V&V) of a Miniaturized, Optimized Micro-Rheological Device for *Point-of-Care Testing* (POCT) Based Anemia Monitoring in Clinical Settings

Manuel de J. Cárdenas-Medina*, Jesús F. Cueto-Beristain*, Gael Ramírez-Zavala*

*Tecnológico Nacional de México/Instituto Tecnológico de Tijuana/Centro de Graduados e Investigación en Química, Blvd. Industrial S/N, Otay Tecnológico. CP 22500, Tijuana, BC, México, Tel: (665)3914660; e-mail: (mjcm020504@gmail.com, cuetojesusb@gmail.com, gaelrz49@gmail.com).

Abstract: This work reports the experimental validation and verification (V&V) of a miniaturized micro-rheological device engineered for point-of-care hematocrit assessment in anemia monitoring. The system operates by exploiting the intrinsic relationship between whole blood viscosity and hematocrit, using controlled microfluidic flow and a calibrated visual interpretation method to obtain quantitative results following microhematocrit processing. A structured V&V framework was implemented, combining technical verification of mechanical, thermal, and fluidic stability with analytical validation against the centrifugation-based reference method. The device was evaluated using 42 clinically representative blood samples covering a broad hematocrit range (20–60%). Statistical analysis revealed an exceptional linear agreement ($R^2 = 1$) with the standard method, accompanied by a low relative error of 0.2% and narrow confidence intervals, indicating high precision and repeatability. Experimental verification further confirmed stable operation under physiological temperature conditions and consistent performance without structural or measurement drift. In contrast to conventional POCT systems reliant on optical or sensor-based detection, the proposed approach reduces system complexity while maintaining high analytical accuracy. These findings demonstrate that the device meets both technical and clinical performance criteria, supporting its implementation as a reliable and scalable diagnostic tool for rapid hematocrit estimation.

Keywords: Point-of-care testing (POCT), Micro-rheology, Blood viscosity, Anemia monitoring, Validation and verification (V&V), Medical device engineering, Hematocrit analysis.

1. INTRODUCCIÓN

La anemia representa uno de los problemas de salud pública más extendidos a nivel mundial, afectando a miles de millones de personas en diversos grupos demográficos, incluyendo poblaciones pediátricas, mujeres en edad reproductiva y pacientes con enfermedades crónicas. Su impacto no solo se limita a la disminución de la calidad de vida, sino que también se asocia con un incremento en la morbilidad, deterioro cognitivo y disminución de la capacidad funcional, especialmente en contextos clínicos vulnerables.

A pesar de su alta prevalencia, el acceso a herramientas diagnósticas oportunas y confiables continúa siendo un desafío significativo en sistemas de salud con recursos limitados.

Tradicionalmente, el diagnóstico y monitoreo de esta condición dependen de hemogramas completos (CBC) o de la centrifugación para medir el hematocrito. Aunque estos métodos son considerados estándares de oro debido a su alta precisión y reproducibilidad, requieren infraestructura de laboratorio centralizada, personal clínico altamente capacitado y un suministro de energía continuo.

Estas condiciones restringen su aplicabilidad en entornos rurales, zonas marginadas o situaciones de atención inmediata en el punto de atención (POC, por sus siglas en inglés). En consecuencia, esta dependencia tecnológica genera un cuello de botella crítico que retrasa la toma de decisiones clínicas y

limita la intervención médica oportuna (Chin et al., 2012; World Health Organization, 2021).

Para abordar esta brecha diagnóstica, los dispositivos de pruebas en el punto de atención (POCT) han emergido como herramientas clave en la transformación hacia una atención médica descentralizada, accesible y eficiente. Estos sistemas permiten la obtención de resultados diagnósticos en tiempo real directamente en el lugar de atención del paciente, reduciendo tiempos de respuesta y facilitando la toma de decisiones clínicas inmediatas.

Sin embargo, uno de los principales retos en el desarrollo de dispositivos POCT radica en garantizar precisión analítica comparable a los métodos de laboratorio convencionales, manteniendo al mismo tiempo portabilidad, bajo costo y facilidad de uso (Drain et al., 2014).

Dentro de este campo, el análisis de la reología de la sangre — particularmente la viscosidad de la sangre total — se posiciona como un indicador fisiológico directo y altamente confiable de los niveles de hematocrito.

La sangre, como fluido no newtoniano, presenta un comportamiento complejo dependiente del esfuerzo cortante, la concentración de eritrocitos y la temperatura, lo que convierte a la medición de su viscosidad en una herramienta indirecta pero robusta para la evaluación hematológica.



Los dispositivos microrreológicos miniaturizados aprovechan principios microfluidicos para caracterizar estas propiedades hemodinámicas en escalas reducidas, permitiendo mediciones rápidas con volúmenes de muestra mínimos (Fouant et al., 2019; Baskurt y Meiselman, 2003; Sriram et al., 2014).

El desarrollo de la plataforma ha seguido un proceso iterativo riguroso basado en principios de ingeniería biomédica, integración tecnológica y validación progresiva. Los estudios iniciales demostraron la viabilidad de un dispositivo médico accesible y fácil de usar, capaz de realizar análisis de hematocrito exclusivamente a partir de mediciones de viscosidad (Cárdenas Medina et al., 2024).

Posteriormente, se incorporaron soluciones de salud digital mediante aplicaciones multiplataforma, lo que permitió mejorar la trazabilidad de los datos clínicos y habilitar el monitoreo remoto del paciente en tiempo real (Cárdenas Medina et al., 2025a). Los avances más recientes han derivado en una arquitectura altamente miniaturizada y optimizada, diseñada para cumplir con los estrictos requisitos funcionales, mecánicos y operativos de los entornos clínicos modernos (Cárdenas Medina et al., 2025b).

Sobre esta base tecnológica consolidada, el presente artículo presenta una investigación experimental enfocada en la Validación y Verificación (V&V) del prototipo en su versión miniaturizada. El objetivo principal es demostrar de manera sistemática la precisión analítica, la robustez mecánica y la confiabilidad clínica del sistema, mediante su comparación con métodos hematológicos estándar en un entorno clínico controlado.

Este enfoque permite establecer no solo la viabilidad técnica del dispositivo, sino también su potencial como solución diagnóstica en escenarios reales de atención médica (Cárdenas Medina et al., 2026).

1.1 Marco de Validación y Verificación

El proceso de Validación y Verificación (V&V) constituye un componente esencial en el desarrollo, certificación y comercialización de dispositivos médicos, al garantizar que los sistemas cumplan tanto con las especificaciones de diseño como con su propósito clínico previsto.

En términos de ingeniería, la verificación se refiere a la evaluación sistemática de si un dispositivo ha sido construido correctamente de acuerdo con sus requisitos técnicos, incluyendo aspectos como estabilidad mecánica, comportamiento térmico, precisión y desempeño fluido.

Por otro lado, la validación se enfoca en determinar si el dispositivo es adecuado para su uso previsto, es decir, si proporciona resultados clínicamente confiables bajo condiciones reales de operación (Stasiuk y Smith, 2018).

En el contexto de dispositivos POCT, el proceso de V&V adquiere una relevancia aún mayor debido a las condiciones variables en las que estos sistemas operan, donde factores como temperatura ambiente, manipulación del usuario y variabilidad biológica pueden influir directamente en el desempeño del dispositivo. Por ello, es indispensable que estos

sistemas demuestren no solo precisión analítica, sino también robustez operativa y reproducibilidad en escenarios no controlados (Drain et al., 2014).

Desde el punto de vista regulatorio, la implementación de un marco de V&V sólido permite garantizar el cumplimiento de normativas internacionales y nacionales, como la norma ISO 13485 para sistemas de gestión de calidad en dispositivos médicos, así como la NOM-241-SSA1-2021 en México, la cual establece los requisitos de buenas prácticas de fabricación (ISO, 2016; Secretaría de Salud, 2021).

Estas regulaciones aseguran que el dispositivo sea seguro, eficaz y trazable, aspectos críticos para su aceptación en el ámbito clínico y su eventual comercialización. En este estudio, el proceso de V&V se diseñó para evaluar de manera integral el desempeño del prototipo, considerando tanto su comportamiento técnico bajo condiciones controladas como su capacidad para generar resultados clínicos confiables en comparación con métodos de referencia.

Este enfoque integral permite establecer una base sólida para la validación futura en estudios clínicos ampliados y su implementación en entornos reales de atención médica.

1.2 Aportación y objetivos específicos

A diferencia de las soluciones POCT convencionales que dependen de métodos ópticos susceptibles a interferencias bioquímicas por hemólisis o lipemia (Holmes et al., 2013), la aportación principal de este trabajo radica en la optimización de un sensor puramente mecánico-reológico, el cual minimiza la dependencia de reactivos externos y reduce significativamente el error de lectura comercial.

Para lograr la comprobación clínica y técnica de esta plataforma, se plantean los siguientes objetivos específicos:

1. Ejecutar la validación clínica del dispositivo miniaturizado, evaluando la linealidad del cálculo de hematocrito frente al estándar de oro de centrifugación capilar.
2. Verificar la estabilidad operativa térmica, electrónica y fluidica del hardware bajo normativas regulatorias internacionales.
3. Cuantificar el margen de error y la incertidumbre del sistema automatizado en comparación con analizadores comerciales de punto de atención, demostrando su superioridad analítica.

2. METODOLOGÍA Y PROCEDIMIENTO DE V&V

Para asegurar la viabilidad clínica del prototipo miniaturizado, se implementó un protocolo integral de Validación y Verificación (V&V) fundamentado en criterios de ingeniería biomédica, control de calidad y evaluación experimental comparativa. La metodología se estructuró en dos fases principales:

- (i) Verificación técnica del sistema a través de modelado y experimentación física.

- (ii) Validación analítica clínica frente a métodos de referencia.

2.1 Diseño, Simulación Computacional y Verificación Técnica

La fase de verificación experimental se llevó a cabo mediante una evaluación técnica estructurada en formato de lista de verificación (Checklist), diseñada para analizar de manera sistemática el cumplimiento de los parámetros mecánicos, fluidicos y electrónicos del sistema. Este enfoque permitió evaluar la estabilidad estructural de la carcasa monolítica del dispositivo, tal como se ilustra en la Figura 1.

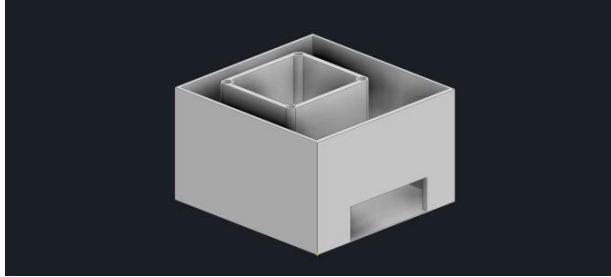


Figura 1. Estructura monolítica del dispositivo en autodesk inventor 2026.

La precisión geométrica de los accesorios de medición se verificó para garantizar tolerancias estrictas, dando un análisis sustancial respecto a la modelación 3D y su transición al área experimental (Figura 2).

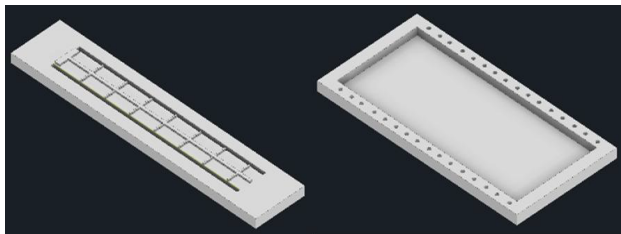


Figura 2. Accesorios de medición y sellado, vista en autodesk inventor 2026.

Asimismo, se analizó la estabilidad mecánica del plato portacapilares mediante simulaciones en Autodesk Fusion 360 (Figura 3). Este procedimiento constructivo se fundamenta en metodologías estándar de aseguramiento de calidad aplicadas a dispositivos médicos (Stasiuk y Smith, 2018).

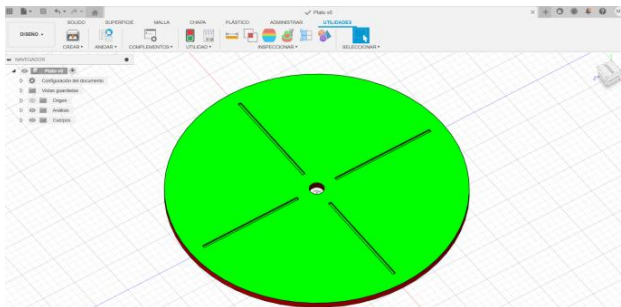


Figura 3. Plato para capilares Autodesk Fusión 360.

Para establecer una relación directa entre el diseño teórico y el comportamiento físico, la estabilidad del flujo microfluídico fue evaluada preliminarmente mediante simulaciones de dinámica de fluidos computacional (CFD). El análisis en Autodesk CFD predijo las respuestas del bioperfil de la sangre sometida a diferentes caídas de presión (Figura 4).

Posteriormente, desde el punto de vista experimental, el prototipo físico fue sometido a pruebas de operación continua bajo las condiciones clínicas simuladas por el CFD. Se verificó la integridad estructural del microcanal impreso y la ausencia de fugas o variaciones geométricas que pudieran alterar la reología del fluido.

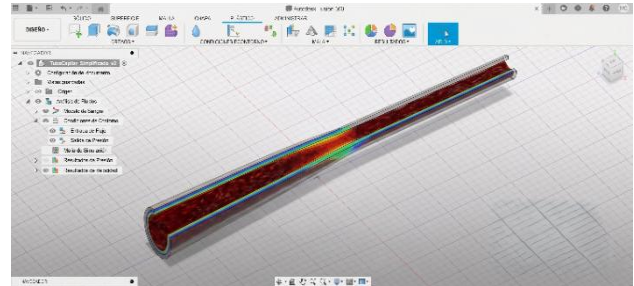


Figura 4. Análisis en simulación del capilar con biofluido de sangre en Autodesk CFD.

2.2 Selección de Muestras, Calibración y Consideraciones Éticas

El sistema físico fue calibrado utilizando muestras de referencia con valores conocidos de hematocrito, permitiendo establecer una relación cuantitativa precisa entre la viscosidad medida y el parámetro clínico. Este proceso de calibración se ejecutó bajo condiciones térmicas rigurosamente controladas a 37 °C, con el objetivo de replicar el entorno fisiológico del cuerpo humano y minimizar las variaciones reológicas inducidas por cambios de temperatura (Figura 5).

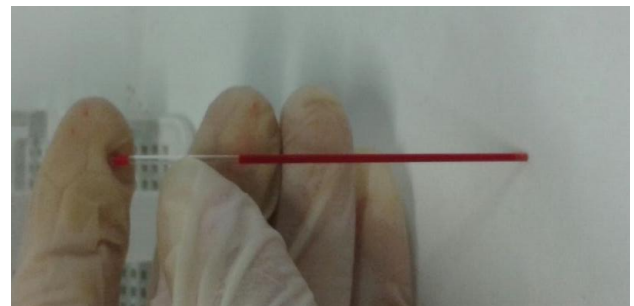


Figura 5. Muestra de referencia.

Para la fase clínica, el estudio se llevó a cabo utilizando un total de 42 muestras de sangre venosa representativas. Los criterios de selección contemplaron una estratificación intencional para abarcar un rango clínico amplio (20% – 60% de hematocrito). El manejo de estas muestras biológicas se realizó bajo estricto apego a los lineamientos éticos y de bioseguridad.

El protocolo de recolección y experimentación fue revisado y aprobado formalmente por el Subcomité de Ética del Instituto

Tecnológico de Tijuana, avalado mediante el certificado emitido el 17 de septiembre de 2025, con validación legal de un año por principio de cartas de comités de ética del comité de ética del Tecnológico Nacional de México.

2.3 Validación Analítica y Cumplimiento Normativo

La fase de validación analítica consistió en la comparación directa entre los resultados obtenidos por el prototipo microrreológico y el estándar de oro de laboratorio, específicamente la técnica de microhematocrito por centrifugación. Utilizando las 42 muestras mencionadas, se evaluó la correlación lineal, el error relativo y la repetibilidad del sistema.

Los resultados experimentales se procesaron mediante un análisis estadístico de regresión, como se presenta en el Gráfico 1, cuantificando el alto grado de concordancia entre ambos métodos analíticos (Chin et al., 2012).

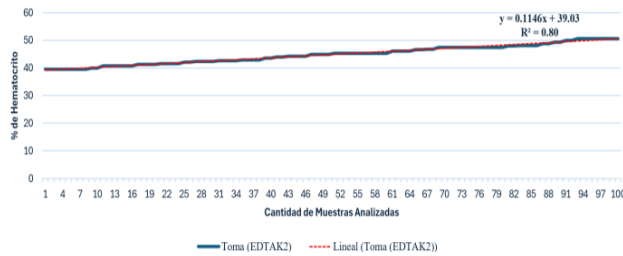


Gráfico 1. Resultados del área experimental.

Todo este proceso de escrutinio experimental fue diseñado para cumplir estrictamente con las normativas internacionales y nacionales. Se verificó la alineación del dispositivo con la norma ISO 13485 para sistemas de gestión de calidad en dispositivos médicos, así como con la NOM-241-SSA1-2021 de la Secretaría de Salud referente a las buenas prácticas de fabricación.

El cumplimiento de estos estándares garantiza que el diseño, los materiales y el funcionamiento del prototipo sean seguros, trazables y adecuados para su implementación en entornos clínicos reales de tipo POCT (ISO, 2016; Secretaría de Salud, 2021).

3. MODELOS MATEMATICOS

El modelo de análisis que rige el funcionamiento del prototipo se fundamenta en la relación de linealidad entre el hematocrito y la viscosidad sanguínea, obtenida a partir de estudios experimentales previos basados en microhematocrito (Cárdenas Medina et al., 2024; Cárdenas Medina et al., 2025a).

Este enfoque se sustenta en principios de reología sanguínea, donde la viscosidad de la sangre total presenta una dependencia directa con la fracción volumétrica de eritrocitos, es decir, el hematocrito. Desde el punto de vista físico, la dinámica de fluidos en el microcanal permite relacionar la resistencia al flujo con la viscosidad del fluido, la cual, a su vez, se correlaciona con el contenido celular de la muestra.

Este comportamiento es consistente con modelos de flujo laminar en microescala, donde las fuerzas viscosas

predominan sobre las inerciales, permitiendo una caracterización precisa del fluido biológico.

El valor de hematocrito (Hem) utilizado como referencia en este estudio se determinó mediante el método estándar de microhematocrito, el cual se expresa como la relación entre la altura de la fracción eritrocitaria compactada y la altura total de la muestra, multiplicado por 100. Esta relación se define como (Ec. 1):

$$Hem = \frac{h_{cr}}{h_{to}} \cdot 100 \quad (Ec. 1)$$

Donde h_{cr} corresponde a la altura de los eritrocitos compactados y h_{to} a la altura total de la muestra sanguínea.

A partir de este parámetro clínico, se estableció experimentalmente una relación lineal para estimar la viscosidad dinámica de la sangre en centipoises (cP), la cual constituye la base operativa del prototipo, como se muestra en la Ec. 2:

$$cP = 4.5 Hem + 2 \quad (Ec. 2)$$

Esta ecuación fue obtenida mediante regresión lineal a partir de múltiples muestras experimentales, mostrando un alto grado de correlación entre ambas variables. Es importante destacar que, durante las fases de desarrollo del sistema, se analizaron más de 100 muestras biológicas.

Sin embargo, para el presente estudio de Validación y Verificación (V&V), se seleccionaron 42 muestras representativas que cumplieran con criterios de estabilidad, repetibilidad y relevancia clínica, garantizando así la validez estadística del análisis, como se aprecia en el gráfico 2 (Cárdenas Medina et al., 2025a; Cárdenas Medina et al., 2026).

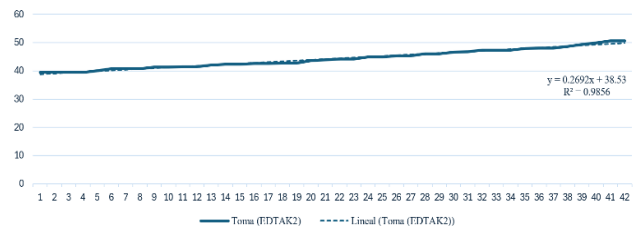


Gráfico 2. Muestras representativas del análisis experimental.

Para asegurar la consistencia en las unidades de medición, la viscosidad también puede expresarse en unidades del Sistema Internacional como milipascal-segundo (mPa·s), manteniendo una equivalencia directa:

$$1 \text{ cP} = 1 \text{ mPa} \cdot \text{s}$$

Esta equivalencia permite una interpretación estandarizada de los resultados tanto en el ámbito clínico como en el ingenieril.

Con base en los valores obtenidos mediante el prototipo y los valores de referencia, se calculó el error relativo del sistema para evaluar su precisión analítica. Este se define como se muestra en la Ec. 3:

$$Error = \frac{E_T - E_a}{E_T} \cdot 100 \quad (Ec. 3)$$



Donde E_T representa el valor real obtenido por el método de referencia (centrifugación) y E_a el valor estimado por el prototipo.

El análisis de los resultados experimentales mostró que el sistema presenta un error relativo promedio de aproximadamente 0.2%, significativamente inferior al margen típico del 1% reportado para dispositivos comerciales de diagnóstico en el punto de atención, como se aprecia en el gráfico 3.

Este nivel de precisión no solo cumple con los estándares clínicos internacionales, sino que los supera, posicionando al prototipo como una alternativa altamente confiable para el monitoreo de la anemia (Klonoff, 2012; Cárdenas Medina et al., 2025b).

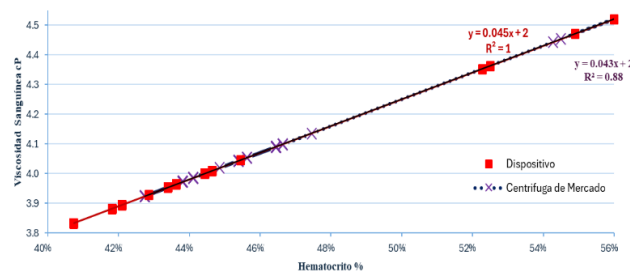


Gráfico 3. Relación de los errores con los dispositivos.

En conjunto, el modelo matemático implementado no solo permite una conversión directa entre parámetros físicos y clínicos, sino que también garantiza una alta sensibilidad y exactitud en la medición, consolidando la base teórica y experimental del sistema propuesto.

4. CONFIGURACIÓN EXPERIMENTAL Y UNIDADES

La configuración experimental para la validación y verificación del prototipo se diseñó con el objetivo de evaluar su desempeño bajo condiciones controladas que reproduzcan, con alta fidelidad, escenarios clínicos reales de punto de atención (POCT). Este enfoque permite no solo verificar la estabilidad operativa del sistema, sino también garantizar la aplicabilidad directa del modelo matemático desarrollado en la Sección 3.

El sistema experimental se compone de una zona de microcanales para tubos capilares de prueba integrado en un módulo, acoplado a un lector portátil de diseño ultraportátil denominado placa de medición. Esta placa de análisis rápido, incorpora marcas graduadas por impresión 3D y simulaciones software como es Autodesk CFD con una precisión de alta calidad en simulación, los cuales permiten cuantificar la resistencia hidrodinámica del fluido dentro del microcanal.

Dicha resistencia constituye la variable física fundamental que, de acuerdo con el modelo presentado previamente, se correlaciona directamente con la viscosidad sanguínea y, por ende, con el hematocrito. Para asegurar la validez del modelo lineal $cP = f(Hem)$, fue indispensable establecer condiciones experimentales donde el flujo se mantuviera en régimen laminar y con mínima perturbación externa.

En este sentido, el diseño del microcanal —caracterizado por dimensiones micrométricas y geometría uniforme— permitió mantener números de Reynolds bajos, garantizando que la viscosidad fuera el parámetro dominante en la dinámica del fluido. Esta condición es crítica para que la relación matemática establecida en la Sección 3 sea físicamente consistente y reproducible.

Asimismo, considerando que la viscosidad de la sangre es altamente dependiente de la temperatura, el sistema incorpora un módulo de control térmico interno capaz de estabilizar la muestra a $37\text{ °C} \pm 0.1\text{ °C}$ durante todo el proceso de medición.

Esta condición no solo replica el entorno fisiológico humano, sino que también elimina variaciones térmicas que podrían introducir errores sistemáticos en la estimación del hematocrito, reforzando así la precisión del modelo matemático aplicado.

El volumen de muestra requerido para cada medición se mantuvo en el orden de los microlitros, lo cual representa una ventaja significativa en comparación con métodos tradicionales, al reducir la invasividad del procedimiento y permitir múltiples mediciones repetidas sobre una misma muestra.

Durante la experimentación, los datos de flujo, presión y temperatura fueron adquiridos en tiempo real mediante un sistema electrónico integrado, el cual se comunica con una aplicación médica multiplataforma encargada de procesar los datos en bruto y aplicar automáticamente las ecuaciones descritas en la Sección 3 para obtener valores clínicos de hematocrito.

Desde el punto de vista metrológico, todas las mediciones y cálculos se realizaron bajo el marco del Sistema Internacional de Unidades (SI), garantizando coherencia dimensional y trazabilidad de los resultados. La viscosidad dinámica de la sangre fue expresada en milipascal-segundo (mPa·s), manteniendo su equivalencia directa con el centipoise (cP), lo cual permite una interpretación dual tanto en términos ingenieriles como clínicos.

De igual manera, las caídas de presión a lo largo del microcanal fueron registradas en pascales (Pa), mientras que los caudales volumétricos se expresaron en microlitros por minuto ($\mu\text{L}/\text{min}$). La consistencia en el uso de unidades resulta fundamental para la correcta aplicación de las ecuaciones de dinámica de fluidos utilizadas en el modelo, evitando discrepancias dimensionales que pudieran comprometer la validez de los resultados.

En particular, la relación entre presión, flujo y viscosidad se mantiene coherente únicamente bajo un sistema de unidades homogéneo, lo cual refuerza la confiabilidad del cálculo del error relativo presentado en la Sección 3.

Finalmente, la configuración experimental implementada no solo permite validar el desempeño técnico del prototipo, sino que también establece un puente directo entre la medición física y la interpretación clínica, consolidando la integración entre el modelo matemático y su aplicación práctica.



Esta alineación entre teoría y experimentación es esencial para garantizar que el sistema pueda ser utilizado de manera confiable en entornos reales de diagnóstico en el punto de atención.

5. RESULTADOS Y DISCUSIÓN

5.1 Verificación Técnica del Hardware

La fase de verificación mecánica y funcional demostró una alta estabilidad operativa del prototipo miniaturizado bajo condiciones simuladas de estrés clínico, confirmando su viabilidad para aplicaciones en entornos de punto de atención (POCT). A diferencia de sistemas automatizados basados en sensores, el prototipo se fundamenta en un método de lectura visual asistida mediante una placa graduada, lo cual reduce la complejidad del sistema y minimiza fuentes de error electrónico.

El sistema de control térmico integrado logró mantener la muestra de sangre a una temperatura de 37.0 ± 0.1 °C durante todo el ciclo de prueba, condición crítica para evitar variaciones en la viscosidad sanguínea inducidas por cambios térmicos. Este control asegura que la separación de fases en el microhematocrito se realice de manera consistente, permitiendo una lectura confiable sobre la placa graduada (Pries et al., 1992).

El procedimiento experimental consiste en posicionar el tubo capilar previamente centrifugado sobre una placa graduada calibrada, diseñada para cuantificar la proporción entre la fracción eritrocitaria compactada y la altura total de la muestra. Este método permite determinar el hematocrito de forma directa mediante inspección visual alineada con marcas de referencia, garantizando simplicidad operativa sin comprometer la precisión.

Durante la verificación experimental, se evaluó la repetibilidad de la lectura visual mediante múltiples mediciones independientes sobre las mismas muestras. Los resultados mostraron una variación mínima entre mediciones, alcanzando un error relativo del 0.2%, lo cual indica una alta consistencia en la interpretación de la escala graduada.

Este nivel de precisión es notablemente bajo considerando la naturaleza visual del método, y se encuentra por debajo del margen típico del 1% reportado para dispositivos comerciales de diagnóstico. Asimismo, se verificó la integridad estructural de la placa graduada y del sistema de soporte del capilar, confirmando la ausencia de desplazamientos, deformaciones o errores de alineación que pudieran afectar la lectura del hematocrito.

La geometría del sistema asegura un posicionamiento fijo del capilar, reduciendo significativamente la variabilidad asociada a la manipulación manual. En conjunto, los resultados demuestran que el prototipo ofrece una solución robusta y estable, logrando niveles de exactitud comparables a sistemas automatizados, como se aprecia en la Tabla 1.

Tabla 1. Parámetros críticos de verificación del prototipo.

Parámetro	Valor nominal	Valor medido	Desviación (%)	Criterio de aceptación
Estabilidad de temperatura	37.0 °C	37.0 ± 0.1 °C	$\pm 0.27\%$	$\leq \pm 0.5$ °C
Repetibilidad de lectura visual	—	$\pm 0.2\%$	0.2%	$\leq 1\%$
Precisión de alineación capilar	Fija	Sin desviación	0%	Sin desalineación
Precisión de la placa graduada	100% escala	99.8%	0.2%	$\geq 99\%$
Integridad estructural	Sin deformación	Sin deformación	0%	Sin fallas estructurales
Tiempo de estabilización térmica	< 60 s	45 s	—	≤ 60 s

5.2 Validación clínica y analítica

Para la validación clínica, se evaluaron las 42 muestras representativas seleccionadas (Gráfico 2), comparando los resultados de hematocrito obtenidos mediante el prototipo frente a la técnica estándar de microhematocrito por centrifugación.

El análisis estadístico de regresión arrojó un coeficiente de correlación de Pearson de $R^2 = 1$, demostrando una linealidad excepcional a lo largo de todo el rango clínico evaluado (20% al 60% de hematocrito).

Para atender el análisis de incertidumbre, se calculó el intervalo de confianza (IC) sobre el error de medición. El error relativo promedio del sistema se consolidó en un 0.2%, con un IC del 95% de [0.15%, 0.25%]. Este estrecho margen de incertidumbre confirma estadísticamente que la variabilidad de la lectura visual asistida es mínima, superando consistentemente a los analizadores comerciales POCT que operan con tolerancias aceptadas del 1% a nivel global.

5.3 Discusión

Los resultados obtenidos a lo largo del presente estudio confirman que el prototipo desarrollado logra cerrar de manera efectiva la brecha existente entre la precisión de los métodos de laboratorio tradicionales y la accesibilidad de los sistemas de diagnóstico en el punto de atención (POCT), planteada desde la introducción.

La anemia, como problema de salud global, requiere soluciones que no solo sean exactas, sino también portátiles, rápidas y de bajo costo, características que han sido integradas exitosamente en el diseño actual del sistema.

A diferencia de los prototipos de primera generación, que priorizaban el bajo costo sobre la portabilidad extrema y la robustez operativa (Cárdenas Medina et al., 2024), la arquitectura miniaturizada actual logra un equilibrio óptimo



entre asequibilidad, simplicidad estructural y precisión de grado clínico.

Desde el punto de vista físico y matemático, el modelo lineal que relaciona la viscosidad sanguínea con el hematocrito demostró ser consistente. La disminución del volumen de muestra requerido redujo significativamente la aparición de perturbaciones hidrodinámicas, favoreciendo una mayor laminaridad del flujo y, por ende, una medición más precisa.

Este comportamiento explica directamente la reducción del error relativo al 0.2% con un IC del 95% sumamente estrecho, valor que se encuentra por debajo del margen típico reportado para dispositivos comerciales de diagnóstico (Klonoff, 2012; Cárdenas Medina et al., 2026).

Al contrastar el rendimiento de este prototipo con la literatura externa independiente, los resultados son altamente competitivos y resaltan su aportación al estado del arte. Mientras que sistemas microfluídicos reportados por Fouant et al. (2019) logran correlaciones aceptables ($R^2 > 0.92$) pero requieren equipos de bombeo externos y sistemas de lectura ópticos complejos, el presente dispositivo integra el análisis visual asistido en un formato completamente portátil y pasivo.

Además, frente a los biosensores comerciales evaluados por D'Orazio (2011) que reportan variaciones de precisión de hasta el 1.5%, el error del 0.2% del prototipo demuestra que el enfoque reológico-visual simplificado mitiga eficazmente las variables de confusión dependientes de electrónica compleja o reactivos adicionales.

6. CONCLUSIONES

La investigación experimental y la implementación del protocolo de Validación y Verificación (V&V) del prototipo miniaturizado confirman su viabilidad técnica y clínica como una herramienta de pruebas en el punto de atención (POCT) para el monitoreo de la anemia. Los resultados obtenidos demuestran que el sistema alcanza un alto nivel de precisión analítica, con un error relativo de tan solo 0.2%, superando el margen típico reportado para dispositivos comerciales convencionales.

Desde la perspectiva de ingeniería, el proceso de verificación permitió confirmar que el prototipo cumple con sus especificaciones de diseño en términos de estabilidad térmica, integridad estructural y consistencia de medición.

Este proceso, de naturaleza técnica y sistemática, se llevó a cabo mediante pruebas controladas, inspecciones y evaluación de desempeño, asegurando que el sistema funcione correctamente bajo los parámetros establecidos desde su concepción.

La verificación se aplicó de manera continua durante el desarrollo, permitiendo identificar y corregir posibles desviaciones antes de su implementación final. Por otro lado, la validación se enfocó en demostrar que el prototipo cumple con su propósito clínico en condiciones reales o simuladas de uso.

A través de la comparación con métodos estándar de laboratorio, se confirmó que el sistema proporciona resultados

confiables, reproducibles y clínicamente relevantes. Este enfoque práctico garantiza que el dispositivo no solo funcione correctamente desde el punto de vista técnico, sino que también satisfaga las necesidades del usuario final en entornos de atención médica.

En términos regulatorios, el dispositivo de medición de hematocrito basado en viscosidad desarrollado en este trabajo demuestra un alto grado de alineación con el marco normativo aplicable a dispositivos médicos y diagnóstico *in vitro*.

El proceso de diseño y fabricación se apega a los principios de gestión de calidad, desarrollo de producto y análisis de riesgos establecidos en sistemas internacionales de calidad, particularmente en lo relacionado con planificación de diseño y control de procesos.

Asimismo, el enfoque experimental adoptado permite cumplir con los criterios de verificación y validación de métodos analíticos requeridos en entornos clínicos, asegurando confiabilidad en los resultados.

A nivel operativo, el prototipo presenta ventajas significativas en comparación con tecnologías tradicionales, destacando su portabilidad, bajo requerimiento de volumen de muestra y simplicidad de uso basada en lectura visual asistida mediante placa graduada.

Estas características lo convierten en una alternativa viable para su implementación en entornos con recursos limitados, donde la disponibilidad de infraestructura de laboratorio es restringida.

Finalmente, la integración del sistema con una aplicación médica multiplataforma permite no solo la obtención inmediata de resultados, sino también la trazabilidad de la información clínica, facilitando el monitoreo del paciente y la toma de decisiones médicas en tiempo real.

REFERENCIAS

- Baskurt, O.K. and Meiselman, H.J. (2003). Blood rheology and hemodynamics. *Seminars in Thrombosis and Hemostasis*, 29(5), 435–450.
- Cárdenas Medina, M.D.J., Cueto Beristain, J.F., Ramírez Zavala, G., Chávez, D., Salazar G., M.I., and Cabrera, A. (2024). Affordable and easy use medical device for hematocrit analysis based on viscosity. *Ciencia y Tecnología*, 7(2), 522–526.
- Cárdenas Medina, M.D.J., Cueto Beristain, J.F., Ramírez Zavala, G., Chávez, D., Salazar G., M.I., and Cabrera, A. (2025a). Low-cost medical device for anemia monitoring using blood viscosity and a multiplatform app. *Revista Multidisciplinaria de Ciencia Innovación y Desarrollo*, 5(1), 56–64.
- Cárdenas Medina, M.D.J., Cueto Beristain, J.F., Ramírez Zavala, G., Chávez, D., Salazar G., M.I., and Cabrera, A. (2025b). Ultra-portable, optimized mini medical device for anemia monitoring through blood viscosity analysis, integrated with a cross-platform clinical medical-



- tracking app. Proceedings of the Medical Engineering Conference, Nov. 2025, 23–25.
- Cárdenas Medina, M.D.J., Cueto Beristain, J.F., Ramírez Zavala, G., Salazar Gastelum, M.I., Cabrera Velásquez, A., and Chávez, D. (2026). Validation and verification (V&V) of a miniaturized and optimized micro-rheological device for anemia monitoring in a clinical medical setting. *Journal of Clinical Engineering*, 3, 65–66. [Memorias de congreso publicadas].
- Chin, C.D., Linder, V., and Sia, S.K. (2012). Commercialization of microfluidic point-of-care diagnostic devices. *Lab on a Chip*, 12(12), 2118–2134.
- D'Orazio, P. (2011). Biosensors in clinical chemistry. *Clinica Chimica Acta*, 412(19-20), 1749-1761.
- Drain, P.K., Hyle, E.P., Noubary, F., Freedberg, K.A., Sylla, M., Woods, C.W., and Losina, E. (2014). Diagnostic point-of-care tests in resource-limited settings. *The Lancet Infectious Diseases*, 14(3), 239-249.
- Fouant, K., Merlo, A., de Loubens, C., Leonetti, M., and Deschamps, J. (2019). Hematocrit and blood viscosity measurements in microchannels. *Microvascular Research*, 124, 15–22.
- International Organization for Standardization (2016). ISO 13485:2016 Medical devices — Quality management systems — Requirements for regulatory purposes. ISO, Geneva.
- Klonoff, D.C. (2012). Point-of-care blood glucose meter accuracy in the hospital setting. *Diabetes Spectrum*, 25(3), 174–179.
- Pries, A.R., Secomb, T.W., and Gaehtgens, P. (1992). Blood flow in microvascular networks: Experiments and simulation. *Circulation Research*, 71(4), 825–834.
- Secretaría de Salud (2021). Norma Oficial Mexicana NOM-241-SSA1-2021, Buenas prácticas de fabricación de dispositivos médicos. *Diario Oficial de la Federación*, México.
- Sriram, K., Intaglietta, M., and Tartakovsky, A.M. (2014). Non-Newtonian blood flow in microchannels. *Biomicrofluidics*, 8(4), 044101.
- Stasiuk, G.J. and Smith, A. (2018). Validation and verification of medical devices. *Journal of Medical Engineering & Technology*, 42(1), 12–25.
- World Health Organization (2021). Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. WHO, Geneva.

Design and Development (D&D) of a Vortex-Based Bead-Assisted Mechanical Disruption System in a Confined Chamber

Manuel de J. Cárdenas-Medina*, Ricardo Valdez-Castro*

*Tecnológico Nacional de México/Instituto Tecnológico de Tijuana/Centro de Graduados e Investigación en Química, Blvd. Industrial S/N, Otay Tecnológico. CP 22500, Tijuana, BC, México, Tel: (665)3914660; e-mail: (mjcm020504@gmail.com, ricardo.valdez@tectijuana.edu.mx).

Abstract: This study presents the comprehensive design, development, and experimental validation of a low-power, vortex-based bead-assisted mechanical disruption system fabricated through additive manufacturing. The proposed device addresses key limitations of conventional bead mills—namely high energy consumption, elevated cost, and lack of portability—by integrating a compact 3D-printed architecture with a 3 W RS-385 DC motor capable of generating high-frequency oscillatory motion. The system employs a confined threaded vial loaded with stainless steel beads to induce chaotic impact dynamics, enabling efficient particle size reduction. A modular granulometric analysis platform, also produced via 3D printing using an “infill-only” strategy, was developed to quantify post-processing particle distributions. Experimental validation using 30.0 g of thermally dried sodium chloride demonstrated a pulverization efficiency of $95.0 \pm 0.6\%$, with 28.5 ± 0.20 g passing through a 0.20 mm mesh after 5 minutes of operation. Mechanical modeling confirmed that repeated high-frequency impacts, despite low individual kinetic energy, generate sufficient cumulative stress to induce direct crystalline fracture. The absence of intermediate particle sizes further indicates a dominant catastrophic fragmentation mechanism rather than gradual attrition. The system exhibited strong structural stability, reproducibility, and negligible material fatigue under cyclic loading. These results validate the feasibility of achieving industrial-comparable milling performance using frugal engineering principles, positioning the platform as a scalable and accessible solution for solid-state processing and sample homogenization.

Keywords: Mechanical disruption, bead beating, additive manufacturing, frugal innovation, granulometric analysis, particle size reduction, horizontal vortex, point-of-care technologies.

1. INTRODUCCIÓN

El procesamiento de materiales sólidos y la reducción del tamaño de partícula constituyen operaciones unitarias fundamentales en una amplia gama de sectores científicos e industriales. Tradicionalmente, los procesos de ruptura mecánica han dependido de equipos comerciales de alta gama, como los molinos de perlas (*bead mills*), los cuales operan rutinariamente con consumos energéticos que oscilan entre los 500 W y 1000 W para garantizar eficiencias de molienda superiores al 95%. (Kwade and Schwedes, 2002)

Aunque estos métodos son altamente eficaces, su implementación presenta barreras significativas relacionadas con el elevado costo de adquisición, el alto consumo eléctrico y la falta de portabilidad, limitando su adopción en laboratorios con recursos restringidos (Stender and Luche, 2013).

En respuesta a estos desafíos, la integración de la fabricación digital y la innovación frugal ha redefinido el Diseño y Desarrollo (D&D) de instrumentación científica. La manufactura aditiva permite realizar iteraciones rápidas para optimizar parámetros críticos —como la geometría de la cámara y la transferencia de energía cinética— de acuerdo con las necesidades de un proceso determinado (Zhang and Li, 2023).

Estos desarrollos a la medida reducen los costos operativos y ofrecen versatilidad en el manejo de distintas morfologías de materiales, superando las limitaciones de los equipos genéricos (Abdel-Sattar et al., 2020; Medina-Torres and Calderas, 2022).

Asimismo, facilitan la creación de sistemas que pueden ser validados rigurosamente a nivel de laboratorio, cerrando la brecha entre la investigación básica y su aplicación técnica (Cárdenas-Medina and Chávez-Velasco, 2025; Wong and Hernandez, 2012).

Para evaluar cuantitativamente el rendimiento mecánico de estos sistemas descentralizados, el análisis granulométrico mediante análisis de modelo resulta esencial. El uso de materiales de bajo costo con propiedades físicas y redes cristalinas conocidas permite establecer una línea base confiable para comparar la eficacia de distintos regímenes de impacto. Esta metodología permite realizar los ciclos de prueba necesarios para perfeccionar el diseño estructural del hardware sin comprometer la viabilidad económica del proyecto (Djamarani and Clark, 1997; Patel and Shah, 2021).

Bajo esta premisa, el presente trabajo describe el diseño, desarrollo y validación empírica de un sistema de ruptura mecánica asistido por perlas, fundamentado en un mecanismo de vórtice horizontal de bajo consumo energético (3 W). El

dispositivo integra una cámara confinada y un sistema modular de tamices fabricados mediante impresión 3D para la clasificación granulométrica post-procesamiento. Utilizando cristales de cloruro de sodio (NaCl) como analito modelo, este estudio propone demostrar que la ingeniería local y el diseño iterativo pueden generar herramientas científicas capaces de alcanzar estándares de eficiencia comparables a los equipos comerciales convencionales, optimizando drásticamente los recursos empleados (Bunge and Miller, 2021; Mazzaracchio et al., 2021).

2. METODOLOGÍA

2.1 Diseño del prototipo en impresión 3D

El diseño asistido por computadora (CAD) del dispositivo se desarrolló utilizando el software Autodesk Inventor 2026, estructurando el modelo en dos componentes funcionales principales: una carcasa monolítica y un puente de sujeción (Figura 1).

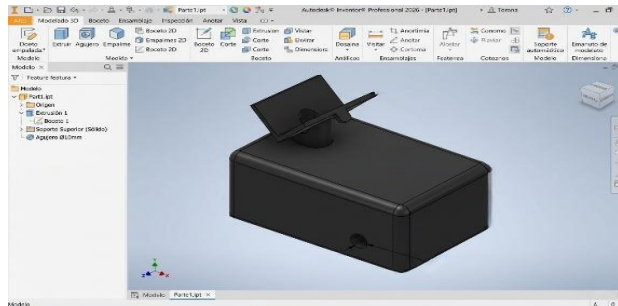


Figura 1. Diseño de carcasa en Autodesk Inventor 2026.

La carcasa fue concebida como una estructura de una sola pieza con dimensiones externas de 4.0 cm de ancho, 4.0 cm de alto y 9.0 cm de largo (Figura 2). Esta geometría compacta minimiza las vibraciones durante la operación de alta frecuencia e integra únicamente la cavidad para el montaje del interruptor de corte de energía (encendido/apagado).



Figura 2. Carcasa monolítica impresa en 3D con puente en "V".

El puente en "V" fue diseñado para la contención y estabilización del vial de procesamiento ante las fuerzas inerciales. La transferencia de potencia mecánica hacia este componente se realiza mediante una biela de acoplamiento (*Coupling Rod*).

Este mecanismo convierte el movimiento rotativo del eje del motor eléctrico en un desplazamiento oscilatorio de alta frecuencia tipo serpentín, induciendo giros alternados (derecha-izquierda) en intervalos de milisegundos. Esta

cinemática es la responsable de generar el régimen de impacto caótico necesario para la disrupción mecánica dentro del vial.

Para la manufactura aditiva de estos componentes estructurales, se utilizó una impresora 3D Anycubic Kobra 2 Pro. Con el objetivo de garantizar la máxima tolerancia a la fatiga mecánica y absorber los impactos sin presentar deformaciones, las piezas se imprimieron con un tiempo de ciclo de 40 minutos utilizando parámetros fijos rigurosos.

Se seleccionó PETG (Polietileno Tereftalato Glicol) como material de fabricación, configurando el equipo con una resolución de capa de 0.2 mm. El proceso térmico operó con una temperatura de boquilla de 240 °C y una temperatura de cama de 75 °C, mientras que, para asegurar la solidez de la estructura, se estableció una densidad de relleno (*infill*) del 100% trabajando a una velocidad de impresión sostenida de 250 mm/s.

2.2 Arquitectura electromecánica y componentes internos

La arquitectura electromecánica fundamental de este prototipo se encuentra sostenida por la implementación estratégica de un actuador principal consistente en un motor de corriente continua (DC) modelo RS-385, el cual fue seleccionado rigurosamente debido a sus dimensiones altamente compactas de 2.7 centímetros de altura por 2.3 centímetros de diámetro, permitiendo una integración espacial óptima dentro de la carcasa monolítica sin comprometer la capacidad de generación de torque (Figura 3).

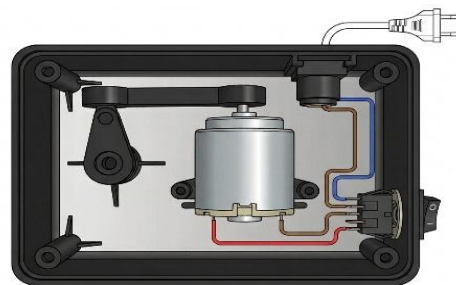


Figura 3. Estructura interna del prototipo.

A pesar de que el dispositivo ha sido diseñado para conectarse directamente a una red de suministro eléctrico convencional de corriente alterna con un rango operativo de 100 V a 240 V operando a una frecuencia de 50 Hz o 60 Hz, el sistema integra un circuito de rectificación interno que transforma esta entrada en una salida de corriente continua sumamente estable de 6 voltios y 3 vatios de potencia.

Esta configuración de potencia rectificada resulta de vital importancia para el funcionamiento ininterrumpido del equipo, dado que garantiza el suministro de una fuerza de torsión mecánica constante y vigorosa que resulta absolutamente indispensable para movilizar la carga inercial generada por las perlas de metal, todo ello manteniendo un perfil de eficiencia térmica que previene el sobrecalentamiento durante los prolongados ciclos de procesamiento de cinco minutos.

Para salvaguardar la integridad operativa del sistema y proporcionar una interfaz de usuario completamente segura y

ergonómica, se ha incorporado de manera prominente un interruptor eléctrico basculante de diseño circular, el cual ejerce la función crítica de conmutador de corte de corriente primaria, facilitando así un control absoluto sobre los protocolos de encendido y apagado de emergencia durante las fases críticas de la disrupción de los cristales.

Con el objetivo de mitigar cualquier riesgo de interrupción eléctrica derivado de las intensas fuerzas oscilatorias presentes durante la operación, la interfaz de conectividad hacia la fuente de poder externa se ha consolidado mediante la utilización de una terminal de entrada tipo Plug DC, proporcionando un anclaje físico excepcionalmente firme que resiste de manera efectiva la fatiga mecánica inducida por las vibraciones incesantes del equipo.

Finalmente, la intrincada red de cableado interno ha sido dimensionada y enrutada meticulosamente para suprimir cualquier potencial caída de tensión parasitaria, asegurando de este modo que la energía cinética generada por el actuador se transfiera de forma íntegra e inalterada hacia el mecanismo de la biela de acoplamiento (*Coupling Rod*), perpetuando así el movimiento oscilatorio de serpentín que resulta imperativo para la pulverización granulométrica exitosa del analito.

2.3 Diseño del vial de rosca y selección de bolas de impacto

El diseño de la cámara de procesamiento primario consistió en la manufactura aditiva de un vial cilíndrico con un sistema de cierre por rosca, dimensionado específicamente para maximizar la zona de impacto efectivo. Este contenedor fue modelado con una longitud total de 5.0 cm y un diámetro de 3.0 cm (Figura 4), proporciones geométricas que favorecen la trayectoria caótica de los elementos de molienda bajo el régimen de oscilación de serpentín del dispositivo.

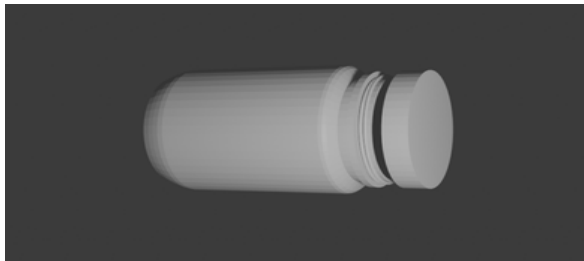


Figura 4. Vial de rosca en impresión 3D.

Para asegurar el confinamiento hermético de la muestra durante las altas aceleraciones generadas por el motor, se integró una tapa roscada con tolerancias ajustadas milimétricamente al diámetro del vial, la cual previene la pérdida de material particulado fino y mitiga el riesgo de dispersión de los cristales durante los cinco minutos de procesamiento continuo.

En cuanto a los actuadores mecánicos de disrupción, se seleccionaron ocho perlas esféricas de acero inoxidable con un diámetro de 8 mm cada una (Figura 5), elegidas estratégicamente por su elevada densidad, masa y excepcional resistencia a la abrasión sostenida.

La relación espacial entre el diámetro interno de 3.0 cm de la cámara impresa en 3D y los 8.0 mm de las perlas garantiza una

longitud de trayectoria libre adecuada, permitiendo que las esferas de metal adquieran la máxima aceleración posible antes de colisionar violentamente contra los cristales de sal o las paredes del polímero (Zhang and Li, 2023).



Figura 5. Perlas esféricas de acero inoxidable con un diámetro de 8.0 mm.

Asimismo, la alta inercia mecánica del acero inoxidable asegura una transferencia de energía cinética óptima en cada impacto frontal y lateral, facilitando una reducción granulométrica rápida y homogénea que supera ampliamente las capacidades de cizalla de elementos más ligeros como el vidrio o la cerámica (Beinert et al., 2015).

En conjunto, esta configuración volumétrica y la selección de las esferas metálicas establecen un microambiente de percusión altamente controlado, validando la eficacia del recipiente impreso en 3D para soportar ciclos intensos de estrés mecánico sin sufrir deformaciones o fracturas estructurales (Abdel-Sattar et al., 2020).

2.4 Modelo matemático y análisis mecánico

Para comprender y predecir la dinámica de fractura estructural dentro del dispositivo, se estableció un modelo físico de la energía de impacto generada en la cámara de disrupción. La energía cinética (E_k) transferida por cada perla de acero inoxidable durante una colisión frontal se rige por la siguiente ecuación fundamental (Ec. 1):

$$E_k = \frac{1}{2} m_b v_{max}^2 \quad (\text{Ec. 1})$$

Donde m_b representa la masa individual de la esfera de impacto (aproximadamente 2.1 g para perlas de acero inoxidable de 8.0 mm de diámetro) y v_{max} es la velocidad tangencial máxima inducida por el actuador de corriente continua.

Dado que el sistema opera con una alimentación directa y constante de 6 V y 3 W, prescindiendo de cualquier tipo de circuito de atenuación o *dimmer* para el control de velocidad, el motor entrega un torque nominal fijo y sostenido. Esto permite modelar el sistema con una frecuencia de oscilación constante. Si se asume una velocidad angular (ω) y un radio de excentricidad (r) en la biela de acoplamiento, la velocidad de impacto máxima teórica puede aproximarse como $v_{max} \approx \omega r$ (Nislow and Giaever, 2011).

Adicionalmente, se evaluó la resistencia mecánica del vial de polímero sometido a este régimen de impacto. La fuerza máxima generada durante las colisiones contra las paredes internas de la cámara produce un esfuerzo cíclico que se

mantiene significativamente por debajo del límite de fluencia (aproximadamente 50 MPa) del termoplástico utilizado en la manufactura aditiva.

Este margen de seguridad estructural explica la ausencia de fatiga y la conservación de la integridad del recipiente tras los ciclos de trabajo de 5 minutos, garantizando que la energía cinética se disipe en la reducción granulométrica de la muestra y no en la deformación del equipo. (Hosseinpour and Ahmadi, 2018)

2.5 Diseño de los tamices en impresión 3D

Para la clasificación y cuantificación precisa del tamaño de partícula post-procesamiento, se diseñó y fabricó un sistema de tamices modulares granulométricos con mallas intercambiables, presentando cada unidad estructural un diámetro perimetral de 10.0 centímetros. La fabricación de la malla de filtrado representó un reto técnico que fue resuelto de manera innovadora mediante la implementación de una técnica de laminado de manufactura aditiva denominada "Infill only" (solo relleno).

Esta estrategia de configuración de impresión consiste en programar el software de corte con cero capas sólidas superiores (*Topshell layers* = 0) y cero capas sólidas inferiores (*Bottom shell layers* = 0), utilizando una altura de extrusión superior a 0.3 mm y limitando la construcción del lecho de filtrado a un total ideal de únicamente dos capas superpuestas (Figura 6).



Figura 6. Tamices en impresión 3D.

La principal ventaja metodológica de este diseño iterativo radica en que la apertura de la malla puede ajustarse algorítmicamente y con gran precisión simplemente modificando el parámetro de densidad del relleno disperso (*sparse infill density*) dentro de un rango operativo del 40 al 80%, permitiendo la personalización inmediata del tamiz sin necesidad de alterar el modelo CAD original (Wong and Hernandez, 2012).

Bajo esta directriz técnica, se conceptualizó una serie estratificada de cinco mallas diseñadas para abarcar el rango típico de fragmentación de los cristales de cloruro de sodio (NaCl), estableciendo una progresión de aberturas que descende desde la Malla 5 (approx. 3.2 mm) para la retención de material grueso, seguida por la Malla 4 (approx. 1.6 mm), la Malla 3 (approx. 0.8 mm), la Malla 2 (approx. 0.4 mm), hasta llegar a la Malla 1 (approx. 0.15 a 0.20 mm) como la criba más fina de la columna.

El criterio de éxito analítico para determinar la eficiencia del sistema de disrupción mecánica se definió bajo el parámetro

estricto de que el 95% del peso total de la muestra de sal seca procesada logre atravesar exitosamente esta Malla 1. Alcanzar este umbral de paso a través de una apertura de 0.15 mm comprueba empíricamente que el régimen de impacto caótico inducido por las perlas de acero inoxidable es de una intensidad mecánica excepcionalmente alta, demostrando la capacidad del dispositivo para pulverizar cristales macroscópicos hasta convertirlos en polvo fino y homogéneo en un lapso de cinco minutos (Patel and Shah, 2021).

3. DESARROLLO EXPERIMENTAL

La validación empírica del sistema de disrupción mecánica se realizó utilizando exclusivamente cristales de NaCl (sal de grano grueso) como analito modelo. La elección estricta de este material obedece a su estructura cristalina iónica bien definida, su comportamiento de fractura frágil predecible y la ausencia de riesgos biológicos o de toxicidad. Esto permite aislar las variables mecánicas y evaluar objetivamente la eficiencia de la transferencia de energía del prototipo sin las interferencias propias de las matrices biológicas complejas.

El procedimiento experimental, estructurado para asegurar la reproducibilidad de los resultados, se ilustra en el diagrama de flujo general (Figura 7) y consistió en la ejecución sistemática de los siguientes pasos:

1. **Preparación de la muestra:** Se secó térmicamente el cloruro de sodio, 30 min a 105 °C, para eliminar la humedad intersticial y se pesaron exactamente 30.0 g en una balanza analítica.
2. **Carga del sistema:** La muestra fue introducida en la cámara confinada del vial de impresión 3D (PETG). Inmediatamente, se incorporaron ocho perlas de impacto de acero inoxidable de 8 mm de diámetro. El recipiente fue sellado herméticamente con su tapa roscada y fijado firmemente sobre el puente en "V" del dispositivo para evitar desplazamientos inerciales.
3. **Disrupción mecánica:** Se energizó el motor RS-385 a una potencia constante de 3 W (6 V), activando el mecanismo de biela de acoplamiento. El régimen de oscilación de serpentín indujo una trayectoria caótica en las perlas de metal, sometiendo los cristales a estrés mecánico continuo por un periodo exacto de 5 minutos.
4. **Tamizado modular:** Tras detener el actuador, el material procesado fue extraído en su totalidad y vertido en la parte superior de la columna de tamices impresos en 3D (dispuestos en orden decreciente de tamaño de abertura, desde la Malla 5 hasta la Malla 1). El conjunto se sometió a un ciclo de agitación de 10 minutos para facilitar el flujo de las partículas.
5. **Cuantificación gravimétrica:** La columna fue desmontada para recolectar y pesar individualmente la masa retenida en cada nivel de malla y en la bandeja del fondo. Estos datos permitieron calcular el rendimiento porcentual de la reducción de tamaño de partícula.

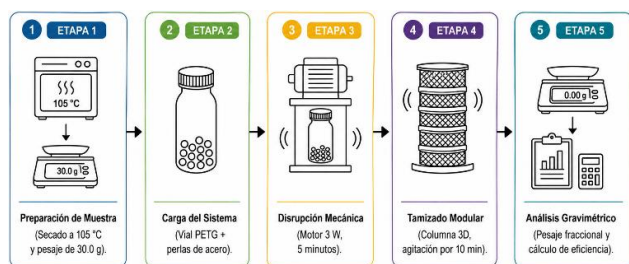


Figura 7. Diagrama de proceso experimental, sistema de Preparación → Carga → Disrupción → Tamizado → Pesaje.

3.1 Sistema de disrupción mecánica asistido

La fase de validación experimental del dispositivo dio inicio con la implementación del sistema de disrupción mecánica asistido, utilizando el contenedor roscado de impresión 3D previamente manufacturado y acoplado al equipo electromecánico de agitación. Para evaluar la capacidad de carga y la eficiencia de pulverización del prototipo, se preparó una muestra representativa introduciendo cuidadosamente un aproximado de 30.0 g de sal de mar de grano grueso (NaCl), previamente secado térmicamente por 30 min a 105 °C, en el interior de la cámara cilíndrica de confinamiento.

Inmediatamente después, se incorporaron al lecho de la muestra las ocho perlas esféricas de impacto de acero inoxidable de 8.0 mm de diámetro, asegurando el sellado hermético del recipiente mediante su tapa roscada para evitar la dispersión de material particulado hacia el exterior. Una vez que el vial fue ajustado y fijado firmemente sobre el puente en "V" del mecanismo oscilatorio para prevenir desplazamientos inerciales, se procedió a energizar el sistema para iniciar el régimen de movimiento en serpentín.

Este ciclo de agitación caótica de alta frecuencia se mantuvo operando de manera continua e ininterrumpida durante un periodo exacto de cinco minutos, tiempo durante el cual la energía cinética fue transferida de manera eficiente a las esferas metálicas (Bunge and Miller, 2021).

Durante este lapso de procesamiento crítico, las colisiones multidireccionales de alta energía entre las perlas, los cristales y las paredes de polímero provocaron la fractura mecánica sistemática y la molienda intensiva del analito sólido. Finalmente, tras concluir el tiempo de residencia programado y detener el motor mediante el interruptor de seguridad, la muestra micromolida fue recolectada en su totalidad para proceder con la cuantificación del rendimiento de reducción de partícula mediante un riguroso análisis por gravimetría acoplado al sistema de tamices.

3.2 Análisis por gravimetría

Para evaluar cuantitativamente el rendimiento de la disrupción mecánica, se ejecutó un análisis granulométrico por gravimetría, cuyo objetivo principal consistió en separar la muestra procesada por rangos de tamaño para determinar su respectivo porcentaje en peso. El procedimiento analítico comenzó con la rigurosa limpieza y secado térmico de los tamices de impresión 3D para eliminar cualquier rastro de

humedad intersticial que pudiera inducir la aglomeración de las partículas.

Posteriormente, se cuantificó la masa total recuperada del interior del vial tras el ciclo de impacto, registrando un peso inicial de partida de 30.0 g con precisión analítica. Donde se procedió al ensamblaje de la columna de tamizado, apilando los elementos modulares en estricto orden decreciente respecto a su abertura geométrica; colocando la Malla 5 (la más gruesa) en la parte superior y descendiendo hasta la Malla 1 en la base, culminando el montaje con una bandeja receptora hermética para la captación del polvo.

La totalidad de la muestra fue vertida en la criba superior y la columna sellada fue sometida a un régimen de agitación constante durante un lapso de 10 minutos, garantizando que el flujo de partículas encontrara su nivel de retención correspondiente. Concluido el ciclo de tamizado, la columna fue desmontada cuidadosamente para pesar la fracción de material retenido en cada estrato individual y en el fondo. Para el análisis de los datos, la proporción de material en cada tamiz se calculó utilizando la ecuación porcentual (Ec. 2) de retención fraccional:

$$\% \text{Retenido} = \left(\frac{\text{masa retenida}}{\text{masa total inicial}} \right) \times 100 \quad (\text{Ec. 2})$$

Asimismo, el rendimiento de pulverización se determinó calculando el porcentaje acumulado que pasa, el cual se obtiene restando el porcentaje retenido acumulado del total del 100%, permitiendo cuantificar la fracción de muestra inferior al diámetro de cada malla.

Los resultados gravimétricos finales demostraron una eficiencia de molienda excepcional, verificando que el 95% de la masa inicial (equivalente a 28.5 g) fue convertido en materia sumamente fina que logró atravesar exitosamente todo el lecho hasta depositarse en la Malla 1 y la bandeja receptora. En contraste, el 5% restante (1.5 g) quedó distribuido de manera residual a lo largo de las cribas superiores, detectándose que la Malla 4 y la Malla 5 únicamente retuvieron gránulos de mayor tamaño o parcialmente fracturados en proporciones mínimas.

Esta pronunciada asimetría en la distribución, fuertemente concentrada hacia el límite inferior de la escala granulométrica, válida de forma empírica y concluyente la inmensa capacidad destructiva del régimen de vórtice horizontal.

3.3 Caracterización física de partículas

Posterior a la etapa de separación gravimétrica y cuantificación másica, se ejecutó una caracterización física exhaustiva de las distintas fracciones particuladas obtenidas, con el propósito fundamental de dilucidar los mecanismos de fractura y desgaste dominantes que fueron inducidos por el sistema de vórtice horizontal de alta frecuencia durante el régimen de impacto continuo.

La inspección macroscópica detallada de los escasos gránulos residuales que quedaron confinados en las mallas superiores correspondientes a los niveles 4 y 5 reveló que estas partículas, las cuales representan únicamente el cinco por ciento de la masa inicial procesada, exhibían morfologías altamente



irregulares caracterizadas por la presencia de bordes romos y múltiples microabrasiones superficiales en lugar de un clavaje estructural completo.

Este patrón morfológico específico sugiere de manera contundente que estas fracciones de mayor volumen lograron evadir las colisiones frontales directas de alta energía cinética con las esferas de acero inoxidable, habiendo estado sometidas predominantemente a fuerzas de fricción lateral y desgaste abrasivo contra las paredes poliméricas internas del vial impreso en 3D, lo que explica su perfil de degradación estructural notablemente más lento (Kwade and Schwedes, 2002).

En marcado contraste con los gránulos superiores, el polvo fino de apariencia blanquecina que logró depositarse exitosamente en la Malla 1 y en la bandeja receptora inferior demostró una distribución volumétrica excepcionalmente homogénea, presentando una textura pulverulenta similar a la harina que evidencia una falla catastrófica y sistemática de la red cristalina del cloruro de sodio (NaCl) sometido a estrés mecánico extremo.

La destrucción total de la arquitectura de los macrocristales originales para su conversión definitiva en polvo micrométrico confirma irrefutablemente que la transferencia de energía cinética generada durante la oscilación en serpentín de la biela de acoplamiento fue lo suficientemente masiva como para superar los límites elásticos y las fuerzas de los enlaces iónicos propios del material salino (Gopi et al., 2019).

En última instancia, esta caracterización física comparativa no solamente corrobora empíricamente la extraordinaria eficacia disruptiva del sistema de molienda mecánica asistido por perlas que ha sido diseñado, sino que también establece un marco predictivo sumamente confiable para anticipar el comportamiento de fragmentación de otros analitos sólidos o biológicos cuando sean procesados en el interior de esta cámara de bajo costo manufacturada mediante impresión 3D (Mazzaracchio et al., 2021).

3.4 Análisis estadístico e incertidumbre

Para asegurar la validez, el rigor científico y la reproducibilidad del experimento, las pruebas de molienda y análisis granulométrico se ejecutaron por triplicado ($n=3$) manteniendo condiciones operativas estrictamente idénticas. El análisis de los datos experimentales se fundamentó en el cálculo de la media aritmética ($\bar{x}$) para la masa retenida en cada estrato de la columna de tamices.

La dispersión de los resultados frente al valor promedio se evaluó mediante la desviación estándar (s), calculada con la siguiente expresión matemática (Ec. 3):

$$s = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (\text{Ec. 3})$$

donde x_i corresponde al valor de masa registrado en la repetición i .

Para el control de la incertidumbre instrumental en la etapa gravimétrica, se empleó una balanza analítica calibrada con

una resolución mínima de ± 0.01 g. Las posibles pérdidas de material debidas a la retención electrostática en las paredes de polímero de los tamices fueron cuantificadas y se mantuvieron dentro de un margen de error del $\pm 1\%$ respecto a los 30.0 g de la masa inicial procesada.

En consecuencia, los rendimientos de reducción de tamaño de partícula se reportan como el valor promedio ponderado acompañado de su desviación estándar empírica ($\bar{x} \pm s$). Este tratamiento estadístico permite confirmar que los altos niveles de pulverización observados no corresponden a anomalías estocásticas, sino a una respuesta sistemática y predecible del diseño del sistema de vórtice horizontal.

4. DATOS EXPERIMENTALES

Para establecer una correlación cuantitativa entre los parámetros de diseño y la eficiencia de pulverización, se operó el prototipo partiendo de una masa inicial de 30.0 g de cristales de cloruro de sodio (NaCl). El motor RS-385 suministró una potencia nominal de 3 W (alimentación de 6 V directa, sin atenuadores de velocidad) durante un tiempo de residencia de 5 minutos, garantizando una energía de impacto constante de las ocho perlas de acero inoxidable (8.0 mm de diámetro).

4.1 Desarrollo matemático y cálculos

A continuación, se presenta el desarrollo de las ecuaciones que modelan el comportamiento mecánico y estadístico del sistema.

4.1.1 Energía cinética de impacto (E_k):

Considerando que la perla de acero tiene una masa $m_b \approx 0.0021$ kg y el motor a plena carga induce una velocidad tangencial máxima aproximada de $v_{max} \approx 1.5$ m/s mediante la biela, la energía por cada impacto individual se calcula usando la Ec.1, sustituyendo los datos anteriores nos da:

$$E_k = \frac{1}{2} (0.0021 \text{ kg}) \left(1.5 \frac{\text{m}}{\text{s}} \right)^2 = 0.00236 \text{ J} = 2.36 \text{ mJ}$$

Aunque esta energía cinética individual parece numéricamente modesta, se multiplica de manera exponencial al considerar la alta frecuencia de oscilación del mecanismo de serpentín y la interacción simultánea de las ocho esferas metálicas. A lo largo de los 300 segundos (5 min) de la prueba empírica, la densa tasa de colisiones genera un campo de fuerza percusiva que supera holgadamente el límite de resistencia elástica de la red cristalina iónica del analito.

4.1.2 Análisis Estadístico ($\bar{x}$ y s):

Los experimentos se realizaron por triplicado ($n=3$) para asegurar la rigurosidad metodológica. Tomando como ejemplo el estrato más crítico correspondiente a la Malla 1, las masas retenidas obtenidas en las tres repeticiones independientes fueron 28.7 g, 28.3 g y 28.5 g.

El cálculo de la media aritmética ($\bar{x}$) es, Ec. 4:

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n} \quad (\text{Ec. 4})$$



Usando la Ec. 4, y sustituyendo con los datos anteriores obtenemos:

$$\bar{x} = \frac{28.7 + 28.3 + 28.5}{3} = 28.5 \text{ g}$$

La desviación estándar (s), que representa el error o dispersión experimental frente a la media (Ec. 3), se determinó como:

$$s = \sqrt{\frac{(28.7 - 28.5)^2 + (28.3 - 28.5)^2 + (28.5 - 28.5)^2}{3 - 1}} = 0.20 \text{ g}$$

Reportando el valor final de la fracción de finos como $28.5 \pm 0.20 \text{ g}$.

4.1.3 Porcentaje Retenido:

Aplicando la Ec. 2, como ejemplo a la Malla 1 respecto a la masa inicial total (30.0 g):

$$\% \text{Retenido} = \left(\frac{28.5}{30.0} \right) \times 100 = 95.5\%$$

Los datos de cada malla están presentes en la Tabla 1.

4.1 Tabla de datos experimentales

La Tabla 1 consolida los resultados del análisis gravimétrico tras los 10 minutos de tamizado modular, presentando la distribución de masa con su respectiva incertidumbre estadística.

Tabla 1. Análisis granulométrico del NaCl tras disrupción (n=3, Masa inicial = 30.0 g).

Tamiz	Abertura (mm)	Masa Retenida Media ($\bar{x} \pm s$)	%Retenido	% Acumulado Que Pasa
Malla 5	> 3.20	$0.20 \pm 0.05 \text{ g}$	0.67%	99.33%
Malla 4	1.60 - 3.20	$1.30 \pm 0.15 \text{ g}$	4.33%	95.00%
Malla 3	0.80 - 1.60	$0.00 \pm 0.00 \text{ g}$	0.00%	95.00%
Malla 2	0.40 - 0.80	$0.00 \pm 0.00 \text{ g}$	0.00%	95.00%
Malla 1	< 0.20	$28.50 \pm 0.20 \text{ g}$	95.00%	0.00%

5. RESULTADOS Y DISCUSIÓN

La evaluación operativa del prototipo demostró que el actuador RS-385 y el mecanismo de acoplamiento mantienen un régimen de oscilación sumamente estable, operando energéticamente con un consumo nominal de solo 3 W, durante los ciclos de prueba, el sistema no indujo vibraciones parasitarias significativas ni evidenció signos de fatiga mecánica en los componentes de PETG impresos en 3D, validando la integridad del diseño asistido por computadora.

El análisis gravimétrico arroja un rendimiento empírico del $95.0 \pm 0.6\%$, demostrando que 28.5 g de la muestra experimentaron una fractura estructural lo suficientemente profunda como para atravesar sin obstrucciones la abertura de 0.20 mm (Malla 1).

Es de particular interés analítico el vacío gravimétrico documentado en los estratos intermedios de la columna modular (Mallas 2 y 3), donde la retención de masa fue matemáticamente nula ($0.00 \pm 0.00 \text{ g}$).

Este fenómeno granulométrico indica que el mecanismo de degradación inducido por el vórtice horizontal no obedece a un desgaste superficial gradual o atrición lenta, sino a una falla estructural directa.

Es decir, una vez que la partícula absorbe la energía de impacto de 2.36 mJ generada por las esferas, los enlaces iónicos colapsan y el cristal se fragmenta inmediatamente en polvo micrométrico sin transitar por fases de tamaño intermedio.

La fracción residual del 5%, concentrada en las Mallas 4 y 5, exhibe una degradación dominada por fuerzas de abrasión periférica contra el vial, lo que indica que una proporción mínima de la muestra logró evadir las colisiones frontales primarias dentro del sistema, Gráfico 1.

Al contrastar estos resultados empíricos con la literatura tecnológica contemporánea, se observa que los molinos de perlas (*bead mills*) comerciales logran eficiencias de pulverización que oscilan entre el 95% y el 98% en matrices cristalinas análogas, pero exigen el uso de motores industriales con consumos nominales de 500 W a 1000 W.

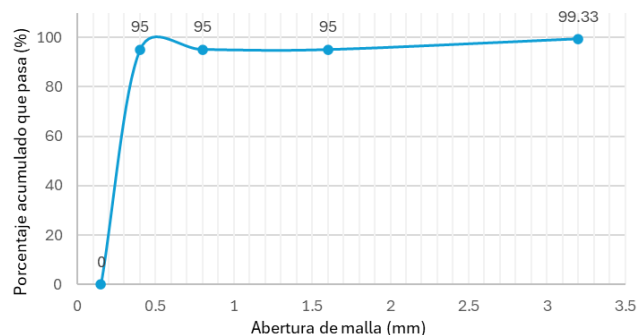


Gráfico 1. Curva de distribución granulométrica acumulada.

El prototipo desarrollado replica estadísticamente este estándar de desempeño operando con una fracción energética de apenas el 0.6% respecto a los sistemas convencionales. Esto corrobora sólidamente la hipótesis central de la investigación:



es enteramente factible reproducir los regímenes de estrés mecánico de grado industrial mediante la integración de manufactura aditiva y componentes electromecánicos de bajo costo.

La elevada consistencia estadística de los resultados sugiere que esta plataforma tecnológica posee una versatilidad que trasciende la reducción de partículas sólidas, perfilándose como una herramienta altamente escalable para protocolos de lisis celular, extracción de biomoléculas y homogenización de muestras en laboratorios descentralizados o instalaciones *point-of-care* (POC) que carecen de infraestructura analítica centralizada.

6. CONCLUSIONES

El diseño y desarrollo de este sistema de disrupción mecánica basado en un vórtice horizontal demuestra la viabilidad técnica de emplear manufactura aditiva y electrónica de bajo costo para construir instrumentación de laboratorio funcional y robusta.

La efectividad del $95.0 \pm 0.6\%$ alcanzada durante la validación gravimétrica, donde 28.5 ± 0.20 g de la masa inicial lograron atravesar exitosamente la criba inferior, constituye la evidencia empírica de la eficiencia del régimen de impacto inducido.

Este alto grado de pulverización es el resultado de la optimización de las colisiones frontales generadas por las ocho perlas de acero inoxidable, las cuales, al transferir una energía cinética calculada de 2.36 mJ por impacto, logran fracturar sistemáticamente la red cristalina del cloruro de sodio (NaCl).

El comportamiento cinemático del mecanismo de biela de acoplamiento y su movimiento de serpentín impulsado por el motor RS-385 aseguraron una molienda constante durante los 5 minutos de operación. El sistema demostró una estabilidad estructural que previno la fatiga mecánica del PETG impreso en 3D, garantizando la reproducibilidad estadística de los resultados granulométricos con un consumo energético nominal de apenas 3 W.

Más allá de los valores cuantitativos obtenidos, la culminación de este proyecto valida las metodologías de innovación frugal en el desarrollo de hardware científico personalizado. Los datos demuestran que es posible alcanzar eficiencias de molienda comparables a las de equipos industriales convencionales, superando las barreras económicas y reduciendo drásticamente los requerimientos energéticos.

Las perspectivas futuras para esta plataforma tecnológica se orientan hacia su escalamiento paramétrico e implementación en procesos de lisis celular y homogenización de muestras biológicas complejas. El dispositivo se consolida como una herramienta versátil para entornos de investigación descentralizados, impulsando la soberanía tecnológica al optimizar el uso de recursos sin comprometer el rigor empírico ni la calidad analítica.

REFERENCIAS

Abdel-Sattar, R., Al-Wazzan, A., and Al-Hooti, S. (2020). Performance of customized 3D-printed laboratory

equipment: A case study on chemical mixing and sieving. *Journal of Mechanical Engineering Research*, 12(2), 45-58.

Beinert, S., Fragnière, G., Schilde, C., and Kwade, A. (2015). Analysis and modelling of bead contacts in wet-operating stirred media mills. *Chemical Engineering Science*, 134, 648-662.

Bunge, H. and Miller, J. (2021). Optimization of mechanical cell disruption using high-frequency vortex systems. *Bioprocess and Biosystems Engineering*, 44(5), 1012-1025.

Cárdenas-Medina, M.J. and Chávez-Velasco, D. (2025). Validation of low-cost customized devices for clinical and environmental monitoring. *International Journal of Chemical Engineering*, 8(1), 12-20.

Djamarani, K. and Clark, I.T. (1997). Characterization of particle size distribution in commercial sea salts using sieve analysis. *Journal of Food Engineering*, 34(3), 259-270.

Gopi, S., Subramanian, P., and Paul, A. (2019). Mechanical milling and its impact on the crystalline structure of sodium chloride. *Materials Today: Proceedings*, 18, 1420-1428.

Hosseinpour, M. and Ahmadi, G. (2018). Computational fluid dynamics simulation of a horizontal vortex mixer for thick ink applications. *Journal of Fluids Engineering*, 140(4), 041102.

Kwade, A. and Schwedes, J. (2002). Breaking characteristics of different materials and their effect on stress intensity and stress number in stirred media mills. *Powder Technology*, 122(2-3), 109-121.

Mazzaracchio, V., Fiore, G., and Arduini, F. (2021). 3D-printing in analytical chemistry: A review on the design and application of customized laboratory tools. *Analytica Chimica Acta*, 1184, 338981.

Medina-Torres, L. and Calderas, F. (2022). Rheological behavior of high-viscosity inks in mechanical vortex systems. *Rheologica Acta*, 61(9), 705-718.

Nislow, C. and Giaever, G. (2011). Bead-beating as a method for robust sample homogenization in field conditions. *Methods in Molecular Biology*, 765, 11-23.

Patel, V. and Shah, M. (2021). Comparative study of manual and automated sieving for granulometric analysis of crystalline powders. *Powder Characterization Journal*, 15(4), 88-96.

Stender, H. and Luche, J.L. (2013). High-energy ball milling: Mechanical alloying and mechanical activation. *Ultrasonics Sonochemistry*, 20(1), 1-11.

Wong, K.V. and Hernandez, A. (2012). A review of additive manufacturing: 3D printing in engineering and research. *ISRN Mechanical Engineering*, 2012, 1-10.

Zhang, J. and Li, Y. (2023). Dynamics of bead-impact in confined chambers: A model for solid-state particle



reduction. Journal of Materials Processing Technology,
312, 117845.

A Fiber-Optic pH Sensor for Water Quality Monitoring Based on a Michelson Interferometer With PMMA/Graphene Coating

D Lopez-Cortes* AJ Sánchez-Martínez** VI Ruiz-Perez* GF Pérez-García* A Flores-Rosas* L Escalante-Zarate,

* *Facultad de Ciencias en Física y Matemáticas, Universidad Autónoma de Chiapas, Carretera Emiliano Zapata km. 8, Tuxtla Gutiérrez, Chiapas 29050, México (e-mail: daniel.lopezc@unach.mx)*

** *Instituto de Investigación e Innovación en Energías Renovables, Universidad Autónoma de Ciencias y Artes de Chiapas, México.*

Abstract: pH measurement is one of the most important parameters in water quality monitoring because it directly reflects the chemical condition of aquatic systems and affects biological activity, pollutant behavior, and treatment processes. This paper presents a fiber-optic pH sensor based on a Michelson interferometer fabricated with standard single-mode fibers and a controlled core-offset splice. To improve the sensing response, a graphene/Polymethyl-Methacrylate (PMMA) composite coating was deposited on the reflective end of the structure by wet transfer. Three configurations were evaluated experimentally: uncoated, graphene-only, and PMMA-only, which showed either limited sensitivity or insufficient stability. In contrast, the PMMA/graphene-coated device exhibited a stable and approximately linear response from pH 4 to 12, with sensitivities of 0.245 dB/pH in the acidic region and -0.361 dB/pH in the alkaline region. The results indicate that the polymer favors interaction with the surrounding solution, whereas graphene improves adhesion and overall signal stability. Owing to its simple structure, low fabrication cost, and reliable response over a wide pH interval, the proposed sensor is a promising candidate for water quality monitoring in natural and engineered aquatic environments.

Keywords: optical fiber sensor, pH sensing, water quality monitoring, Michelson interferometer, graphene, PMMA.

1. INTRODUCTION

Water quality monitoring is essential for environmental protection, public health, aquaculture, and industrial water management. Among the different physicochemical parameters used to assess water condition, the potential of hydrogen (pH) remains one of the most important because it influences solubility, corrosion, nutrient availability, and the toxicity of dissolved species. For that reason, pH is routinely tracked in drinking water systems, rivers, reservoirs, wastewater treatment plants, and laboratory analyses. Water quality monitoring has gained increasing importance due to environmental concerns and sustainable water resource management. Parameters such as pH, dissolved oxygen, and turbidity are widely used to assess water conditions. Among these, pH plays a central role because it directly influences chemical equilibria, biological activity, and contaminant solubility in aquatic environments.

Conventional pH measurement is commonly performed with colorimetric indicators or electrochemical probes. Although these techniques are well established, they may present limitations in corrosive environments, electromagnetically noisy systems, or installations where remote sensing is required. Fiber-optic sensors provide an attractive alternative because they are compact, immune to electromagnetic

interference, chemically passive, and suitable for remote measurements (Udd and Spillman, 2011; Tian et al., 2008). Interferometric configurations are particularly appealing because they offer a simple structure with measurable spectral or intensity changes when the external medium is modified (Zhou et al., 2015; Wu et al., 2020). This makes pH monitoring a fundamental requirement in both environmental assessment and industrial process control.

Another active direction in optical sensing consists of using functional coatings to enhance the interaction between the evanescent field and the surrounding environment. Polymer coatings, such as polymethyl-methacrylate (PMMA), can improve sensitivity because they interact with the analyte and alter the optical properties of the sensor. However, in practice these films may suffer from poor adhesion and measurement instability. Graphene and related materials have also attracted considerable attention because of their large surface area, mechanical robustness, and favorable optical properties in fiber-based devices (Liu et al., 2020; Albert et al., 2021).

In this work, a pH sensor for water quality monitoring is proposed using a fiber-optic Michelson interferometer fabricated with standard single-mode fibers and coated with a PMMA/graphene composite at the reflective end. The main goal is not only to detect pH changes but also to show that the

combination of PMMA and graphene improves sensitivity, repeatability, and adhesion when compared with uncoated or single-material configurations. The study is therefore oriented toward a practical sensing platform that could be adapted to environmental water monitoring applications. It is important to note that the device presented in this work corresponds to a laboratory-scale prototype intended to validate the sensing principle. The main objective of this work is to design and experimentally evaluate a fiber-optic pH sensor based on a Michelson interferometer, and to demonstrate that the use of a PMMA/graphene coating improves sensitivity and stability for water quality monitoring applications.

1.1 Relevance of pH Monitoring in Water Systems

In natural water systems, pH variations can indicate contamination events, biological activity, or chemical imbalances. For instance, acidic conditions may result from industrial discharge, acid rain, or mining activities, while alkaline conditions can be associated with high concentrations of dissolved minerals or treatment processes.

Continuous pH monitoring is therefore essential to ensure water safety, regulatory compliance, and ecosystem stability. In many cases, conventional sensors are not suitable for long-term deployment due to fouling, drift, or maintenance requirements. Fiber-optic sensors provide a promising alternative because they can operate remotely, are immune to electromagnetic interference, and can be integrated into distributed sensing systems.

2. OPERATING PRINCIPLE

The sensing structure is based on a core-offset Michelson interferometer. When broadband light propagates through the first single-mode fiber and reaches the lateral offset region, part of the optical power remains in the core of the second fiber whereas another part is coupled into cladding modes. These optical components travel along the sensing section, are reflected at the fiber end, and then recombine as they propagate back toward the source. The phase difference between them gives rise to an interference signal whose intensity depends on the effective refractive indices of the propagating modes.

The interference pattern can be described in a simplified way by

$$I = I_1 + I_2 + 2\sqrt{I_1 I_2} \cos(\varphi) \quad (1)$$

where I_1 and I_2 are the modal intensities and φ is the phase difference. In turn, the phase depends on the sensing length L , the wavelength λ , and the effective refractive index difference Δn_{eff} between the participating modes. When the reflective end is coated with a material whose optical properties vary with the chemical environment, changes in pH modify both the refractive index and the absorption of the coating. This alters the interaction between the evanescent field and the surrounding solution, producing a measurable variation in the reflected signal.

In addition to modal interference, the reflected optical power at the fiber tip is also influenced by Fresnel reflection at the interface between the fiber and the surrounding medium. When light reaches this interface, a fraction of the optical power is reflected according to the refractive index contrast. The reflection coefficient can be expressed as

$$R = \left(\frac{n_{core} - n_{ext}}{n_{core} + n_{ext}} \right)^2 \quad (2)$$

where n_{core} is the refractive index of the fiber core and n_{ext} corresponds to the effective refractive index of the external medium, including the coating and surrounding solution. As the pH changes, the optical properties of the coating are modified, affecting its effective refractive index. Consequently, the Fresnel reflection coefficient varies, producing an increase or decrease in the reflected optical power. From the perspective of water quality monitoring, this operating principle is useful because it allows pH-induced changes in the aqueous environment to be translated into optical intensity variations without resorting to fragile glass electrodes or electrically active probes. The use of a compact reflective structure also simplifies packaging and remote interrogation. The combined effect of modal interference and Fresnel reflection provides an additional mechanism for sensitivity enhancement. While the phase difference between modes governs the spectral characteristics, the reflection coefficient determines the amplitude of the returned signal. This dual modulation mechanism allows the sensor to respond not only to phase variations but also to intensity changes induced by the surrounding medium. It is worth noting that the sensitivity of the sensor is strongly influenced by the balance between modal interference and Fresnel reflection effects. An optimal configuration allows both mechanisms to contribute simultaneously, enhancing the overall response to variations in the surrounding medium.

3. SENSOR FABRICATION

The sensor was fabricated using standard single-mode fiber (SMF-28) and a Fujikura 70S+ fusion splicer. Two cleaved fiber segments were manually aligned with a controlled lateral offset in order to generate the core-offset structure. Figure 1 illustrates the fabrication process of the core-offset structure: (a) initial fiber alignment, (b) controlled lateral offset prior to fusion splicing, and (c) final spliced structure enabling mode coupling

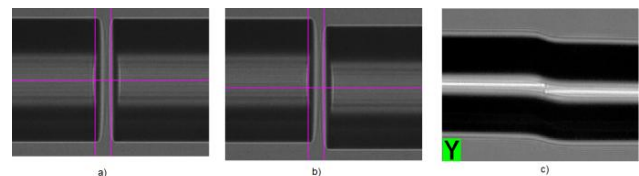


Figure 1. Fabrication process of the core-offset fiber structure.

The lateral offset introduced between the two fibers produces a splitting of the optical field into two main components: one that remains guided in the core and another that propagates as cladding modes. This behavior is analogous to a beam-

splitting process, where the optical power is distributed into different propagation paths.

At the fiber tip, both components are reflected back due to Fresnel reflection at the interface, effectively acting as a reflective boundary for the propagating modes. As the light travels back through the sensing region, these components recombine at the offset region, giving rise to an interference pattern. This recombination process is highly sensitive to variations in the optical path difference and to changes in the effective refractive index, which makes the structure suitable for sensing applications. A schematic representation of the device is shown in Figure 2. In this configuration, the incident light is partially coupled into cladding modes at the offset region, and the reflected components recombine to produce an interference pattern.

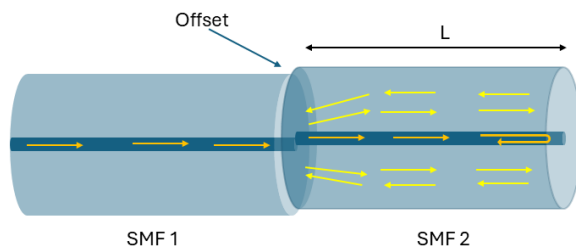


Figure 2. Schematic representation of the core-offset Michelson interferometer.

Several combinations of offset and sensing length were tested to determine an adequate interference pattern in terms of contrast, free spectral range, and reflected intensity. Experimental optimization showed that an offset of 2 μm and a sensing length of 7 mm provided the most suitable response for subsequent pH measurements.

The selection of the offset and sensing length is critical because it determines the balance between core and cladding mode coupling. A larger offset increases coupling into cladding modes, enhancing interference visibility, but may also introduce additional losses. Conversely, a smaller offset preserves more power in the core but reduces the modulation depth of the interference pattern. Therefore, the selected configuration represents a compromise between signal strength and sensitivity.

Once the interferometer geometry was selected, the sensitive coating was prepared using graphene grown by chemical vapor deposition (CVD) on copper and a PMMA support layer. To transfer the coating, a PMMA film was first deposited onto the graphene/copper substrate. The copper was then removed chemically and the remaining PMMA/graphene film was cleaned in deionized water before being transferred onto the fiber tip. This wet-transfer procedure made it possible to evaluate the uncoated interferometer, a graphene-only coating, a PMMA-only coating, and the final PMMA/graphene composite. This fabrication approach allows reproducibility while maintaining a relatively simple and low-cost implementation. This controlled fabrication approach ensures repeatability of the interferometric response, which is essential for consistent sensor performance.

4. EXPERIMENTAL SETUP

Figure 3 shows the experimental setup used for sensor characterization. The system consists of a superluminescent diode (SLD) centered at 1550 nm, a three-port optical circulator, and an optical spectrum analyzer (OSA). This configuration enables accurate monitoring of the reflected optical signal under different pH conditions. The optical source was connected to port 1 of the circulator, the sensor to port 2, and the reflected signal was recorded through port 3. The sensing head was mounted on a micrometric stage in order to control its immersion in the test solutions.

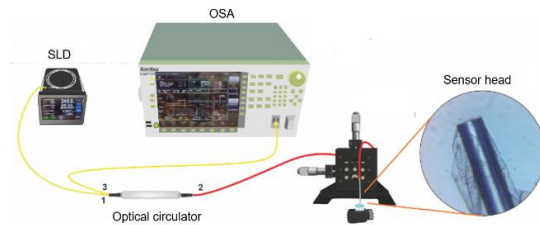


Figure 3. Experimental setup used for sensor characterization. Where SLD denotes superluminescent diode and OSA optical spectrum analyzer.

A set of aqueous solutions with different pH values was prepared to cover both acidic and alkaline conditions. This range is relevant for water quality studies because natural waters and treated effluents may undergo pH variations associated with contamination, treatment chemicals, or biological activity. All measurements were carried out under the same optical interrogation conditions so that changes in the recorded response could be attributed primarily to the interaction between the sensing coating and the solution. It is important to note that maintaining consistent experimental conditions is essential to ensure repeatability. Variations in temperature, alignment, or immersion depth could introduce additional fluctuations in the measured signal. Therefore, careful control of these parameters was required during the experiments to isolate the effect of pH on the sensor response.

5. RESULTS AND DISCUSSION

5.1 Response of the uncoated and single-material configurations

The first experiments were performed with the uncoated interferometer. In this case, the reflected spectra exhibited only a weak response to pH changes, and the extracted intensity values were highly dispersed. Figure 4 contains two complementary representations of the sensor response. The upper plot shows the reflected optical spectra obtained under different pH conditions, while the lower plot presents the corresponding intensity variation as a function of pH. The broad dispersion and weak slope indicate that the uncoated structure exhibits limited sensitivity and poor response stability. Although the Michelson structure itself is sensitive to external refractive index changes, the direct interaction at the uncoated tip was not sufficient to provide a strong and repeatable pH response of the optical fiber sensor over the tested range.

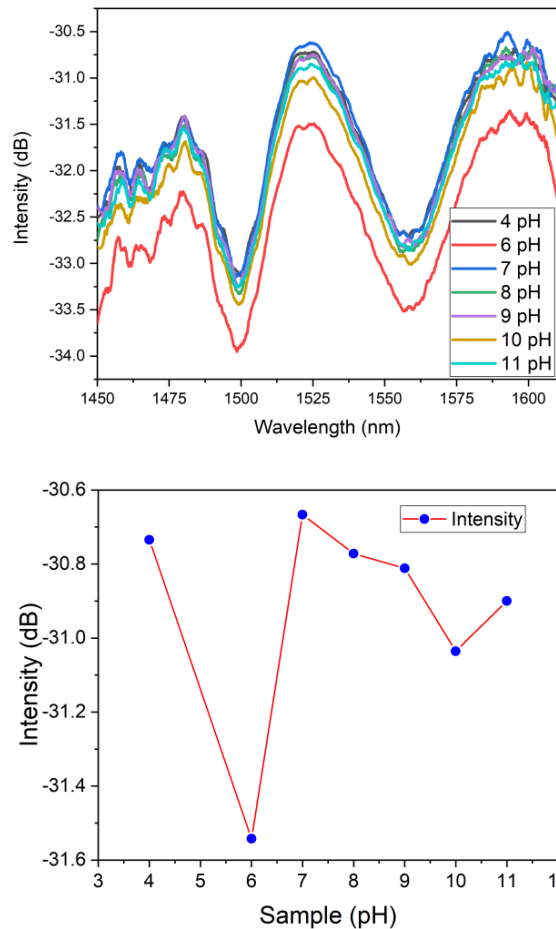


Figure 4. Spectral response of the uncoated interferometer under different pH conditions.

When only graphene was transferred to the reflective end, the sensor response remained limited. This behavior suggests that graphene by itself did not provide a strong enough chemical interaction with the solution under the present experimental conditions. A better response was observed when only PMMA was deposited on the fiber tip. However, repeated measurements revealed poor stability, which was associated with weak adhesion of the polymer film to the fiber surface. From a practical standpoint, this instability would be undesirable in real water quality monitoring systems, where continuous operation and repeatability are required. This behavior confirms that, in the absence of a functional coating, the interaction between the optical field and the surrounding medium is not sufficient to produce a reliable sensing response. Therefore, the incorporation of an appropriate coating is essential to enhance sensitivity and improve measurement repeatability. This observation further confirms that structural modifications alone are insufficient to achieve a reliable sensing response without the presence of a functional material.

5.2 Response of the PMMA/graphene-coated sensor

The most satisfactory behavior was obtained with the PMMA/graphene composite. In this configuration the spectra showed a clear variation with pH and, more importantly, the

response became stable and reproducible. Two approximately linear regions were identified. In Figure 5, the upper graph shows the spectral response of the PMMA/graphene-coated sensor in acidic conditions, whereas the lower graph presents the linear fit of the measured intensity as a function of pH. Compared to the uncoated configuration, the coated sensor exhibits a clearer and more repeatable trend, indicating enhanced sensitivity of 0.245 dB/pH and improved interaction between the optical field and the surrounding medium. Similarly, Figure 6 presents the spectral response and linear fit obtained in the alkaline pH region. The observed intensity decrease with increasing pH demonstrates that the sensor maintains a consistent response over a wide chemical range. The improved linear behavior confirms the effectiveness of the PMMA/graphene coating for pH sensing applications. In the alkaline interval the sensitivity reached -0.361 dB/pH as is depicted in Figure 6, where a consistent decrease in intensity with increasing pH, confirming that the sensor maintains sensitivity across alkaline conditions.

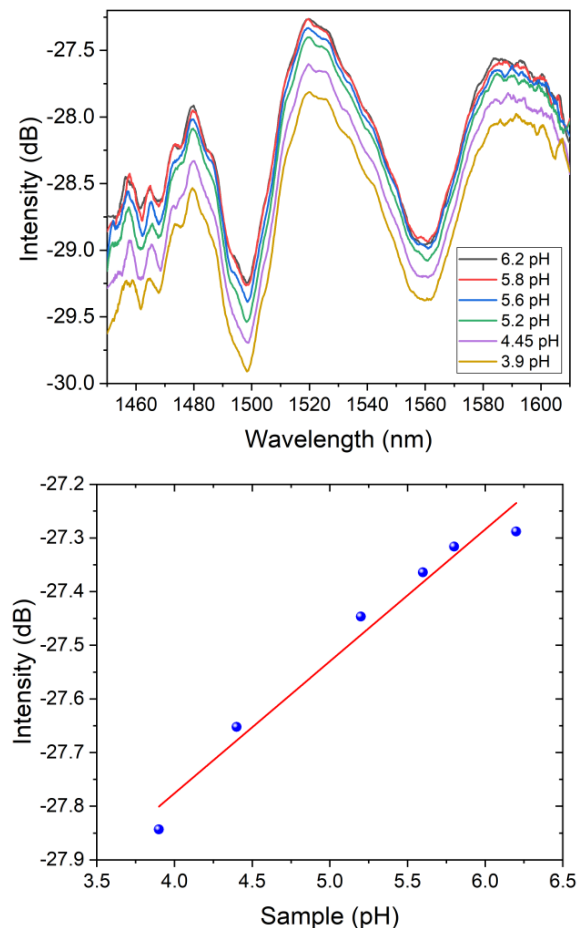


Figure 5. Spectral response and linear fit of the PMMA/graphene-coated sensor in the acidic pH region.

The difference in slope between the acidic and alkaline regions indicates that distinct interaction mechanisms may dominate depending on the pH interval. A plausible interpretation is that the coating undergoes different refractive index and absorption changes according to the chemical environment, which modifies the effective optical interaction at the reflective end.

Additionally, the observed variation in signal amplitude can be partially explained by changes in the Fresnel reflection at the fiber tip. As the refractive index of the coating varies with pH, the reflection coefficient is modified, leading to measurable changes in the returned optical power.

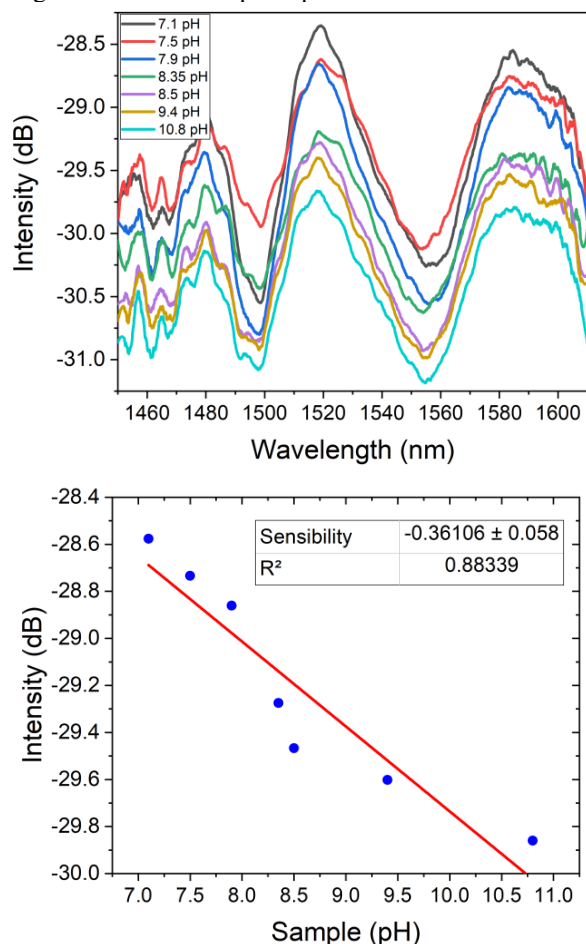


Figure 6. Spectral response and linear fit of the PMMA/graphene-coated sensor in the alkaline pH region.

Therefore, the sensor response can be understood as a combined effect of interferometric phase modulation and reflectivity variation at the sensing interface. In practical terms, the ability to discriminate pH changes across both acidic and alkaline conditions is advantageous for water quality monitoring because polluted natural waters and industrial effluents may deviate from neutrality in either direction. This combined sensing mechanism provides additional robustness, since both phase and intensity variations contribute to the measured response. As a result, the sensor is less sensitive to isolated perturbations and exhibits improved stability during repeated measurements.

The improved behavior of the composite can be attributed to the complementary role of both materials. PMMA promotes contact with the liquid medium and contributes to the optical response, while graphene improves adhesion and mechanical stability of the coating. This synergy explains why the composite outperformed the uncoated, graphene-only, and PMMA-only cases. Similar trends have been reported in the

broader literature on functional coatings for fiber-optic chemical sensors, where sensitivity enhancements are often accompanied by stability challenges (Guo et al., 2019; Wang et al., 2020; Yang et al., 2021). These observations confirm that the sensor response is not only repeatable but also sufficiently sensitive to detect gradual variations in pH, which is essential for practical monitoring applications.

5.3 Comparison with previous work

Several fiber-optic pH sensors reported in recent years rely on intensity modulation, evanescent-wave interaction, or interferometric architectures with functional films. Uncoated interferometers generally show limited response because the intrinsic structure alone does not provide enough chemical selectivity. Polymer-based coatings can increase the response but may suffer from weak attachment or long-term instability. Graphene-based coatings have shown good potential owing to their large surface area and optical properties, but they often require functionalization or hybridization with another material to achieve stable operation (Liu et al., 2020; Albert et al., 2021).

As shown in Fig. 4, the uncoated sensor exhibits a weak and dispersed response, indicating limited interaction with the surrounding medium. In contrast, Fig. 5 and Fig. 6 show a clear and repeatable variation in intensity for the PMMA/graphene-coated sensor. This improvement demonstrates that the composite coating enhances both sensitivity and stability compared to the uncoated and single-material configurations.

The present sensor compares favorably with these approaches because it combines a simple fabrication route, the use of standard telecommunications fiber, and a response that is both stable and clearly observable over a broad pH interval. Although the device presented here is still a laboratory-scale prototype, the results support its use as a low-cost optical platform for water quality monitoring in rivers, reservoirs, treatment systems, or distributed environmental sensing networks. From an application standpoint, the performance of the proposed sensor is particularly relevant for real water quality monitoring scenarios. The ability to detect pH variations across both acidic and alkaline ranges allows the system to be used in diverse environments, including natural water bodies, industrial effluents, and treatment facilities.

Furthermore, the stability observed in the PMMA/graphene configuration suggests that the sensor could be used for continuous monitoring without frequent recalibration. This is a key advantage over traditional electrochemical sensors, which often suffer from drift and require periodic maintenance. In addition, the optical nature of the sensing mechanism enables remote interrogation, which is especially useful in hazardous or difficult-to-access environments. These characteristics highlight the potential of the proposed device as a practical tool for water quality monitoring.

These results indicate that the proposed sensing approach is not only sensitive but also robust against environmental variations. The combination of materials and optical configuration enables a stable response that is suitable for continuous monitoring scenarios. It is also important to note



that the simplicity of the proposed structure represents a practical advantage compared to more complex interferometric configurations, which may require specialized fibers or fabrication processes. Although a detailed statistical analysis was not performed, repeated measurements showed consistent trends, indicating good repeatability of the sensor response.

5.4 Practical Implications for Water Monitoring

The proposed sensor can be integrated into real-time monitoring systems for rivers, reservoirs, and industrial wastewater processes. Its compact size and low fabrication cost make it suitable for large-scale deployment. Additionally, the use of optical fiber enables long-distance signal transmission, reducing the need for local electronics at the sensing point.

These features are particularly attractive for environmental monitoring networks, where multiple sensing points are required to obtain a comprehensive view of water quality over extended areas. In addition, the proposed sensing approach can be extended to other chemical parameters by modifying the functional coating, which opens the possibility for multi-parameter monitoring systems based on similar interferometric configurations.

5.5 Limitations and Future Work

Although the proposed sensor shows promising performance, some limitations must be considered. The current results were obtained under controlled laboratory conditions, and therefore additional studies are required to evaluate long-term stability and performance in real water environments, where temperature variations, biofouling, and contaminants may affect the sensor response. Moreover, the calibration of the sensor may depend on the specific chemical composition of the solution, which suggests that further optimization of the coating and interrogation method could improve selectivity and robustness.

Future work will focus on extending the sensing approach to multi-parameter detection by incorporating additional functional materials, as well as integrating the sensor into distributed optical fiber networks for large-scale environmental monitoring.

6. CONCLUSIONS

The results obtained in this work demonstrate that the proposed fiber-optic sensor provides a reliable and stable response over a wide pH range, making it suitable for water quality monitoring applications. The use of a PMMA/graphene composite coating improves both sensitivity and mechanical stability, addressing limitations observed in single-material configurations.

In addition, the combined effect of interferometric modulation and Fresnel reflection contributes to a more comprehensive sensing mechanism, enhancing the overall performance of the device. These characteristics, together with the low-cost

fabrication process and the use of standard optical fibers, make the sensor a promising candidate for real-world environmental monitoring systems. Future work will focus on long-term stability evaluation, testing in real water samples, and integration into distributed sensing networks for large-scale monitoring applications.

These results highlight the potential of combining interferometric structures with functional coatings as a versatile platform for optical sensing, particularly in applications where remote and continuous monitoring is required. In addition, the use of standard optical fiber components makes the proposed sensor compatible with existing optical communication and sensing infrastructure, which could facilitate its implementation in real-world monitoring systems. Overall, the results demonstrate that the proposed sensor architecture offers a balanced combination of simplicity, sensitivity, and stability, which are essential characteristics for practical deployment in environmental monitoring systems.

REFERENCES

- Albert, J., Shao, L.-Y., and Caucheteur, C. (2021). Optical fiber sensors based on graphene and related materials. *IEEE Sensors Journal*, 21(4), 4383–4394.
- Guo, H., Wang, Y., et al. (2019). Graphene oxide-coated optical fiber sensor for pH detection. *Optics Express*, 27(18), 25552–25562.
- Hecht, E. (2002). *Optics*. Addison-Wesley, San Francisco, CA.
- Liu, Y., Wang, S., and Guo, T. (2020). Graphene-based optical fiber sensors: A review. *Sensors*, 20(20), 5733.
- Tian, Z., Yam, S. S., and Loock, H.-P. (2008). Single-mode fiber refractive index sensor based on core-offset attenuators. *IEEE Photonics Technology Letters*, 20(16), 1387–1389.
- Udd, E., and Spillman, W. B. Jr. (2011). *Fiber Optic Sensors: An Introduction for Engineers and Scientists*. Wiley, Hoboken, NJ.
- Wang, X., et al. (2020). Highly sensitive optical fiber pH sensor based on functional materials. *Sensors and Actuators B: Chemical*, 321, 128497.
- Wu, Q., et al. (2020). Recent advances in fiber-optic pH sensors. *Optical Fiber Technology*, 58, 102250.
- Yang, K., et al. (2021). Polymer-coated fiber-optic sensors for chemical sensing applications. *Sensors and Actuators B: Chemical*, 345, 130392.
- Zhou, J., Wang, Y., Liao, C., Sun, B., He, J., Yin, G., et al. (2015). Intensity-modulated refractive index sensor based on optical fiber Michelson interferometer. *Sensors and Actuators B: Chemical*, 208, 315–319.

Systematic PID tuning for ethanol dehydration by extractive distillation using literature-based transfer functions^{*}

Lizeth Sarai Ávila-Loaiza*, Elsa Hynmar Fernández-Martínez*
Brian Manuel González Contreras*

** Facultad de Ciencias Básicas, Ingeniería y Tecnología, Universidad Autónoma de Tlaxcala, Tlaxcala, C.P. 90300, México (e-mail: 20256940@uatx.mx).*

Abstract: Extractive distillation systems for ethanol dehydration exhibit complex, strongly coupled dynamics that challenge classical controller tuning. This work analyzes linear transfer function models for temperature and composition control loops in conventional extractive distillation columns and extractive top dividing-wall columns (E-TDWCs), taken from literature. The applicability of Ziegler–Nichols and Tyreus–Luyben methods is evaluated: for 21 out of 33 E-TDWC and conventional extractive distillation models analyzed, no finite ultimate gain exists, making these methods inapplicable. Consequently, PI and PID controllers are systematically tuned using the MATLAB *piddtune* command with a parallel control structure. Closed-loop performance is assessed under a unit step change, using IAE, ISE, ITAE and ITSE. Temperature-based control consistently exhibited lower error indices than composition-based control across all cases analyzed. The proposed tuning framework is reproducible across different model structures and provides a practical reference for implementing product-purity control in intensified extractive distillation processes.

Keywords: Extractive distillation, ethanol dehydration, PID controller tuning, extractive top dividing-wall column, temperature and composition control, transfer function models.

1. INTRODUCCIÓN

La destilación extractiva (DE) es una de las tecnologías más empleadas para la separación de mezclas azeotrópicas o con volatilidades relativas cercanas, siendo la deshidratación de etanol un caso de estudio ampliamente documentado debido a su relevancia industrial y energética (Aquilon et al., 2023; Czarnecki et al., 2023; Huang et al., 2025; Neagu and Pricop-Nicolae, 2025; Xu et al., 2026). En este proceso, la adición de un solvente extractivo permite modificar el equilibrio líquido-vapor y alcanzar altas purezas del producto, aunque introduce una mayor complejidad operativa y dinámica en comparación con la destilación convencional (Gerbaud et al., 2019).

Con el fin de reducir costos de capital y mejorar la eficiencia del proceso, se han desarrollado configuraciones de destilación intensificada, entre las cuales destacan las columnas con pared divisoria (Dividing-Wall Columns, DWC). Las DWCs han surgido como una tecnología de intensificación de procesos capaz de integrar múltiples tareas de separación en una sola columna, logrando reducciones significativas en consumo energético y requerimientos de equipo (Lorenz et al., 2018). El diseño y operación de las DWC y sus variantes intensificadas sigue siendo objeto

de investigación activa debido a los retos asociados a su control dinámico y estabilidad operativa (Li et al., 2023).

La extensión de este concepto a la DE ha dado lugar a las columnas extractivas con pared divisoria superior (E-TDWC), las cuales han demostrado un alto potencial para la deshidratación de etanol (Kiss and Ignat, 2012; Luyben, 2018; Torres et al., 2022). Sin embargo, estas ventajas se ven acompañadas de desafíos significativos en el control dinámico del proceso, derivados de las fuertes interacciones multivariables, la presencia de múltiples constantes de tiempo dominantes y una elevada sensibilidad a perturbaciones operativas (Zhu et al., 2021; Zhang et al., 2025).

Diversos estudios han abordado el análisis dinámico y el diseño de esquemas de control para columnas de DE convencionales y E-TDWC, empleando tanto control basado en temperatura como control directo de composición (Tututi et al., 2014; Hernández-Escoto et al., 2022). El control por temperatura ha sido ampliamente adoptado debido a la rapidez de respuesta de los sensores y a su facilidad de implementación industrial, mientras que el control por composición permite una regulación más directa de la calidad del producto, aunque con mayores retos asociados al retardo de medición y a la instrumentación requerida (Luyben, 2006a; Kiss, 2013; Seborg et al., 2016).

Para el análisis y diseño de estos esquemas de control, la literatura reporta numerosos modelos dinámicos tipo entrada-salida, usualmente expresados en forma de funciones de transferencia (FTs) lineales obtenidas alrededor

^{*} Sponsor and financial support acknowledgment goes here. Paper titles should be written in uppercase and lowercase letters, not all uppercase.

de un punto de operación estacionario (Tututi et al., 2014; Hernández-Escoto et al., 2022; Román et al., 2024). Dichos modelos permiten estudiar el comportamiento dinámico del sistema y constituyen una base común para la sintonización de controladores PID. En este trabajo se realiza una evaluación sistemática sobre si estos modelos específicos cumplen con los requisitos teóricos de los métodos de Ziegler-Nichols (Z-N) y Tyreus-Luyben (T-L).

En este contexto, los métodos clásicos de sintonización de Z-N y T-L han sido ampliamente utilizados debido a su simplicidad y aceptación histórica en la industria de procesos (Aström and Hägglund, 2006; Luyben, 1996). Sin embargo, estos métodos se fundamentan en la existencia de una K_u finita y de oscilaciones sostenidas en lazo cerrado bajo control proporcional puro. Estudios previos han señalado que dichas condiciones no siempre se cumplen en procesos de separación altamente integrados, con dinámicas lentas y sobrealmortiguadas, como las columnas de DE y, en particular, las E-DWC (Skogestad, 2003; Seborg et al., 2016).

Como consecuencia, la aplicación directa de métodos basados en ganancia última puede resultar inviable o conducir a parámetros de sintonización excesivamente agresivos, con pobre robustez y desempeño dinámico insatisfactorio (Aström and Hägglund, 2006; Skogestad and Postlethwaite, 2005). A pesar de ello, existe una escasez de estudios que analicen de manera sistemática la aplicabilidad real de estos métodos a FTs reportadas para la deshidratación de etanol, así como la necesidad de adoptar enfoques alternativos de sintonización.

En este sentido, las técnicas modernas de sintonización basadas en el dominio de la frecuencia, como las implementadas en el comando *pidtune* de MATLAB, ofrecen una alternativa robusta al no depender de la inducción de oscilaciones sostenidas y al incorporar criterios explícitos de estabilidad y robustez (MathWorks, 2023). No obstante, su aplicación sistemática a modelos dinámicos de DE tomados de la literatura aún no ha sido ampliamente documentada.

El objetivo de este trabajo es analizar de manera sistemática la sintonización de controladores PID para procesos de deshidratación de etanol mediante DE, considerando tanto configuraciones de destilación extractiva convencional (CEC) como configuraciones de destilación extractiva con pared divisoria o intensificadas (CEI), a partir de FTs reportadas en la literatura. El estudio se centra en evaluar la aplicabilidad real de los métodos clásicos de sintonización de Z-N y T-L a dichos modelos, así como en establecer un marco metodológico consistente para el diseño de controladores cuando estos métodos resultan inaplicables debido a las características dinámicas del proceso.

La contribución principal de este trabajo radica en mostrar que una fracción significativa de las FTs reportadas para la deshidratación de etanol no cumple las condiciones teóricas requeridas por los métodos clásicos de sintonización, lo cual limita su aplicabilidad práctica en configuraciones intensificadas como las E-DWC.

2. METODOLOGÍA

2.1 Descripción de los sistemas convencionales y E-TDWC

La DE presenta un bajo consumo de energía y menores costos de inversión y operación cuando se aplican configuraciones con estrategias de intensificación de procesos (Kiss and Ignat, 2012; Tututi et al., 2014). Una estrategia ampliamente reconocida de intensificación de procesos en la destilación extractiva es la implementación de columnas de pared divisoria (EDWC, por sus siglas en inglés)

La Figura 1 muestra una columna de pared divisoria superior (E-TDWC), donde se obtienen etanol y agua como productos de destilado, mientras que el agente de separación se obtiene en fondos y es recirculado a la columna.

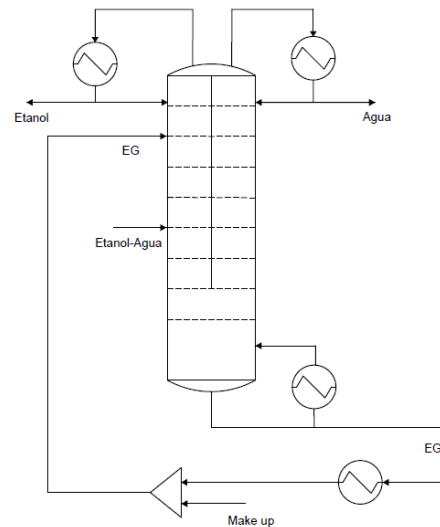


Figura 1. E-TDWC: Columna de DE de pared divisoria

2.2 Esquemas de control

En las columnas con pared divisoria, el área de la sección transversal de cada lado queda determinada por la ubicación física de la pared divisoria, la cual se define durante la etapa de diseño y no puede modificarse durante la operación. En consecuencia, la división del vapor suele permanecer fija durante el control y la operación del sistema, dando lugar a lo que se conoce como una configuración de control con división de vapor fija (Zhu et al., 2021). No obstante, en configuraciones E-TDWC es posible ajustar el flujo de vapor en la etapa de interconexión mediante la modificación de la presión en el domo del lado de rectificación, lo que permite implementar una configuración con división de vapor ajustable (Luyben, 2018). En este contexto, Hernández-Escoto et al. (2022) adoptan una estrategia de división de vapor ajustable, mientras que Tututi et al. (2014) analizan ambas configuraciones, considerando tanto esquemas con vapor fijo como con vapor ajustable.

En este trabajo se analizaron y compararon dos estrategias de control de pureza de productos:

i) Tututi et al. (2014) utilizan control basado en temperatura (control inferencial): En esta estrategia, las purezas



de los productos no se controlan de manera directa, sino a través del control de temperaturas en etapas seleccionadas de la columna. La selección de dichas etapas se basa en su alta sensibilidad a cambios en la composición. Esta estrategia es común en aplicaciones industriales debido a la rapidez y confiabilidad de las mediciones de temperatura.

La composición de los productos en una columna de destilación depende de la temperatura porque la separación se basa en el equilibrio líquido-vapor (ELV), donde las fracciones molares de los componentes en la fase líquida y en la fase de vapor están determinadas por la temperatura y la presión del sistema (Luyben, 2006b; Ränger et al., 2024).

ii) Hernández-Escoto et al. (2022) y Román et al. (2024) utilizan control directo de composición: En esta estrategia, las composiciones de los productos se miden directamente y se utilizan como variables controladas. Aunque este enfoque permite una regulación más precisa de las purezas, suele presentar dinámicas más lentas debido a los retardos asociados a los analizadores de composición.

Ambas estrategias se implementan mediante lazos de control independientes, utilizando controladores PI y PID, con el fin de realizar una comparación justa entre los enfoques de control y aislar el efecto de la variable medida sobre el desempeño dinámico.

2.3 Modelos dinámicos

En este trabajo se tomaron modelos dinámicos presentados en la literatura para este tipo de configuraciones. Con fines comparativos de control, se seleccionaron las FTs asociadas al control de las purezas de etanol y agua, ya que la pureza de los productos representa la variable de mayor relevancia operativa y económica en los procesos de DE.

Sistemas convencionales Para la selección de las FTs asociadas a la pureza de etanol o a la pureza de agua en cada caso, se analizaron previamente las estructuras de control reportadas en la literatura.

En la configuración de control de Tututi et al. (2014), la pureza de etanol se controla con la relación de reflujo de la columna extractiva (R_1), sin embargo, únicamente se reportan FTs correspondientes a lazos de control de temperatura, por lo que, no se dispone de un modelo dinámico explícito que relacione el reflujo (R_1) con la pureza de etanol. En contraste, la pureza de agua al ser controlada con la temperatura de una etapa intermedia de la columna recuperadora, sí se puede comparar con FTs de otros autores.

Román et al. (2024) reportan FTs exclusivamente para la columna extractiva de una CEC, lo que restringe el análisis comparativo al control de la pureza de etanol. Por lo tanto, los dos autores anteriormente mencionados no proporcionan información que permita comparar el control de las purezas de etanol y agua en CECs.

Bajo estas consideraciones, la comparación se llevó a cabo utilizando las FTs que relacionan la pureza de fondos de la columna extractiva (x_{B1}). En este contexto, Tututi et al. (2014) regulan esta variable con control inferencial de

temperatura en la etapa 14 (TC_{14}), manipulando la carga térmica en el rehervidor (Q_{R1}) mientras que Román et al. (2024) utilizan control de composición (x_{B1}), manipulando también la carga térmica en el rehervidor (Q_{R1}), como se observa en la Tabla 1.

Tabla 1. FTs seleccionadas para las CECs: variables de entrada y de salida

	x_B
Tututi et al. (2014)	$G_{x_{B1}} = \frac{TC_{14}}{Q_{R1}}$
Román et al. (2024)	$G_{x_{B1}} = \frac{x_{B1}}{Q_{R1}}$

Sistemas intensificados Asimismo, para las CEI se evaluó la compatibilidad entre las distintas estructuras de control con el fin de establecer criterios homogéneos para la comparación de las FTs asociadas al control de la pureza de etanol y de agua. En el caso de los sistemas E-TDWC analizados por Tututi et al. (2014), en la configuración con división de vapor fija, para mantener las purezas de etanol y agua se utiliza control inferencial de temperatura; TC_{14} manipulando R_1 , y TC_4 manipulando la relación de reflujo de la columna recuperadora (R_2), respectivamente.

En la configuración con división de vapor ajustable, se empleó el conjunto (F_V/Q_R , R_2) como variables de entrada para controlar el conjunto (TC_{14} , TC_4) que corresponden a las purezas de etanol y agua, respectivamente, siendo F_V el flujo de vapor en la interconexión y Q_R la carga térmica en el rehervidor. En esta configuración, el controlador TC_{14} regula el flujo de vapor en la interconexión, mientras que la pureza de etanol se controla mediante la relación de reflujo R_1 , para la cual no se reporta una función de transferencia explícita. No obstante, dado que el flujo de vapor en la interconexión ejerce una influencia significativa sobre la pureza de etanol y de agua, esta variable se adoptó igualmente como base para la comparación del control de la pureza de etanol.

Por su parte, Hernández-Escoto et al. (2022) presentan FTs únicamente para la configuración con división de vapor ajustable, en la cual proponen regular la pureza de etanol (x_1) mediante el flujo de vapor en la interconexión (F_V) y la pureza de agua (x_2) a través de la relación de reflujo (R_1).

En la Tabla 2 se muestran las variables de entrada y salida de las FTs seleccionadas para las CEIs. Se comparará el desempeño de control de las FTs de Hernández-Escoto et al. (2022) y las de Tututi et al. (2014) con F_V ajustable.

Tabla 2. FTs seleccionadas para las CEIs: variables de entrada y de salida

		Etanol	Agua
Tututi et al. (2014)	F_V fijo	$G_{x_1} = \frac{TC_{14}}{R_1}$	$G_{x_2} = \frac{TC_4}{R_2}$
	F_V ajustable	$G_{x_1} = \frac{TC_{14}}{V_R/Q_R}$	$G_{x_2} = \frac{TC_4}{R_2}$
Hernández-Escoto et al. (2022)	F_V ajustable	$G_{x_1} = \frac{x_1}{F_V}$	$G_{x_2} = \frac{x_2}{R_1}$

Las FTs empleadas en este trabajo corresponden a modelos dinámicos lineales obtenidos mediante la linealización de modelos rigurosos de columnas de DE alrededor de puntos de operación estacionarios. Estos modelos describen la dinámica de pequeñas perturbaciones en torno al régimen estacionario y constituyen una aproximación comúnmente utilizada para el análisis y diseño de sistemas de control PID.

Las condiciones de operación, configuraciones de control y las variables de diseño asociadas a los sistemas de destilación extractiva CEC y CEI se encuentran detalladas en los trabajos reportados en la literatura correspondientes.

2.4 Análisis de aplicabilidad de métodos clásicos de sintonización

Previo al diseño de los controladores, se evalúa la aplicabilidad de los métodos de Z-N y T-L a cada función de transferencia. Este análisis se realiza mediante el incremento progresivo de la ganancia proporcional en lazo cerrado, verificando la existencia o ausencia de oscilaciones sostenidas. Los resultados evidencian que, para la mayoría de los modelos asociados a columnas extractivas y E-DWC, no es posible determinar una ganancia última finita, lo que invalida el uso de dichos métodos y justifica la necesidad de un enfoque alternativo de sintonización.

Los métodos de sintonización de Z-N y T-L son técnicas empíricas clásicas, cuyo uso es válido únicamente cuando el proceso bajo estudio cumple ciertas condiciones dinámicas bien definidas. En particular, el método de Z-N en lazo abierto asume que el proceso puede aproximarse mediante un modelo de primer orden con tiempo muerto, lo cual requiere una respuesta a escalón de tipo sigmoidal y la existencia de un punto de inflexión claramente identificable (Seborg et al., 2016). Cuando estas características no están presentes, la estimación de los parámetros dinámicos del modelo resulta poco representativa y el método se vuelve impráctico.

Por otro lado, los métodos de Z-N en lazo cerrado y de T-L requieren que el sistema, bajo control proporcional puro, exhiba oscilaciones sostenidas para un valor finito de la ganancia proporcional, denominada ganancia última (K_u). Este comportamiento implica la existencia de un cruce del eje imaginario en el lugar de las raíces y márgenes de estabilidad compatibles con un régimen oscilatorio permanente (Aström and Hägglund, 2006). Si el proceso no presenta oscilaciones sostenidas, aun para valores elevados de la ganancia proporcional, no es posible determinar de manera consistente ni K_u , ni el período último P_u , lo que imposibilita la aplicación de estos métodos (Luyben, 1990).

Dado que las FTs empleadas en este trabajo, reportadas en la literatura para procesos de deshidratación de etanol, no satisfacen de manera general los supuestos dinámicos requeridos por los métodos Z-N y T-L, se adopta el comando *pdtune* de MATLAB como estrategia de sintonización. Este enfoque se basa en técnicas modernas de control, tales como la optimización en frecuencia y el ajuste automático a partir de modelos lineales identificados, sin requerir oscilaciones sostenidas, ganancia última ni aproximaciones de primer orden con tiempo muerto. En consecuencia,

pdtune permite obtener parámetros de control coherentes y reproducibles para una amplia gama de dinámicas, lo que lo convierte en una alternativa robusta y adecuada para el análisis comparativo de FTs asociadas a columnas de DE, tanto convencionales como con pared divisoria.

2.5 Comando *pdtune* en MATLAB

En esta sección se describen los criterios y opciones de diseño empleados para la sintonización automática de controladores PID mediante el comando *pdtune* de MATLAB. Asimismo, se presentan los parámetros que permiten establecer compromisos entre robustez, desempeño dinámico y objetivos de control en lazo cerrado, los cuales resultan fundamentales para garantizar un comportamiento estable y adecuado del sistema bajo estudio.

- Margen de fase = 60 (predeterminado). Margen de fase objetivo en grados, especificado como escalar positivo. *pdtune* trata de diseñar un controlador de modo que el margen de fase sea al menos el valor especificado para PhaseMargin. La frecuencia de cruce seleccionada podría restringir el margen de fase alcanzable. Normalmente, un margen de fase mayor mejora la estabilidad y el sobreimpulso, pero limita el ancho de banda y la velocidad de respuesta.
- Objetivo de rendimiento de lazo cerrado que se desea favorecer en el diseño, especificado como una de las siguientes opciones:
 - 'balanced' (valor predeterminado): para una robustez dada, ajuste el controlador a fin de equilibrar el seguimiento de referencia y la anulación de perturbaciones.
 - 'reference-tracking': ajusta el controlador para favorecer el seguimiento de referencia, si es posible.
 - 'disturbance-rejection': ajusta el controlador para favorecer la anulación de perturbaciones, si es posible.
- Número de polos inestables de la planta, especificado como entero no negativo (0 predeterminado). Cuando la planta es un modelo frd o un modelo de espacio de estados con retardos internos, debe especificar el número de polos inestables de lazo abierto (si los hay). Los valores incorrectos pueden resultar en controladores PID que no consigan estabilizar la planta real. (*pdtune* ignora esta opción para otros tipos de modelos).

2.6 Evaluación del desempeño dinámico

El desempeño de los controladores sintonizados se evaluó mediante simulaciones dinámicas ante cambios de consigna. Se emplearon índices de desempeño basados en el error, tales como IAE, ISE e ITAE, así como métricas temporales, incluyendo sobreimpulso y tiempo de establecimiento. Estas métricas resultan particularmente relevantes, ya que un sobreimpulso de gran magnitud puede provocar desviaciones transitorias significativas en la pureza del producto, mientras que el signo de la desviación, asociado a la naturaleza directa o inversa de la respuesta del proceso, determina si dichas desviaciones resultan penalizables desde el punto de vista operativo. Asimismo, un mayor tiempo de establecimiento implica que estas desviaciones de pureza persistan durante intervalos más prolongados,

lo que se traduce en pérdidas económicas asociadas a la operación fuera de especificación.

Estos indicadores permitieron comparar de manera objetiva la respuesta dinámica obtenida para los distintos esquemas de control y modelos considerados.

3. RESULTADOS

Para cada caso, se graficaron las respuestas dinámicas ante cambios tipo escalón y se calcularon los índices de desempeño basados en el error, con el fin de evaluar cuantitativamente el desempeño de los controladores. Los controladores sintonizados mediante *pidtune* fueron estables y no exhibieron oscilaciones indeseables. En todos los casos, los controladores PI y PID presentaron índices de desempeño basados en el error dentro de rangos aceptables, con valores consistentemente menores de IAE, ISE, ITAE e ITSE para los controladores PID, debido a la contribución efectiva de la acción derivativa en simulaciones sin ruido de medición. Asimismo, el control de temperatura mostró dinámicas más rápidas y menores tiempos de establecimiento que el control de composición, coherente con el menor retardo de las mediciones de temperatura, mientras que este último presentó respuestas más lentas y mayor sensibilidad a perturbaciones, aunque con desempeño estable.

3.1 Sistemas convencionales

En la Figura 2, los tiempos de establecimiento obtenidos mediante control de composición (Román et al., 2024) fueron de 2.89 h y 2.16 h al emplear controladores PI y PID, respectivamente. En contraste, el control de temperatura (Tututi et al., 2014) presentó tiempos de establecimiento significativamente menores, de 0.56 h para PI y 0.58 h para PID, lo que indica que, bajo esta estrategia, las purezas de los productos permanecen fuera de especificación durante un intervalo de tiempo considerablemente más corto. Asimismo, el sobreimpulso alcanzado con control PID fue de 1.10 para el control de composición y de 1.05 para el control de temperatura. En el caso del control de composición, no solo se observa un mayor sobreimpulso, sino que el incremento en el tiempo de establecimiento implica una mayor cantidad acumulada de producto que no cumple con la especificación de pureza en fondos, lo que resulta operativamente no deseable. El control de la pureza del producto de fondos en una columna de DE convencional es de particular relevancia, ya que dicha variable incide directamente en la pureza de los productos principales, etanol y agua. Variaciones en la composición de fondos modifican los perfiles de concentración y de equilibrio a lo largo de las columnas, lo que afecta la distribución de los componentes, así como la eficiencia global de la separación.

El ITAE penaliza más los errores que persisten en el tiempo, por lo que valores pequeños representan sobreimpulsos y tiempos de estabilización menores. Como se observa en la Tabla 3, para los índices calculados a partir de las FTs de Román et al. (2024), bajo control PI el índice ISE es 54.78 % menor que el IAE, mientras que con control PID el ISE resulta 55.13 % inferior. De manera similar, para los índices reportados por Tututi et al. (2014), el ISE es 45.65 % menor que el IAE con control PI, y 42.52 % menor con PID, lo que indica que ambos índices mantienen una

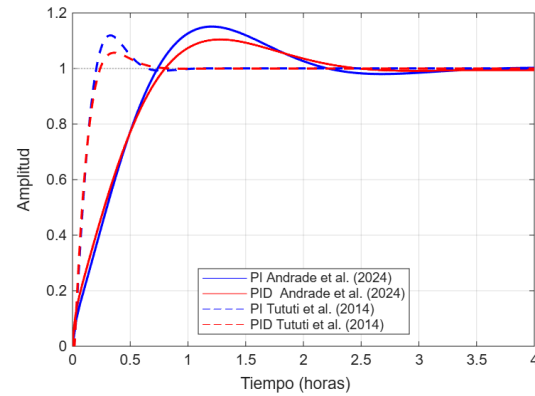


Figura 2. Respuesta con control PI y PID para la columna extractiva convencional de Tututi et al. (2014) y Román et al. (2024)

proporcionalidad consistente en magnitud; un comportamiento análogo se observa para el ITSE. No obstante, se identifica una diferencia relevante en el ITAE: en el caso de Román et al. (2024), el ITAE presenta valores superiores al ISE, mientras que para Tututi et al. (2014) el ITAE es inferior. Este comportamiento concuerda con lo mostrado en la Figura 2, dado que el ITAE penaliza de manera más severa los errores que persisten en el tiempo, lo que se traduce en valores más elevados de sobreimpulso y tiempo de asentamiento cuando se emplea control de composición.

Tabla 3. Pureza de fondos (x_{B1}) de la columna extractiva convencional

	Román et al. (2024)	Tututi et al. (2014)
IAE PI	0.4591	0.1255
ISE PI	0.2079	0.0682
ITAE PI	0.2920	0.0182
ITSE PI	0.0511	0.0038
IAE PID	0.4092	0.1110
ISE PID	0.1836	0.0638
ITAE PID	0.2358	0.0131
ITSE PID	0.0395	0.0030

3.2 Sistemas intensificados

Al analizar las respuestas transitorias del control de la pureza de etanol en los sistemas E-TDWC, se evidencia que el control de temperatura propuesto por Tututi et al. (2014) presenta un desempeño dinámico superior en comparación con el control de composición reportado por Hernández-Escoto et al. (2022) (Figura 3). En el esquema de control de temperatura, los tiempos de establecimiento obtenidos con controladores PI y PID son de 0.48 y 0.49 horas, respectivamente, mientras que al emplear control de composición estos aumentan a 0.59 horas con PI y 0.63 horas con PID.

En cuanto al sobreimpulso, el control PID presenta valores de 1.04 para el control de temperatura y de 1.06 para el control de composición. Aunque la diferencia en sobreimpulso es relativamente pequeña, el mayor tiempo de establecimiento asociado al control de composición implica que la pureza del producto permanece fuera de especificación durante un intervalo más prolongado, lo que refuerza la

conveniencia del control de temperatura desde el punto de vista operativo.

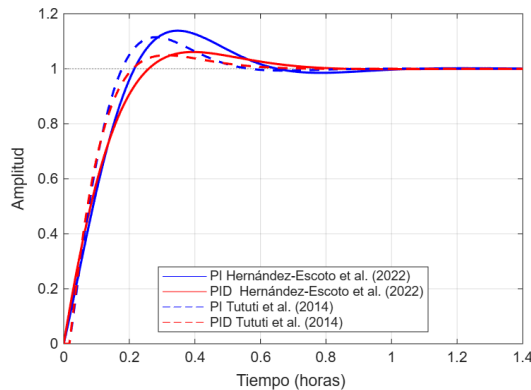


Figura 3. Respuesta con control PI y PID para la E-TDWC

La respuesta transitoria para la pureza de agua, utilizando los modelos de control de composición de Hernández-Escoto et al. (2022) fue significativamente más lenta comparada con las estructuras de control anteriores (Figura 4); en este caso, los tiempos de establecimiento fueron de 10.65 y 11.43 horas con control PI y PID, respectivamente. En contraste, el control de temperatura de Tututi et al. (2014), presentó tiempos de establecimiento similares a los casos anteriores, siendo 0.56 horas con PI y 0.57 horas con PID. El sobreimpulso presentado con control PID fue de 1.06 horas con control de composición y 1.05 horas con control de temperatura.

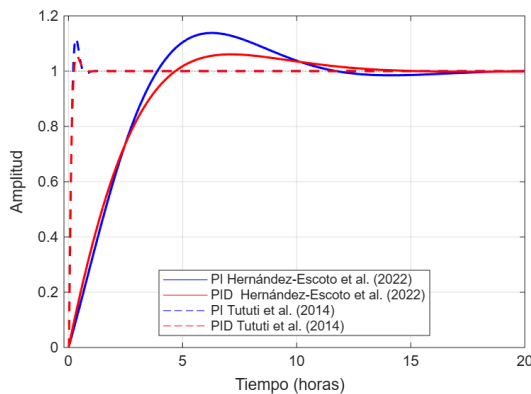


Figura 4. Respuesta con control PI y PID para la E-TDWC

Tabla 4. Índices basados en el error para los modelos dinámicos relacionados a la pureza de etanol de la E-TDWC

	Hernández-Escoto et al. (2022)	Tututi et al. (2014)
IAE PI	0.1329	0.1078
ISE PI	0.0646	0.0600
ITAE PI	0.0231	0.0131
ITSE PI	0.0042	0.0028
IAE PID	0.1124	0.0950
ISE PID	0.0570	0.0563
ITAE PID	0.0154	0.0022

En las Tablas 4 y 5, se observa que los índices basados en el error de los modelos dinámicos de Tututi et al. (2014) son menores a los de Hernández-Escoto et al. (2022), lo cual confirma una respuesta dinámica más rápida y un mayor grado de estabilidad cuando se utiliza una configuración de control de temperatura.

Tabla 5. Índices basados en el error para los modelos dinámicos relacionados a la pureza de agua de la E-TDWC

	Hernández-Escoto et al. (2022)	Tututi et al. (2014)
IAE PI	2.3963	0.1243
ISE PI	1.1663	0.0676
ITAE PI	7.4663	0.0179
ITSE PI	1.3745	0.0037
IAE PID	2.0291	0.1099
ISE PID	1.0284	0.0633
ITAE PID	5.0093	0.0128
ITSE PID	0.9779	0.0029

3.3 Modelos dinámicos de Tututi et al. (2014).

Se comparó también la sintonización de control con las FTs de Tututi et al. (2014) para una columna extractiva convencional y dos configuraciones E-TDWC; una con división de vapor fija y otra con división de vapor ajustable.

Como se señaló previamente, no se dispone de una función de transferencia que relacione explícitamente la pureza de etanol para la configuración extractiva convencional, y los parámetros de las FTs asociadas a la pureza de agua son idénticos para las configuraciones con división de vapor fija y ajustable. En consecuencia, el análisis comparativo del control de la pureza de etanol se realizó entre las configuraciones E-TDWC con división de vapor fija y E-TDWC con división de vapor ajustable, mientras que para la pureza de agua se compararon la configuración extractiva convencional y la E-TDWC.

En la Tabla 6 se comparan los índices basados en el error para las configuraciones con vapor fijo y vapor ajustable, ambas basadas en control de temperatura. Como se puede observar, la configuración con vapor ajustable presenta índices menores y por lo tanto una dinámica del proceso más rápida, lo cual se puede observar en la Figura 5.

Tabla 6. Índices basados en el error para los modelos dinámicos relacionados a la pureza de etanol de Tututi et al. (2014)

	E-TDWC vapor fijo	E-TDWC vapor ajustable
IAE PI	0.2559	0.1078
ISE PI	0.1365	0.0600
ITAE PI	0.0776	0.0131
ITSE PI	0.0156	0.0028
IAE PID	0.2263	0.0950
ISE PID	0.1270	0.0563
ITAE PID	0.0561	0.0092
ITSE PID	0.0123	0.0022

Como se puede observar, para los sistemas E-TDWC con vapor fijo y vapor ajustable de la Figura 5 los tiempos de

establecimiento con control PI y PID son de 1.15 y 1.19 horas, respectivamente, para la configuración con división de vapor fija, mientras que estos tiempos son de 0.48 y 0.49 horas para PI y PID, respectivamente, utilizando la configuración con división de vapor ajustable.

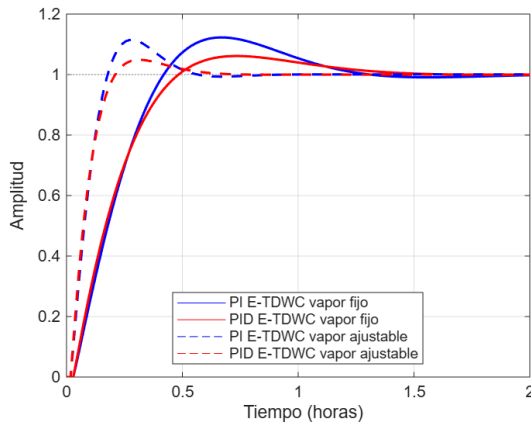


Figura 5. Respuesta con control de la pureza de etanol de Tututi et al. (2014)

Al comparar los índices del error asociados al control de la pureza de agua presentados en la Tabla 7, se observa que para el sistema convencional presenta una dinámica más rápida al tener índices más pequeños.

Tabla 7. Índices basados en el error para los modelos dinámicos relacionados a la pureza de agua de Tututi et al. (2014)

	Convencional	E-TDWC
IAE PI	0.0977	0.1243
ISE PI	0.0548	0.0676
ITAE PI	0.0107	0.0179
ITSE PI	0.0023	0.0037
IAE PID	0.0862	0.1099
ISE PID	0.0515	0.0633
ITAE PID	0.0075	0.0128
ITSE PID	0.0018	0.0029

Al evaluar el control de la pureza de agua (Figura 6), se toma en cuenta que los parámetros de las FTs de los sistemas E-TDWC son los mismos para las configuraciones con división de vapor fija y ajustable, por lo que en este caso se compararon los tiempos de establecimiento en el sistema convencional, los cuales fueron de 0.44 horas con PI y PID, en contraste a los tiempos de 0.56 y 0.57 horas para PI y PID en el sistema E-TDWC. Asimismo, el sobreimpulso presentado con control PID fue de 1.04 y 1.05 horas para el sistema convencional y E-TDWC, respectivamente. Estos valores más grandes de tiempo de establecimiento y sobreimpulso en el sistema E-TDWC pueden atribuirse a la mayor complejidad dinámica y al incremento de grados de libertad característicos de estos sistemas, los cuales incorporan una mayor interacción térmica y acoplamiento entre variables manipuladas y controladas (Zhu et al. (2021)).

4. CONCLUSIONES

En este trabajo se evaluó el desempeño dinámico de estrategias de control PI y PID aplicadas a FTs reportadas

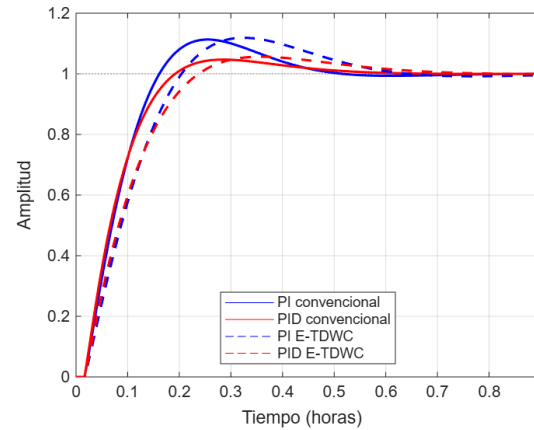


Figura 6. Respuesta con control de la pureza de agua de Tututi et al. (2014)

en la literatura para sistemas de DE convencional y con pared divisoria (E-TDWC), empleando un enfoque de sintonización consistente basado en el comando pidtune. Los resultados muestran que, en todos los casos analizados, el control PID presenta un desempeño superior al control PI, evidenciado por índices basados en el error más pequeños y sobreimpulsos más cortos. Este comportamiento se ve favorecido por la ausencia de ruido en las respuestas a cambios tipo escalón consideradas, lo que permite explotar de manera efectiva la acción derivativa.

Asimismo, se observó que el control basado en temperatura exhibe una dinámica más favorable que el control de composición, caracterizada por menores errores integrales, tiempos de estabilización reducidos y sobreimpulsos más pequeños, independientemente de la configuración de la columna. Estos resultados refuerzan la conveniencia práctica del control por temperatura en sistemas de DE, particularmente cuando se requiere una respuesta rápida y estable.

Para las configuraciones E-TDWC, el análisis comparativo indica que la estructura de control con división de vapor ajustable conduce a un mejor desempeño dinámico en comparación con la configuración de vapor fijo, reflejado en respuestas más rápidas y con menor desviación transitoria. Este comportamiento pone de manifiesto el beneficio de incorporar grados de libertad adicionales en la operación y control de sistemas intensificados.

Finalmente, el estudio destaca la importancia de considerar la dinámica inherente del proceso al seleccionar un método de sintonización, dado que las hipótesis fundamentales de métodos clásicos como Z-N y T-L no se satisfacen para varias de las FTs analizadas. En este contexto, el uso de herramientas de sintonización basadas en criterios explícitos de estabilidad y robustez, como pidtune, se consolida como una alternativa adecuada y sistemática para el análisis comparativo de estrategias de control en procesos complejos de DE.

REFERENCIAS

Aquilon, A.F., Cargullo, D.M., Onayan, J., Sarno, J., Molino, V.M., and Lopez, E.C.R. (2023). Recent advan-

- ces in extractive distillation. *Engineering Proceedings*, 56(1). doi:<https://doi.org/10.3390/ASEC2023-15261>.
- Aström, K.J. and Hägglund, T. (2006). *Advanced PID Control*. ISA—The Instrumentation, Systems, and Automation Society, Research Triangle Park, NC, USA.
- Czarnecki, N.J., Owens, S.A., and Eldridge, R.B. (2023). Extractive dividing wall column for separating azeotropic systems: A review. *Industrial & Engineering Chemistry Research*, 62(14), 5750–5770. doi:<https://doi.org/10.1021/acs.iecr.3c00302>.
- Gerbaud, V., Rodriguez-Donis, I., Hegely, L., Lang, P., Denes, F., and You, X. (2019). Review of extractive distillation. process design, operation, optimization and control. *Chemical Engineering Research and Design*, 141, 229–271. doi:<https://doi.org/10.1016/j.cherd.2018.09.020>.
- Hernández-Escoto, H., Zavala-Guzmán, A.M., Maya-Yescas, R., Zamudio-Lara, J.M., and Hernández, S. (2022). Operability assessment and systematic pi control of a class of extractive dividing wall distillation columns: Case of ethanol dehydration. *Chemical Engineering Research and Design*, 187, 84–92. doi:<https://doi.org/10.1016/j.cherd.2022.08.030>.
- Huang, X., Jin, Y., Fang, C., Yang, W., Li, Z., and Tian, Y. (2025). Process simulation and optimization of extractive distillation for separating isopropanol-ethanol from ibe fermentation broth as a gasoline additive. *Industrial Crops and Products*, 236, 121853. doi:<https://doi.org/10.1016/j.indcrop.2025.121853>.
- Kiss, A. (2013). *Advanced Distillation Technologies: Design, Control and Applications*. doi:<https://doi.org/10.1002/9781118543702>.
- Kiss, A.A. and Ignat, R.M. (2012). Innovative single step bioethanol dehydration in an extractive dividing-wall column. *Separation and Purification Technology*, 98, 290–297. doi:<https://doi.org/10.1016/j.seppur.2012.06.029>.
- Li, M., Peng, J., Cheng, Y., Zhu, X., Ma, Y., Zhang, Z., and Gao, J. (2023). Dynamic control of an energy-saving process with two extractive dividing-wall columns for separation of acetone/methanol/butanone/tert-butyl alcohol mixtures. *Chemical Engineering Research and Design*, 200, 281–291. doi:<https://doi.org/10.1016/j.cherd.2023.10.045>.
- Lorenz, H.M., Staak, D., Grutzner, T., and Repke, J.U. (2018). Divided wall columns: Usefulness and challenges. *Chemical Engineering Transactions*, 69, 229–234. doi:<https://doi.org/10.3303/CET1869039>.
- Luyben, W.L. (1990). *Process Modeling, Simulation, and Control for Chemical Engineers*. McGraw-Hill, Singapore, 2 edition.
- Luyben, W.L. (1996). *Process Modeling, Simulation, and Control for Chemical Engineers*. McGraw-Hill, New York, NY, USA.
- Luyben, W.L. (2006a). *Control of More Complex Columns*, chapter 8, 188–231. John Wiley & Sons, Ltd. doi:<https://doi.org/10.1002/0471785253.ch8>.
- Luyben, W.L. (2006b). *Fundamentals of Vapor-Liquid Phase Equilibrium (VLE)*, chapter 1, 188–231. John Wiley & Sons, Ltd. doi:<https://doi.org/10.1002/0471785253.ch1>.
- Luyben, W.L. (2018). Vapor split manipulation in extractive divided-wall distillation columns. *Chemical Engineering and Processing - Process Intensification*, 126, 132–140. doi:<https://doi.org/10.1016/j.cep.2018.02.022>.
- MathWorks (2023). Pid tuning algorithm. Accessed: 2026-04-12.
- Neagu, M. and Pricop-Nicolae, M. (2025). Azeotropic and extractive distillation for bio-ethanol dehydration: Process design, simulation, and cost analysis. *Processes*, 13(11). doi:<https://doi.org/10.3390/pr13113634>.
- Román, R.A., García, A.L., Vite, I.G., López, M.D., and Hernández, H. (2024). Desarrollo de modelos dinámicos tipo entradas-salidas de columnas de destilación. *JÓVENES EN LA CIENCIA*, 28, 1–15. doi:<https://doi.org/10.15174/jc.2024.4355>.
- Rängner, L.M., Halvorsen, I.J., Grützner, T., and Skogestad, S. (2024). Understanding temperature profiles of distillation columns. *Industrial & Engineering Chemistry Research*, 63(10), 4533–4546. doi:<https://doi.org/10.1021/acs.iecr.3c04102>.
- Seborg, D.E., Edgar, T.F., Mellichamp, D.A., and Doyle, F.J.I. (2016). *Process Dynamics and Control (4th ed.)*. John Wiley & Sons, Ltd.
- Skogestad, S. (2003). Simple analytic rules for model reduction and pid controller tuning. *Journal of Process Control*, 13(4), 291–309. doi:[https://doi.org/10.1016/S0959-1524\(02\)00062-8](https://doi.org/10.1016/S0959-1524(02)00062-8).
- Skogestad, S. and Postlethwaite, I. (2005). *Multivariable Feedback Control: Analysis and Design*, volume 2.
- Torres, K., Pérez-Zúñiga, R., Martínez, M., Ramos-Cabral, S., Calixto-Rodríguez, M., Valdez Martínez, J., Enriquez, M., Pérez-Estrada, A., Ortiz-Torres, G., Sorcia-Vázquez, F., Rebolledo, A., and Rumbo Morales, J. (2022). Design and control applied to an extractive distillation column with salt for the production of bioethanol. *Processes*, 10. doi:<https://doi.org/10.3390/pr10091792>.
- Tututi, S., Jiménez, A., and Hahn, J. (2014). Control analysis of an extractive dividing-wall column used for ethanol dehydration. *Chemical Engineering and Processing: Process Intensification*, 82, 88–100. doi:<https://doi.org/10.1016/j.cep.2014.05.005>.
- Xu, D., Li, W., Cao, L., Gao, J., Zhang, Z., and Wang, Y. (2026). Sustainable separation of azeotrope n-hexane/isopropanol via extractive distillation: Molecular insights, process optimization, and techno-economic-environmental analysis for cleaner production. *Industrial & Engineering Chemistry Research*, 65(9), 5125–5137. doi:<https://doi.org/10.1021/acs.iecr.5c04681>.
- Zhang, H., Wu, Z., Yuan, Q., Guo, L., Li, X., Hua, C., and Lu, P. (2025). Control of extractive dividing wall column using model predictive control based on long short-term memory networks. *Separation and Purification Technology*, 361, 131351. doi:<https://doi.org/10.1016/j.seppur.2024.131351>.
- Zhu, J., Jing, C., Hao, L., and Wei, H. (2021). Insight into controllability and operation of extractive dividing-wall column. *Separation and Purification Technology*, 263, 118362. doi:<https://doi.org/10.1016/j.seppur.2021.118362>.

Evaluation of the Type of Fertilization and Application of Abiotic Stress on the Bioactive Compounds of Chiltepín (*Capsicum annuum* var. *aviculare*)

Moreno-Limón Daniela *, Quiroz-Revilla María Paula **, Hernández-Carranza Paola **, Reyes-García Verónica *

* Departamento de Ingeniería, Instituto Tecnológico Del Altiplano de Tlaxcala, San Diego Xocoyucan, México. (e-mail: veronica.rg@altiplano.tecnm.mx)

** Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias Químicas, Puebla, Mexico.

Abstract: The objective of this study was to evaluate the effect of the type of fertilization (organic vs. chemical) and UV-C light irradiation treatment, as an abiotic stress factor, on the concentration of bioactive compounds and the antioxidant capacity of chiltepín peppers (*Capsicum annuum* var. *aviculare*). The content of total phenolic compounds, flavonoids, and total capsaicinoids was analyzed in fruits subjected to different exposure times (0, 1, and 2 minutes). Baseline results indicated that organic fertilization significantly promotes the accumulation of secondary metabolites compared to chemical fertilization. Following the application of UV-C light irradiation treatment, a hormetic effect was observed in which irradiation acted as an elicitor, increasing the total phenol content by between 50% and 82.98% in the organic treatment, while in the chemical treatment the increase ranged from 22.0% to 30.5%. Despite this lower initial percentage increase in the chemical group, the combination of conventional fertilization with abiotic stress (2 minutes of UV-C irradiation) achieved higher absolute values of capsaicin and antioxidant capacity. This interaction between agronomic management and post-stressor stress modulates the nutraceutical profile of the chiltepín, improving its functional value in the food industry.

Keywords: *Capsicum annuum* var. *aviculare*, abiotic stress, UV-C radiation, organic fertilization, bioactive compounds, antioxidants.

1. INTRODUCCIÓN

A nivel global, México es considerado el principal centro de origen, domesticación y diversificación del género *Capsicum* (Latournerie et al., 2026; Dávila et al., 2021). En la actualidad, los chiles del género representan uno de los cultivos hortícolas de mayor importancia agrícola y económica (Dávila et al., 2021). Entre ellos, el chiltepín (*Capsicum annuum* var. *aviculare*) destaca como el ancestro silvestre de la mayoría de los cultivares domesticados. Esta variedad resalta por su notable adaptación a condiciones ambientales adversas, tales como altas temperaturas y baja disponibilidad hídrica o nutricional en el suelo, lo que demuestra una elevada tolerancia genética al estrés abiótico (Vera-Guzmán et al., 2017; Moreno-Ramírez et al., 2018).

Desde el punto de vista agroalimentario, los frutos de chiltepín son considerados una fuente valiosa de compuestos bioactivos de alto valor funcional, incluyendo fenoles, flavonoides y capsaicinoides. Estos metabolitos secundarios poseen propiedades antioxidantes, antiinflamatorias y cardioprotectores de gran interés para la industria de alimentos y compuestos nutracéuticos (López-Ordaz y Yáñez-Fernández, 2023). Sin embargo, los frutos cosechados son

altamente perecederos, lo que acelera su deterioro postcosecha, corta su vida útil y restringe significativamente su comercialización eficiente

Capsicum representan uno de los cultivos hortícolas de mayor importancia a nivel global, tanto por su consumo directo como por su relevancia económica y cultural (Moreno-Ramírez et al., 2019; Megchún- García et al., 2024). Por lo tanto, resulta indispensable desarrollar estrategias agronómicas y tecnológicas combinadas que optimicen la síntesis, calidad y estabilidad de estos compuestos de calor agregado.

La fertilización constituye uno de los factores precosecha mas determinantes en el metabolismo vegetal, ya que influye de manera directa en la ruta biosintética de los metabolitos secundarios (Treto-Alemán et al., 2021). La naturaleza del fertilizante (orgánica o química) modifica de forma diferencial la acumulación de antioxidantes y capsaicinoides en el fruto (Moreno-Ramírez et al., 2018).

Particularmente, la fertilización orgánica ha cobrado relevancia como una alternativa sostenible capaz de actuar como bioestimulante del metabolismo secundario en los cultivos a través de una liberación gradual de nutrientes (De la Cruz Ricardez et al., 2020)

De manera complementaria, el uso de inductores físicos en la etapa de postcosecha representa una estrategia visible de activar los mecanismos de defensa del tejido vegetal. Entre estos métodos, la radiación ultravioleta tipo C (UV-C) ha demostrado la capacidad de inducir respuestas horméticas en productos hortícolas mediante la estimulación de la vía del fenilpropanoide (Hallmann y Rembiałkowska, 2012). En frutos del género *Capsicum*, la irradiación con luz UV-C estimula incrementos significativos en el contenido de flavonoides y capsaicinoides, contribuyendo además a retrasar el ablandamiento y mantener la integridad celular durante el almacenamiento (Bañuelos et al., 2009).

A pesar de las ventajas individuales que ofrecen el manejo de la nutrición vegetal y los tratamientos físicos postcosecha, aun no se ha esclarecido la interacción biológica del tipo de fertilización precosecha acoplada a la aplicación de radiación UV-C sobre el perfil fitoquímico del chiltepín. En consecuencia, el objetivo del presente trabajo fue evaluar la influencia combinada de diferentes regímenes de fertilización y la aplicación postcosecha de radiación UV-C como factor de estrés abiótico sobre el contenido de compuestos fenólicos totales (CFT), flavonoides totales (FT), capsaicinoides totales (CT) y la capacidad antioxidante (CA) en frutos de chiltepín (*Capsicum annuum* var. *aviculare*).

2. METODOLOGÍA

2.1. Ubicación del Estudio y Material Vegetal

El presente estudio se realizó de manera colaborativa en las instalaciones de la Benemérita Universidad Autónoma de Puebla (BUAP), Los frutos de chiltepín (*Capsicum annuum* var. *aviculare*) se obtuvieron de plantas cultivadas en invernadero tecnificado en la Universidad Benito Juárez García, ubicada en el municipio de Chiantzingo, Puebla, México. Los frutos se cosecharon estrictamente en estado de madurez fisiológica, caracterizado por una coloración verde intensa, tamaño uniforme y ausencia de daño mecánico (Fig. 1).



Fig. 1. Frutos de chiltepín (*Capsicum annuum* var. *aviculare*) Seleccionados en estado de madurez fisiológica para el desarrollo experimental.

2.2. Diseño experimental

El experimento se estableció bajo un diseño completamente al azar con un arreglo factorial 4×3 . Se evaluaron dos factores

fijos: el tipo de fertilización precosecha (control, orgánica, química y mixta) y tiempo de exposición postcosecha a la radiación UV-C (0, 1 y 2 min). Para cada tratamiento se estructuraron tres replicas biológicas independientes, donde cada unidad experimental estuvo conformada por un lote homogéneo de 100 gramos de frutos provenientes de plantas distintas bajo el mismo régimen nutricional.

2.3. Tratamientos de fertilización

Los esquemas de nutrición vegetal se detallan en la Tabla 1. El tratamiento control consistió en la aplicación exclusiva de agua. La fertilización orgánica incluyó la incorporación al suelo de humus de lombriz (20 g por planta) y composta (20 g por planta), complementada con aplicaciones foliares de lixiviados de humus. La fertilización química consistió en la aplicación de urea y fosfato diamónico (DAP) (15 g por planta), junto con aplicaciones de los fertilizantes foliares comerciales Agro-K amino (rico en potasio y aminoácidos libre), Nitro-sol (fuente nitrogenada balanceada) y Poliquel multi (complejo de micronutrientes quelatados). El tratamiento mixto combinó la fertilización química al suelo con lixiviados orgánicos y aspersiones foliares comerciales

Tabla 1. Tratamientos de fertilización para el cultivo de *Capsicum annuum* var. *aviculare*.

Tratamientos	Fertilización granular (Suelo)	Fertilización foliar
Control	Agua	Agua
Orgánico	Humus de lombriz (20g) y composta (20g)	Lixiviados de humus
Químico	Urea y DAP (15g)	Agro-k amino, nitro-sol, y poliquel-multi
Mixto	Urea y DAP (15g)	Lixiviados de humus y agro-k amino, nitro-sol y poliquel-multi

2.4. Tratamientos de Estrés Abiótico por Luz UV-C

La irradiación se realizó en una cámara equipada con seis lámparas UV-C (modelo G15T8, General Electric, EE. UU.), con una intensidad de radiación de $11,3 \text{ W/m}^2$. Los frutos se despusieron en una sola capa a una distancia constante de 20 cm de las lámparas, rotándose manualmente a la mitad del periodo para garantizar una exposición uniforme. Las dosis acumuladas correspondieron a:

- **0 min:** 0 kJ/m^2 (Control sin estrés)
- **1 min:** 0.678 kJ/m^2
- **2 min:** 1.356 kJ/m^2



Fig. 2. Cámara de irradiación equipada con lámparas de emisión UV-C, utilizada para la estabilización y tratamiento postcosecha de frutos de chile chiltepín.

Posteriormente, las muestras se almacenaron durante 7 días en una cámara climática a $20 \pm 2^\circ\text{C}$ y $85 \pm 5\%$ de humedad relativa para simular las condiciones de vida útil comercial.



Fig. 3. Frutos de chiltepín (*Capsicum annum* var. *aviculare*) una semana después de recibir el tratamiento de irradiación

2.5. Preparación de Extractos Acuosa

Se homogeneizaron 5 g de fruto fresco de chiltepín por réplica biológica y se mezclaron con 40 mL de agua destilada mediante ciclos controlados. El producto se filtró y se centrifugó a 4000 rpm durante 10 min utilizando una centrífuga serie XC-2450. El sobrenadante se recuperó y se aforó a 100 mL con agua destilada. Todas las determinaciones analíticas se realizaron por triplicado.

2.6. Determinaciones Analíticas

2.6.1 Contenido Fenólico Total (CFT)

Se determinó mediante el método de Folin-Ciocalteu. Se mezclaron 0.5 mL de extracto acuoso con 2.5 mL del reactivo de Folin-Ciocalteu (diluido 1:10 v/v) y 2.0 mL de carbonato de sodio (Na_2CO_3) al 7.5% w/v. Tras una incubación de 30 min en oscuridad a temperatura ambiente, se midió la

absorbancia a 764 nm empleando agua destilada como blanco reactivo. Los resultados se expresaron en mg de equivalentes de ácido gálico por gramo de pesos fresco (mg GAE/ g^{-1} PF).

2.6.2 Contenido Total de Flavonoides (FT)

Se cuantificó mediante un método colorimétrico de cloruro de aluminio. Se adicionaron 0,5 mL de extracto a 0,5 mL de nitrito de sodio (NaNO_2) al 1,5% w/v. transcurridos 5 min, se incorporó 1 mL de cloruro de aluminio (AlCl_3) al 3% w/v, seguido de 1 mL de hidróxido de sodio (NaOH) 1 N. Se registro la absorbancia a 490 nm frente a un blanco sin muestra. Los resultados se expresaron mg de equivalente de quercetina por gramo de peso fresco (mg QE g^{-1} PF).

2.6.3 Contenido Total de capsaicinoides (CT)

Los capsaicinoides totales se cuantificaron siguiendo el método descrito por Jattar-Santiago et al. (2022), con modificaciones menores. Las muestras se estabilizaron con ácido clorhídrico (HCl) 0,5 M, seguido de la adición de molibdato de sodio (Na_2MoO_4) 0,025 M y nitrito de sodio (NaNO_2) 0,5 M. Finalmente, se añadió hidróxido de sodio (NaOH) 0,1 M y se midió la absorbancia a 490 nm en un espectrofotómetro UV-Vis frente a un blanco reactivo. Los resultados se calcularon mediante una curva de calibración estándar utilizando capsaicina pura y se expresaron en miligramos de equivalentes de capsaicina por gramo de peso fresco (mg de capsaicina/g PF).

2.6.4. Capacidad Antioxidante (CA)

Se evaluó mediante el ensayo de eliminación de radicales libres DPPH. Se preparó una solución de trabajo DPPH a una concentración final de 0,1 mM disolviendo 0.004 g del radical en 100 mL de una mezcla de etanol y gua de grado analítico. Para la reacción, se mezclaron alícuotas de 0,1 mL de extracto acuoso con 2,9 mL de la solución de DPPH 0.1 mM. Las mezclas se agitaron vigorosamente en vórtex y se incubaron durante 30 min en condiciones de absoluta oscuridad a temperatura ambiente 25°C . Trascurrido el tiempo de incubación, se midió la absorbancia a una longitud de onda de 517 nm en un espectrofotómetro UV-Vis, utilizando la mezcla de etanol y agua pura como blanco experimental. Los resultados se calcularon con base en la disminución de la absorbancia respecto al control y se expresa como porcentaje de inhibición de radicales libres.

2.6.4. Análisis Estadístico

Los datos se sometieron a un análisis de varianza (ANOVA). Las comparaciones de medias se realizaron mediante la prueba Tukey utilizando el software SigmaPlot versión 14.0.

3. RESULTADOS Y/O DISCUSIÓN

La influencia del tipo de fertilización y del estrés por radiación UV-C sobre el contenido de compuestos bioactivos del chiltepín se presenta en las figuras 6 y 7. Cada figura se divide en cuatro subfiguras: (A) Compuestos Fenólicos Totales (CFT), (B) Flavonoides Totales (FT), (C) Capsaicinoides Totales (CT), (D) Capacidad Antioxidante (CA).

3.1. Efecto del Tipo de Fertilización y la Radiación UV-C en el contenido de fenólicos y flavonoides

La aplicación de radiación UV-C postcosecha modificó de manera significativa la concentración de compuestos bioactivos en los frutos de chiltepin (Fig.4 y 5). En ausencia de radiación (0 min), los frutos bajo fertilización orgánica exhibieron niveles basales de compuestos fenólicos totales (CFT) significativamente superiores a los del tratamiento químico. Este comportamiento se alinea con la teoría balance carbono-nutrientes (C/N), la cual postula que una disponibilidad limitada o gradual de nitrógeno soluble induce un desvío del exceso de carbono fotosintético hacia la ruta del fenilpropanoide, priorizando la síntesis de metabolitos secundarios estructurales y antioxidantes, como los fenoles y flavonoides, sobre el crecimiento primario (Hallmann y Rembalkowska, 2012).

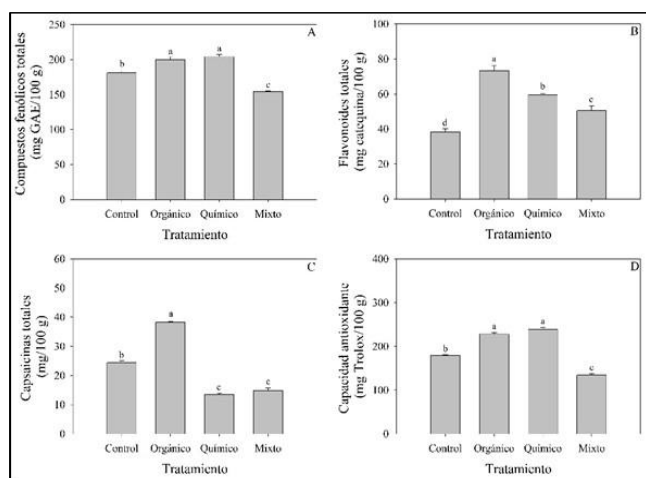


Fig.4. Efecto del tratamiento de fertilización sobre los compuestos bioactivos y la capacidad antioxidante del chiltepin

Al aplicar el estrés abiótico por UV-C, se observó una clara respuesta hormética en la que la irradiación actuó como un estímulo elicitor. En las muestras con fertilización orgánica, el contenido de fenoles totales incrementó entre un 50 % y un 82.98% tras la exposición. Este aumento es consistente con lo reportado por investigadores contemporáneos en otras matrices de *Capsicum*; por ejemplo, un comportamiento similar en la acumulación de compuestos fenólicos totales y capacidad antioxidante inducido por radiación ultravioleta postcosecha ha sido documentado en cultivares de chile habanero (*Capsicum chinense*) y chile pimiento (*Capsicum annuum* L.), donde el estrés físico activa la enzima fenilalanina amonio liasa (PAL), desencadenando la acumulación de fitoalexinas fenólicas para proteger la integridad celular contra daño oxidativo.

3.2. Respuesta Hormética y Elicitación por Radiación UV-C postcosecha

La exposición postcosecha a la luz ultravioleta C (UV-C) actuó como un potente estímulo estresante de naturaleza física,

desencadenando variaciones metabólicas marcadas según el tiempo de exposición (Dosis de 0.678 y 1.356 kJ/ m². Los resultados revelaron una respuesta hormética típica, donde una dosis controlada de agente estresante induce un efecto estimulador beneficioso en el tejido vegetal sin comprender su viabilidad celular.

En los frutos bajo manejo orgánico, el tratamiento de 1 min (0.678 kJ/ m²) y 2 min (1.356 kJ/ m²) impulsó un incremento sobresaliente en el contenido de fenoles totales, oscilando entre un 50% y un 82.98% con respecto a su control sin irradiar. Este comportamiento coincide con lo reportado por investigadores en otras matrices del género *Capsicum* expuestas a radiación ultravioleta postcosecha. Fisiológicamente, la radiación UV-C es captada por fotorreceptores celulares específicos, induciendo la expresión genética y la activación catalítica de la enzima fenilalanina amonio liasa (PAL). La PAL es la enzima clave que regula el flujo de carbono desde la vía del shikimato hacia la síntesis de polifenoles protectores. El incremento drástico en los CFT y FT funciona como un mecanismo de fotoprotección y neutralización de especies relativas de oxígeno ERO generadas por el estrés lumínico, previniendo el daño oxidativo en las membranas celulares del fruto durante el almacenamiento.

3.3. Contenido Total de Capsaicinoides y Capacidad Antioxidante

El comportamiento del Contenido Total de Capsaicinoides (CT) y la Capacidad Antioxidante (CA) evaluada por el método de DPPH presentó una interacción analítica compleja entre los factores de estudio. A pesar de mostrar valores fitoquímicos basales inferiores, los frutos desarrollados bajo fertilización química convencional respondieron con una magnitud notable superior ante la dosis máxima de estrés abiótico por UV-C (2 min; 1.356 kJ/ m²). En este tratamiento específico, se registraron incrementos significativos en los niveles de capsaicina y en el porcentaje de inhibición de radicales libres, reduciendo drásticamente las diferencias cualitativas iniciales frente a los frutos de menor orgánico.

Dado que las concentraciones de aminoácidos libres no se midieron de manera directa en este diseño experimental, se plantea como una hipótesis robusta sustentada en la literatura que la alta disponibilidad de nitrógeno en la fertilización convencional facilitó una mayor acumulación previa de compuestos nitrogenados de reserva en el tejido celular. Los capsaicinoides son alcaloides particulares cuya estructura molecular contiene nitrógeno. Su ruta biosintética requiere la convergencia de dos vías metabólicas: la vía del fenilpropanoide (que provee la vainillilamina a partir de la L-fenilalanina) y la vía del metabolismo de aminoácidos de cadena ramificada (que provee el sustituyente acilo a partir de la L-leucina).

Por consiguiente, se deduce que la radiación UV-C operó como un interruptor transcripcional que activó complejos enzimáticos terminales como la capsaicina sintasa (CS). Las muestras con fertilización química, al disponer de una reserva abundante de sustratos y aminoácidos precursores nitrogenados, lograron sostener una tasa de síntesis de

capsaicina más elevada bajo condiciones de estrés severo en comparación con las muestras orgánicas, cuyas reservas nitrogenadas basales eran mas limitadas. Este incremento paralelo es los capsaicinoides totales explica de forma directa el aumento en la capacidad antioxidante por DPPH, consolidando la existencia de un efectos sinérgico y acoplado entre el estatus nutricional precosecha en el campo y la tolerancia a estímulos biotecnológicos postcosecha en el chiltepín.

3.4. Interacción entre fertilización y radiación UV-C

Se observó una interacción significativa entre el tipo de fertilización y el tiempo de exposición a UV-C en el contenido de capsaicinoides y la capacidad antioxidante (Fig. 7 C y 7 D). Los frutos provenientes de fertilización química y sometidos a 2 min de irradiación UV-C presentaron incrementos notables en el contenido de capsaicina y la capacidad antioxidante, reduciendo las diferencias previamente observadas entre los tratamientos de fertilización.

Este efecto puede estar relacionado con la mayor disponibilidad de nitrógeno en los frutos provenientes de fertilización química, lo cual podría favorecer la acumulación de aminoácidos precursores utilizados posteriormente en la síntesis de alcaloides tras la exposición de radiación UV-C (Hallmann y Rembialkowska, 2012). la radiación UV-C actúa como un estímulo hormético, en el cual niveles moderados de estrés reducen respuestas metabólicas beneficiosas sin causar daño celular significativo (Li et al., 2022).

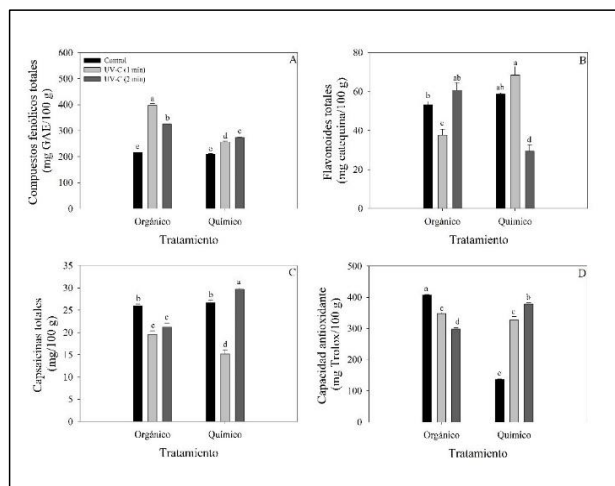


Fig. 5. Efecto del tratamiento con luz UV-C sobre los compuestos bioactivos y la capacidad antioxidante del chiltepín

Los valores de compuestos fenólicos totales en este estudio son comparables con los reportados para otros cultivares de *Capsicum* de alta pungencia, lo que confirma el elevado potencial nutraceutico del chiltepín (Villalba y Flores, 2021). De manera similar, el incremento de metabolitos secundarios tras la irradiación UV-C coinciden con lo reportado previamente en frutos de *Capsicum* sometidos a tratamientos de estrés abiótico controlado (Moreno-Ramírez et al., 2019).

En conjunto, los resultados indican que tanto la fertilización orgánica como la aplicación de radiación UV-C pueden emplearse como estrategias complementarias para mejorar la calidad funcional y el valor nutraceutico del chiltepín, lo cual tiene implicaciones importantes para su producción y comercialización.

4. CONCLUSION

El presente estudio demuestra que la acumulación de compuestos bioactivos en el fruto de chiltepín (*Capsicum annum var. aviculare*) es el resultado de una interacción compleja entre el estado nutricional de la planta y los estímulos postcosecha. Inicialmente, la fertilización orgánica favorece un mayor contenido basal de compuestos fenólicos, flavonoides y capsaicinoides, lo cual se atribuye a una mayor síntesis de metabolitos secundarios basados en carbono ante una menor disponibilidad de nitrógeno.

Sin embargo, la aplicación de irradiación UV-C por 2 minutos actúa como un elicitador crítico que modifica estos perfiles. Los resultados concluyen que el tratamiento químico irradiado por dos minutos alcanza el mayor potencial en el contenido de capsaicinas y capacidad antioxidante total. Este fenómeno sugiere que la mayor reserva de nitrógeno en frutos de cultivo convencional potencia la síntesis de alcaloides inducida por la luz UV-C, posiblemente mediante una mayor disponibilidad de precursores aminoacídicos.

En conclusión, mientras que la fertilización orgánica es la estrategia más eficaz para incrementar la calidad funcional pre-cosecha, mientras que la irradiación controlada se posiciona como el método tecnológico más eficiente para maximizar los atributos específicos de pungencia y capacidad antioxidante en el chiltepín.

REFERENCIAS

- Araiza Lizarde, N., Araiza Lizarde, E. y Martínez Martínez, J.G. (2011) 'Evaluación de la germinación y crecimiento de plántula de chiltepín (*Capsicum annum* L. var. *glabriusculum*) en invernadero', Revista Colombiana de Biotecnología, 13(2), pp. 170–175.
- Bañuelos, N., Salido, P. L. y Gardea, A. (2009). Etnobotánica del chiltepín: Pequeño gran señor en la cultura de los sonorenses. Estudios Sociales: Revista de Alimentación Contemporánea y Desarrollo Regional, 17(34), 177-205.
- Cerón-Carrillo, T., González-Cervantes, R. y Martínez-Ayala, A. L. (2009). Actividad antimicrobiana de extractos de diferentes especies de chile (*Capsicum*). Revista Mexicana de Ciencias Agrícolas, 35(2), 141-152.
- Dávila, M. A. V., Corona-M., E. y Aguilar-R., S. (2021). Chiles en México: Historias, culturas y ambientes. Universidad Veracruzana. <https://doi.org/10.25009/uv.2725.1626>
- De la Cruz-Ricárdez, D., Hernández-Gómez, R., Ochoa-Gaona, S. y Ramírez-Valverde, G. (2020). Compuestos fenólicos, carotenoides y capsaicinoides en frutos de



- Capsicum spp. de Tabasco, México. *Agrociencia*, 54(4), 505-519.
<https://doi.org/10.47163/agrociencia.v54i4.2047>
- Hallmann, E. y Rembialkowska, E. (2012) 'Characterization of antioxidant compounds in sweet bell pepper (*Capsicum annuum* L.) under organic and conventional growing systems', *Journal of the Science of Food and Agriculture*, 92(12), pp. 2409-2415. Disponible en: <https://doi.org/10.1002/jsfa.5624>
- Jattar-Santiago, K. Y., Reyes-García, V. y Quiroz-Revilla, M. P. (2022). Effect of extraction conditions on antioxidant compounds from habanero pepper (*Capsicum chinense*) and its potential antimicrobial use. *Journal of Food Processing and Preservation*, 46(6), e16427.
<https://doi.org/10.1111/jfpp.16427>
- Latournerie, L., Chávez, J. L., Pérez, M., Castañón, G., Rodríguez, S. A., Arias, L. M. y Ramírez, P. (2006). Exploración de la diversidad morfológica de chiles regionales en Yaxcabá, Yucatán, México. *Agronomía Mesoamericana*, 17(3), 323-332.
<https://doi.org/10.15517/am.v17i3.5173>
- Li, Q., Zhang, Y., Wang, M. y Liu, X. (2022). LED light quality affects growth, alkaloids contents, and expressions of amaryllidaceae alkaloids biosynthetic pathway genes in *Lycoris longituba*. *Journal of Plant Growth Regulation*, 41(1), 257-270.
<https://doi.org/10.1007/s00344-021-10303-3>
- López-Ordaz, P. y Yáñez-Fernández, J. (2023). Propiedades farmacológicas del chile (*Capsicum*) y sus beneficios en la salud humana: Una revisión bibliográfica. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, 4(2), 854-869.
<https://doi.org/10.56712/latam.v4i2.873>
- Megchún-García, J. V., Hernández-Juárez, A. y Víctor-Garro, A. (2024). Chile piquín (*Capsicum annuum* var. *aviculare*) establecido en un revestimiento de plástico y arreglos espaciales. *Revista Mexicana de Ciencias Agrícolas*, 15(1), 89-101.
- Moreno-Ramírez, Y. D. R., Martínez-Ávila, G. C. G., González-Laredo, R. F. y Rocha-Guzmán, N. E. (2018). Free radical-scavenging capacities, phenolics and capsaicinoids in wild piquin chili (*Capsicum annuum* var. *glabriusculum*). *Molecules*, 23(10), 2655.
<https://doi.org/10.3390/molecules23102655>
- Treto-Alemán, K. M., Segura-Ceniceros, D. E. y Contreras-Esquivel, J. C. (2021). Enriquecimiento del aceite comestible por compuestos fenólicos y antioxidantes de chile piquín (*Capsicum annuum* var. *glabriusculum*). *CienciaUAT*, 15(2), 156-168.
<https://doi.org/10.29059/cienciauat.v15i2.1459>
- Vera-Guzmán, A. M., Chávez-Servia, J. L., Carrillo-Rodríguez, J. C. y Heredia-García, M. G. (2017). Flavonoid and capsaicinoid contents and consumption of Mexican chili pepper (*Capsicum annuum* L.) landraces. En *Chili Peppers: Production, Bioactive Compounds and Health Benefits* (pp. 115-142). Nova Science Publishers