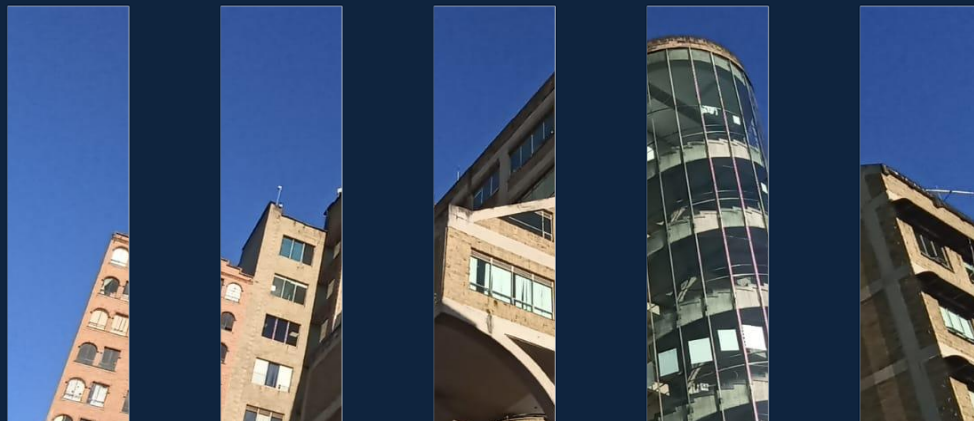


PROCEEDINGS OF THE FACULTY OF ENGINEERING, UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS



ISSN: 3028-6727 (en línea)

Vol: 4

No.: 1

Year: 2025

Annual Volume



**UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS**

Doctorado
en Ingeniería

<http://doctoradoingenieria.udistrital.edu.co>

Juan Carlos Figueroa-García Ph.D. – Ed.

Elvis Eduardo Gaona García Ph.D. – Ed.



PROCEEDINGS OF THE FACULTY OF ENGINEERING,
UNIVERSIDAD DISTRITAL FRANCISCO JOSÉ DE CALDAS
VOL 4 - 2025

Memorias de los eventos de la Facultad de Ingeniería
Universidad Distrital Francisco José de Caldas
ISSN: 3028-6727 (en línea)

Universidad Distrital Francisco José de Caldas
Rector: Dr. Giovanni Tarazona

Vicerrector Académico: Ignacio Rodríguez Molano Ph.D.

Vicerrector Administrativo: Elverth Santos Romero

Decano Facultad de Ingeniería: Ing. Mg, Edilberto Suárez Torres

Director Doctorado en Ingeniería: Germán Mendez Giraldo Ph.D.

Juan Carlos Figueroa-García Ph.D.

Elvis Eduardo Gaona-García Ph.D.

Comité Científico

Germán Hernández Ph.D., Universidad Nacional de Colombia

Carlos Franco Ph.D., Universidad Distrital Francisco José de
Caldas

Lindsay Álvarez Pomar Ph.D., Universidad Distrital Francisco
José de Caldas

Diego León Ph.D. (C), Universidad Externado

Árbitros Volumen 4

Germán Mendez Giraldo Ph.D., Universidad Distrital Francisco
José de Caldas

Yesid Díaz M.Sc., Universidad de La Salle

Dr. Roman Neruda, Informatics Dept. Charles University Prague,
Czech Republic

Feizar Javier Rueda-Velazco Ph.D., Universidad Distrital
Francisco José de Caldas

Rafael Bello Ph.D., Universidad “Martha Abreu” de las Villas

Jesus Lopez Ph.D., Universidad Autónoma de Occidente

John Leonardo Vargas Mesa M.Sc., Universidad del Rosario

Diana Guzmán Ph.D., Universidad Distrital Francisco José de
Caldas

Dr. Petra Vidnerova, Informatics Dept. Charles University
Prague, Czech Republic

Nelson Díaz Aldana Ph.D., Universidad Distrital Francisco José
de Caldas

Carlos Franco Ph.D., Universidad Distrital Francisco José de
Caldas

Diego León Ph.D. (C), Universidad Externado

Gestión Editorial

Juan Carlos Figueroa-García Ph.D.

Corrección de Estilo

Elvis Eduardo Gaona-García Ph.D.

Diagramación y montaje de cubierta

Juan Carlos Figueroa-García Ph.D.

Imagen de portada

Juan Carlos Figueroa-García Ph.D.

Periodicidad

Anual

Dirección: Cra 7 # 40B- 53 – Semisótano administrativo, Bogotá D.C. – República de Colombia

Teléfono: (+57) 6013239300 Ext: 1419 – 1402

email: doctoradoing@udistrital.edu.co

Sitio web: <https://doctoradoingenieria.udistrital.edu.co/memorias>

Prefacio

El incremento en las aplicaciones prácticas de técnicas de IA y de modelos de simulación han sido consistentes durante la última década, por lo que el componente computacional en la resolución de problemas de ingeniería es hoy en día parte estructural de la sociedad. Es en esta vía que los eventos organizados por la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas en el año 2025 han contribuido de manera positiva al desarrollo científico en temas como IA, optimización, simulación, matemáticas, computación, bioingeniería, métodos y modelos de investigación de operaciones y sus aplicaciones. Todos los trabajos resumidos aceptados fueron revisados por un equipo de revisores y han pasado por un proceso editorial con el fin de compilar este cuarto volumen que recopila dichos resultados de investigación.

Agradecemos al comité científico y el equipo de revisores por su ayuda incondicional en el proceso de revisión y aceptación de los trabajos. También agradecemos a los autores quienes son la piedra angular de este volumen de abstracts y representan a nuestra comunidad académica.

Diciembre 2025

Juan Carlos Figueroa-García Ph.D.
Elvis Eduardo Gaona-García Ph.D.
Editores

Índice general

Inteligencia Artificial

Beyond IVR Touch-Tones: Customer Intent Routing using LLMs.....	1
<i>Sergio Rojas-Galeano</i>	
ProfileXAI: User-Adaptive Explainable AI	6
<i>Gilber Gallego</i>	
Revisión de literatura de la IA y la gobernanza en el mejoramiento de servicios en establecimientos de crédito.....	10
<i>Melissa Cotera, David Longas, Jenny Moscoso, Yony Ceballos</i>	
Generacion de Images Aplicando Tecnicas de Autoencoder y SMOTE....	15
<i>Juan Miranda Gómez, Carlos Andres Ferro Sanchez, Reinel Reinel Tabares Soto</i>	
Development and Comparative Evaluation of Machine Learning and Deep Learning Models for Energy Consumption Forecasting in Colombia's Non Interconnected Zones	20
<i>Johan Cuellar, Elvis Gaona, Camilo Lopez</i>	
Artificial Intelligence in Non-Interconnected Zones: Challenges and Solutions for Resource Management	24
<i>Alejandro Marquez, Sandra Carvajal, Dahiana López</i>	
Evaluation of a customized Unet model for estimating cephalometric landmarks using heat maps on a standard deviation of intensity projection (STDIP) dataset of CBCT images.	29
<i>Juan Giraldo</i>	

Aplicaciones

Modelado para Optimización de Horarios en Restaurantes de Medellín ...	34
<i>Juan Galeano Florez, Yony Ceballos, Jorge Martinez Arenas</i>	
Autonomous Microalgae Control with AI and IoT for CO2 Capture and Renewable Energy	39
<i>Diana Montealegre, Nobel Castellanos, Agustin Marulanda</i>	
FPGA-Based Real-Time Heat Solver for Navier-Stokes-Inspired UWOC Modeling	44
<i>Miguel Solarte-Sanchez, David Marquez Vilora, Isnardo Arenas-Navarro, Neil Guerrero-Gonzalez</i>	

Una implementación de la propuesta de modelo extendido de administración de tierras (LADM-COL) para el inventario vial.....	48
<i>Ariana Cruz, Luz Rocha, Wilmar Fernández</i>	
Techno-Economic Analysis of Afro descendent Energy Communities to Empower Rural Entrepreneurships	53
<i>Yuri Lopez, Alberto Coronado, José María Yusta Loyo, Sofia Rojas Rueda, Maximilianne Acevedo</i>	
Gestión de la demanda para el aplanamiento de curva de consumo eléctrico en poblaciones densas.....	59
<i>Francisco Moya</i>	
Propuesta Metodológica para Evaluación Técnico-Económica en la Producción de Amoníaco Verde Multipropósito en Colombia.....	64
<i>Carlos Nieto</i>	

Simulación

Análisis de la dinámica de los embarazos no deseados y los abortos inseguros en Colombia mediante dinámica de sistemas	68
<i>Natalia Chacón, Carlos Franco, Diana Guzman</i>	
Comprendiendo la dinámica del cultivo de coca en Colombia: un enfoque desde la dinámica de sistemas	71
<i>Sara García, Laura Cárdenas</i>	
Dynamic Modeling of the Natural Gas Market in Colombia in the Framework of a Sustainable Energy Transition.....	75
<i>Derlyn Franco, Juan Carlos Osorio, Diego Fernando Manotas</i>	
Identificación de Técnicas de Analítica para la mejora de la precisión en modelos de Simulación, una revisión crítica de literatura	80
<i>Alejandro Álvarez, Yony Ceballos</i>	
Desafíos en el Modelado de Comportamientos Colectivos: Revisión de la Literatura sobre Metodologías de Simulación Basada en Agentes	84
<i>Danna Espitia Cano, Lindsay Álvarez Pomar</i>	
Dinámica de la cadena de suministro Laboral sujeta a Disrupciones. Caso Cacao	89
<i>Carlos Arango, Dany Guerra, Andres Hualpa,</i>	
Persistence Bias and the Bullwhip Effect: An Agent- Based Model of a Behavioral Driver in Supply Chains	96
<i>Edgar Cerón, Claudia Estévez-Mujica, Eva Hernández-Gress</i>	

Immersive Virtual Reality vs. Scale-Model Simulation in Emergency Medical Response: A Study on Stress and Training Performance	101
<i>Manuela Gonzalez, Daniel Morales, Oscar Caldas, Edilberto Mejia</i>	

Beyond IVR Touch-Tones: Customer Intent Routing using LLMs

Sergio Rojas-Galeano

Facultad de Ingeniería

Universidad Distrital Francisco José de Caldas

Bogotá, Colombia

srojas@udistrital.edu.co

Abstract—Widespread frustration with rigid touch-tone Interactive Voice Response (IVR) systems for customer service underscores the need for more direct and intuitive language interaction. While speech technologies are necessary, the key challenge lies in routing intents from user phrasings to IVR menu paths, a task where Large Language Models (LLMs) show strong potential. Progress, however, is limited by data scarcity, as real IVR structures and interactions are often proprietary. We present a novel LLM-based methodology to address this gap. Using three distinct models, we synthesized a realistic 23-node IVR structure, generated 920 user intents (230 base and 690 augmented), and performed the routing task. We evaluate two prompt designs: descriptive hierarchical menus and flattened path representations, across both base and augmented datasets. Results show that flattened paths consistently yield higher accuracy, reaching 89.13% on the base dataset compared to 81.30% with the descriptive format, while augmentation introduces linguistic noise that slightly reduces performance. Confusion matrix analysis further suggests that low-performing routes may reflect not only model limitations but also redundancies in menu design. Overall, our findings demonstrate proof-of-concept that LLMs can enable IVR routing through a smoother, more seamless user experience—moving customer service one step ahead of touch-tone menus.

Index Terms—LLMs, Touch-tone IVR menus, Natural Language Call Routing, Prompt Engineering, Synthetic Data Generation, Data Augmentation.

I. INTRODUCTION

THE modern customer service landscape is increasingly shaped by automation, with Interactive Voice Response (IVR) systems acting as the entry point for millions of daily interactions. Originally intended to streamline call routing and reduce costs, traditional touch-tone IVRs have become a source of widespread user frustration. Callers often struggle with navigating complex menus, repeatedly pressing buttons, enduring long wait times, and failing to reach the right department. These pain points translate into inefficient call handling, longer agent interactions, and ultimately, reduced customer satisfaction. This underscores the growing need for a more natural, efficient, and user-friendly alternative.

This demand has driven interest in natural language-based call routing. While research in Natural Language Processing (NLP) has long explored conversational interfaces using rule-based and machine learning approaches, recent advances in Large Language Models (LLMs) offer a transformative opportunity. LLMs can interpret and generate human-like text with

high accuracy, positioning them as central to the Natural Language Understanding (NLU) component of next-generation IVR systems. When integrated with speech recognition and synthesis, LLMs promise to move beyond the limitations of rigid menu trees and toward flexible, conversational interfaces.

However, evaluating LLMs for IVR call routing presents significant challenges, especially around data availability. Unlike many NLP tasks supported by large public datasets, IVR systems are typically proprietary. Their hierarchical menu structures are confidential, and real user interaction data is rarely accessible due to privacy and legal constraints. This lack of transparent, domain-specific data limits both development and rigorous evaluation of natural language call routing solutions.

To address these challenges, we propose a synthetic, LLM-driven experimental methodology that simulates a realistic IVR environment and enables controlled evaluation. Our pipeline uses three LLMs to construct the experimental framework. The first generates a 23-node IVR structure for a hypothetical telecom company. The second produces a dataset of 920 user intents, combining a base set of 230 distinct intents with 690 linguistically varied augmentations. The third model performs the routing task by matching each intent to the correct IVR path.

We compare routing performance under two context representations: a *Descriptive IVR Menu* (a plain-text hierarchical description) and *Flattened IVR Paths* (a list of terminal nodes identified by their Dual-Tone Multi-Frequency—DTMF—sequences). Our findings highlight how the structure and clarity of context input can significantly affect routing accuracy and offer practical insights for future implementations of LLM-based IVR systems.

II. RELATED WORK

The transition from rigid touch-tone IVR systems to more natural, conversational interfaces has long been a goal in automated customer service. Early studies, such as [10], demonstrated the advantages of natural language call routing over traditional DTMF systems in terms of user experience and efficiency. This effort has driven decades of research in NLP for intent recognition and dialogue management, spanning rule-based approaches, statistical models, and traditional machine learning techniques.

Recent advances in Large Language Models (LLMs) such as GPT-3, InstructGPT, and GPT-4 [2], [7], [6] have significantly advanced conversational AI. These models can interpret free-form user input with unprecedented fluency and contextual awareness. However, their application to IVR systems—particularly for directly mapping unstructured user complaints to structured IVR menu paths—remains largely unexplored. Existing work on IVR optimization often employs machine learning but stops short of using LLMs for the routing task itself. For example, [4] and [1] apply statistical models and classifiers to improve triage and routing accuracy based on IVR data. Similarly, [8], [9] discuss the promise of transformer-based models like BERT and GPT for enhancing intent detection and NLU in chatbots and IVR applications. Yet, none report concrete implementations where LLMs perform direct complaint-to-path mapping in a phone menu context.

Other contributions, such as [5], propose AI-enhanced IVR architectures incorporating speech recognition and NLP. These works often focus on system-level design or high-level integration strategies but do not provide empirical evaluations of LLMs applied to the routing task. A recent scientometric review by [3] confirms this gap, surveying decades of IVR research without identifying any study that evaluates LLMs for path selection based on user complaints.

In summary, while prior research acknowledges the potential of advanced NLP to improve IVR routing, it lacks direct exploration of LLMs for mapping natural language input to structured IVR menu paths. This absence highlights the novelty and relevance of our proposed methodology.

III. METHODOLOGY

A. Experimental Design Overview

To address the data scarcity problem in developing and evaluating natural language IVR systems, we propose a novel, LLM-driven experimental framework. This approach simulates a realistic environment for assessing LLM performance in intent routing, without relying on proprietary IVR data or sensitive customer interactions.

The methodology is structured as a three-stage pipeline, with distinct LLMs assigned to each stage: (1) generating a synthetic IVR menu structure, (2) producing a linguistically diverse set of user intents, and (3) executing natural language intent routing and performance evaluation. This design supports a controlled, reproducible proof-of-concept, allowing us to systematically test the impact of different IVR context representations on routing effectiveness. Table I summarizes the specific LLMs used in each stage of the experiment. IVR menus and user intents datasets are available at: <https://tinyurl.com/Beyond-Touch-Tones-LLMs>.

B. Synthetic IVR Menu Generation

To simulate a realistic customer service environment, we required a detailed IVR phone menu structure. Due to the proprietary nature of real-world IVRs, LLM1 was prompted to synthesize a 23-node hierarchical menu for a fictitious telecom provider, “AgentNet”. The prompt instructed LLM1 to emulate

TABLE I
LLM ASSIGNMENTS FOR EXPERIMENTAL PIPELINE

Id	Model	Role
LLM1	gpt-3.5-turbo	Synthetic IVR menu generation
LLM2	gpt-4o-mini	User intent dataset creation
LLM3	gpt-4.1-mini	Intent interpretation and routing

the structure of complex enterprise IVRs, covering service areas such as billing, technical support, account management, and new services. The output included structured node descriptions, each detailing its purpose and the available sub-options, with varying depth and complexity across the menu hierarchy.

From this synthetic IVR, we derived two distinct context representations for use with LLM3:

- **Descriptive IVR Menu:** A plain-text, hierarchical outline of the entire IVR structure, preserving all node descriptions and branching logic. This format mimics how a user might comprehend the IVR tree if presented with a full menu listing (see Appendix A).
- **Flattened IVR Paths:** A list of all terminal paths expressed as explicit DTMF sequences (e.g., 1-2-3), each linked to a specific service destination. This format offers a concise, navigable view of the IVR endpoints, abstracting away intermediate layers (see Appendix B).

C. User Intent Dataset Creation

To evaluate LLM-driven call routing, a diverse and precisely labeled dataset of user intents was essential. Given the absence of public IVR datasets that map free-form complaints to multi-level navigation paths, we employed LLM2 to synthesize and augment such data in a controlled setting with known ground truths. The dataset generation proceeded in two stages:

- 1) **Base Intent Generation:** LLM2 was prompted to create 10 distinct user complaints or queries for each of the 23 terminal nodes in the AgentNet IVR. These 230 base intents were designed to clearly and uniquely correspond to a specific endpoint, ensuring unambiguous routing paths.
- 2) **Intent Augmentation:** To increase linguistic variability, each base intent was paraphrased into three alternate versions using LLM2. Prompts explicitly instructed the model to preserve semantic equivalence and ground truth mappings while introducing variation in wording, length, and tone. Controlled noise was also induced, including interjections, filler phrases, and minor grammatical deviations, to better simulate real-world inputs.

The final dataset totaled 920 user complaints (230 base + 690 augmented), offering both linguistic diversity and routing precision for evaluating LLM3.

D. Natural Language Intent Routing

This subsection outlines the experimental setup and prompting strategies used to evaluate LLM3 in natural language call routing. The goal was to assess how effectively LLM3 could map diverse user intents from our synthetic dataset to their correct IVR destinations, given different context formats.

Two experimental conditions were tested, differing only in how the AgentNet IVR menu was presented to LLM3:

- 1) **Routing with Descriptive IVR Menu:** LLM3 received each user intent alongside the full plain-text hierarchical depiction of the IVR system. Prompts instructed the model to analyze the user query and return the corresponding DTMF path. The template used was:

```
Given the following IVR menu structure, identify the
exact DTMF path (e.g., '1-2-3') that best
addresses the user's query. Output only the path.

IVR Menu:
[Full Descriptive IVR Menu Text]

User Query:
[User's natural language intent]
```

- 2) **Routing with Flattened IVR Paths:** In this condition, LLM3 was shown a concise list of terminal paths, each annotated with a brief service description. The prompt instructed the model to select the most appropriate DTMF sequence for the given intent. The template was:

```
Select the most appropriate DTMF path from the list
below that corresponds to the user's query.
Output only the path.

Available Paths:
[List of Flattened IVR Paths, e.g., '1-1-1: Check
Balance', '1-2-1: Pay Bill']

User Query:
[User's natural language intent]
```

In both conditions, LLM3 processed all 920 user intents. Prompts included strict formatting instructions to ensure the model returned only the DTMF path (e.g., 1-2-3) with no additional text. This constraint was essential for automated evaluation. All calls were executed with LLM3 using its default inference settings.

E. Mitigation of Data Contamination Bias

To prevent data contamination—where evaluation models benefit from prior exposure to task artifacts—we ensured a strict separation of LLM roles throughout the pipeline. The IVR menu was synthesized with help from LLM1 but finalized manually, ensuring a realistic yet unseen structure. User intents and their paraphrases were generated by LLM2, based only on the terminal node labels, with no role in menu creation. Routing was handled by a third model, LLM3, which only received the IVR context in natural language form and had no exposure to prior stages. This separation across generation, augmentation, and evaluation stages minimizes leakage risk and ensures LLM3’s performance reflects genuine generalization.

F. Evaluation Metrics

LLM3’s routing predictions were evaluated using two metrics. **Accuracy** measured the proportion of exact matches between predicted and ground truth DTMF paths (e.g., 1-2-3). **Confusion matrices** were used to visualize errors across all 23 terminal paths, highlighting systematic misclassifications. Evaluations were conducted in Python using pandas and scikit-learn.

IV. RESULTS

This section presents the empirical results of the natural language intent routing experiments conducted with LLM3. We compare routing performance across two key factors: the IVR context format (Descriptive IVR Menu vs. Flattened IVR Paths) and the user intent dataset composition (Base Only vs. Augmented).

A. Overall Routing Accuracy

The primary evaluation metric was exact-match classification accuracy. Table II summarizes LLM3’s performance across all four experimental conditions.

TABLE II
LLM3 ROUTING ACCURACY BY IVR CONTEXT AND DATASET TYPE

IVR Context	Dataset	Accuracy (%)	N
Flattened Paths	Base Only	89.13	230
Flattened Paths	Augmented	86.52	920
Descriptive Menu	Base Only	81.30	230
Descriptive Menu	Augmented	77.07	920

Across both datasets, LLM3 achieved higher accuracy when provided with the Flattened IVR Paths, with a peak performance of 89.13% on the Base Only dataset. This indicates that a concise, list-based representation of menu options better supports accurate routing than a verbose hierarchical structure.

In both IVR context types, accuracy was higher on the Base Only dataset compared to its Augmented counterpart. This suggests that while linguistic augmentation introduces useful variation, it may also increase ambiguity or surface edge cases that challenge precise intent classification.

B. Confusion Matrix Insights

To gain a finer-grained view of LLM3’s routing behavior, confusion matrices were generated for each of the four experimental conditions, capturing classification outcomes across all 23 terminal IVR nodes.

Analysis of the matrices (Figures 1–4) reveals consistent patterns in both high-performing and error-prone routing cases.

Several paths exhibited near-perfect F1-scores (e.g., 1-1, 2-1-2, 3-4, 3-2, 3-3), indicating that LLM3 accurately classified intents corresponding to these frequently encountered or semantically distinctive nodes.

Conversely, certain paths posed consistent challenges. Most notably, 2-2-3 showed particularly low recall across all conditions—dropping to 0.12 with the Descriptive IVR Menu and 0.20 with the Flattened IVR Paths under the Augmented dataset. This suggests that user intents targeting 2-2-3 were frequently misclassified, potentially due to semantic ambiguity or overlap with nearby options. Path 1-9 also exhibited low recall in several settings.

Interestingly, the Flattened IVR Paths improved LLM3’s ability to identify some difficult nodes. For instance, recall for 2-2-1 increased to 1.00 in the Flattened + Augmented condition, despite lower precision in the Descriptive IVR Menu setting. This indicates that the flattened representation

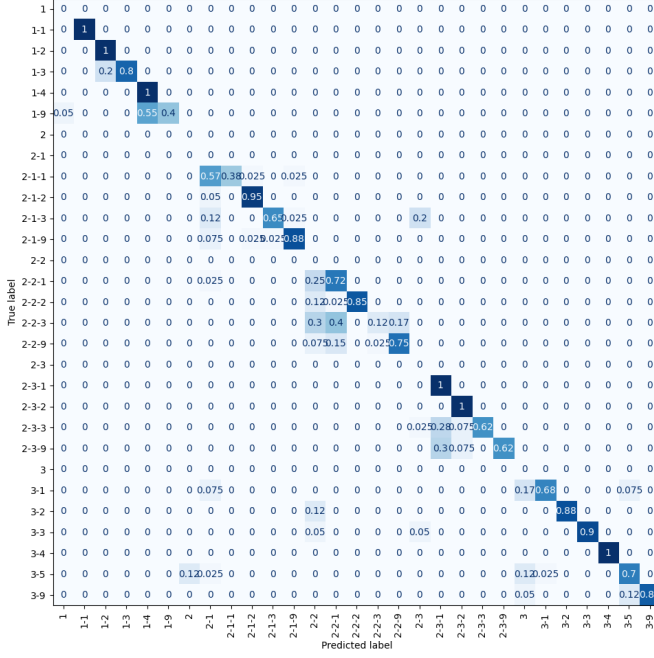


Fig. 1. Confusion matrix for Descriptive IVR Menu with Augmented Dataset.

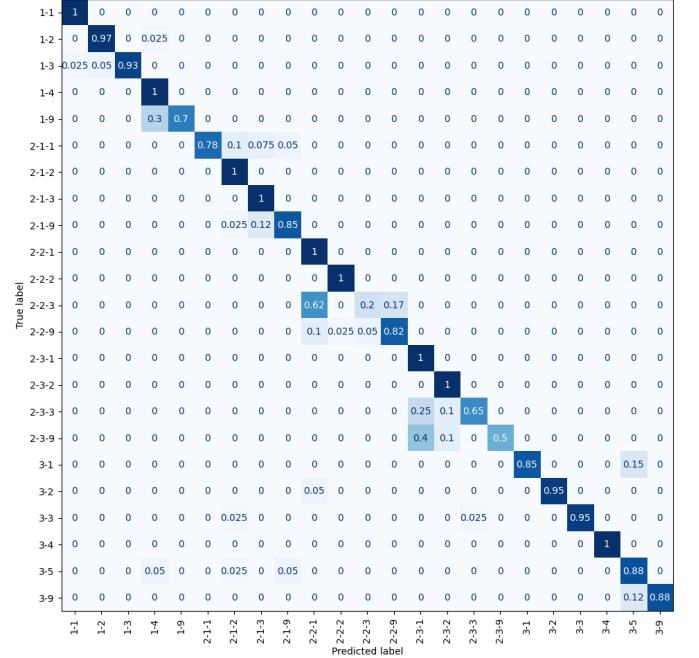


Fig. 3. Confusion matrix for Flattened IVR Paths with Augmented Dataset.

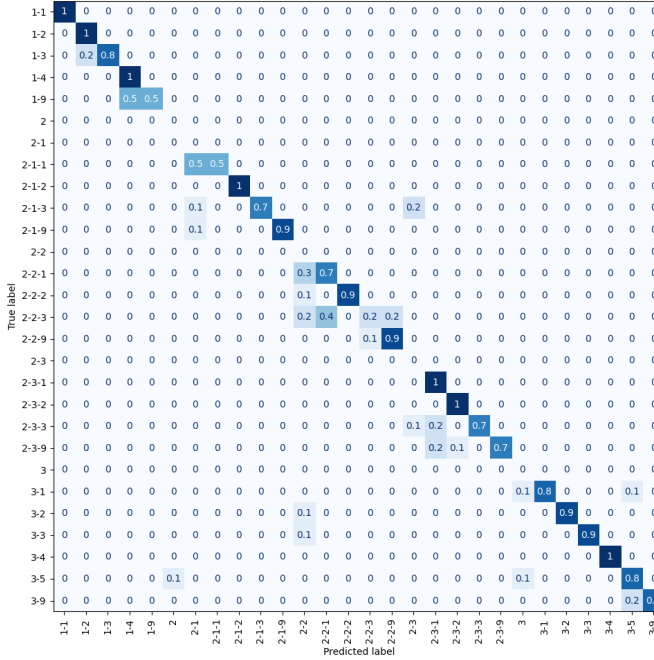


Fig. 2. Confusion matrix for Descriptive IVR Menu with Base Only Dataset.

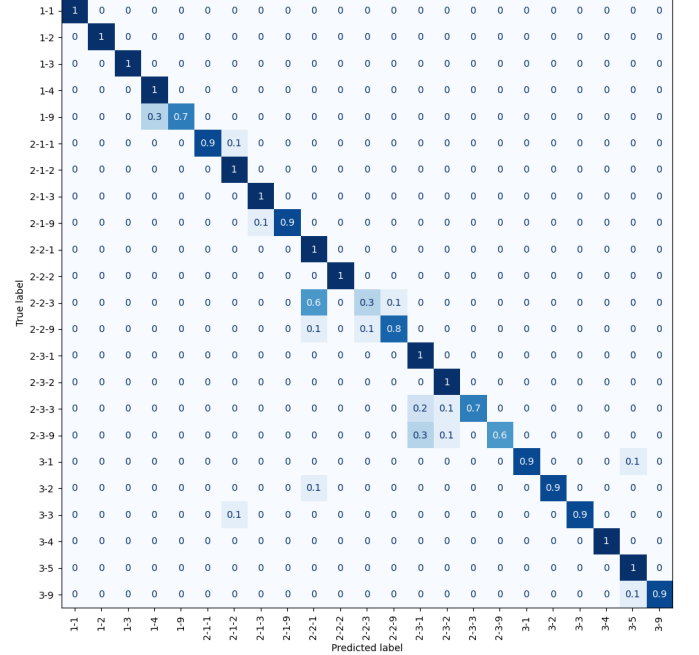


Fig. 4. Confusion matrix for Flattened IVR Paths with Base Only Dataset.

V. DISCUSSION

A. Key Insights

Our results show that LLMs route intents more accurately when given flattened IVR paths (up to 89.13%) than verbose menu descriptions (as low as 77.07%), suggesting that concise, structured prompts reduce noise and align better with the routing task. This supports the principle that clarity and brevity enhance LLM performance in classification settings. Importantly, transforming menus into flattened paths is a simple,

helped the model locate the correct path more reliably, even if it occasionally overpredicted that target.

Overall, these confusion matrix patterns reveal not only the LLM’s strengths in routing clear-cut intents, but also its limitations in disambiguating semantically similar nodes—especially when exposed to linguistically noisy data. The effect of IVR context formatting plays a non-trivial role in mitigating or exacerbating these issues.

automatable process for real-world use.

Contrary to expectations, performance dropped when using augmented data—despite preserved semantics—likely due to increased lexical variability that made intent mapping noisier. This highlights a trade-off: while augmentation may boost generality, it can reduce precision in tasks requiring exact mappings. Confusion matrix analysis further revealed that certain menu paths consistently underperformed, pointing to ambiguous or overlapping intent definitions. These findings suggest that LLMs can also serve as diagnostic tools for IVR design flaws, such as redundancy or overly granular distinctions that confuse both users and models.

Overall, the study underscores the value of structured prompts and clean data for LLM-based routing, and the potential of LLMs not only as classifiers but as evaluators of IVR menu quality.

B. Limitations and Future Work

This work used synthetic data and a single IVR structure, limiting generalizability. Real-world phrasings, LLM model and provider diversity, likewise more varied IVRs menus should be explored. We also used exact-match metrics and single-turn routing, whereas real systems often involve near-miss handling and multi-turn interactions. Finally, practical deployment concerns like latency and cost were not assessed and remain important areas for future research.

VI. CONCLUSION

This study evaluated how context representation and dataset design affect LLM-based intent routing in IVR systems. We found that concise, flattened representations of IVR paths led to significantly higher accuracy than verbose menu descriptions, and that dataset augmentation—despite increasing linguistic variety—introduced noise that slightly degraded performance. These findings highlight the value of prompt clarity and data precision in LLM applications. Future work should test these patterns on real-world data, explore multi-turn scenarios, and examine how LLMs internally handle structured prompts.

REFERENCES

- [1] TS. Aswin, Himanshu Batra, and Mathangi Ramachandran. Comparative analysis of machine learning algorithms on IVR data. *EAI Endorsed Transactions on Energy Web*, 5(18), 2018.
- [2] Tom B Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in Neural Information Processing Systems*, 33, 2020.
- [3] Ecaterina Coman. IVR systems used in call center management: a scientometric analysis of the literature. *Frontiers in Computer Science*, 7, Apr 2025.
- [4] N. Ilk, Guangzhi Shang, and Paulo B. Góes. Improving customer routing in contact centers: An automated triage design based on text analytics. *Journal of Operations Management*, 66:553–577, Jul 2020.
- [5] Gassyrbek Kosherbay and Nurgissa Apbaz. AI-Based IVR. *arXiv preprint arXiv:2408.10549*, 2024.
- [6] OpenAI. GPT4 Technical Report. *arXiv preprint arXiv:2303.08774*, 2023.
- [7] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Karian Slama, Alex Ray, and et al. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems*, 35:27735–27750, 2022.

- [8] Puneet Singh. AI-Powered IVR and Chat: A New Era in Telecom Troubleshooting. *SSRN Electronic Journal*, Jan 2025.
- [9] Puneet Singh. Streamlining Telecom Customer Support with AI-Enhanced IVR and Chat. *SSRN Electronic Journal*, Apr 2025.
- [10] Bernhard Suhm, Josh Bers, Dan McCarthy, Barbara Freeman, David Getty, Katherine Godfrey, and Pat Peterson. A comparative study of speech in the call center: natural language call routing vs. touch-tone menus. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 283–290, 2002.

APPENDIX A. AGENTNET IVR MENU STRUCTURE

```

IVR Menu Name: "AgentNet IVR"
--- Root Menu ---
IVR Message:
Welcome to AgentNet! Connecting you with what you need.
For Billing and Account Management, press 1.
For Technical Assistance, press 2.
For New Services and Upgrades, press 3.
To hear these options again, press 0.

--- Branch 1: Billing and Account Management (DTMF: 1) ---
IVR Message:
You've selected Billing and Account Management.
To check your balance, press 1.
To request your current invoice, press 2.
To make a payment, press 3.
To dispute recent charges, press 4.
To speak with a billing representative, press 9.
To return to the main menu, press 0.

--- Branch 2: Technical Assistance (DTMF: 2) ---
IVR Message:
You've selected Technical Assistance. We're here to help.
For Internet service issues, press 1.
For Mobile phone service issues, press 2.
For Video Streaming issues, press 3.
To return to the main menu, press 0.

--- Sub-Menu for Internet Issues (DTMF: 2-1) ---
IVR Message:
You've selected Internet issues.
To troubleshoot your modem or router, press 1.
For slow internet speeds, press 2.
For service outages, press 3.
To speak with an Internet Technical Representative, press 9.
To return to the previous menu, press 0.

--- Sub-Menu for Mobile Phone Service Issues (DTMF: 2-2) ---
IVR Message:
You've selected Mobile phone service issues.
For call and text issues, press 1.
To buy a new phone or upgrade, press 2.
For mobile device support, press 3.
To speak with a Mobile Technical Representative, press 9.
To return to the previous menu, press 0.

--- Sub-Menu for Video Streaming Issues (DTMF: 2-3) ---
IVR Message:
You've selected Video Streaming issues.
For quality problems, press 1.
For app or login errors, press 2.
For content or subscriptions, press 3.
To speak with a Video Streaming Technical Representative, press 9.
To return to the previous menu, press 0.

--- Branch 3: New Services and Upgrades (DTMF: 3) ---
IVR Message:
Welcome to New Services and Upgrades.
To inquire about new Internet plans, press 1.
To inquire about new Mobile phone plans, press 2.
To inquire about Video Streaming packages, press 3.
To add new lines to an existing account, press 4.
To upgrade your current service, press 5.
To speak with a sales representative, press 9.
To return to the main menu, press 0.

```

APPENDIX B. AGENTNET IVR TERMINAL PATHS

DTMF Path	Description	Type
1-1	Billing and Account Management → Check Balance	Self-service
1-2	Billing and Account Management → Request Current Invoice	Self-service
1-3	Billing and Account Management → Make Payment	Self-service
1-4	Billing and Account Management → Dispute Recent Charges	Self-service
1-9	Billing and Account Management → Billing Representative	Agent Handoff
2-1-1	Technical Assistance → Internet Issues → Modem / Router Troubleshooting	Self-service
2-1-2	Technical Assistance → Internet Issues → Slow Speed	Self-service
2-1-3	Technical Assistance → Internet Issues → Service Outages	Self-service
2-1-9	Technical Assistance → Internet Issues → Technical Representative	Agent Handoff
2-2-1	Technical Assistance → Mobile Phone Service Issues → Connection Issues	Self-service
2-2-2	Technical Assistance → Mobile Phone Service Issues → Buy a New Phone	Self-service
2-2-3	Technical Assistance → Mobile Phone Service Issues → Mobile Device Support	Self-service
2-2-9	Technical Assistance → Mobile Phone Service Issues → Technical Representative	Agent Handoff
2-3-1	Technical Assistance → Video Streaming Issues → Quality Problems	Self-service
2-3-2	Technical Assistance → Video Streaming Issues → App / Login Errors	Self-service
2-3-3	Technical Assistance → Video Streaming Issues → Content / Subscriptions	Self-service
2-3-9	Technical Assistance → Video Streaming Issues → Technical Representative	Agent Handoff
3-1	New Services and Upgrades → New Internet Plans	Self-service
3-2	New Services and Upgrades → New Mobile Phone Plans	Self-service
3-3	New Services and Upgrades → New Video Streaming Packages	Self-service
3-4	New Services and Upgrades → Add New Lines to Existing Account	Agent Handoff
3-5	New Services and Upgrades → Upgrade Current Service	Agent Handoff
3-9	New Services and Upgrades → Sales Representative	Agent Handoff

ProfileXAI: User-Adaptive Explainable AI

Gilber A. Corrales, Carlos Andrés Ferro Sánchez, Reinel Tabares-Soto, Jesús Alfonso López Sotelo, Gonzalo A. Ruz, and Johan Sebastian Piña Durán

Abstract—ProfileXAI is a model- and domain-agnostic framework that couples post-hoc explainers (SHAP, LIME, Anchor) with retrieval - augmented LLMs to produce explanations for different types of users. The system indexes a multimodal knowledge base, selects an explainer per instance via quantitative criteria, and generates grounded narratives with chat-enabled prompting. On Heart Disease and Thyroid Cancer datasets, we evaluate fidelity, robustness, parsimony, token use, and perceived quality. No explainer dominates: LIME achieves the best fidelity-robustness trade-off (Infidelity ≤ 0.30 , $L < 0.7$ on Heart Disease); Anchor yields the sparsest, low-token rules; SHAP attains the highest satisfaction ($\bar{x} = 4.1$). Profile conditioning stabilizes tokens ($\sigma \leq 13\%$) and maintains positive ratings across profiles ($\bar{x} \geq 3.7$, with domain experts at 3.77), enabling efficient and trustworthy explanations.

Index Terms—Explainable AI, Large Language models, User-adaptive explanations

I. INTRODUCTION

Artificial intelligence (AI) permeates most processes [1], [2]. As model architectures and parameter counts soar, their

Manuscript received July 27, 2025; accepted August 13, 2025. This work was supported in part by the graduate assistantship scholarship granted through Resolution No. 7906 by the Vice-Rectorate for Research and the Universidad Autónoma de Occidente; the National Agency for Research and Development (ANID), Applied Research Subdirection/IDeA I+D 2023 Grant [folio ID23110357]; *Classification of Alzheimer's stages using Nuclear Magnetic Resonance Imaging and clinical data from Deep Learning techniques* [873–139], Universidad Autónoma de Manizales, Manizales, Colombia; ORIGEN 0011323, Sistema General de Regalías (SGR)—Asignación para la Ciencia, Tecnología e Innovación project BPIN 2021000100368; *Technological platform for the classification of Alzheimer's disease stages using nuclear magnetic resonance imaging, clinical data and Deep Learning techniques* PRY-89, Universidad de Caldas, Manizales, Colombia; *Estrategia didáctica interactiva virtual para el fomento de habilidades TIC y su relación con el pensamiento computacional* PRY-121, Universidad de Caldas, Manizales, Colombia. G. A. R. thanks ANID FONDECYT 1230315, ANID-MILENIO-NCN2024_103, ANID PIA/BASAL AFB240003, and Centro de Modelamiento Matemático (CMM) FB210005, BASAL funds for centers of excellence from ANID-Chile. (Corresponding author: Gilber A. Corrales.)

G. A. Corrales is with the Facultad de Ingeniería y Ciencias Básicas, Universidad Autónoma de Occidente, Cali 760000, Colombia; and with the Escuela de Gobierno, GobLab, Universidad Adolfo Ibáñez, Santiago de Chile 8320000, Chile. (e-mail: gacorrales@uao.edu.co).

C. A. Ferro Sánchez is with the Facultad de Ingeniería y Ciencias Básicas, Universidad Autónoma de Occidente, Cali 760000, Colombia.

R. Tabares-Soto is with the Escuela de Gobierno, GobLab, Universidad Adolfo Ibáñez, and the Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez, Santiago de Chile 8320000, Chile; and with the Departamento de Electrónica y Automatización, Universidad Autónoma de Manizales, and the Departamento de Sistemas e Informática, Universidad de Caldas, Manizales 170003, Colombia.

J. A. López Sotelo is with the Facultad de Ingeniería y Ciencias Básicas, Universidad Autónoma de Occidente, Cali 760000, Colombia.

G. A. Ruz is with the Facultad de Ingeniería y Ciencias, Universidad Adolfo Ibáñez; Millennium Nucleus for Social Data Science (SODAS); and Center of Applied Ecology and Sustainability (CAPES), Santiago de Chile 8320000, Chile.

J. S. Piña Durán is with the Escuela de Gobierno, GobLab, Universidad Adolfo Ibáñez, Santiago de Chile 8320000, Chile; and with Universidad Autónoma de Manizales, Manizales 170003, Colombia.

decision mechanisms become opaque, effectively turning them into black-box systems [3], [4]. Explainable AI (XAI) aims to restore transparency while maintaining predictive accuracy [5]; however, existing techniques often fail to adapt their explanations to audiences with heterogeneous expertise [6], [7]. Integrating classical XAI with large language models (LLMs) has recently emerged as a promising [8].

Initial case studies utilized ChatGPT to generate SHAP and counterfactual outputs for student-risk analytics and the Iris benchmark [9]. Spitzer et al. later demonstrated that context-augmented prompting yields higher user satisfaction than retrieval-based prompting when explaining a deep-learning cost predictor [10]. In the networking domain, a fully automated 6G framework integrates XGBoost-SHAP with Llama 2 to diagnose SLA-latency anomalies, thereby increasing operator trust while revealing occasional decision errors [11]. Complementary methodological work formalises evaluation metrics—soundness, completeness, and fluency—and demonstrates that human readers prefer narrative SHAP summaries [12]. A recent survey synthesises these advances but highlights persistent issues of coherence and factuality [13]. Operational prototypes illustrate the practical upside: TalkToModel embeds GPT-J/3.5 within an interactive dialogue engine that reformats attribution-based explanations on demand, and most clinicians and ML professionals prefer it to conventional dashboards [14]. In recommender systems, LLM-generated justifications significantly enhance perceived transparency across feature, collaborative, and knowledge-based pipelines [15]. Finally, explanation-consistency finetuning improves the logical alignment of LLM summaries by approximately 10 % without degrading task accuracy [16].

We present a domain- and model-agnostic framework that advances this literature along three key axes. First, it adapts output granularity and style to distinct user profiles — machine learning experts, domain experts, and non-technical users — thereby maximizing relevance. Second, an interactive chat module enables stakeholders to refine queries and resolve residual uncertainties in real-time. Third, a dynamic engine selects the most suitable XAI method and enriches it through a multimodal retrieval-augmented generation pipeline, producing grounded, audience-specific narratives.

II. METHODOLOGY

A. Pipeline

Figure 1 illustrates the ProfileXAI architecture. The user—a machine-learning (ML) engineer—provides three inputs:

- **Knowledge base**, which supplies contextual information used to enrich the explanations, and this can be multimodal.

- **Black-box model** whose behaviour is to be explained (ex., Support vector machine, Multilayer perceptron, Random forest).
- **Dataset** (or a subset thereof) on which the model operates.

The information-extraction module then processes the knowledge base in a multimodal manner, identifies the most relevant components (extracting images or text from different types of documents), and stores them in a vector database. When an instance is submitted, the Retrieval-Augmented Generation (RAG) subsystem retrieves relevant fragments from the database to compose the generation prompt, enabling the system to generate explanations in natural language with context. The Explanation Engine, based on the instance entered by the user, executes three interpretability methods: SHAP [17], LIME [18], and Anchor [19], and automatically selects the most suitable one for each instance according to predefined metrics based on some metrics of [20].

The resulting explanation is produced in natural language and tailored to three user profiles:

- **ML engineer:** technical details, performance metrics, and raw model and explanation outputs.
- **Domain expert:** a translation of the explanatory content into terminology aligned with the application domain.
- **Non-technical user:** accessible language with illustrative examples and minimal jargon.

If the user poses follow-up questions, the interactive chat module enables a deeper exploration of any aspect of the generated explanation.

B. Experiments

We conducted experiments on two public datasets: Heart Disease with 13 features [21] and Differentiated Thyroid Cancer Recurrence with 16 features [22]. The knowledge base comprised the articles [23], [24]. We trained a multilayer perceptron (MLP) on the first dataset and a Random Forest on the second. Our evaluation comprised three blocks:

- 1) **XAI-metric analysis.** We adapted and assessed three standard interpretability metrics —Infidelity [25], Lipschitz [26], and Effective complexity [27] —across 100 instances of each dataset. For every explanation method, we report the mean and standard deviation of each metric Table I.
- 2) **Token consumption.** We recorded the number of tokens consumed per user profile (ML engineer, domain expert, non-technical) and per explanation method on 200 instances of each dataset. Table II summarises the averages.
- 3) **Satisfaction simulation.** Following the Hoffman survey [28], a simulated LLM scored seven explanation-quality items on a 1–5 scale (1 = very low, 5 = very high). We assessed 200 instances per Dataset, stratifying the results by user profile and explanation method. We thus obtained an average satisfaction score for each profile Table III.

III. RESULTS AND ANALYSIS

Regarding Table I across the three criteria—*robustness* (Local Lipschitz), *parsimoniousness* (Effective Complexity) and *fidelity* (Infidelity)—. LIME attains the best trade-off: the lowest Infidelity (≈ 0.08 – 0.30) and the strongest robustness ($L < 0.7$ Infidelity ≤ 0.30 , $L < 0.7$ to Heart Disease dataset), at the cost of a moderate complexity of 4–5 features. Anchor produces the most parsimonious explanations: out of the 13 (or 16) available features, the model typically needs only 3–4 (Effective Complexity) to alter its prediction. SHAP attains low-infidelity, high-fidelity explanations—consistent with the results reported by [29]—yet this advantage comes at the cost of diminished robustness ($L \approx 1.7$ – 2.0) and greater explanatory complexity (≈ 8 features). Both drawbacks become more pronounced as the feature space expands from 13 to 16 variables. In short, LIME offers the best balance, Anchor excels when brevity is paramount, and SHAP is preferable when capturing rich feature interactions outweighs stability considerations.

Table II quantifies the total token budget (features *plus* narrative) that a reader must process. For ML engineers Anchor is consistently the most concise 1131 tokens $\pm$ 1205 tokens, followed by SHAP and finally LIME, mirroring the relative verbosity of each method’s textual wrapper. For domain experts and non-technical users the pattern depends on the dataset: in the larger Dataset B Anchor again minimises cognitive load (7%–12% fewer tokens than SHAP, 11%–14% fewer than LIME), whereas for the Dataset A, LIME required the fewest tokens overall. Standard deviations confirm that token counts remain stable across the 200 instances ($\sigma \leq 13\%$ of the mean), indicating predictable effort requirements. Overall, if brevity is paramount for technical stakeholders Anchor is preferable, while LIME trades additional tokens for slightly richer contextualization.

Across the seven Hoffman items (Table III), SHAP receives the highest mean satisfaction in both tasks ($\bar{x}_{\text{meth}} = 3.9$ on Dataset A, 4.1 on Dataset B). LIME trails by ≈ 0.2 points, while Anchor ranks last yet very close (< 0.1 from LIME). Differences between user profiles are modest: Domain experts are the most critical, with an average score of 3.77, whereas non-technical users rate explanations marginally higher, especially on Dataset B. Taken together, all three XAI methods achieve solid upper-neutral acceptance (≥ 3.7), but SHAP enjoys a small, systematic advantage in perceived explanatory quality.

IV. CONCLUSION

We introduced ProfileXAI, a model- and domain-agnostic framework that couples classical post-hoc explainers with retrieval-augmented LLMs to *dynamically tailor* explanations to three distinct user profiles. On two medical benchmarks the system automatically chooses between SHAP, LIME and Anchor, verbalises the selected output at a suitable technical depth, and supports follow-up queries via chat.

The quantitative study confirms that no single explainer dominates every axis. LIME offers the best fidelity–robustness balance (Infidelity ≤ 0.30 , $L < 0.7$ to Heart Disease dataset);

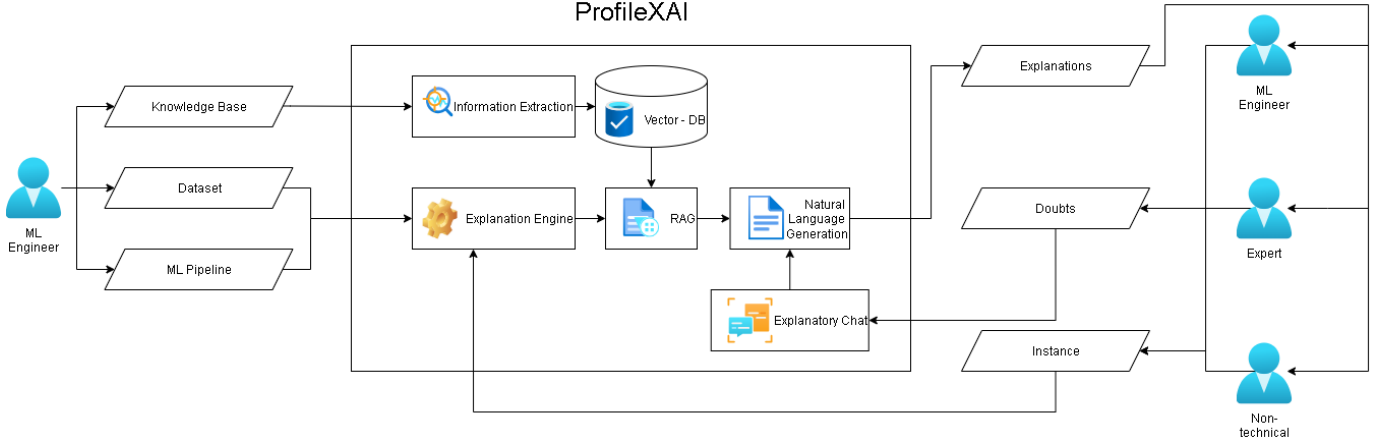


Fig. 1. ProfileXAI system architecture.

TABLE I

MEAN $\pm$ STANDARD DEVIATION OF THREE XAI METRICS—INFIDELITY, LIPSCHITZ, AND EFFECTIVE COMPLEXITY (EFFCOMP)—COMPUTED OVER 100 INSTANCES FOR EACH EXPLANATION METHOD (ANCHOR, LIME, SHAP) ON THE HEART DISEASE DATASET (DATASET A) AND THE DIFFERENTIATED THYROID CANCER RECURRENCE DATASET (DATASET B). CELLS MARKED “—” INDICATE THAT THE METRIC IS NOT WELL-DEFINED FOR ANCHOR.

Method	Dataset A			Dataset B		
	Infidelity	Lipschitz	EffComp	Infidelity	Lipschitz	EffComp
Anchor	—	0.88 ± 0.29	3.48 ± 4.69	—	1.49 ± 0.66	4.51 ± 6.15
LIME	0.30 ± 0.41	0.16 ± 0.08	4.16 ± 5.22	0.08 ± 0.04	1.65 ± 0.58	5.22 ± 6.67
SHAP	0.36 ± 0.40	1.76 ± 1.25	8.57 ± 5.60	0.23 ± 0.09	1.97 ± 0.46	7.56 ± 7.68

Note—The metric does not apply to Anchor because its rule-based output does not provide continuous feature importances, unlike the other two methods.

TABLE II

MEAN $\pm$ STANDARD DEVIATION OF TOTAL (INPUT + OUTPUT) TOKEN CONSUMPTION PER EXPLANATION OVER 200 INSTANCES FOR THREE USER PROFILES (ML ENGINEER, DOMAIN EXPERT, NON-TECHNICAL) AND THREE EXPLANATION METHODS (ANCHOR, LIME, SHAP) ON THE HEART DISEASE DATASET (DATASET A) AND THE DIFFERENTIATED THYROID CANCER RECURRENCE DATASET (DATASET B).

Method	Dataset A			Dataset B		
	ML	Domain	Non	ML	Domain	Non
Anchor	1131 ± 133	2289 ± 481	2314 ± 480	1205 ± 63	3358 ± 287	3398 ± 293
LIME	1347 ± 32	2017 ± 206	2029 ± 200	1626 ± 42	3781 ± 258	3782 ± 234
SHAP	1216 ± 37	2104 ± 410	2110 ± 443	1419 ± 43	3598 ± 245	3607 ± 218

TABLE III

SATISFACTION RATINGS (HOFFMAN SCALE: 1–5) FOR EACH EXPLANATION METHOD (SHAP, LIME, ANCHOR) AND USER PROFILE (ML ENGINEER, DOMAIN EXPERT, NON-TECHNICAL) OVER 200 INSTANCES ON THE HEART DISEASE DATASET (DATASET A) AND THE DIFFERENTIATED THYROID CANCER RECURRENCE DATASET (DATASET B). COLUMNS 1–7 CORRESPOND TO THE QUESTIONNAIRE ITEMS; $\bar{x}_{\text{prof}}$ IS THE AVERAGE PER PROFILE, AND $\bar{x}_{\text{meth}}$ THE AVERAGE PER METHOD.

Method	Profile	Dataset A							$\bar{x}_{\text{prof}}$	$\bar{x}_{\text{meth}}$	Dataset B							$\bar{x}_{\text{prof}}$	$\bar{x}_{\text{meth}}$
		1	2	3	4	5	6	7			1	2	3	4	5	6	7		
SHAP	ML	4.0	3.9	3.5	3.1	3.9	4.2	4.0	3.8		4.1	4.0	4.0	3.6	4.2	4.8	4.2	4.1	
	Domain	4.2	4.0	3.7	3.4	4.0	4.6	4.4	4.0	3.9	4.1	4.0	3.8	3.3	4.1	4.3	4.1	3.9	4.1
	Non	4.1	4.0	3.3	3.1	4.0	4.0	4.0	3.8		4.6	4.3	3.9	3.6	4.4	4.1	4.0	4.1	
LIME	ML	4.0	3.9	3.5	3.0	4.0	4.3	3.9	3.8		4.2	4.1	3.8	3.6	4.1	4.6	4.2	4.0	
	Domain	4.0	3.7	3.3	2.9	3.7	3.8	3.7	3.6	3.7	4.0	3.8	3.4	2.9	3.9	3.9	3.8	3.7	3.9
	Non	4.2	4.0	3.5	3.3	4.1	3.9	4.0	3.8		4.6	4.1	3.7	3.7	4.3	4.1	4.0	4.1	
Anchor	ML	3.9	3.5	3.2	2.8	3.6	3.7	3.5	3.5		4.0	3.9	3.7	3.1	3.9	4.3	3.9	3.8	
	Domain	4.1	3.8	3.4	3.0	3.7	4.4	4.2	3.8	3.7	4.1	3.7	3.4	2.9	3.8	4.3	4.1	3.7	3.8
	Non	4.3	4.1	3.5	3.2	4.1	3.7	3.9	3.9		4.2	4.0	3.6	3.4	4.1	3.8	3.8	3.8	

Anchor yields the sparsest rules and lowest token load; SHAP trades brevity for richer detail and thus achieves the highest Hoffman score ($\bar{x} = 4.1$). Profile-conditioned prompts keep token use stable ($\sigma \leq 13\%$) and satisfaction solidly positive

($\bar{x} \geq 3.7$), even though domain experts simulations remain the most demanding ($\bar{x} = 3.77$).

These findings substantiate the value of user-adaptive narration: by aligning explanatory granularity with audience needs,

ProfileXAI reconciles interpretability, cognitive economy and stakeholder satisfaction. Future work will extend the framework to multimodal data, incorporate additional explainers (e.g. counterfactual and concept-based), and validate with human participants to refine the simulated assessments.

REFERENCES

- [1] M. Ashok, R. Madan, A. Joha, and U. Sivarajah, "Ethical framework for artificial intelligence and digital technologies," *International Journal of Information Management*, vol. 62, p. 102433, 2022. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0268401221001262>
- [2] C. Cath, "Governing artificial intelligence: ethical, legal and technical opportunities and challenges," *Philosophical Transactions of the Royal Society A*, vol. 376, p. 20180080, 2018. [Online]. Available: <http://doi.org/10.1098/rsta.2018.0080>
- [3] A. Manure, S. Bengani, and S. S, *Transparency and Explainability*. Berkeley, CA: Apress, 2023, pp. 61–106. [Online]. Available: https://doi.org/10.1007/978-1-4842-9982-1_3
- [4] J. Henriques, T. Rocha, P. de Carvalho, C. Silva, and S. Paredes, "Interpretability and explainability of machine learning models: Achievements and challenges," in *International Conference on Biomedical and Health Informatics 2022*, E. Pino, R. Magjarević, and P. de Carvalho, Eds. Cham: Springer Nature Switzerland, 2024, pp. 81–94.
- [5] D. Minh, H. X. Wang, Y. F. Li *et al.*, "Explainable artificial intelligence: a comprehensive review," *Artificial Intelligence Review*, vol. 55, pp. 3503–3568, 2022. [Online]. Available: <https://doi.org/10.1007/s10462-021-10088-y>
- [6] M. Frasca, D. La Torre, G. Pravettoni *et al.*, "Explainable and interpretable artificial intelligence in medicine: a systematic bibliometric review," *Discover Artificial Intelligence*, vol. 4, p. 15, 2024. [Online]. Available: <https://doi.org/10.1007/s44163-024-00114-7>
- [7] M. Valiente Fernández, A. Lesmes González de Aledo, F. Delgado Moya *et al.*, "Shap model explainability in emco-pal mortality prediction: a critical analysis," *Intensive Care Medicine*, vol. 49, p. 1559, 2023. [Online]. Available: <https://doi.org/10.1007/s00134-023-07252-z>
- [8] P. Mavrepis, G. Makridis, G. Fatouros, V. Koukos, M. M. Separdani, and D. Kyriazis, "Xai for all: Can large language models simplify explainable ai?" 2024. [Online]. Available: <https://arxiv.org/abs/2401.13110>
- [9] T. Susnjak, "Beyond predictive learning analytics modelling and onto explainable artificial intelligence with prescriptive analytics and chatgpt," *International Journal of Artificial Intelligence in Education*, vol. 34, pp. 452–482, 2024. [Online]. Available: <https://doi.org/10.1007/s40593-023-00336-3>
- [10] P. Spitzer, S. Celis, D. Martin, N. Köhl, and G. Satzger, "Looking through the deep glasses: How large language models enhance explainability of deep learning models," in *Proceedings of Mensch Und Computer 2024*, ser. MuC '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 566–570. [Online]. Available: <https://doi.org/10.1145/3670653.3677488>
- [11] A. Mekrache, M. Mekki, A. Ksentini, B. Brik, and C. Verikoukis, "On combining xai and llms for trustworthy zero-touch network and service management in 6g," *IEEE Communications Magazine*, vol. 63, no. 4, pp. 154–160, 2025.
- [12] A. Zyteck, S. Pidò, and K. Veeramachaneni, "Llms for xai: Future directions for explaining explanations," 2024. [Online]. Available: <https://arxiv.org/abs/2405.06064>
- [13] A. Bilal, D. Ebert, and B. Lin, "Llms for explainable ai: A comprehensive survey," 2025. [Online]. Available: <https://arxiv.org/abs/2504.00125>
- [14] D. Slack, S. Krishna, H. Lakkaraju, and S. Singh, "Talktomodel: Explaining machine learning models with interactive natural language conversations," 2023. [Online]. Available: <https://arxiv.org/abs/2207.04154>
- [15] S. Lubos, T. N. T. Tran, A. Felfernig, S. Polat Erdeniz, and V.-M. Le, "Llm-generated explanations for recommender systems," in *Adjunct Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization*, ser. UMAP Adjunct '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 276–285. [Online]. Available: <https://doi.org/10.1145/3631700.3665185>
- [16] Y. Chen, C. Singh, X. Liu, S. Zuo, B. Yu, H. He, and J. Gao, "Towards consistent natural-language explanations via explanation-consistency finetuning," 2024. [Online]. Available: <https://arxiv.org/abs/2401.13986>
- [17] S. M. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems*, I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, and R. Garnett, Eds., vol. 30. Curran Associates, Inc., 2017. [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2017/file/8a20a8621978632d76c43dfd28b67767-Paper.pdf
- [18] M. T. Ribeiro, S. Singh, and C. Guestrin, "'why should i trust you?': Explaining the predictions of any classifier," 2016. [Online]. Available: <https://arxiv.org/abs/1602.04938>
- [19] —, "Anchors: High-precision model-agnostic explanations," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 32, no. 1, Apr. 2018. [Online]. Available: <https://ojs.aaai.org/index.php/AAAI/article/view/11491>
- [20] M. Pawlicki, A. Pawlicka, F. Uccello, S. Szelest, S. D'Antonio, R. Kozik, and M. Choraś, "Evaluating the necessity of the multiple metrics for assessing explainable ai: A critical examination," *Neurocomputing*, vol. 602, p. 128282, 2024. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0925231224010531>
- [21] S. W. P. M. Janosi, Andras and R. Detrano, "Heart Disease," UCI Machine Learning Repository, 1989, DOI: <https://doi.org/10.24432/C52P4X>.
- [22] S. Borzooei and A. Tarokhian, "Differentiated Thyroid Cancer Recurrence," UCI Machine Learning Repository, 2023, DOI: <https://doi.org/10.24432/C5632J>.
- [23] R. C. Detrano, A. János, W. Steinbrunn, M. E. Pfisterer, J.-J. Schmid, S. Sandhu, K. H. Guppy, S. Lee, and V. Froelicher, "International application of a new probability algorithm for the diagnosis of coronary artery disease," *The American journal of cardiology*, vol. 64 5, pp. 304–10, 1989. [Online]. Available: <https://api.semanticscholar.org/CorpusID:23545303>
- [24] S. Borzooei, G. Briganti, M. Golparian *et al.*, "Machine learning for risk stratification of thyroid cancer patients: a 15-year cohort study," *European Archives of Oto-Rhino-Laryngology*, vol. 281, pp. 2095–2104, 2024. [Online]. Available: <https://doi.org/10.1007/s00405-023-08299-w>
- [25] C.-K. Yeh, C.-Y. Hsieh, A. S. Suggala, D. I. Inouye, and P. Ravikumar, *On the (in)fidelity and sensitivity of explanations*. Red Hook, NY, USA: Curran Associates Inc., 2019.
- [26] D. Alvarez-Melis and T. S. Jaakkola, "Towards robust interpretability with self-explaining neural networks," in *Proceedings of the 32nd International Conference on Neural Information Processing Systems*, ser. NIPS'18. Red Hook, NY, USA: Curran Associates Inc., 2018, p. 7786–7795.
- [27] A. phi Nguyen and M. R. Martínez, "On quantitative aspects of model interpretability," 2020. [Online]. Available: <https://arxiv.org/abs/2007.07584>
- [28] R. R. Hoffman, S. T. Mueller, G. Klein, and J. Litman, "Measures for explainable ai: Explanation goodness, user satisfaction, mental models, curiosity, trust, and human-ai performance," *Frontiers in Computer Science*, vol. Volume 5 - 2023, 2023. [Online]. Available: <https://www.frontiersin.org/journals/computer-science/articles/10.3389/fcomp.2023.1096257>
- [29] M. A. Qureshi, A. A. Noor, A. Manzoor, M. D. Mazhar Qureshi, W. Rashwan, and A. Younus, "Explainability in action: A metric-driven assessment of five xai methods for healthcare tabular models," *medRxiv*, 2025. [Online]. Available: <https://www.medrxiv.org/content/early/2025/05/21/2025.05.20.25327976>

Revisión de literatura de la IA y la gobernanza en el mejoramiento de servicios en establecimientos de crédito.

Melissa Coteri Pernet, David Longas Castro, Jenny Moscoso Escobar, Yony Fernando Ceballos

UNIVERSIDAD DE ANTIOQUIA

Abstract—La Inteligencia Artificial (IA) se ha convertido en un tema determinante para la toma de decisiones en las organizaciones. Ello, incluye, a los establecimientos de ahorro y crédito que han tenido que apropiarse esta nueva tecnología, con el fin de mejorar sus procesos de acompañamiento a sus stakeholders (léase también como públicos de interés o estratégicos de una organización), dado que los servicios son un tema estratégico y misional que permite mejorar el desempeño organizacional.

Por ello, el propósito de esta revisión sistemática de la literatura es examinar cómo la apropiación de la inteligencia artificial, ligada con una gobernanza corporativa eficaz, influye en el mejoramiento de los servicios en establecimientos de ahorro y crédito.

Finalmente, este estudio aborda un tema de investigación no solo actual, sino también pertinente, dada la escasez de análisis conjuntos de estas variables en el contexto estudiado. La revisión identificó 33 artículos que cumplieron con los criterios establecidos, sugiriendo que la temática está en fase de construcción y resalta la necesidad de profundizar en su análisis.

Index Terms - Inteligencia artificial, Gobernanza corporativa, Servicios financieros

I. INTRODUCCIÓN

La inteligencia artificial (IA) ha sido una tecnología innovadora en diversas industrias en el mundo, incluido el sector financiero, la emersión de ésta en los establecimientos de ahorro y crédito ha permitido la mejora en la experiencia de clientes, la optimización de procesos y el fortalecimiento de toma de decisiones estratégicas [1]. En este escenario, el papel de la gobernanza corporativa desempeña un papel fundamental para definir una normatividad que garantice el uso seguro de estas tecnologías [2].

Es por ello, que la apropiación de la IA en el sector financiero, en este caso, los establecimientos de ahorro y crédito va más allá de la recepción de herramientas automatizadas; implica un

cambio tanto a nivel organizacional como cultural para maximizar el potencial [3].

Por tanto, este trabajo tiene como propósito realizar una revisión sistemática de literatura sobre la apropiación de la inteligencia artificial y su gobernanza en el mejoramiento de los servicios financieros de los establecimientos de ahorro y crédito. Se pretende analizar las tendencias, oportunidades y retos que enfrenta este sector al adoptar tecnologías inteligentes, además de proporcionar una base teórica y conceptual que proporcione la implementación eficiente, transparente y responsable.

II. OBJETIVOS

Objetivo general:

Elaborar la revisión sistemática de literatura a partir de fuentes secundarias provenientes de bases de datos especializadas sobre los resultados de la apropiación de la IA y la gobernanza en los servicios financieros de las cooperativas de ahorro y crédito.

Objetivos específicos:

1. Diseñar estrategia de búsqueda a través de operadores (and, or, not, etc) de los hallazgos de estudios sobre la adopción de la inteligencia artificial y la gobernanza en los servicios financieros de establecimientos de ahorro y crédito.
2. Seleccionar estudios relevantes en las bases de datos especializadas (Scopus, Science Direct, Emerald) sobre la relación entre la inteligencia artificial y la gobernanza en los servicios financieros de establecimientos de ahorro y crédito.

Este trabajo académico es el desarrollo de la práctica en modalidad investigativa de Melissa Coteri Pernet, estudiante de la Facultad de Ingeniería, Universidad de Antioquia, Medellín, Colombia. (correo electrónico: melissa.coteri@udea.edu.co).

David Longas Castro está realizando una pasantía de investigación en el mismo proyecto, estudiante de Administración de empresas, Universidad de Antioquia, Medellín, Colombia (correo electrónico: david.longasc@udea.edu.co).

Jenny Moscoso Escobar, profesora titular de la Facultad de Ciencias Económicas, Universidad de Antioquia, Investigadora Grupo de Investigación en Finanzas – GIF (correo electrónico: jenny.moscoso@udea.edu.co).

Yony Fernando Ceballos profesor titular de la Facultad de ingeniería, Universidad de Antioquia e investigador del Grupo ingeniería y sociedad (Correo electrónico: yony.ceballos@udea.edu.co).

El trabajo de grado se enmarca en el proyecto de investigación “Apropiación de la IA y la gobernanza en el mejoramiento de servicios en establecimientos de crédito. El caso de las cooperativas de ahorro y crédito en Antioquia”. Universidad de Antioquia – Universidad Pontificia Bolivariana.

3. Sintetizar los hallazgos recopilados sobre búsqueda objeto de estudio.
4. Realizar el informe final de la revisión sistemática de literatura investigado.

III. MARCO TEÓRICO

Las tecnologías de la información y las comunicaciones (TIC) han alcanzado un alto grado de desarrollo lo que les permite tener una mayor participación y de carácter decisivo en ciertos procesos productivos y de servicios, así como en las diferentes relaciones entre las personas y las organizaciones, revolucionando la sociedad moderna. Tras una fase que se focalizó en la informatización y digitalización, se da paso a una nueva etapa en la que las TIC transforman todos los procesos socioeconómicos. En este contexto, la inteligencia artificial (IA) emerge como un elemento clave del cambio. [4]

Por otro lado, la mayoría de los autores apuntan a que la IA en la actualidad es uno de los desarrollos importantes que ha impactado a la sociedad. Según [5, p. 1]: “La inteligencia artificial (IA) ha evolucionado hasta convertirse en una tecnología madura y sofisticada, entró silenciosamente en nuestras vidas e hizo un gran salto en el último año. Los modelos generativos de IA han demostrado que la IA ha pasado, en tan solo unos meses, prácticamente de la ciencia ficción a convertirse en una parte esencial de la vida cotidiana de cientos de millones de personas en todo el mundo.”

Dada la creciente influencia de la IA en distintas disciplinas, su estudio es pertinente. En este sentido,

La IA se está aplicando a numerosas actividades realizadas por los seres humanos y se destacan entre otras las siguientes líneas de investigación científicas: la robótica, la visión artificial, técnicas de aprendizaje y la gestión del conocimiento. Estas dos últimas aplicaciones de la IA son las que más directamente se emplean al campo de las finanzas, debido a que en este campo existe una fuerte motivación orientada a la construcción de sistemas de información que incorporen conocimiento, y que permitan a los decisores de las organizaciones tomar decisiones eficientes y oportunas en el ámbito de la gestión financiera empresarial [6, p. 5].

Debido a que las técnicas de aprendizaje se caracterizan según [7, p. 44] por: “ser una amplia colección de algoritmos utilizados para crear modelos que aprendan de los datos históricos, con el objetivo de hacer predicciones o conocer las relaciones que pueden existir entre variables de entrada y salida (ocurrencia de un evento)”. Así mismo, la IA se puede definir como una subdisciplina de la informática centrada en automatizar el comportamiento inteligente [8].

Además, estamos inmersos en la Cuarta Revolución Industrial, en donde la IA desempeña un papel clave por 5 características: la transparencia en los sistemas de inteligencia artificial suele pasar desapercibida durante su uso; su dificultad radica en que trabajan principalmente con seres humanos como referencia; su adaptabilidad se manifiesta en su dependencia del contexto, lo que otorga múltiples formas; la transversalidad, ya que actualmente ningún sector escapa a su influencia; y, finalmente, requiere una constante actualización y mejora, algo característico de las tecnologías de la información y la

comunicación, generando una demanda única de respuestas inmediatas [9].

Debido a la IA y a la utilización de esta dentro de las organizaciones surgió un concepto que es la gobernanza de la IA que en pocas palabras son las políticas, prácticas y procesos que las organizaciones utilizan para gestionar responsablemente el desarrollo y la implementación de la IA garantizando que se desarrollen y utilicen de manera ética, transparente y responsable dentro de una organización. Los marcos de gobernanza de la IA incluyen prácticas internas y externas, como la gestión del riesgo, los comités de ética, las medidas de transparencia y la participación de las partes interesadas. [5]

Por su parte, la gestión del conocimiento se vincula con la Industria 4.0, dado que hace referencia a la revolución industrial impulsada por la transformación digital, enmarcados en nueve tecnologías que permiten a las organizaciones realizar cambios en la producción, las cuales son big data y análisis, robots autónomos, simulación, integración horizontal y vertical de sistemas, internet industrial de las cosas, redes de seguridad, nube, fabricación aditiva y realidad aumentada [10]. La gestión adecuada del conocimiento es fundamental para que las organizaciones mejoren sus servicios. La integración exitosa de la IA en los procesos de gestión del conocimiento no solo optimiza la manera en que se adquiere, almacena y utiliza la información, sino que también influye directamente en los resultados de la toma de decisiones [11].

Entre tanto, y en relación con la gobernanza y citando a [12], definen la gobernanza como la realización de relaciones políticas entre diversos actores involucrados en el proceso de decidir, ejecutar y evaluar decisiones sobre asuntos de interés público, proceso que puede ser caracterizado como un juego de poder, en el cual competencia y cooperación coexisten como reglas posibles; y que incluye instituciones, tanto formales como informales. La gobernanza se asume como el ejercicio de dar participación a los actores del sector público, privado y de la sociedad, en los asuntos que los afectan y empoderarlos para que tomen las decisiones que correspondan a la solución de aquellos [13].

En adición, la gobernanza corporativa adquiere una relevancia crucial dentro del ámbito empresarial y no debe ser ignorada ni subestimada, ya que su propósito es asegurar el funcionamiento eficiente y la sostenibilidad económica de cualquier organización a lo largo del tiempo [14].

Es importante tener en cuenta la definición de gobernanza corporativa desde la perspectiva organizacional es

el conjunto de prácticas, formales e informales, que gobiernan las relaciones entre los administradores y todos aquellos que invierten recursos en la empresa, principalmente accionistas y acreedores [...] En estas circunstancias, unas buenas prácticas son clave para el acceso de las empresas a los mercados de capital. [15]

Finalmente, aunque el término gobernanza es aplicable a todo tipo de organizaciones, tanto públicas como privadas, de él se desprenden distintas interpretaciones. En el ámbito privado, suele emplearse el concepto de gobernanza o gobierno corporativos como equivalente de gobernanza, para referirse al

modelo de gestión empresarial, sin importar la forma legal o el tamaño de la organización, pese a que originalmente esta noción fue concebida para describir la administración de grandes corporaciones [16].

VII. METODOLOGÍA

Para el desarrollo de la revisión sistemática de literatura sobre la apropiación de la inteligencia artificial y la gobernanza en el mejoramiento de servicios en establecimientos de ahorro y crédito, la metodología aplicada se centró en la búsqueda, identificación y selección de artículos académicos relevantes. Esta revisión se apoyó en una estrategia mixta, integrando enfoques cualitativos y cuantitativos, siguiendo los lineamientos propuestos por Kitchenham para revisiones sistemáticas [17].

Desde el enfoque cualitativo, se realizó una búsqueda exhaustiva en bases de datos especializadas, aplicando criterios de inclusión y exclusión previamente definidos con el objetivo de garantizar la calidad y pertinencia de las fuentes seleccionadas. Esta etapa permitió reunir estudios clave que abordan la relación entre inteligencia artificial, gobernanza y servicios financieros.

Por su parte, el enfoque cuantitativo complementó el análisis mediante la identificación de patrones y tendencias en la adopción de la inteligencia artificial y su influencia en los modelos de gobernanza en este tipo de entidades. Además, se aplicaron métricas de impacto que facilitaron la evaluación de la relevancia de los estudios más influyentes. Estas herramientas son clave en un contexto donde la transformación digital, impulsada por tecnologías como la inteligencia artificial, está remodelando los servicios financieros, tal como lo destacan Brynjolfsson y McAfee [18].

En cuanto al proceso de búsqueda, se realizaron cuatro indagaciones en bases de datos académicas reconocidas: Emerald, Science Direct y Scopus. Tras el análisis de los resultados obtenidos en la tercera búsqueda, se decidió continuar únicamente con Emerald y Science Direct, dado que estas ofrecieron una mayor cantidad de artículos pertinentes y de alto impacto relacionados con la temática de investigación. En contraste, Scopus fue descartada debido al bajo número de artículos indexados identificados en el área específica abordada.

La combinación de ambos enfoques fortalece la comprensión de lo requerido, proporcionando una visión integral sobre cómo la inteligencia artificial puede contribuir al fortalecimiento de la gobernanza y a la mejora de los servicios en los establecimientos de ahorro y crédito.

V. RESULTADOS PRELIMINARES

En la cuarta búsqueda se determina la codificación de las diferentes combinaciones, nombrándose de la siguiente manera (también ver tabla 1):

- B1: Artificial Intelligence+ Corporate Governance + Finance Services + Credit Unions
- B2: Artificial Intelligence + Corporate Governance + Credit Unions

- B3: Artificial Intelligence + Finance Services + Credit Unions
- B4: Corporate Governance + Finance Services + Credit Unions

RECORDS				
Bases de datos	B1	B2	B3	B4
Emerald	1	15	4	4
Science Direct	1	67	156	121

Tabla 1. Documentos encontrados.

Los artículos identificados se depuraron según dos filtros definidos. Estos son los siguientes:

El filtro 1 permite conservar los artículos de acceso gratuito ("open access") relacionados con el análisis de información de la apropiación de la inteligencia artificial en establecimientos de crédito (IC1); este filtro también permite conservar las publicaciones disponibles en formato electrónico publicadas entre 2014 y 2025 (IC2) y los artículos redactados en inglés y castellano (IC4). Se descartan los artículos sin acceso a su archivo digital (EC1), también se descartan los artículos bibliográficos, trabajos de grados, tesis no investigadas y artículos de revista digital. (EC2).

El filtro 2 permite conservar las publicaciones cuyo texto completo estaba disponible (IC3). También permite eliminar artículos duplicados, como se visualiza en la siguiente imagen:

La figura 1 presenta los resultados obtenidos tras la aplicación de los filtros establecidos en el proceso de selección de literatura. En una primera etapa, se identificaron un total de 369 publicaciones, de las cuales 24 procedían de la base de datos Emerald y 345 de Science Direct.

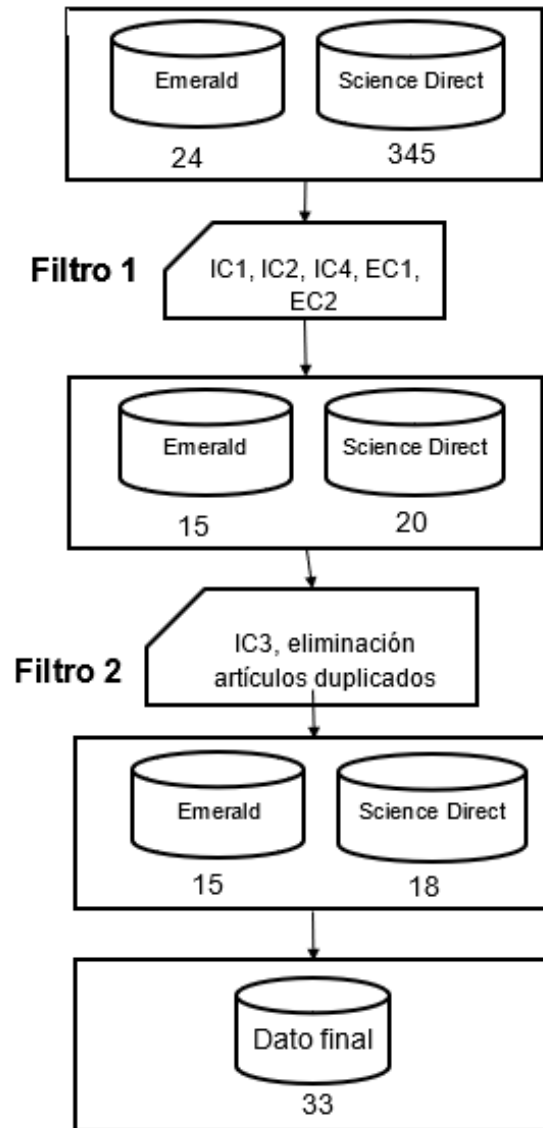


Figura 1. Diagrama de flujo del proceso de selección de artículos. Fuente: Elaboración propia

La implementación del Filtro 1 permitió conservar aquellos estudios que abordaban el análisis de información sobre la apropiación de la inteligencia artificial en establecimientos de crédito, limitando además los resultados a publicaciones disponibles en formato electrónico y comprendidas entre los años 2014 y 2025. Como resultado, se seleccionaron 15 artículos de Emerald y 20 de Science Direct. Posteriormente, el Filtro 2 se aplicó para eliminar duplicados y excluir aquellos documentos cuyo texto completo no estaba disponible, lo que redujo el conjunto a 33 publicaciones: 15 de Emerald y 18 de Science Direct. Finalmente, estos 33 artículos fueron considerados pertinentes para el desarrollo del estudio.

VI. RESULTADOS ESPERADOS

Los resultados esperados de la revisión sistemática literatura va a permitir que se proporcione una comprensión completa sobre como la IA ha sido adoptada en los establecimientos de ahorro

y crédito, identificando sus tendencias, oportunidades y retos en la mejora de los procesos financieros. Además, se examinará los marcos de gobernanza que rigen su implementación, permitiendo evaluar las mejores prácticas que influyen en los triunfos o fracasos en los servicios financieros en establecimientos de ahorro y crédito.

Así mismo, estos resultados servirán como referencia para otras investigaciones relacionadas con el estudio desde las organizaciones, la economía de la empresa, las finanzas y la inteligencia artificial.

Aportes Académicos

Caracterización de la población estudiada en la investigación.

Comprender cómo la apropiación de la IA y la gobernanza contribuyen al mejoramiento de los servicios en los diferentes establecimientos de ahorro y crédito del sector real.

Aportes al Sector Real

Los resultados esperados del trabajo de investigación tienen como objetivo difundir el conocimiento a las comunidades académicas y del sector real, con el fin de generar un impacto en estos actores de la sociedad y mejorar su proceder frente al tema estudiado.

VII. CONCLUSIÓN

En resumen, el trabajo previo proporcionó un avance en las etapas de la revisión sistemática de la literatura, donde se evidencia que la adopción de la inteligencia artificial, junto a una gobernanza empresarial efectiva, constituyendo un elemento crucial para el fortalecimiento de los servicios financieros en las instituciones de ahorro y crédito. A través del análisis de artículos elegidos según criterios estrictos en bases de datos académicas de gran relevancia como Emerald y Science Direct, se encontró información significativa que muestra un interés creciente en el efecto de la IA sobre la eficiencia operativa, la experiencia del cliente y la toma de decisiones estratégicas en el ámbito financiero.

Entre las principales limitaciones del estudio se encuentra la escasa disponibilidad de literatura reciente y específica para este sector. Sin embargo, esta situación enfatiza aún más la necesidad de seguir investigando en esta área.

VIII. REFERENCIAS

- [1] S. Russell y P. Norvig, Inteligencia artificial: Un enfoque moderno (4.^a ed.). Pearson, 2021.
- [2] J. I. C. F. J. V. López, Inteligencia artificial y ética en la gestión pública, Venezuela: Colección Experiencias Escuela CLAD, 2021.
- [3] C. Contreras, "El impacto de la inteligencia artificial en la industria financiera," *Instituto de Analistas*, abr. 2024. [En línea].
- [4] R. Bello Pérez, «La inteligencia artificial en la transformación digital de la sociedad,» *TEMAS - cultura ideología sociedad*, vol. 121, n° 2, pp. 11-17, 2025.
- [5] A. Herrera-Poyatos, J. Del Ser, M. López de Prado, F.-Y. Wang, E. Herrera-Viedma y F. Herrera, «RESPONSIBLE ARTIFICIAL INTELLIGENCE SYSTEMS: A ROADMAP TO SOCIETY'S TRUST THROUGH TRUSTWORTHY AI, AUDITABILITY, ACCOUNTABILITY, AND GOVERNANCE,» 2025. [En línea].
- [6] M. D. C. S. Sierra, Inteligencia artificial en la gestión financiera empresarial. *Pensamiento & Gestión*, (23), 153-186, 2007.
- [7] J. Ospina-Gutiérrez y E. Aristizábal, Aplicación de inteligencia artificial y técnicas de aprendizaje automático para la evaluación de la susceptibilidad por movimientos en masa: *Revista Mexicana de Ciencias Geológicas*, v. 38, núm. 1, p. 43-54, 2021.
- [8] C. Langton y K. Shimohara, «Artificial life,» *Cambridge: The MIT Press*, 1997.
- [9] J. L. Verdegay, M. T. Lamata, D. Pelta y C. Cruz, «Inteligencia artificial y problemas de decisión: la necesidad de un contexto ético,» 2021. [En línea]. DOI: <https://doi.org/10.14349/sumneg/2021.V12.N27.A2>
- [10] R. Blanco, J. Fontodrona y C. Poveda, La industria 4.0: El estado de la cuestión. *Economía Industrial*, 406, 151-164, 2017.
- [11] L. Leoni, G. Gueli, M. Ardolino, M. Panizzon y S. Gupta, «AI-empowered KM processes for decision-making: empirical evidence from worldwide organisations,» 2024. [En línea]. DOI: <https://doi.org/10.1108/JKM-03-2024-0262>
- [12] C. J. Termeer, S. Drimie, J. Ingram, L. Pereira y M. J. Whittingham, A diagnostic framework for food system governance arrangements: The case of South Africa. *NJAS-Wageningen Journal of Life Sciences*, 84, 85-93, 2018.
- [13] L. Lotero, N. Molina, J. Vallejo, G. Hincapié y L. Botero, Territorios socialmente responsables. Una mirada desde el índice sintético del desarrollo sostenible para los municipios del departamento de Antioquia-Colombia, Medellín: Editorial Universidad Pontificia Bolivariana, 2019.
- [14] R. & D. B. B. Royae, «Role of Corporate Governance in Organization. *GSTF International Journal on Business Review (GBR)*, 2(3), 84-89. http://dx.doi.org/10.5176/2010-4804_2.3.224,» 2023.
- [15] Corporación Andina de Fomento, «Lineamientos para un Código Latinoamericano de Gobierno Corporativo,» 2013. [En línea].
- [16] M. J. Romero, «REVISIÓN SISTEMÁTICA DEL CONCEPTO GOBERNANZA COOPERATIVA,» 2020. [En línea]. DOI: <https://doi.org/10.24215/23143738e083>
- [17] B. Kitchenham, «Procedures for Performing Systematic Reviews,» Keele University, 2004.
- [18] E. B. a. A. McAfee, Machine, Platform, Crowd: Harnessing Our Digital Future., New York: W. W. Norton & Company, 2017.

Generación de Imágenes Aplicando Técnicas de Autoencoder y SMOTE.

Juan Manoel Miranda Gómez^{*15}, Carlos Andres Ferro Sanchez²³, Reinel Reinel Tabares-Soto⁴⁵⁶

¹ Instituto Federal do Maranhão, Campus Monte Castelo, Avenida Getúlio Vargas, 04, Monte Castelo, São Luís, 65030-005, Maranhão, Brasil.

²Facultad de Ingeniería y Ciencias Básicas, Universidad Autónoma de Occidente, Cali, 760030, Valle, Colombia.

³Escuela de Ingeniería Eléctrica y Electrónica (EIEE, Facultad de Ingeniería, Universidad del Valle, Cali, 760030, Valle, Colombia.

⁴Departamento de Electrónica y Automatización, Universidad Autónoma de Manizales, Manizales, 170001, Caldas, Colombia.

⁵Departamento de Sistemas e Informática, Universidad de Caldas, Manizales, 170001, Caldas, Colombia.

⁶GobLab Escuela de Gobierno, Universidad Adolfo Ibáñez, Santiago, 7941169, Santiago, Chile.

*Corresponding author(s). Email(s): juan.miranda41303@ucaldas.edu.co

Contributing authors: cferro@uao.edu.co

Resumen — Este documento propone un mecanismo novedoso para la generación de imágenes sintéticas orientado a reducir el desbalance de clases en diferentes tipos de imágenes, incluyendo biomédicas, dígitos manuscritos y objetos naturales. El método desarrollado ofrece una alternativa al aumento de datos tradicional mediante la combinación de un autoencoder convolucional optimizado (VGG5-Autoencoder) y la técnica SMOTE aplicada en el espacio latente del modelo. La propuesta destaca por su facilidad de implementación, alta interpretabilidad y un menor costo computacional respecto a métodos complejos como las Redes Generativas Antagónicas (GANs). Adicionalmente, permite generar imágenes sintéticas de alta calidad, preservando características semánticas relevantes, lo cual se valida mediante métricas cuantitativas (índice SSIM) y comparaciones experimentales en conjuntos de datos públicos y clínicos, mostrando así su potencial aplicación en contextos reales con limitaciones severas de disponibilidad de datos, especialmente en el área biomédica.

Términos clave — Autoencoder, VGG, SMOTE, SSIM, Adam, AdamW, GANs, aumento de datos.

I. INTRODUCCION

La generación de datos sintéticos ha emergido como una estrategia fundamental para mitigar el desbalance de clases en múltiples dominios de aprendizaje automático, especialmente en contextos donde la recolección de datos reales es limitada, costosa o éticamente restringida. Las Redes Generativas Antagónicas (GANs) han sido ampliamente utilizadas en este propósito, mostrando resultados notables en términos de realismo visual. Sin embargo, presentan

desventajas importantes como su complejidad arquitectónica, inestabilidad en el entrenamiento y elevados requerimientos computacionales, lo que limita su aplicabilidad en entornos con recursos limitados. [1], [2].

En este contexto, el presente trabajo propone una alternativa más eficiente, basada en una arquitectura optimizada de autoencoder convolucional (denominada VGG5-Autoencoder), combinada con la técnica de sobre muestreo SMOTE (Synthetic Minority Over-sampling Technique) [3]. La propuesta se enfoca en aplicar SMOTE no sobre las imágenes originales, sino sobre el espacio latente comprimido generado por el autoencoder, aprovechando su capacidad para capturar representaciones semánticas relevantes y reducir la dimensionalidad de entrada.

Este enfoque busca generar imágenes sintéticas de alta calidad con menor costo computacional y mayor interpretabilidad, superando las limitaciones de métodos tradicionales que aplican transformaciones aleatorias o interpolaciones visuales sin control semántico. Además, se presenta un marco metodológico que incluye evaluación cuantitativa mediante el índice SSIM, comparación de optimizadores y funciones de pérdida, y validación experimental con datasets públicos y clínicos. Este trabajo tiene como objetivo demostrar que es posible aumentar datos visuales de manera efectiva y eficiente, con aplicaciones potenciales en escenarios reales como el análisis biomédico, donde el desbalance de clases es crítico.

II. MATERIALES Y MÉTODOS

Combinando arquitecturas de codificador-decodificador convolucionales con técnicas de submuestreo estadístico destinadas a superar las limitaciones de eficiencia computacional de los métodos generativos convencionales, el presente trabajo desarrolla y valida una metodología alternativa para la generación de imágenes sintéticas. Con una reducción significativa en los requisitos de recursos computacionales y el tiempo de procesamiento, el enfoque metodológico se basa en la hipótesis de que la aplicación de SMOTE (Técnica de Submuestreo de Minorías Sintéticas) [3] en espacios de representación latente de alta dimensión puede generar muestras sintéticas de calidad comparable a las producidas por Redes Generativas Antagónicas (GANs).

Cuatro componentes metodológicos principales definen el diseño experimental: (1) la evaluación cruzada utilizando conjuntos de datos de referencia y especializados, incluidos conjuntos de datos estándar de visión por computadora y un conjunto clínico de proyecciones CBCT; (2) el desarrollo de una arquitectura óptima de autoencoder (VGG5Autoencoder) con configuraciones específicas para diferentes dominios de imagen; (3) la implementación adaptada de SMOTE [3] para operar en espacios latentes comprimidos; y (4) la evaluación cuantitativa utilizando métricas de similitud estructural (SSIM) [4] y análisis comparativo de eficiencia computacional. Bajo condiciones controladas de validación experimental, este marco metodológico permite la caracterización metódica del rendimiento del enfoque propuesto.

A. Selección y Caracterización de Conjuntos de Datos

Se emplearon tres conjuntos de datos conocidos en visión por computadora, cuidadosamente elegidos según sus características morfológicas y complejidad representacional. CIFAR-10 proporciona una excelente variabilidad intraclase con una complejidad textural modesta mediante 60,000 imágenes a color ($32 \times 32 \times 3$ píxeles) distribuidas en 10 clases equilibradas Fig. 1. Considerado un referente fundamental para la evaluación de técnicas de generación sintética, MNIST consiste en una baja complejidad estructural y una destacada consistencia morfológica de 70,000 imágenes de dígitos manuscritos en escala de grises ($28 \times 28 \times 1$ píxel) Fig. 2. Usando 10 clases de prendas de vestir, Fashion-MNIST preserva la estructura dimensional de MNIST, pero añade complejidad intermedia y variabilidad morfológica moderada, permitiendo así la evaluación de la robustez del método propuesto frente a varios grados de complejidad semántica Fig. 3. Mantener la distribución normal de 60,000 muestras de entrenamiento y 10,000 muestras de validación garantiza la comparabilidad con investigaciones anteriores en la literatura.

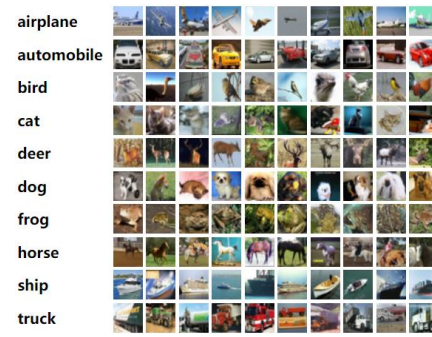


Figura. 1. Cifar 10

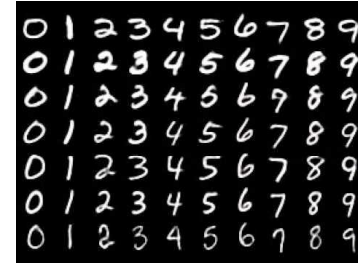


Figura. 2. MNIST



Figura. 3. Fashion MNIST

A demás de estas bases de datos públicas, se utilizó una base de datos propia que consistió en 500 proyecciones 2D de la desviación estándar de intensidad tomadas de volúmenes de CBCT, las cuales fueron corregidas en posición natural de la cabeza (PNC), que mostraban desequilibrio de clases entre necesidad de extracción y no extracción en tratamientos de ortodoncia. Sumado a esto, las imágenes cuentan con una buena resolución espacial que permite la distinción entre tejidos. La selección de este conjunto de datos responde a la necesidad de validar la metodología propuesta en entornos del mundo real con severas limitaciones en la disponibilidad de datos, una condición común en aplicaciones biomédicas donde obtener muestras presenta una gran dificultad tecnológica y ética. La adquisición de esta base de datos fue aprobada por el comité de ética correspondiente (Universidad Autónoma de Occidente, Acta No. 01-2024, 12 de febrero de 2024), y todos los datos de los pacientes fueron anonimizados antes del análisis, en cumplimiento con las normas de la Declaración de Helsinki. Fig. 4.

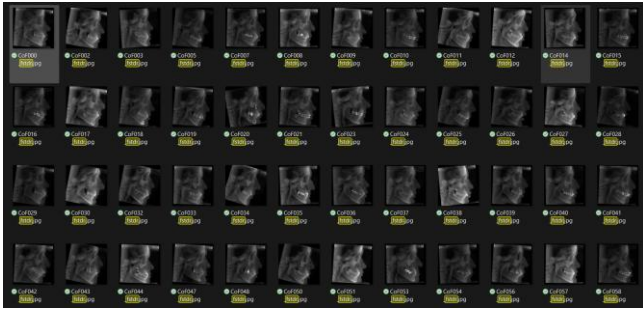


Figura. 4. CBCT

Una vez seleccionadas las bases de datos, a esta se les realizo un preprocesamiento que consistió en:

- 1) **Normalización:** se estandarizo una distribución homogénea en los conjuntos de datos en el rango 0 y 1, para optimizar el proceso de entrenamiento del autoencoder convolucional escalando las intensidades mediante una transformación lineal. [5].
- 2) **Preprocesamiento específico:** para el conjunto de datos clínicos, ya que las proyecciones 2D de los volúmenes de CBCT revelaron una heterogeneidad dimensional inicial, se realizó el cambio de tamaño a una resolución estándar de 512×512 píxeles.

B. Desarrollo de Arquitectura Autoencoder Optimizada: VGG5Autoencoder

- 1) **Diseño Convolutivo:** La arquitectura propuesta, denominada VGG5-Autoencoder, como se muestra en la TABLA I, se desarrolló como una adaptación optimizada de la topología VGG16 [6]. Este diseño se fundamenta en principios de reducción dimensional progresiva mediante operaciones de convolución y pooling en la etapa del encoder; esta implementa una secuencia de 5 bloques convolucionales con un incremento progresivo en el número de filtros. La compresión dimensional se logra mediante MaxPooling, generando un espacio latente comprimido de dimensionalidad variable según las dimensiones de entrada. Por otra parte, el decoder emplea una topología simétrica inversa, usando operaciones convoluciones transpuestas para recuperar la resolución espacial original.

TABLA I
ESTRUCTURA DE LOS MODELOS

DataSet	Encoder					Espacio Latente	Decoder						
MNIST	16	32	64	128	128	128	128	128	64	32	16	4	1
Fashion MNIST	16	32	64	128	128	128	128	128	64	32	16	4	1
CIFAR10	16	32	64	128	128	128	128	128	64	32	16	4	3
FSTD	32	64	128	256	256	36864	256	256	128	64	32	5	1
	Convs 2D						Transpose Convs 2D				Convs 2D		

El espacio latente creado es la base que guarda las características semánticas importantes, que luego se utiliza en la técnica de SMOTE [3] para la creación de datos sintéticos.

2) Entrenamiento y optimización

La optimización del modelo se realizó mediante la implementación de una función de perdida basada en similitud estructura (SSIM) [4], seleccionada por su capacidad de evaluación de características morfológicas. El optimizador utilizado fue AdamW [7] debido a su estabilidad en términos de convergencia y rendimiento consistente, también se ajustó AdamW [7] con los hiperparametros de la TABLA II.

TABLA II
HIPERPARAMETROS ADAMW

AdamW Parametros	Values
learning_rate	0,001
wight_decay	0,004
beta_1	0,9
beta_2	0,999
epsilon	1,00E-07
amsgrad	FALSE
use_ema	FALSE
ema_momentum	0,99

Los criterios de convergencia se establecieron mediante un umbral de SSIM Loss < 0.1 [4], determinado como un indicador empírico de calidad para una reconstrucción fidedigna de la imagen. En la TABLA III muestra los resultados de entrenamiento, donde vemos que las imágenes de MNIST, Fashion MNIST y las imágenes CBCT tuvieron mejores resultados.

TABLA III
Resultados de entrenamiento

Imágenes	Optimizer	Loss Function	Epochs	Batch Size	Delta Time MM:SS	Result Loss Funcion	MSE
MNIST	AdamW	SSIMLoss	32	512	02:54	0,0409	0,0059
Fashion MNIST	AdamW	SSIMLoss	32	512	02:35	0,0779	0,0128
CIFAR 10	AdamW	SSIMLoss	32	512	04:50	0,2473	0,0131
CBCT	AdamW	SSIMLoss	8	1	02:48	0,0659	0,00082

C. Implementación Adaptada de SMOTE en Espacios Latentes Comprimidos

La metodología novedosa propuesta en este estudio se centra en la adaptación de la técnica SMOTE (Synthetic Minority Over-sampling Technique) [3], para operar en espacios de representación latente de alta dimensionalidad generados por la VGG5-autoencoder, aprovechando las propiedades de la compresión semántica del encoder para facilitar la creación de vectores sintéticos mediante la interpolación controlada entre K-vecinos más cercanos en el espacio latente comprimido.

Este proceso se estructuro en 4 fases secuenciales: 1) transformación de las imágenes originales al espacio latente., 2) análisis de baja densidad para identificar regiones de baja representación correspondientes a clases minoritarias, 3) aplicación específica de SMOTE [3] mediante interpolación K-

NN Fig. 5) decodificación de los vectores sintéticos para reconstrucción de las imágenes. Fig. 6.

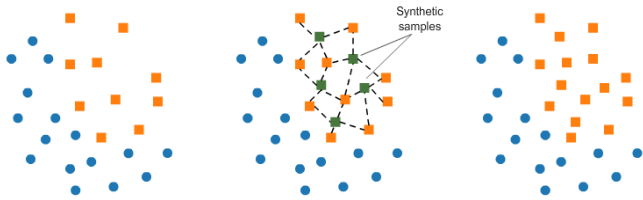


Fig. 5. Balanceo de Clases utilizando Smote

D. Evaluación Cuantitativa y Protocolo de Validación Experimental

1) Evaluación de similitud estructural

La principal herramienta de evaluación para la evaluación cuantitativa fue el Índice de Similitud Estructural (SSIM), que combina la evaluación del brillo, el contraste y los elementos estructurales para generar una evaluación perceptivamente significativa de la calidad reconstructiva. Operando dentro del rango $[0, 1]$, esta estadística permite la cuantificación de la retención de patrones morfológicos y cualidades texturales en las imágenes sintéticas producidas [4].

2) Metodologías de Análisis Comparativo y Validación

El diseño experimental utilizó una técnica de validación cruzada a través de la creación controlada de falsos desequilibrios en los conjuntos de prueba, lo que permitió una evaluación objetiva de la capacidad del procedimiento para producir muestras sintéticas de clases minoritarias en entornos experimentalmente controlados [8]. En comparación con los métodos generativos convencionales, los requisitos de rendimiento establecieron niveles aceptables con $SSIM > 0.9$ [4] para aplicaciones de aumento general y duraciones de generación de menos de 5 minutos para 1,000 imágenes sintéticas, mostrando así una mejor eficiencia computacional con un factor entre 5 y 10 veces

III. RESULTADOS

Tras el entrenamiento y evaluación de los modelos generativos, se obtuvo una calidad destacable en las imágenes generadas para los conjuntos de datos **MNIST** y **Fashion-MNIST**, logrando una reproducción visual clara y coherente con las clases objetivo. En contraste, las imágenes correspondientes al conjunto **CIFAR-10** presentaron un grado mayor de borrosidad, atribuible tanto a su mayor complejidad visual como a la resolución de entrada del dataset.

Respecto a las imágenes generadas a partir de proyecciones 2D de la desviación estándar de intensidad extraídas de volúmenes de **CBCT**, se observó una ligera pérdida de nitidez. Sin embargo, estas imágenes resultaron ser suficientemente claras para los fines del problema planteado, permitiendo distinguir patrones relevantes asociados a las clases objetivo.

Se llevó a cabo una evaluación cualitativa de las imágenes generadas, centrada en verificar si la clase asignada por la técnica **SMOTE** [3] se correspondía visualmente con la imagen

sintetizada. Esta validación consideró tanto la nitidez como el realismo percibido en las muestras. Para ello, se analizaron manualmente imágenes aleatorias de cada clase, identificando en la mayoría de los casos una coherencia satisfactoria. Se plantea como trabajo futuro el desarrollo de clasificadores automáticos para realizar esta evaluación de forma sistemática, en especial sobre imágenes cuya clasificación no coincida con la clase generada por **SMOTE** [3].

Adicionalmente, se analizaron los efectos de los **optimizadores** y **funciones de pérdida** en la estabilidad del entrenamiento y en la calidad de las imágenes generadas. El optimizador **Lion** presentó comportamientos menos predecibles, como convergencia prematura o sobreajuste en etapas tempranas llegan a un SSIM Loss en el validation de 0.5 lo que lo hace generar imágenes con baja nitidez o imágenes completamente negras para los de **MNIST**. En cambio, **Adam** [9] y **AdamW** [7] mostraron una mayor estabilidad y permitieron una reproducción más consistente de los resultados. Algo similar se observó con la función de pérdida **MSELoss**, la cual condujo a una convergencia más lenta y a imágenes de menor calidad visual, llegando a después de 32 épocas llegar con un Loss de 3, o en mejores casos de 1.5, y su contraparte **SSIMLoss** en 32 épocas llegaba a valores < 0.1 .

Tras la selección adecuada de las combinaciones de modelo, función de pérdida y optimizador, se logró generar imágenes aceptables para cada uno de los conjuntos de datos utilizados. La Fig. 6 presenta ejemplos representativos de los resultados obtenidos.

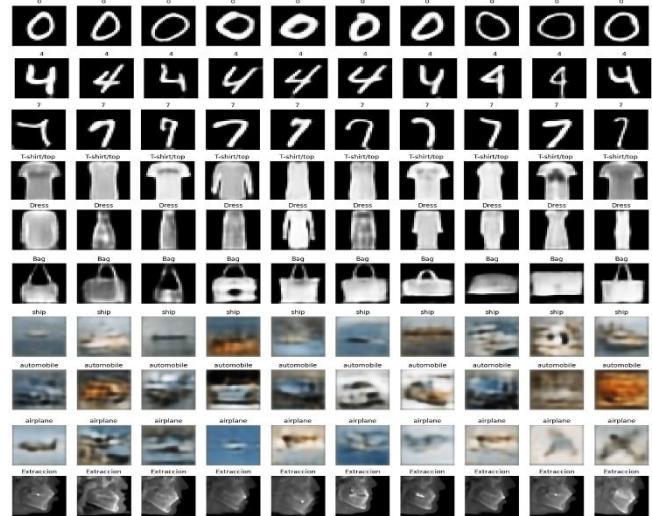


Fig. 6. Imágenes Generadas

IV. CONCLUSIÓN

La combinación de un autoencoder convolucional optimizado (VGG5-Autoencoder) con **SMOTE** en el espacio latente permite generar imágenes sintéticas de alta fidelidad y control semántico, preservando rasgos morfológicos clave y logrando un balanceo de clases dirigido. Este enfoque ofrece una calidad visual comparable a la de los métodos basados en GANs, pero con una reducción de hasta diez veces en el coste computacional, lo que facilita su aplicación en entornos con recursos limitados. Su simplicidad e interpretabilidad lo hacen fácilmente integrable en flujos de trabajo existentes, resultando especialmente valioso en contextos biomédicos donde la

disponibilidad de datos está restringida por razones éticas o técnicas. Como líneas futuras, se recomienda validar el método en imágenes volumétricas 3D, desarrollar métricas automáticas de evaluación de fidelidad mediante clasificadores supervisados y explorar variantes de autoencoders (por ejemplo, variacionales) junto con técnicas de sobremuestreo avanzadas para enriquecer aún más la diversidad de los datos generados.

REFERENCIAS

- [1] A. Radford, L. Metz, and S. Chintala, “Unsupervised representation learning with deep convolutional generative adversarial networks,” *arXiv preprint arXiv:1511.06434*, 2015.
- [2] M. Lucic, K. Kurach, M. Michalski, S. Gelly, and O. Bousquet, “Are GANs created equal? A large-scale study,” *arXiv preprint arXiv:1711.10337*, 2017.
- [3] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic minority over-sampling technique,” *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.
- [4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [5] Y. LeCun, L. Bottou, G. B. Orr, and K.-R. Müller, “Efficient backprop,” in *Neural networks: Tricks of the trade*, Springer, 1998, pp. 9–50.
- [6] K. Simonyan and A. Zisserman, “Very deep convolutional networks for large-scale image recognition,” *arXiv preprint arXiv:1409.1556*, 2014.
- [7] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2019. [Online]. Available: <https://arxiv.org/abs/1711.05101>
- [8] R. Kohavi, “A study of cross-validation and bootstrap for accuracy estimation and model selection,” in *Proc. Int. Joint Conf. Artificial Intelligence (IJCAI)*, 1995, pp. 1137–1145.
- [9] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint, arXiv:1412.6980*, 2014. [Online]. Available: <https://arxiv.org/abs/1412.6980>
- [10] M. Frid-Adar, E. Klang, M. Amitai, J. Goldberger, and H. Greenspan, “GAN-based synthetic medical image augmentation for increased CNN performance in liver lesion classification,” *Neurocomputing*, vol. 321, pp. 321–331, 2018.



Juan Manoel Miranda Gómez nació en Nova Lima, Minas Gerais, Brasil, el 26 de abril de 2004. Recibió su título de técnico en informática del Instituto Federal de Ciência e Tecnologia do Maranhão (IFMA), Brasil, en 2022. Actualmente, es estudiante de ingeniería en sistemas en la Universidad de Caldas, Colombia. Él trabaja como Desarrollador Freelance y Auxiliar Contable. Sus intereses de investigación actuales y anteriores se centran en la inteligencia artificial, siendo miembro del semillero en IA de la Universidad de Caldas desde marzo de 2024.

El Sr. Miranda Gómez fue miembro del equipo ganador del hackathon del Hackday 2024, patrocinado por la empresa Indra. No ha reportado membresías en sociedades profesionales distintas a IEEE ni trabajo para comités o publicaciones de IEEE.

Carlos Andres Ferro Sanchez, photograph and biography not available at the time of publication.



REINEL TABARES-SOTO recibió el primer título de Ingeniero Electrónico de la Universidad Nacional de Colombia en 2009, el segundo título de Ingeniero en Sistemas y Computación de la Universidad de Caldas, Colombia, en 2016, la maestría en Ingeniería Electrónica de la Universidad Nacional de Colombia en 2017, y el doctorado en Ingeniería de la Universidad Autónoma de Manizales, Colombia, en 2021. Actualmente es Profesor en la Facultad de Inteligencia Artificial e Ingenierías de la Universidad de Caldas y la Universidad Autónoma de Manizales, Colombia. Sus principales intereses de investigación incluyen el esteganálisis, el aprendizaje automático, el aprendizaje profundo, la bioinformática y la computación de alto rendimiento.

Development and Comparative Evaluation of Machine Learning and Deep Learning Models for Energy Consumption Forecasting in Colombia's Non-Interconnected Zones

Johan Nicolas Cuellar Salinas, Elvis Eduardo Gaona García, Andrés Camilo López García

Abstract—Non-Interconnected Zones (ZNI) in Colombia present significant challenges for reliable and sustainable energy access, primarily due to reliance on often intermittent and costly generation sources. Accurate energy consumption forecasting is critical for effective energy planning and the integration of sustainable solutions in these regions. However, conventional forecasting models often struggle with predictions on systems with unique characteristics like the ones of the ZNI, including distinct consumption patterns and possible data limitations. While Machine Learning (ML) and Deep Learning (DL) offer advanced predictive capabilities, this research addresses the technical challenge of specifically adapting and evaluating these models for robust energy consumption forecasting given the unique characteristics and possible data limitations of isolated Colombian ZNI communities.

This research focuses on the development, adaptation, and evaluation of an energy consumption forecasting model for Colombian ZNI, primarily utilizing Long Short-Term Memory (LSTM) networks. The methodology involves characterizing ZNI consumption dynamics, identifying key predictive variables (climatic, socio-economic, operational), and implementing specialized data handling and feature engineering techniques suited to the ZNI data environment. The LSTM model will be specifically adapted to these conditions, and its predictive performance (e.g., using MAE, RMSE) will be rigorously evaluated and compared against traditional simple linear forecasting methods to assess its relative effectiveness and robustness. The core technical objective is to engineer and validate an LSTM based model tailored to achieve high predictive performance within the ZNI context. The anticipated outcome is a technically sound predictive tool capable of generating accurate energy consumption forecasts, thereby providing a more reliable foundation for local energy system planning and the dimensioning of energy infrastructure.

Index Terms—Non-Interconnected Zones (ZNI), Colombia, energy consumption forecasting, Machine Learning, Deep Learning, Long Short-Term Memory (LSTM), energy planning

I. INTRODUCTION AND CONTEXT

Reliable and sustainable access to electricity remains a significant challenge in Colombia's Non-Interconnected Zones (ZNI), which are characterized by geographical isolation and fragmented, limited energy infrastructure. Communities residing in these areas often depend on costly, intermittent, and unsustainable generation systems, restricting local development and creating uncertainty in the planning of more efficient and context appropriate energy solutions.

In this scenario, the ability to accurately forecast energy consumption, as highlighted in previous studies on rural and isolated systems, [1], is crucial for guiding strategies related to the planning, operation, and sizing of alternative electrical systems, such as the integration of renewable technologies, hybrid systems, or community microgrids. However, predicting energy demand in these zones is particularly challenging due to distinctive consumption patterns, limited availability and quality of data, and high variability in contextual factors related to climate, socioeconomic, and operational aspects.

Traditional forecasting methodologies tend to show limitations when faced with the complexity of ZNI communities as shown in the study [2], where consumption dynamics often respond to nonlinear relationships and scenarios of incomplete information. In response to these challenges, approaches based on machine learning and deep learning—especially those designed for time series analysis, such as LSTM neural networks [3]—offer opportunities to increase predictive capacity and provide forecasts more closely aligned with local realities.

This research proposes to address these issues through the development and adaptation of predictive models based on machine learning and deep learning, considering the particularities of energy consumption in isolated communities, the identification of functional and data requirements, and the selection and adjustment of suitable model architectures for the limited data environment characteristics of ZNI communities.

Furthermore, the study seeks to rigorously assess the capability of these models to generate sufficiently accurate and detailed forecasts to support decision making in the operational planning and sizing of rural energy systems.

In this way, the work aims to contribute to strengthening the technical tools available for energy management in Colombia's Non-Interconnected Zones, providing informed decision making that supports sustainable solutions tailored to local needs.

The proposed research will be conducted using a structured three-phase methodology, as visualized in the flowchart below. The process begins with *Phase 1*, which involves data collection and preprocessing to create a robust dataset that captures the unique context of the ZNI. *Phase 2* focuses on the development and comparative evaluation of state-of-the-art machine learning models, such as LSTM and Random Forest, to identify the most accurate predictor. Finally, *Phase 3* consists of a rigorous validation of the best-performing model to ensure its practical utility and reliability as a tool for effective energy planning

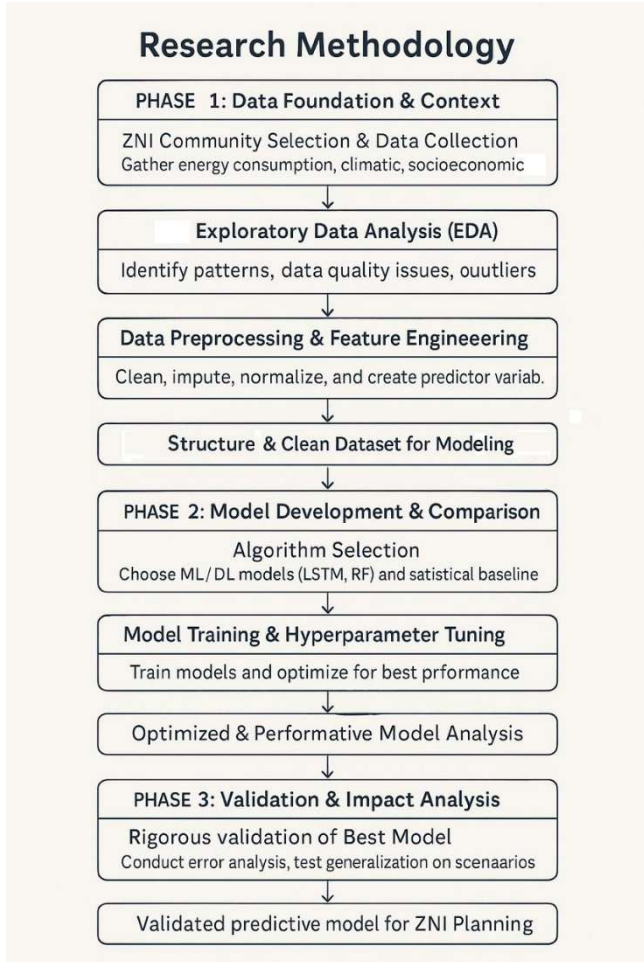


Fig 1. Flowchart of the research methodology

II. DATA COLLECTION AND ANALYSIS

The identification and selection of relevant data for energy

consumption forecasting in Colombia's Non-Interconnected Zones (ZNI) is currently being guided by a comprehensive review of the state of the art literature. This process aims to determine which data sources and variables are most instrumental for developing accurate predictive models tailored to the distinct context and constraints of isolated communities.

Potentially useful data for energy consumption prediction in ZNI settings encompass several categories:

- **Historical energy consumption records:** Hourly, daily, or monthly consumption data, even if limited, serve as the primary foundation for temporal modeling and trend analysis.
- **Climatic and meteorological variables:** Local temperature, humidity, solar irradiation, wind speed, and precipitation influence both the availability of renewable resources and consumption patterns, especially in communities with variable climates or dependence on electric heating/cooling [4], [5].
- **Socioeconomic indicators:** Population size, household income levels, education, and appliance ownership can affect overall demand and usage profiles.
- **Operational and infrastructure characteristics:** Details about the local generation mix (e.g., diesel, solar, hybrid systems), storage capacity, system outages, energy tariffs, and energy losses provide context for interpreting consumption fluctuations and modeling system constraints.
- **Behavioral and event-based data:** Records of community activities, agricultural cycles, market days, or local events may explain periodic variations or anomalies in energy usage.
- **Geographical and accessibility factors:** Data on remoteness, road connectivity, and proximity to service centers may help contextualize distinct usage patterns or infrastructure challenges [6].

The availability and quality of these data types can vary significantly across different ZNI communities due to limitations in measurement infrastructure and record keeping. Accordingly, the data collection strategy prioritizes variables that are both impactful for forecasting accuracy and realistically obtainable in the ZNI context. Ongoing literature review and expert consultation are informing the iterative selection and refinement of relevant features for subsequent modeling stages.

III. MACHINE LEARNING AND DEEP LEARNING MODELS FOR ENERGY CONSUMPTION FORECASTING

A wide range of machine learning (ML) and deep learning (DL) models have been explored in the literature for the task of energy consumption forecasting as highlighted in the study [7]. The choice of model typically depends on the characteristics of the available data, the temporal resolution required, and the

complexity of consumption patterns in the target environment. Conventional approaches include regression-based techniques such as linear regression, autoregressive integrated moving average (ARIMA) [1], and exponential smoothing methods. While these models can effectively capture linear trends and periodicities, their performance is often limited when dealing with non-linear or highly variable energy demand profiles, such as those found in remote or data constrained communities.

To overcome these limitations, various ML models have been increasingly adopted, including:

- **Decision Tree based methods:** Models such as Random Forest (RF) and Gradient Boosting Machines (GBM) offer robust performance for handling heterogeneous data and capturing non-linear relationships between predictors and energy demand.
- **Support Vector Regression (SVR):** Effective for regression tasks with limited data and high dimensional feature spaces.
- **k Nearest Neighbors (kNN):** Utilized in some contexts for short term load forecasting due to its simplicity and adaptability to local patterns.

With the increasing availability of more extensive and higher frequency datasets, deep learning models have gained prominence for energy consumption forecasting, particularly for their ability to automatically extract complex temporal and contextual patterns. Notable DL models used include:

- **Recurrent Neural Networks (RNN):** Designed for sequence data, though traditional RNNs may suffer from vanishing gradient issues with long sequences.
- **Long Short Term Memory Networks (LSTM):** A specialized form of RNN, LSTM networks [3] are widely employed due to their effectiveness in learning long term dependencies and temporal structures in energy consumption time series.
- **Gated Recurrent Units (GRU):** Similar to LSTM but with a simplified architecture, GRUs are also applied in load forecasting scenarios.
- **Convolutional Neural Networks (CNN):** Occasionally integrated or combined with RNN/LSTM architectures to extract local temporal features in multi-dimensional energy datasets.

The current study is leveraging insights from this body of work to identify, select, and adapt ML/DL models that are most suitable for the specific challenges of energy consumption prediction in Colombia's ZNI, considering factors such as data limitations, variability in consumption patterns, and the operational context of isolated communities. Models are evaluated not only for predictive accuracy, but also for their interpretability and practical feasibility in the ZNI context.

IV. CONCLUSIONS

This work presents the ongoing stages of a research project aimed at addressing the challenges of forecasting energy consumption in the Non-Interconnected Zones (ZNI) of Colombia. By conducting a thorough state of the art review and systematically identifying the most relevant data sources and predictive variables, we lay a methodological foundation for developing models tailored to the unique conditions and data limitations of isolated communities.

While the implementation, adaptation, and evaluation of various machine learning and deep learning models are currently underway, the central hypothesis guiding this research remains clear: if a predictive model based on machine learning or deep learning is specifically adapted to the available data and contextual characteristics of the ZNI communities, it will be possible to achieve estimations of electricity consumption with sufficient accuracy and practical utility. Such improvements in predictive capability will, in turn, expand the quality and quantity of data available for planning and sizing local energy solutions.

The anticipated outcomes of this research aspire to provide a validated and context-sensitive methodology for energy consumption forecasting in data scarce and operationally complex environments. Successfully demonstrating the hypothesis would represent a significant contribution to the informed design and implementation of efficient, sustainable, and context appropriate energy systems for Colombia's isolated communities.

Further work will focus on completing the modeling, rigorous performance evaluation, and practical validation phases to confirm or refine these initial premises.

V. ACKNOWLEDGMENTS

The Universidad Distrital Francisco José de Caldas provides financial support to the project Artificial Intelligence for the Identification of Energy Potentials in Off-Grid Communities in Colombia, Code : PIICI-2024-2032.

VI. REFERENCES

- [1] Ahmed, S. Ali, A. Raza, I. Hussain, A. Bilal, N. L. Fitriyani, Y. Gu, and M. Syafrudin, "Novel deep neural network architecture fusion to simultaneously predict short-term and long-term energy consumption," *PLoS ONE*, vol. 20, 1 2025.
- [2] [2] M. Matos, J. Almeida, P. Gonçalves, F. Baldo, F. J. Braz, and P. C. Bartolomeu, "A machine learning-based electricity consumption forecast and management system for renewable energy communities," *Energies*, vol. 17, 2 2024.
- [3] [3] H. Alizadegan, A. Radmehr, H. Karimi, and M. A. Ilani, "Comparative study of long short-term memory (lstm), bidirectional lstm, and traditional machine learning approaches for energy consumption prediction," 5 2024. [Online]. Available: <https://www.preprints.org/manuscript/202405.1020/v1>
- [4] K. Bogner, F. Pappenberger, and M. Zappa, "Machine learning techniques for predicting the energy consumption/production and its uncertainties driven by meteorological observations and forecasts," *Sustainability (Switzerland)*, vol. 11, 2019.

- [5] H. C. Bloomfield, D. J. Brayshaw, M. Deakin, and D. Greenwood, "Hourly historical and near-future weather and climate variables for energy system modelling," *Earth System Science Data*, vol. 14, pp. 2749–2766, 6 2022.
- [6] A. B. K. Kpomonè, A. B. K. Kpomonè, A. B. K. Kpomonè, P. E. T. Gnadi, P. E. T. Gnadi, P. E. T. Gnadi, B. Yao, B. Yao, B. Yao, K. Dosseh, and N. Komla, "Multiple linear regression to predict electrical energy consumption based on meteorological data: Application to some sites supplied by the ceb in togo," *American Journal of Applied Sciences*, vol. 21, pp. 15–27, 1 2024.
- [7] D. S. Khafaga, E. S. M. El-Kenawy, A. A. Alhussan, and M. M. Eid, "Forecasting energy consumption using a novel hybrid dipper throated optimization and stochastic fractal search algorithm," *Intelligent Automation and Soft Computing*, vol. 37, pp. 2117–2132, 2023.

Artificial Intelligence in Non-Interconnected Zones: Challenges and Solutions for Resource Management

A. Márquez, S. Carvajal, and D. López

Abstract— In remote and off-grid regions, the challenge of unstable energy supply due to the unpredictability of renewable resources and variable demand is significant. This research investigates the potential of artificial intelligence (AI) to enhance resource management and frequency control in these isolated areas, aiming to propose solutions that improve system stability. A systematic literature review paired with system dynamics analysis was conducted to identify pertinent AI techniques and assess their feasibility in settings marked by technical and socioeconomic limits. The findings indicate that AI can facilitate the forecasting of generation and demand, optimize energy storage, and adapt in real-time to fluctuations, which collectively bolster frequency stability. Nevertheless, the application of AI faces obstacles, including insufficient historical data and inadequate monitoring and control systems. The system dynamics model further demonstrated that incorporating AI fosters positive cycles for both frequency stability and resource management. Ultimately, the research concludes that AI methods strike a suitable balance between accuracy and practicality. However, the scalability of these solutions is contingent upon various factors that emphasize technical, social, and economic interoperability, which are crucial for achieving sustainable energy transitions in isolated regions.

Index Terms—artificial intelligence, frequency control, isolated areas, resource management, system dynamics.

I. INTRODUCTION

IN isolated and off-grid areas, the lack of infrastructure and the variability of renewable sources cause instability in energy supply particularly in rural and remote communities where infrastructure is limited or nonexistent. In such regions, traditional energy solutions, such as diesel generators or standalone photovoltaic systems, are commonly used due to their ease of deployment. However, these systems often struggle with inefficiency, high operating costs, and significant environmental impact, which has driven interest in more sustainable alternatives like the integration of microgrids and renewable energy sources [1], [2].

Unlike microgrids which are designed to manage multiple sources of generation, storage, and demand participation isolated areas typically rely on decentralized generation, such as individual solar photovoltaic systems or small generators [3]. This disconnect poses unique challenges for system stability, especially regarding frequency regulation, which is essential to ensure a reliable electricity supply.

In this context, the variability of renewable sources like solar and wind, along with the absence of real-time monitoring

infrastructure, limits the effectiveness of conventional control strategies. This has opened the door to advanced approaches, including the implementation of artificial intelligence (AI), which may offer adaptive solutions to optimize operations in these environments [4].

However, applying AI in off-grid settings involves three major challenges: (1) lack of historical data, (2) limited communication infrastructure, and (3) high variability in generation and demand [4]. Furthermore, the socioeconomic and geographic conditions of these regions such as low population density and dispersed users—make conventional solutions difficult to implement [5]. Despite these obstacles, AI emerges as a promising tool to overcome the constraints of traditional systems. Through predictive control strategies and machine learning, AI can optimize the use of local energy resources, improve accuracy in demand-generation forecasting, and help maintain supply stability even in scenarios with low continuity or intermittent renewable generation [6], [7].

II. AI MODELS FOR ENERGY MANAGEMENT IN ISOLATED AREAS

The non-interconnected zones of Colombia and Latin America encompass vast territories, often covering areas comparable to small provinces or even entire departments. In many cases, low population density and limited accessibility hinder the delivery of basic services. These areas range from scattered rural communities to regions requiring complex resource management across large geographic expanses. They also face high levels of unmet energy needs. Ensuring a continuous electricity supply in these territories presents a valuable opportunity to drive educational, healthcare, and economic development, ultimately improving quality of life and reducing social inequalities [3].

Isolated zones present technical challenges, including frequency instability and resource management within these systems [8]. These areas face unique issues such as the absence of historical data, limited communication infrastructure, and high variability in both generation and demand [6], [9]. Frequency control in standalone systems, in particular, relies on adaptive strategies to compensate for renewable intermittency and the low inertia of diesel generators. Moreover, conventional control models—designed for stable environments—often fail to handle the nonlinear dynamics typical of off-grid systems, such as abrupt load fluctuations or equipment failures [10],

[11].

Current approaches rarely consider the socioeconomic constraints of these regions, where low population density and geographic dispersion increase the cost of monitoring and maintenance. This reality limits the scalability of machine learning-based solutions, which require simplified frameworks that strike a balance between accuracy and operational feasibility [5], [12].

In these contexts, an energy resource management system is essential for coordinating operations, optimizing power use, managing loads, and monitoring the system [12]. When integrated into energy management, AI enhances operational efficiency and improves the sustainability of hybrid systems. It also enables smarter scheduling of operations for example, by predicting demand for activities such as water pumping [13]. Additionally, AI algorithms support the operation of storage systems, intelligently managing battery charge and discharge cycles to align with demand and renewable energy availability, thereby ensuring a more reliable energy supply [14]. Incorporating AI techniques makes these systems more robust and efficient overall, covering everything from resource assessment to system operation and maintenance [15].

Therefore, to maintain system stability and frequency control in isolated areas, advanced AI-based control techniques prove useful in hybrid energy systems [16], [17]. These strategies allow systems to adapt to changing conditions, uphold grid stability, improve resilience to sudden changes, and enhance overall reliability through distributed control functions, predictive modeling, daily scheduling, and optimization algorithms [13], [18].

III. METHODOLOGY

Figure 1 illustrates the systematic search strategy applied to academic databases, using filters designed to identify literature on artificial intelligence (AI) applied to resource management and frequency control in rural electrification and remote areas. This approach enabled the selection of relevant studies addressing the technical and operational challenges of isolated systems.

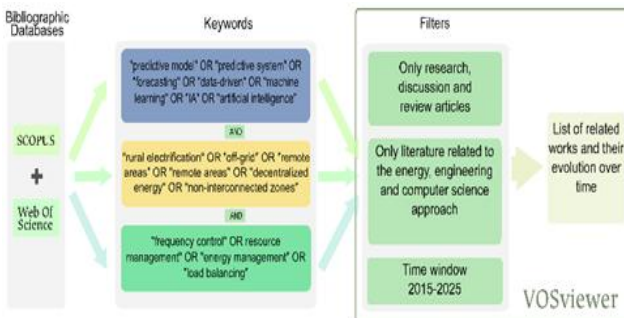


Fig. 1. Systematic search strategy in academic databases.

This structured methodology ensured a focused review of key publications, prioritizing studies that bridge AI applications with the unique demands of rural electrification. The selected literature forms the foundation for analyzing technical gaps and viable solutions in decentralized energy systems.

IV. RESULTS OF THE STATE-OF-THE-ART REVIEW

The literature analysis demonstrates that AI can also incorporate information about the condition of generation equipment, such as the degradation of solar panels or the efficiency of wind turbines, to anticipate reductions in production capacity before critical failures occur. Techniques such as reinforcement learning enable models not only to predict generation and demand, but also to learn how to optimize energy management over time and enhance responsiveness to unexpected variations in the grid [7]. This advanced prediction capability not only helps prevent frequency imbalances but also allows for efficient scheduling of energy storage and distribution strategies. Consequently, it reduces the need for manual interventions and improves the system's operational stability [7].

When properly trained, these AI models can make autonomous real-time decisions to maintain the balance between generation and demand [19]. For instance, if the system predicts an increase in consumption due to a heatwave, it can anticipate this by activating energy storage systems or redistributing load across different generation sources before a frequency imbalance occurs. Likewise, predictive AI-based control can generate adjustment signals for inverters and microgrid regulators, ensuring that generated energy is distributed optimally without compromising system stability [20]. For instance, if the system predicts an increase in consumption due to a heatwave, it can anticipate this by activating energy storage systems or redistributing load across different generation sources before a frequency imbalance occurs. Likewise, predictive AI-based control can generate adjustment signals for inverters and microgrid regulators, ensuring that generated energy is distributed optimally without compromising system stability [21].

Various approaches have been proposed to address the frequency stability challenge by combining optimization techniques, predictive control, and machine learning. In particular, AI-based models have gained relevance due to their ability to analyze historical and real-time data, enabling efficient energy management and rapid response to fluctuations in generation and demand [22], [23].

A comparative analysis is performed of eight frequency control techniques and their use in each type of frequency control. Primary control is understood as the rapid, local response (milliseconds to seconds) to stabilize frequency after disturbances. Secondary control is defined as the restoration of frequency to nominal values (seconds to minutes) through coordinated adjustments. Tertiary control is defined as the economic and global management (minutes to hours) that readjusts setpoints, optimizes resources, and coordinates reserves [24], [25].

a) RNN/LSTM

Recurrent Neural Networks (RNNs) and their LSTM variant are designed to process sequential data, maintaining the influence of previous states on current outputs. They are ideal for tertiary control in predictive energy management and secondary control for setpoint adjustments. Their key strength

lies in analyzing complex time-series patterns, but they require high computational power and large training datasets [26], [27].

b) *MPC (Model Predictive Control)*

Model Predictive Control (MPC) uses mathematical models to predict system behavior and optimize real-time actions. It excels in primary control (fast response) and secondary control (precise system restoration). While it efficiently handles multiple variables, its implementation can be complex [20], [28].

c) *Fuzzy Logic*

Fuzzy Logic employs rule-based decision-making instead of mathematical models, using degrees of truth for human-like reasoning. It is effective in primary control (robust response in low-inertia systems) and secondary control (environmental adaptability). Its flexibility in uncertain environments is a major advantage, though design complexity can be a challenge [17], [29].

d) *Reinforcement Learning*

This method enables systems to learn frequency regulation through trial and error, optimizing actions based on rewards/penalties. It performs well in secondary control (adaptive restoration) and tertiary control (global optimization). While its dynamic adaptability is a strength, training can be slow and costly [7], [23].

e) *Genetic Algorithms*

Inspired by natural evolution, these algorithms optimize solutions via selection, crossover, and mutation. They are useful in secondary control (controller tuning) and tertiary control (resource optimization). They do not require exact mathematical models but have high computational demands [30], [31].

f) *Multi-Agent Control*

This approach uses autonomous agents that collaborate or compete to achieve shared or individual goals. It is ideal for secondary control (distributed restoration) and tertiary control (collaborative resource management). It distributes computational load and enhances resilience, though agent synchronization can be challenging [16], [22].

g) *Fuzzy Logic + Evolutionary Algorithms*

This hybrid combines Fuzzy Logic (for uncertainty handling) with Evolutionary Algorithms (for optimization), efficiently tuning fuzzy rules. It is versatile in secondary control but has limited primary/tertiary applications. While highly adaptable, it requires significant processing time [17], [19], [32].

h) *DNN (Deep Neural Networks)*

Deep Neural Networks (DNNs) learn complex nonlinear data representations through multiple hidden layers. They are applied in secondary control (adaptive regulation) and tertiary control (nonlinear optimization). They handle dynamic conditions effectively but demand substantial computational resources and large datasets [33], [34].

This comparison shows that there is no single ideal method

for non-interconnected zones: while techniques like MPC or DNN offer high precision, their complexity may be incompatible with low-resource environments. In contrast, approaches such as PID Control or Fuzzy Logic prioritize simplicity, though with limitations in dynamic scenarios. The selection should be based on a balance between technical capacity, data availability, and operational costs [35].

Figure 2 shows the applications of each AI technique in the control types based on the specific results in Table 1. In summary, for primary control: MPC and Fuzzy Logic (for speed and simplicity). Secondary only: Reinforcement Learning and Multi-Agent Control (adaptability and distribution). Tertiary only: RNN/LSTM (predictive analysis). Suggested combinations: MPC (primary + secondary), RL (secondary + tertiary).

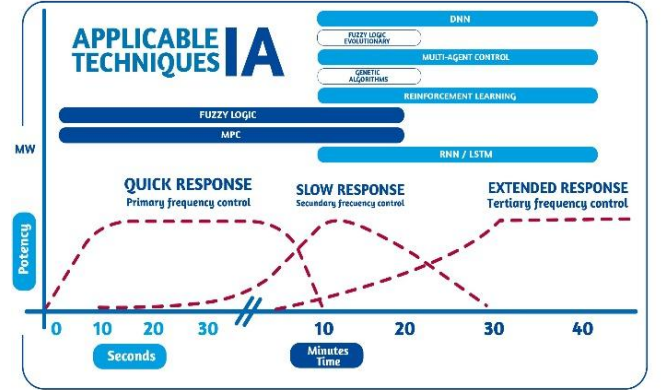


Fig. 2. AI methods applicable in isolated areas, with their advantages and limitations. Adapted from [37].

As shown in Figure 2, each AI technique has an optimal niche in frequency control based on its capabilities and limitations. Therefore, the choice depends on the balance between speed, accuracy, computational cost, the hierarchical level of the control being implemented, and the types of resources involved in such control, since the conditions in each isolated area differ [36].

These techniques may enable the implemented model not only to predict generation and demand but also to learn how to optimize energy management over time and improve its responsiveness to unexpected variations in the grid [7]. This advanced predictive capability not only helps prevent frequency imbalances but also allows for the efficient scheduling of energy storage and distribution strategies, reducing the need for manual interventions and enhancing the system's operational stability [12].

V. DISCUSSION: USE OF AI TECHNIQUES IN ISOLATED AREA

To comprehensively understand the relationships between variables affecting electrification in isolated areas, an analysis using system dynamics has been conducted. This modeling approach captures the nonlinear complexity and feedback interactions characteristic of these contexts. Unlike traditional linear and static models, it does not assume independence between technical, socioeconomic, and technological variables

but considers their interaction. It also supports the integration of artificial intelligence (AI) into energy resource management by analyzing both problems and potential solutions.

A central challenge in electrifying isolated areas is the integration of renewable energy sources, whose intermittency can destabilize low-inertia systems through frequency fluctuations. If unmanaged, these can degrade supply quality and lead to blackouts, affecting key activities such as agriculture. Additionally, limited infrastructure restricts real-time data collection, which is crucial for accurate AI performance. The problem becomes circular: poor data limits AI effectiveness; without AI, renewable integration remains difficult, reinforcing reliance on expensive and polluting diesel generation.

As shown in Figure 3, the system model includes two main feedback loops. The green loop represents a socioeconomic dynamic where a more stable supply enhances productive activities, notably agriculture, increasing system profitability and, eventually, energy demand. This rising demand can affect frequency stability. The second, technical loop emphasizes the role of AI in managing variable generation and demand. This, in turn, requires improved data quality and monitoring infrastructure, reinforcing system control and stability.

The model shows how external factors like intermittent generation and fluctuating demand drive frequency instability, which necessitates more AI-based control. When effectively applied, AI fosters a virtuous cycle: improved resource management reduces instability and facilitates higher renewable integration.

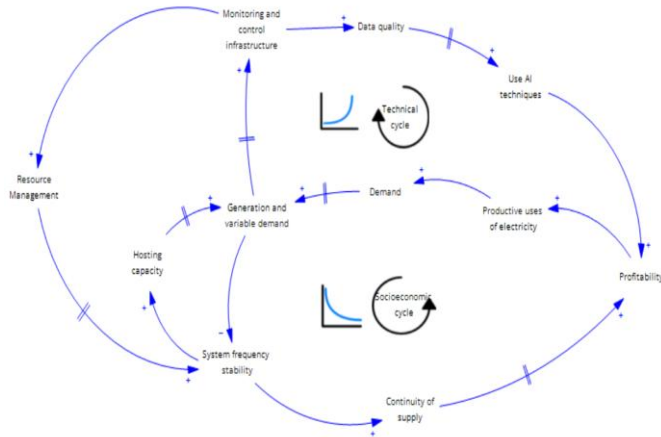


Fig. 3. System dynamics of integrating AI techniques in isolated areas.

AI enables real-time decision-making based on historical and predictive data. It also supports decentralized, modular energy systems key in remote areas by allowing microgrids to operate autonomously and collaboratively. This architecture reduces latency and increases system resilience. Scalable, low-cost AI models, such as simplified algorithms or decentralized controllers, can mitigate infrastructure limitations and promote sustainable electrification. For instance, AI-coordinated solar inverters can predict minute-by-minute generation and adjust energy flow, preventing sudden frequency drops.

The proposed system dynamics model offers a pathway to

break the negative cycle through sustainable intervention. Implementing AI highlights the need to strengthen infrastructure with affordable sensors and IoT communication, building a continuous data stream that enhances AI performance. This supports both technical stability and a socioeconomic feedback loop: as reliability improves, investment and productive activity grow, fostering further renewable deployment.

A viable solution would involve using simplified algorithms, such as federated learning models or lightweight techniques that require less computational power and can operate with partial or decentralized data. Additionally, adopting low-cost sensors and basic IoT networks would enable the collection of critical information without relying on advanced infrastructure. These strategies, combined with the proposed system dynamics approach, would facilitate a scalable transition toward smart microgrids even in resource-constrained environments.

VI. CONCLUSIONS

Artificial intelligence (AI) emerges as a promising solution to improve energy management in isolated areas, offering the potential to address critical challenges such as forecasting variable generation and demand. Through predictive techniques, AI optimizes the balance between supply and demand, reduces dependence on fossil fuels, and enhances operational efficiency. However, its implementation requires specific adaptations, as these remote areas face challenges in their electrical infrastructure that limit the development of adaptive techniques.

The system dynamics analysis revealed that integrating AI in isolated regions not only resolves technical issues but also generates positive socioeconomic impacts. By stabilizing the energy supply, productive activities are stimulated, increasing demand and thereby system profitability, which in turn justifies expanding renewable generation capacity. Nonetheless, to ensure sustainability, it is essential to develop decentralized, low-cost AI models tailored to rural contexts, prioritizing interoperability and community participation. The combination of technical and social approaches is vital for transforming these systems into long-term sustainable solutions.

ACKNOWLEDGMENT

The authors would like to acknowledge the support for the project entitled: Identificación de Estrategias y Criterios para la Electrificación Sostenible en comunidades Rurales Aisladas de Economías Emergentes, HERMES code 62627, executed through the Vicedecanatura de Investigación y Extensión of the Facultad de Ingeniería y Arquitectura.

REFERENCES

- [1] M. M. Alam, M. J. Hossain, M. A. Zamee, and A. Al-Durra, "Design and operation of future low-voltage community microgrids: An AI-based approach with real case study," *Appl Energy*, vol. 377, p. 124523, Jan. 2025, doi: 10.1016/j.apenergy.2024.124523
- [2] A. Arango-Manrique, D. López-García, C. Arango-Lemoine, and S. X. Carvajal-Quintero, "Business Model Proposal for Energy Transition towards Operational and Economic Sustainability for Rural Electrifi-

- cation: Colombian Case,” *Sustainability*, vol. 13, no. 8, p. 4308, Apr. 2021, doi: 10.3390/su13084308
- [3] A. Márquez-Bedoya, S. X. Carvajal-Quintero, and D. López-García, “Energía sostenible en zonas aisladas de América Latina: revisión de la integración del enfoque WEF Nexus para impulsar el desarrollo rural,” *Información tecnológica*, vol. 36, no. 1, pp. 1–14, Feb. 2025, doi: 10.4067/s0718-07642025000100001
 - [4] R. F. Colmenares-Quintero, G. Maestre-Gongora, O. C. Valderrama-Riveros, M. Baquero-Almazo, and K. E. Stansfield, “A Data-Driven Architecture for Smart Renewable Energy Microgrids in Non-Interconnected Zones: A Colombian Case Study,” *Energies (Basel)*, vol. 16, no. 23, p. 7900, Dec. 2023, doi: 10.3390/en16237900
 - [5] I. Granit, “What makes Colombia’s indigenous peoples adopt microgrids? Social acceptance and financial constraints in renewable energy diffusion,” *Energy Res Soc Sci*, vol. 101, p. 103132, Jul. 2023, doi: 10.1016/j.erss.2023.103132
 - [6] A. Shahid, R. Ahmadihangar, A. Rosin, A. Bli-nov, T. Korötko, and D. Vinnikov, “Leveraging the machine learning techniques for demand-side flexibility – A comprehensive review,” *Electric Power Systems Research*, vol. 238, p. 111185, Jan. 2025, doi: 10.1016/j.epsr.2024.111185
 - [7] Y. Shang et al., “Stochastic dispatch of energy storage in microgrids: An augmented reinforcement learning approach,” *Appl Energy*, vol. 261, p. 114423, Mar. 2020, doi: 10.1016/j.apenergy.2019.114423
 - [8] E. E. Gaona, C. L. Trujillo, and J. A. Guacaneme, “Rural microgrids and its potential application in Colombia,” *Renewable and Sustainable Energy Reviews*, vol. 51, pp. 125–137, Nov. 2015, doi: 10.1016/j.rser.2015.04.176
 - [9] L. Lv, Z. Wu, L. Zhang, B. B. Gupta, and Z. Tian, “An Edge-AI Based Forecasting Approach for Improving Smart Microgrid Efficiency,” *IEEE Trans Industr In-form*, vol. 18, no. 11, pp. 7946–7954, Nov. 2022, doi: 10.1109/TII.2022.3163137
 - [10] K. Mahmud, S. Azam, A. Karim, S. Zobaed, B. Shanmugam, and D. Mathur, “Machine Learning Based PV Power Generation Forecasting in Alice Springs,” *IEEE Access*, vol. 9, pp. 46117–46128, 2021, doi: 10.1109/ACCESS.2021.3066494
 - [11] D. Naware and A. Mitra, “Data-Driven Technology Applications in Planning, Demand-Side Management, and Cybersecurity for Smart Household Community,” *IEEE Transactions on Artificial Intelligence*, vol. 5, no. 10, pp. 4868–4883, Oct. 2024, doi: 10.1109/TAI.2024.3417389
 - [12] M. M. Ahmed, H. M. Bawayan, M. A. Enany, M. M. Elymany, and A. A. Shaier, “Modern advancements of energy storage systems integrated with hybrid renewable energy sources for water pumping application,” *Engineering Science and Technology, an International Journal*, vol. 62, p. 101967, Feb. 2025, doi: 10.1016/j.jestch.2025.101967
 - [13] P. Arévalo, D. Ochoa-Correa, and E. Villa-Ávila, “Optimizing Microgrid Operation: Integration of Emerging Technologies and Artificial Intelligence for Energy Efficiency,” *Electronics (Basel)*, vol. 13, no. 18, p. 3754, Sep. 2024, doi: 10.3390/electronics13183754
 - [14] G. Talluri, G. M. Lozito, F. Grasso, C. Iturrino Garcia, and A. Luchetta, “Optimal Battery Energy Storage System Scheduling within Renewable Energy Communities,” *Energies (Basel)*, vol. 14, no. 24, p. 8480, Dec. 2021, doi: 10.3390/en14248480
 - [15] K. Ukoba, K. O. Olatunji, E. Adeoye, T.-C. Jen, and D. M. Madyira, “Optimizing renewable energy systems through artificial intelligence: Review and future prospects,” *Energy & Environment*, vol. 35, no. 7, pp. 3833–3879, Nov. 2024, doi: 10.1177/0958305X241256293
 - [16] M. Azeroual et al., “Advanced energy management and frequency control of distributed microgrid using multi-agent systems,” *International Journal of Emerging Electric Power Systems*, vol. 23, no. 5, pp. 755–766, Oct. 2022, doi: 10.1515/ijeeps-2021-0298
 - [17] F. Asghar, M. Talha, and S. H. Kim, “Fuzzy logic-based intelligent frequency and voltage stability control system for standalone microgrid,” *International Transactions on Electrical Energy Systems*, vol. 28, no. 4, p. e2510, Apr. 2018, doi: 10.1002/etep.2510
 - [18] A. Martínez and P. Arévalo, “Distributed Peer-to-Peer Optimization Based on Robust Reinforcement Learning with Demand Response: A Review,” *Computers*, vol. 14, no. 2, p. 65, Feb. 2025, doi: 10.3390/computers14020065
 - [19] A. Abazari, M. G. Dozein, and H. Monsef, “An optimal Fuzzy-logic based frequency control strategy in a high wind penetrated power system,” *J Franklin Inst*, vol. 355, no. 14, pp. 6262–6285, Sep. 2018, doi: 10.1016/j.jfranklin.2018.06.012
 - [20] S. Shahzad, M. A. Abbasi, M. A. Chaudhry, and M. M. Hussain, “Model Predictive Control Strategies in Microgrids: A Concise Revisit,” *IEEE Access*, vol. 10, pp. 122211–122225, 2022, doi: 10.1109/ACCESS.2022.3223298
 - [21] U. K. Das et al., “Forecasting of photovoltaic power generation and model optimization: A review,” *Renewable and Sustainable Energy Reviews*, vol. 81, pp. 912–928, Jan. 2018, doi: 10.1016/j.rser.2017.08.017
 - [22] S. Weckx, R. D’Hulst, and J. Driesen, “Primary and Secondary Frequency Support by a Multi-Agent Demand Control System,” *IEEE Transactions on Power Systems*, vol. 30, no. 3, pp. 1394–1404, May 2015, doi: 10.1109/TPWRS.2014.2340582
 - [23] W. Cui, Y. Jiang, and B. Zhang, “Reinforcement Learning for Optimal Primary Frequency Control: A Lyapunov Approach,” *IEEE Transactions on Power Systems*, vol. 38, no. 2, pp. 1676–1688, Mar. 2023, doi: 10.1109/TPWRS.2022.3176525
 - [24] S. X. Carvajal, J. Serrano, and S. Arango, “Colombian ancillary services and international connections: Current weaknesses and policy challenges,” *Energy Policy*, vol. 52, pp. 770–778, Jun. 2013, doi: 10.1016/j.enpol.2012.10.041
 - [25] H. Armghan, Y. Xu, N. Ali, and U. Farooq, “Co-ordination control in hybrid energy storage based microgrids providing ancillary services: A three-layer control approach,” *J Energy Storage*, vol. 93, p. 112221, Jul. 2024, doi: 10.1016/j.est.2024.112221
 - [26] M. Aslam, J.-M. Lee, H.-S. Kim, S.-J. Lee, and S. Hong, “Deep Learning Models for Long-Term Solar Radiation Forecasting Considering Microgrid Installation: A Comparative Study,” *Energies (Basel)*, vol. 13, no. 1, p. 147, Dec. 2019, doi: 10.3390/en13010147
 - [27] H. Bevrani, H. Golpira, A. R. Messina, N. Hatziairgyriou, F. Milano, and T. Ise, “Power system frequency control: An updated review of current solutions and new challenges,” *Electric Power Systems Research*, vol. 194, p. 107114, May 2021, doi: 10.1016/j.epsr.2021.107114
 - [28] M. Jayachandran and G. Ravi, “Decentralized model predictive hierarchical control strategy for is-landed AC microgrids,” *Electric Power Systems Research*, vol. 170, pp. 92–100, May 2019, doi: 10.1016/j.epsr.2019.01.010
 - [29] S. Sahoo, N. K. Jena, A. Anurag, A. K. Naik, B. K. Sahu, and P. C. Sahu, “Frequency Regulation of Microgrid using Fuzzy-logic based controllers,” in *2023 International Conference in Advances in Power, Signal, and Information Technology (APSIT)*, IEEE, Jun. 2023, pp. 120–123, doi: 10.1109/APSIT58554.2023.10201752
 - [30] K. Sabahi, M. Tavan, and A. Hajizadeh, “Adaptive type-2 fuzzy PID controller for LFC in AC microgrid,” *Soft comput*, vol. 25, no. 11, pp. 7423–7434, Jun. 2021, doi: 10.1007/s00500-021-05704-6
 - [31] Rekha Regar, Raunak Jangid, and Kapil Parikh, “Load Frequency Control of Two Area System Using Genetic Algorithm,” *International Journal of Engineering Research and*, vol. V6, no. 04, Apr. 2017, doi: 10.17577/IJERTV6IS040198
 - [32] M. Dehghani and H. Nazari-pouya, “An Optimized Fuzzy Adaptive Distributed Secondary Controller for Micro Grids,” in *2023 North American Power Symposium (NAPS)*, IEEE, Oct. 2023, pp. 1–5, doi: 10.1109/NAPS58826.2023.10318552
 - [33] M. Mishra, J. Nayak, B. Naik, and A. Abraham, “Deep learning in electrical utility industry: A comprehensive review of a decade of research,” *Eng Appl Artif Intell*, vol. 96, p. 104000, Nov. 2020, doi: 10.1016/j.engappai.2020.104000
 - [34] F. Liang, C. Shen, W. Yu, and F. Wu, “Towards Optimal Power Control via Ensembling Deep Neural Networks,” *IEEE Transactions on Communications*, vol. 68, no. 3, pp. 1760–1776, Mar. 2020, doi: 10.1109/TCOMM.2019.2957482
 - [35] S. X. Carvajal Quintero, D. López García, A. Arango Manrique, J. D. Marín Jiménez, and C. Arango Lemoine, “Servicios complementarios en la transición energética,” 2023. Available: <https://repositorio.unal.edu.co/handle/unal/84692>
 - [36] T. A. Rajaperumal and C. C. Columbus, “Transforming the electrical grid: the role of AI in advancing smart, sustainable, and secure energy systems,” *Energy Informatics*, vol. 8, no. 1, p. 51, Apr. 2025, doi: 10.1186/s42162-024-00461-w
 - [37] E. Obando, D. Lopez, and S. Carvajal, *Active Demand and its Integration into the Distribution System with the Use of Metaheuristic Techniques: Opportunities, Conceptual Evolution, and Application Challenges*. 2024, doi: 10.22541/au.171420173.36135829/v1

Evaluation of a customized Unet model for estimating cephalometric landmarks using heat maps on a standard deviation of intensity projection (STDIP) dataset of CBCT images.

Juan E. Giraldo-Reyes, Carlos A. Ferro-Sánchez, Mauricio J. Mozo-Terán, Cristian O. Diaz, Reinel T. -Soto, Gilber A. Corrales,

Abstract— The application of deep learning techniques to medical imaging has shown significant promise in automating the detection of cephalometric landmarks, offering precision levels comparable to expert clinicians. This study presents a novel deep learning-based framework for automatic landmark localization using customized U-Net architecture enhanced with attention mechanisms and dropout regularization. The dataset consists of 2D lateral cephalometric radiographs from a mestizo population; each annotated with 26 anatomical landmarks. Ground truth heatmaps were generated for spatial supervision, and localization performance was evaluated using clinically relevant metrics: Localization Error (LE), Successful Detection Rate (SDR), and a custom pixel error threshold.

Index Terms— Cephalometric landmarks, Cephalometry, Cone-beam computed tomography (CBCT), Deep learning, Heat maps, Landmark detection, Medical image analysis, Orthodontics, Standard deviation of intensity projection (STDIP), U-Net.

I. INTRODUCTION

IN recent years, advances in deep learning (DL) have significantly transformed the automated analysis of medical images, particularly in the localization of cephalometric landmarks. Through the use of convolutional neural networks (CNNs), it has been possible to replicate—and in some cases approach—the performance of human experts under specific conditions. These technologies have shown considerable potential to improve diagnostic accuracy and reduce clinical workload in orthodontics and maxillofacial surgery.

Recent meta-analyses and systematic reviews consistently report average localization errors within a 2 mm threshold, with success rates ranging from 80% to 90% [1], [2]. However, these studies also highlight concerns regarding methodological bias, limited dataset diversity, and insufficient evidence of generalization to heterogeneous populations. Furthermore, some authors [3] have questioned the clinical validity of using a fixed 2 mm threshold as the primary evaluation metric, emphasizing the need for clinically interpretable criteria. Others, such as [4], demonstrated that modern architectures like YOLOv3 outperform earlier models, such as SSD and YOLOv2, in terms of precision, although no model consistently surpasses the accuracy of experienced orthodontists. [2] found success rates of 81% within 2 mm and 96% within 4 mm, confirming promising results, albeit with considerable variability across models and anatomical landmarks. [3] questioned the clinical validity of the 2 mm threshold as a universal success criterion, while [4] emphasized the superiority of architectures such as YOLOv3, which outperformed earlier models like SSD and YOLOv2 in terms of precision.

Collectively, the literature indicates that well-trained DL systems achieve average errors between 1–2 mm and detection rates exceeding 85% in most studies. Nonetheless, limitations remain: reliance on homogeneous public datasets, reduced precision in complex landmarks, scarce clinical validation, and non-standardized evaluation metrics. Future research is oriented toward improving generalizability through multicenter data, validating clinical impact in real-world settings, and

Manuscript received July 28, 2025; revised —. (Corresponding author: Juan Esteban Giraldo Reyes; e-mail: juan_esteban.giraldo@uao.edu.co, juanestebangiraldoreyes0@gmail.com.)

Carlos Andrés Ferro Sánchez is with the Faculty of Engineering and Basic Sciences, Universidad Autónoma de Occidente, Cali 760030, Valle, Colombia (e-mail: cferro@uao.edu.co). Juan Esteban Giraldo Reyes is with the Faculty of Engineering and Basic Sciences, Universidad Autónoma de Occidente, Cali 760030, Valle, Colombia (e-mail: juan_esteban.giraldo@uao.edu.co). Mauricio José Mozo Terán is with the Faculty of Engineering and Basic Sciences, Universidad Autónoma de Occidente, Cali 760030, Valle, Colombia (e-mail: mauricio.mozo@uao.edu.co).

Cristian Orlando Díaz Laverde is with the Faculty of Engineering and Basic Sciences, Universidad Autónoma de Occidente, Cali 760030, Valle, Colombia; and with the Faculty of Health and the Faculty of Engineering, Universidad del Valle, Cali 760030, Valle, Colombia (e-mail: cristian_orl.diaz@uao.edu.co). Reinel Tabares-Soto is with the Department of Electronics and Automation, Universidad Autónoma de Manizales, Manizales 170001, Caldas, Colombia; the Department of Systems and Informatics, Universidad de Caldas, Manizales 170001, Caldas, Colombia; and GobLab School of Government, Universidad Adolfo Ibáñez, Santiago 7941169, Chile (e-mail: reinel.tabares@ucaldas.edu.co). Gilber A. Corrales is with the Faculty of Engineering and Basic Sciences, Universidad Autónoma de Occidente, Cali 760030, Valle, Colombia; and also with Universidad Adolfo Ibáñez, Santiago 8320000, Chile (e-mail: gacorrales@uao.edu.co).

extending these techniques to 3D cephalometry. The integration of DL as an assistive tool holds promise for reducing clinical workload and enhancing diagnostic consistency in orthodontics, reinforcing its role as a complementary technology rather than a replacement for expert judgment.

In response to these limitations, the present study proposes a system based on a modified U-Net architecture, enhanced with attention mechanisms and dropout regularization, for the automatic detection of cephalometric landmarks in a mestizo population. The methodology includes the generation of heatmaps for spatial supervision, evaluation through clinically relevant metrics (LE, SDR), and interpretability via critical error metrics. Additionally, the workflow was migrated to scalable computing environments (Google Colab Pro) to ensure computational feasibility.

This approach aims not only to improve the precision in the localization of key anatomical landmarks but also to contribute to the clinical applicability of deep learning by integrating interpretability, robustness, and generalizability as fundamental pillars of the experimental design.

Regarding original studies, [5] implemented a CNN that outperformed traditional methods in detecting landmarks within 2 mm, setting a precedent for subsequent research. [6] and [7] compared modern object detection models such as YOLOv3 and SSD, demonstrating that YOLOv3 achieved lower error dispersion and accuracy comparable to that of human orthodontists. [8] integrated an automated cephalometric system into a web platform, while [9] introduced Bayesian neural networks capable of estimating uncertainty in each prediction, thereby improving clinical interpretability. Other innovative approaches, such as CephaNN [10] with attention mechanisms, or 2D–3D hybrid models based on CBCT projections [11], have further expanded the scope and robustness of current solutions.

II. MATERIALS

This study presents a deep learning-based system for the automatic detection of cephalometric landmarks using a custom U-Net architecture enhanced with attention mechanisms and dropout regularization. The methodology encompasses data preparation, model architecture, training strategy, evaluation protocol, and migration considerations for scalable computation.

A. Data Acquisition and Preprocessing

The dataset consists of grayscale lateral cephalometric radiographs, each annotated with 26 anatomical landmarks representing critical skeletal and dental structures. These annotations were extracted from structured Excel files, with each record providing the x and y coordinates of the landmarks in pixels. Corresponding radiographic images were provided in .jpg format. The landmark points were acquired by two expert orthodontists using standard deviation of intensity projections (STDIP) from CBCT volumes, specifically in sagittal slices, previously corrected to natural head position [12]. This study was approved by the corresponding ethics committee (Universidad Autónoma de Occidente, Acta No. 01-2024,

February 12, 2024), and all patient data were anonymized prior to analysis, in accordance with the principles of the Declaration of Helsinki.

All images were uniformly resized to 256×256 pixels to ensure compatibility with the neural network input. Landmark coordinates were rescaled accordingly using the resizing factors along both axes. Additionally, original pixel sizes (in millimeters) for each radiograph were provided in a separate metadata file, allowing for accurate conversion of localization errors from pixel units to millimeters.

To prepare the data for training, we applied a normalization transform that scaled pixel intensity values to the range $[-1, 1]$, using a mean of 0.5 and standard deviation of 0.5. For each set of landmark coordinates, ground truth heatmaps were generated using 2D Gaussian distributions centered at each landmark location, with a fixed standard deviation of $\sigma = 5$ pixels. These heatmaps serve as spatial supervision targets for the network.

Finally, two completely independent datasets were created: one used exclusively for training and validation through cross-validation, and another reserved solely for the final evaluation of the model (test). This strict separation was implemented to prevent overfitting, avoid any data leakage between the training and testing phases, and ensure an objective and realistic assessment of the system's overall performance.

B. Model Architecture

The proposed model is a modified U-Net architecture designed specifically for heatmap regression, as illustrated in Fig. 1, which shows the complete U-Net structure. It follows an encoder–latent space–decoder design:

- The encoder consists of sequential convolutional layers with ReLU activations and max pooling operations to progressively reduce spatial dimensions while increasing feature abstraction. Dropout layers with rates of 0.1 to 0.15 were inserted to prevent overfitting.
- The bottleneck incorporates an attention mechanism consisting of a 1×1 convolution followed by a sigmoid activation, which modulates the bottleneck features. A higher dropout rate (0.2) is used at this stage to improve generalization.
- The decoder uses transposed convolutions (upsampling) and additional convolutional layers to recover spatial resolution. Dropout is gradually reduced as spatial resolution increases. The final output layer applies a 1×1 convolution to generate 26 heatmaps, each corresponding to a landmark.

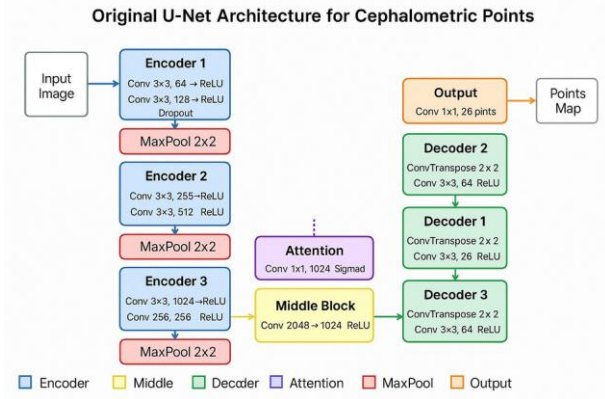


Fig 1. Unet Architecture.

C. Training Strategy

The table 1 summarizes the main hyperparameters used during model training. The Adam optimizer was employed, offering a slight advantage by combining the benefits of AdaGrad and RMSProp, allowing for efficient adaptation of the learning rate for each parameter. The learning rate was set to 0.0001, and the Mean Squared Error (MSE) loss function was used to evaluate the discrepancy between predicted and ground-truth heatmaps. Regularization was applied using dropout (0.1 in the encoder/decoder and 0.2 in the bottleneck), and an early stopping strategy was implemented, with a maximum of 50 epochs and termination if no improvement in validation loss was observed over 10 consecutive epochs. Training was conducted with a batch size of 8 samples per iteration, ensuring stable weight updates.

TABLE I
Hyperparameter

Hyperparameter	Value / Description
Optimizer	Adam
Learning Rate	0.0001
Loss Function	Mean Squared Error (MSE) between predicted and ground-truth heatmaps
Dropout Rate	0.1 in encoder/decoder, 0.2 in bottleneck
Epochs	Up to 50, with early stopping after 10 epochs without validation loss improvement
Batch Size	8 samples per iteration

D. Evaluation Protocol

The table 2 has results demonstrate that the model consistently achieves mean errors below the clinically accepted 2 mm gold standard for cephalometric landmark localization [13]. This threshold is critical for ensuring diagnostic reliability in orthodontics. Notably, Fold 3 achieved the best performance, with a mean error of 1.3870 mm, indicating strong consistency and precision. Overall, these findings confirm the model's

clinical validity and its potential as a reliable tool for automated landmark detection.

TABLE II
Mean error and standard deviation per fold

Fold	Mean error (mm)	Standard deviation (mm)
1	1.5488	± 2.4513
2	1.4225	± 1.9270
3	1.3870	± 1.8681
4	1.4471	± 2.8392
5	1.4498	± 2.0668

Table 3 shows that the model achieves a high Successful Detection Rate (SDR) within the clinically accepted 2.0 mm threshold [13], reaching 82.83%, indicating reliable performance for orthodontic diagnosis. Moreover, a progressive increase in detection rates is observed as the error threshold widens, reaching up to 96.58% within 4.0 mm. This suggests that, with further refinement, many of the localized points could fall within the clinical gold standard, reinforcing the accuracy and robustness of the proposed model for automatic cephalometric landmark detection.

TABLE III
SDR Thresholds

SDR Threshold	Valor (%)
SDR 2.0 mm	82.83
SDR 2.5 mm	90.21
SDR 3.0 mm	93.59
SDR 4.0 mm	96.58

Table 4 presents the mean error and standard deviation values for each evaluated cephalometric landmark. Most points show mean errors below 1.5 mm, which is within the clinically acceptable range. However, some landmarks—such as ENP (3.11 mm), apx_1 (2.31 mm), apx_up (2.05 mm), Go (2.03 mm), and Ar (1.47 mm)—exhibit higher errors and greater variability.

This behavior is related to the anatomical complexity and low radiographic visibility of these points, due to overlapping structures or individual anatomical variations. In contrast, landmarks such as U1, Sn, Bb, and Pn show minimal errors (≈ 0.84 mm) and low dispersion, reflecting higher model accuracy.

TABLE IV
Mean Error and Standard Deviation per Cephalometric Landmark

Landmark	Mean Error (mm)	Standard Deviation (mm)
U1	0.8468	± 0.5751
Sn	0.8474	± 0.5024
Bb	0.8477	± 0.4321

Pn	0.8679	± 0.6911
Ll	0.9424	± 0.4988
Pmn	1.0059	± 0.6816
Sts	1.0312	± 0.6192
Ul	1.0363	± 0.6482
Pog	1.1314	± 1.5037
L1	1.1685	± 0.7676
Gn	1.1757	± 0.7021
O	1.1967	± 0.6113
S	1.2016	± 0.9020
Sti	1.2648	± 1.2506
M	1.3226	± 0.9845
B	1.3329	± 0.7978
N	1.3521	± 1.3975
Pogb	1.3706	± 1.3950
A	1.4223	± 1.0094
Ar	1.4736	± 3.4837
ENA	1.7793	± 1.3673
Po	1.9170	± 1.7808
Go	2.0397	± 2.0003
apx_up	2.0595	± 4.7354
apx_l	2.3127	± 4.2706
ENP	3.1141	± 2.5231

Figure 2 shows the output from the test set of Fold 3, displaying a lateral cephalometric radiograph with the localization of key anatomical landmarks. Blue dots represent the model’s predictions, while pink crosses indicate the ground truth annotations. Overall, a strong alignment is observed between predicted and actual points, reflecting high model accuracy—consistent with Fold 3 having the lowest mean error (1.3870 mm) among all evaluated folds.

However, some landmarks show noticeable misalignment, particularly in regions such as the posterior nasal spine (ENP), the root apices (apx_1 and apx_up), and the Gonion (Go), where the distance between prediction and reference is larger. This visual discrepancy aligns with the numerical results, which reported higher errors for these specific landmarks.

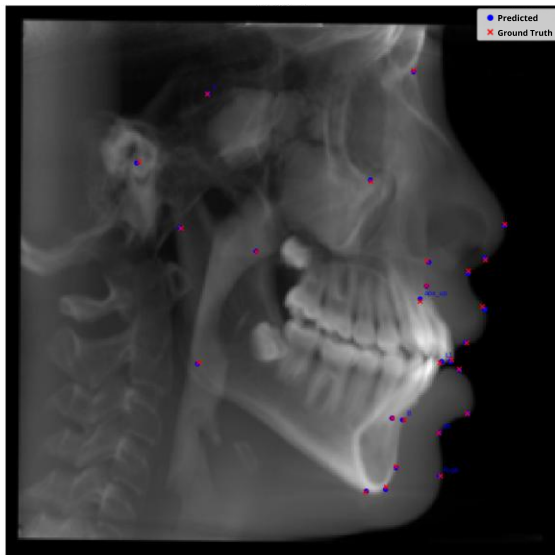


Fig 2. Test Image Fold 3

IV. DISCUSSION

Table 4 shows the mean error and standard deviation values for each evaluated cephalometric landmark. In general, most points exhibit mean errors below 1.5 mm, which is considered clinically acceptable. However, some landmarks show significantly higher errors, such as ENP (3.1141 mm), apx_1 (2.3127 mm), apx_up (2.0595 mm), Go (2.0397 mm), and Ar (1.4736 mm), along with high standard deviations, indicating greater dispersion and less consistency in automatic prediction.

This behavior is mainly explained by the anatomical complexity and inherent difficulties in radiographic visualization of these points. For example, Go (Gonion) is located at the mandibular angle, an area where overlapping mandibular branches make precise localization difficult. Ar (Articulare), located at the intersection of the mandibular condyle and the cranial base, can be easily confused due to artifacts such as cephalostat ear rods. The points apx_1 and apx_up, which correspond to the root apices of the incisors, show greater variability due to dental inclination, anatomical differences between patients, and overlapping bone structures. Finally, ENP (Posterior Nasal Spine) is a deep structure that is difficult to delineate in lateral projections.

In contrast, landmarks such as U1, Sn, Bb, and Pn, which are located in anatomically clear regions with high radiographic contrast, show minimal errors (≈ 0.84 – 0.87 mm) and low variability, demonstrating higher reliability in their automatic detection.

In summary, although the model shows solid performance across most landmarks, the highest errors are associated with points that present high anatomical variability, low radiographic visibility, and structural overlap. This highlights the importance of considering such limitations in clinical settings and future development of AI models for cephalometric analysis.

V. CONCLUSION

The present study demonstrates that the final landmark localization model achieves competitive performance compared to current state-of-the-art approaches within the evaluated domain, showing promising proximity to gold-standard accuracy benchmarks. While high-precision metrics are not yet fully attained, the results highlight the model’s solid baseline capabilities and its potential for future clinical or research applications.

Notably, the heatmap-based regression approach employed here entails substantial computational demands, particularly during multi-fold training and inference. This underscores the opportunity to explore lighter-weight architectures or alternative training strategies that could reduce computational overhead without compromising predictive precision. Approaches such as fine-tuning with domain-specific data,

model pruning, or knowledge distillation are promising directions for optimization.

Importantly, this project remains in an intermediate development phase. Continued methodological refinement, integration of more advanced loss functions, and targeted improvement of underperforming landmarks are expected to enhance model performance in subsequent iterations. Overall, the study lays a robust foundation for future work and contributes valuable insights toward the development of automated, precise, and computationally efficient landmark detection systems.

REFERENCES

- [1] M. Schwendicke et al., “Artificial intelligence in dental diagnostics: A scoping review,” *J. Dent.*, vol. 103, Art. no. 103497, Jan. 2021, doi: 10.1016/j.jdent.2020.103497.
- [2] D. Londoño, F. Martínez, and C. Rueda, “Performance of deep learning models for cephalometric landmark detection: A systematic review and meta-analysis,” *Int. J. Comput. Dent.*, vol. 26, no. 3, pp. 221–234, 2023.
- [3] G. Polizzi and R. Leonardi, “Clinical evaluation metrics for automated cephalometry: Are 2 mm enough?,” *Orthod. Craniofac. Res.*, vol. 27, no. 1, pp. 1–8, 2024, doi: 10.1111/ocr.12578. R. Fardel, M. Nagel, F. Nuesch, T. Lippert, and A. Wokaun, “Fabrication of organic light emitting diode pixels by laser-assisted forward transfer,” *Appl. Phys. Lett.*, vol. 91, no. 6, Aug. 2007, Art. no. 061103.
- [4] A. Ribas-Sabartés, J. Roig, and M. Romero, “Comparison of deep learning architectures for cephalometric landmark detection,” *Comput. Biol. Med.*, vol. 168, Art. no. 106031, 2024, doi: 10.1016/j.combiomed.2024.106031.
- [5] S. Ö. Arık, J. P. Ibarz, C. Wen, and T. Pfister, “Fully automated cephalometric analysis using convolutional neural networks,” in *Proc. IEEE Int. Symp. Biomed. Imaging (ISBI)*, 2017, pp. 1285–1289.
- [6] J. Park, H. Lee, and M. Kim, “Evaluation of YOLO-based object detectors for cephalometric landmark detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2019, pp. 2025–2032.
- [7] J. Hwang, S. Park, and D. Kim, “Automated cephalometric landmark detection using deep learning: Performance comparison of SSD and YOLOv3,” *Diagnostics*, vol. 10, no. 8, Art. no. 607, 2020, doi: 10.3390/diagnostics10080607.
- [8] J. Kim, H. Kwon, and M. Cho, “Web-based automatic cephalometric analysis using deep convolutional networks,” *Korean J. Orthod.*, vol. 50, no. 3, pp. 168–176, 2020.
- [9] J. H. Lee, Y. J. Lee, and H. S. Lee, “Bayesian deep learning for uncertainty-aware cephalometric landmark detection,” *Sci. Rep.*, vol. 10, no. 1, Art. no. 17247, 2020, doi: 10.1038/s41598-020-74375-3.
- [10] Y. Qian, X. Fan, and Q. Li, “CephaNN: Deep neural network for cephalometric landmark detection with attention mechanisms,” *Comput. Methods Programs Biomed.*, vol. 195, Art. no. 105607, 2020.
- [11] S. M. Lee, J. W. Lee, and Y. Park, “2D–3D hybrid deep learning model for cephalometric landmark detection from CBCT projections,” *J. Clin. Med.*, vol. 8, no. 9, Art. no. 1391, 2019.
- [12] C. A. Ferro-Sánchez, C. O. D. Laverde, S. E. N. Rodríguez, G. A. C. Gallego, J. F. Aristizábal, and O. I. C. Salazar, “Methodology for the correction of a CBCT volume from the skull to the natural head position,” *MethodsX*, vol. 13, Art. no. 103073, 2024, doi: 10.1016/j.mex.2024.103073.
- [13] J.-H. Lee, H.-J. Yu, M. Kim, J.-W. Kim, and J. Choi, “Automated cephalometric landmark detection with confidence regions using Bayesian convolutional neural networks,” *BMC Oral Health*, vol. 20, no. 1, Art. no. 270, 2020, doi: 10.1186/s12903-020-01256-7.

Modelado para Optimización de Horarios en Restaurantes de Medellín

Juan José Galeano Flórez *Universidad de Antioquia*
Medellín, Colombia

juan.j.galeano@udea.edu.co

Yony Fernando Ceballos *Universidad de Antioquia*
Medellín, Colombia

yony.cebillos@udea.edu.co

Jorge Enrique Martínez Arenas *Investigador Independiente*
Medellín, Colombia

jorge.martinez@bitcubo.com

Abstract—This research proposes a hybrid mathematical model to optimize staff scheduling in Medellín’s restaurants, integrating mixed-integer linear programming (MILP) and adaptive heuristics. The approach combines two phases: initial assignment of employees to time slots and roles, and continuity adjustment to minimize operational rotations.

Tested in a pilot restaurant with 56 employees and 30 workstations, the model achieved the same coverage as the expert’s manual assignment while reducing weekly scheduling time from 2 hours to 20 minutes. In a hypothetical scenario with 10 employees and 6 time slots, it demonstrated computational efficiency (30 seconds) and full compliance with labor policies.

Key outcomes include saving 2 weekly managerial hours per venue, enabling reallocation of human resources to strategic tasks. The model balances productivity and ethical labor conditions by enforcing continuous work limits and mandatory rest periods. By replicating manual scheduling accuracy with enhanced speed, this solution offers a scalable framework for improving operational efficiency in dynamic hospitality environments.

Index Terms—Staff Scheduling, Mathematical Optimization, Hybrid Model, Operation Research, Operational Efficiency, Restaurant Industry

I. INTRODUCCIÓN

LA programación eficiente de horarios en el sector restaurador de Medellín enfrenta desafíos críticos, como la fluctuación de demanda, el cumplimiento normativo laboral y la gestión de recursos humanos. Tradicionalmente, esta tarea recae en gerentes que dedican horas a asignaciones manuales, generando ineficiencias y márgenes de error significativos. Este trabajo propone un modelo matemático híbrido que integra técnicas de optimización para automatizar la asignación de personal, garantizando cobertura operativa, continuidad en turnos y adherencia al marco legal colombiano.

El estudio se justifica ante la necesidad de reducir costos operativos y mejorar la competitividad de restaurantes emergentes, donde la planificación manual limita la escalabilidad. A través de un enfoque de dos etapas—asignación inicial y ajuste de continuidad—, el modelo logra soluciones equivalentes a las de expertos humanos, pero con 84% menos tiempo

de procesamiento. Adicionalmente, incorpora variables clave como habilidades del personal, días de descanso y restricciones legales, ofreciendo un marco replicable para la industria de servicios.

II. MARCO TEÓRICO

La programación óptima de horarios en restaurantes de Medellín enfrenta desafíos multidimensionales, integrando teoría de programación, técnicas de optimización, normativa laboral y particularidades sectoriales. Estudios previos abordan estos aspectos de forma aislada, pero su articulación es crítica para diseñar soluciones alineadas con las necesidades operativas y legales [1], [2]. Este marco se estructura en cuatro pilares: fundamentos teóricos de programación de horarios, técnicas de optimización aplicadas, marco legal colombiano y contexto local del sector restaurador.

A. Fundamentos teóricos

La programación de horarios se conceptualiza como un proceso modular [1], que incluye: modelado de demanda (basado en tareas, pronósticos o turnos), asignación de días libres, diseño de turnos, construcción de líneas de trabajo (secuencias cíclicas/acíclicas), distribución de tareas y asignación individual de personal. La complejidad varía según el nivel de integración entre módulos, el tipo de demanda (estocástica o determinística) y las unidades de trabajo (horas, turnos o tareas). Enfoques como la programación de tripulaciones aéreas (generación-optimización-asignación) evidencian la diversidad metodológica en la literatura [1].

B. Técnicas de optimización aplicadas

Se emplean métodos heurísticos (SSL, SRS), meta-heurísticos (búsqueda tabú, algoritmos genéticos) y modelos matemáticos (MILP, CP) [2], [3]. La programación con restricciones (CP) y modelos de colas son efectivos para sistemas estocásticos, mientras enfoques híbridos (MILP con búsqueda local) equilibran eficiencia y precisión [4]. Los objetivos incluyen minimizar costos laborales, maximizar calidad de

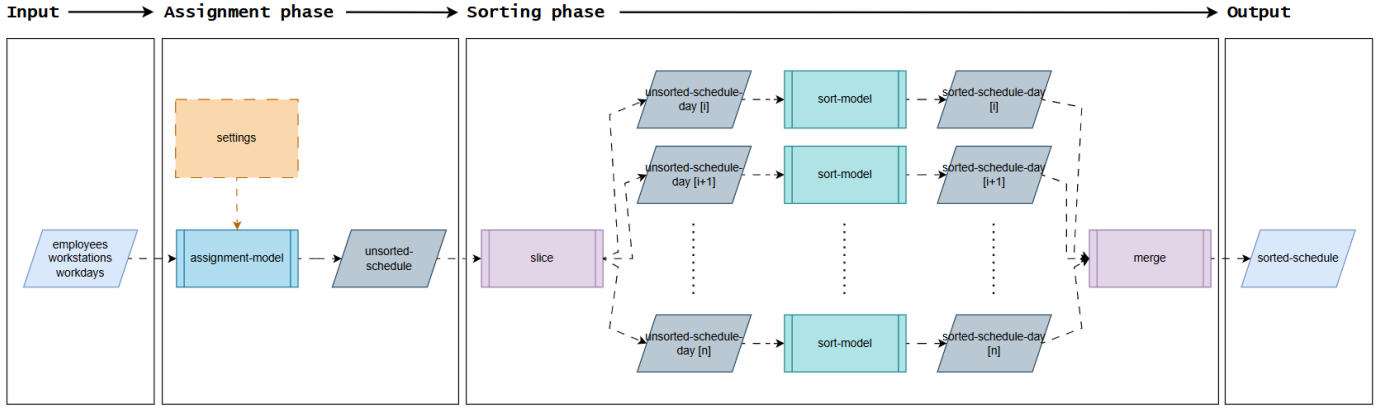


Fig. 1. Modelo híbrido: Interacción entre los modelos

servicio (tiempos de espera, tasas de abandono) y garantizar equidad en turnos [5]. Modelos multi-objetivo priorizan metas mediante pesos, considerando *trade-offs* entre costos, cumplimiento legal y preferencias de empleados [2].

C. Marco legal

El Código Sustantivo del Trabajo colombiano [6] establece: jornada máxima de 42 horas/semana, límites de horas extras (2 diarias/12 semanales), recargos del 35% por trabajo nocturno (9:00 p. m. – 6:00 a. m.) y 75% por dominicales/festivos. Además, exige descansos obligatorios (24 horas dominicales remuneradas) y compensaciones por trabajo en días libres. Jornadas flexibles (4-9 horas/día) y turnos sucesivos (hasta 56 horas/semanales en operación continua) añaden complejidad al modelado [6]. Incumplir estas normas genera costos adicionales y sanciones, requiriendo su integración explícita en los modelos de optimización.

D. Contexto local

El sector restaurador es clave en Medellín: tercero en participación empresarial (18.7% de cierres en 2023) y generador de 160.711 empleos en 2024 (+23.9% vs. 2023) [7], [8]. Aunque contribuye al 1.8% del PIB nacional [9], enfrenta altos costos operativos (18.5% ideal en nómina vs. realidades superiores). Esta dualidad (crecimiento económico vs. fragilidad operativa) subraya la necesidad de modelos de programación que optimicen costos laborales, cumplan la normativa y adapten a fluctuaciones de demanda, factores críticos para la sostenibilidad sectorial [8].

III. METODOLOGÍA

Este estudio adoptó un enfoque mixto, combinando entrevistas con gerentes de puntos de venta y la formulación de modelos matemáticos (Figura 1). Primero, se identificaron variables clave mediante entrevistas. Luego, se desarrolló un modelo híbrido en dos etapas: uno para asignación de oficios y otro para garantizar la continuidad operativa. La validación incluyó datos sintéticos y pruebas piloto para evaluar escalabilidad y sensibilidad.

A. Identificación de variables

Las variables identificadas en conjunto con los expertos fueron:

- Permisos o ausentismos
- Descansos
- Habilidad del personal
- Horas trabajadas por persona
- Requerimiento de personal
- Variación de la demanda

B. Selección del modelo

Dado que los modelos MILP presentan limitaciones en entornos de alta complejidad, se optó por un enfoque híbrido, dividiendo el problema en dos modelos.

C. Modelado matemático

1) *Modelo de asignación*: respeta disponibilidad, habilidades y demanda.

Conjuntos

- E : Conjunto de empleados.
- D : Conjunto de días.
- S : Conjunto de turnos.
- P : Conjunto de franjas horarias.
- W : Conjunto de oficios.

Parámetros

- $\rho_{dpw} \in \mathbb{N}$: Cantidad de empleados necesarios el día d en el oficio w en la franja horaria p .
- $\alpha_{edw} \in \{0, 1\}$: Disponibilidad del empleado e en el día d en el oficio w .
- $\sigma_{sp} \in \{0, 1\}$: Pertenencia de la franja horaria p al turno s .
- $\varphi_{ds} \in \mathbb{R}^+$: Costo de asignar a un empleado en el día d en el turno s .
- $\lambda \in \mathbb{R}^+$ con $\lambda > \max_{d \in D, s \in S} (\varphi_{ds})$: Costo por cada empleado faltante por cada oficio por cada franja horaria por cada día.
- $\gamma_w \in \mathbb{R}^+$ con $\gamma_w \ll \lambda$: Costo de asignar a un empleado en el oficio w .

Variables:

$$x_{edpw} = \begin{cases} 1 & \text{si } e \text{ es asignado a } w \text{ en } d \text{ y } p \\ 0 & \text{en caso contrario} \end{cases}$$

$$y_{eds} = \begin{cases} 1 & \text{si } e \text{ trabaja en } s \text{ el día } d \\ 0 & \text{en caso contrario} \end{cases}$$

$m_{dpw} \in \mathbb{N}$: Cantidad de empleados faltantes

Función objetivo:

$$\min Z = \sum_{e,d,s} \varphi_{ds} y_{eds} + \sum_{d,p,w} \lambda m_{dpw} + \sum_{e,d,p,w} \gamma_w x_{edpw} \quad (1)$$

Restricciones:

$$\sum_w x_{edpw} \leq 1 \quad \forall e, d, p \quad (2)$$

$$\sum_s y_{eds} \leq 1 \quad \forall e, d \quad (3)$$

$$\sum_e (x_{edpw} \cdot \alpha_{edw}) + m_{dpw} \geq \rho_{dpw} \quad \forall d, p, w \quad (4)$$

$$\sum_s (y_{eds} \cdot \sigma_{sp}) \geq x_{edpw} \quad \forall e, d, p, w \quad (5)$$

2) *Modelo de continuidad*: busca mantener un mismo oficio en franjas consecutivas.

Conjuntos

- E : Conjunto de empleados.
- P : Conjunto de franjas horarias.
- W : Conjunto de oficios.
- $E_{pw} \subseteq E$: Subconjunto de empleados que durante p están asignados en w
- $P_e \subseteq P$: Subconjunto de franjas horarias asignadas al empleado e .
- $W_e \subseteq W$: Subconjunto de oficios asignados al empleado e .

Parámetros

- $\tau_e \in P_e$: Última franja horaria asignada al empleado e
- $\mu_{pw} \in \mathbb{N}$: Cantidad de empleados asignados durante p en w

Variables:

$$x_{epw} = \begin{cases} 1 & \text{si } e \text{ realiza } w \text{ en } p \\ 0 & \text{en caso contrario} \end{cases}$$

$$y_{ep} = \begin{cases} 1 & \text{si hay cambio de oficio entre } p \text{ y } p+1 \\ 0 & \text{en caso contrario} \end{cases}$$

Función objetivo:

$$\min Z = \sum_e \sum_{p \in P_e \setminus \{\tau_e\}} y_{ep} \quad (6)$$

Restricciones:

$$y_{ep} \geq x_{epw} - x_{e(p+1)w} \quad \forall e, p, w \quad (7)$$

$$y_{ep} \geq x_{e(p+1)w} - x_{epw} \quad \forall e, p, w \quad (8)$$

$$\sum_{w \in W_e} x_{epw} \leq 1 \quad \forall e, p \quad (9)$$

$$\sum_{e \in E_{pw}} x_{epw} = \mu_{pw} \quad \forall p, w \quad (10)$$

D. Validación del modelo**Dimensiones evaluadas:**

- Su **escalabilidad** frente a la complejidad horaria.
- La **sensibilidad** a parámetros operativos clave (franjas horarias, oficios).

Hipótesis:

- 1) El tiempo crece exponencialmente con franjas horarias.
- 2) El número de oficios también influye exponencialmente.
- 3) Segmentación adecuada reduce espacio de búsqueda.
- 4) Más soluciones factibles implican mayor tiempo de cómputo.

Metodología: Se probaron 9 combinaciones:

- Franjas por turno: 5, 10, 20
- Oficios por empleado: 1, 2, 5

Con 10 réplicas por combinación. Todas las hipótesis fueron respaldadas, confirmando la validez teórica y práctica del modelo.

E. Selección del piloto

Se eligió un restaurante con **dificultad media-alta**, evaluando:

- Cantidad de empleados
- Cantidad de oficios
- Horario de servicio
- Requerimientos operativos

Características del restaurante piloto:

- **Operación:**
 - Brunch (08:00-12:00), Almuerzo (12:00-18:00), Cena (18:00-21:30)
- **Personal:** 56 empleados, 30 oficios.
- **Turnos:**
 - Mañana: 06:00–14:00 / 06:00–15:00
 - Tarde: 10:30–19:00 / 10:30–20:00 / 12:00–20:00
 - Noche: 12:00–20:30 / 12:00–21:30
- **Segmentación:**
 - 12 intervalos de 30 min
 - 5 intervalos de 1 h
 - 3 intervalos de 1.5 h

F. Configuración técnica

- Procesador: Intel Core i5-12600 (6 núcleos, 12 hilos, 4.8 GHz turbo).
- Memoria RAM: 16 GB DDR4 @ 3600 MHz (dual-channel).
- Tarjeta gráfica: AMD Radeon RX 6600 (8 GB GDDR6).

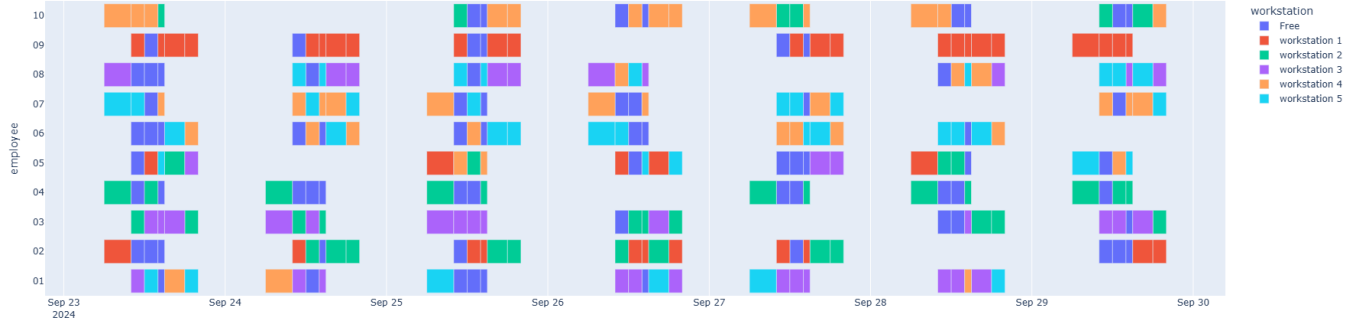


Fig. 2. Resultado del modelo de asignación

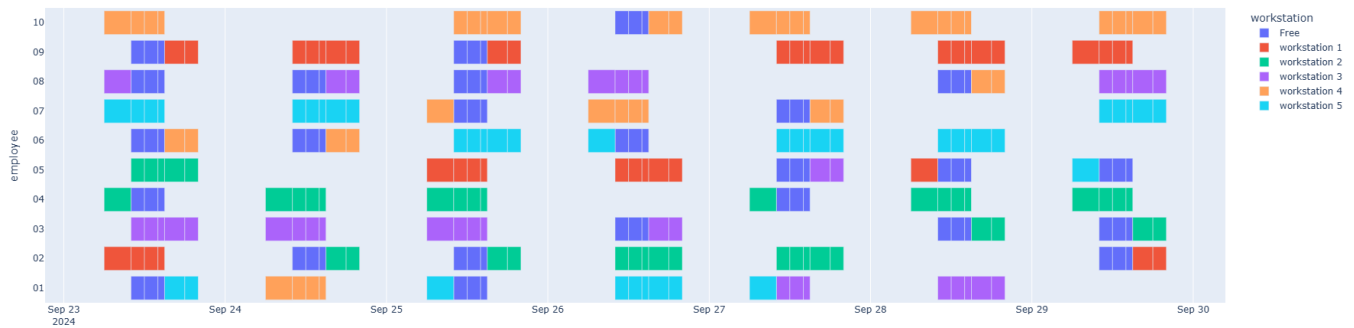


Fig. 3. Resultado del modelo de continuidad

- Placa madre: Gigabyte B560M Aorus Pro (chipset Intel B560).
- Lenguaje de programación: Python 3.11.4
- Librería de optimización: Google OR-Tools 9.8.
- Sistema operativo: Windows 11 Pro 64-bit (22H2).

IV. RESULTADOS

El modelo computacional replicó las asignaciones de turnos diseñadas por expertos humanos, logrando una reducción del 84% en el tiempo de planificación. Mientras la programación manual requiere **2 horas** para organizar una semana laboral, el sistema automatizado generó soluciones equivalentes en solo **20 minutos**. Este proceso se divide en dos fases: una asignación inicial de empleados (15 minutos) basada en habilidades y horarios, y un ajuste de continuidad entre turnos (5 minutos). La eficiencia obtenida no solo iguala los estándares de calidad humana, sino que libera aproximadamente 2 horas semanales de gestión administrativa por punto de venta.

La optimización impacta directamente en la operación estratégica: al reducir la carga administrativa, las gerentes pueden reorientar esfuerzos hacia actividades de mayor valor agregado, como análisis de desempeño o mejora de servicio al cliente. Además, la escalabilidad del sistema permite replicar estos beneficios a nivel corporativo, garantizando consistencia

en las asignaciones y cumplimiento normativo en múltiples locaciones simultáneamente, un desafío crítico en cadenas de retail o gastronomía.

Caso de estudio: Restaurante emergente

Se simuló un escenario con 10 empleados polivalentes, horario de 14 horas (06:00–20:00) y 6 franjas horarias variables. Cada empleado contaba con habilidades específicas (Tabla I), excepto en su día de descanso obligatorio. La configuración incluía turnos de 1 a 4 horas, con demanda mínima de 30 asignaciones diarias (5 oficios por franja). Los costos asociados priorizaban penalizar faltantes ($\lambda = 10$) sobre redundancias ($\gamma_w = 1.5$), incentivando cobertura precisa sin overstaffing.

El modelo cumplió el 100% de los requerimientos operativos, evitando excesos de personal (overstaffing = 0%) y respetando días de descanso. La solución combinó una asignación inicial (Figura 2) que distribuía empleados según demanda, y un ajuste de continuidad (Figura 3) para optimizar secuencias de turnos. Este enfoque bidimensional aseguró no solo cobertura, sino también consistencia en las jornadas laborales, crítico para mantener la productividad y satisfacción del equipo.

La ejecución del modelo tomó **30 segundos** en hardware

TABLA I
OFICIOS DISPONIBLES POR EMPLEADO

Empleado	Oficios disponibles
01	[0, 3, 4, 5]
02	[0, 1, 2]
03	[0, 2, 3]
04	[0, 2]
05	[0, 1, 2, 3, 4, 5]
06	[0, 4, 5]
07	[0, 4, 5]
08	[0, 3, 4, 5]
09	[0, 1]
10	[0, 2, 4]

estándar, demostrando eficiencia incluso con múltiples franjas horarias y restricciones complejas. Este rendimiento resalta la viabilidad técnica del sistema para entornos operativos dinámicos, donde la rapidez de respuesta es esencial. Adicionalmente, se verificó el cumplimiento estricto de políticas laborales y competencias, validando la solución como técnica y operativamente robusta.

Los resultados posicionan al sistema como una herramienta estratégica para la gestión de personal. Al integrar precisión en asignaciones, reducción de costos administrativos y adaptabilidad escalable, el modelo supera las limitaciones de la planificación manual. Su capacidad para equilibrar cobertura operativa con optimización de recursos humanos lo hace aplicable no solo en gastronomía, sino en cualquier sector con demandas cíclicas y necesidad de agilidad en la gestión de turnos.

V. CONCLUSIONES

La investigación desarrollada propone un modelo matemático que democratiza el acceso a estrategias avanzadas de programación horaria, permitiendo que restaurantes emergentes alcancen niveles de eficiencia operativa comparables a los obtenidos por expertos humanos. Este enfoque no solo contribuye a la optimización de costos operativos, reduciendo gastos innecesarios, sino que también fortalece la competitividad en mercados con márgenes de ganancia reducidos, donde la eficiencia es un factor determinante para la sostenibilidad.

Asimismo, la implementación de un sistema de optimización basado en métricas cuantitativas logra minimizar los sesgos inherentes a la asignación manual de turnos, promoviendo una estandarización en la toma de decisiones críticas como una cobertura efectiva durante los picos de demanda, garantizando un servicio eficiente sin comprometer la disponibilidad del personal.

Un aspecto destacado del modelo es que no solo optimiza la productividad, sino que también establece un marco ético que respeta las condiciones laborales al integrar límites máximos de horas continuas y periodos mínimos de descanso. Esta característica demuestra que la automatización no está reñida con el cuidado de los empleados, sino que puede ser una herramienta para promover un equilibrio entre los

objetivos operativos y el bienestar del personal.

Finalmente, la arquitectura del sistema se diseñó con un enfoque flexible, lo que le confiere un alto potencial de adaptabilidad. Mediante ajustes paramétricos, el modelo podría extenderse a otros sectores de servicios con necesidades similares, como hotelería o centros de atención al cliente. Esta característica abre la puerta a futuras investigaciones y aplicaciones prácticas, consolidando el modelo como una solución escalable y de amplio alcance para la gestión eficiente de recursos humanos en diversos contextos operativos.

REFERENCIAS

- [1] A. T. Ernst, H. Jiang, M. Krishnamoorthy, and D. Sier, "Staff scheduling and rostering: A review of applications, methods and models," *European Journal of Operational Research*, vol. 153, pp. 3–27, 2004, timetabling and Rostering. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S037722170300095X>
- [2] J. V. den Bergh, J. Beliën, P. D. Bruecker, E. Demeulemeester, and L. D. Boeck, "Personnel scheduling: A literature review," *European Journal of Operational Research*, vol. 226, pp. 367–385, 2013. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0377221712008776>
- [3] M. Defraeye and I. V. Nieuwenhuysse, "Staffing and scheduling under nonstationary demand for service: A literature review," *Omega*, vol. 58, pp. 4–25, 2016. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0305048315000754>
- [4] S. M. Al-Yakoob and H. D. Sherali, "Mathematical models and algorithms for a high school timetabling problem," *Computers & Operations Research*, vol. 61, pp. 56–68, 2015. [Online]. Available: <https://doi.org/10.1016/j.cor.2015.02.011>
- [5] F. L. Burgert, M. Windhausen, M. Kehder, N. Steireif, S. Mütze-Niewöhner, and V. Nitsch, "Workforce scheduling approaches for supporting human-centered algorithmic management in manufacturing: A systematic literature review and a conceptual optimization model," *Procedia Computer Science*, vol. 232, pp. 1573–1583, 2024, 5th International Conference on Industry 4.0 and Smart Manufacturing (ISM 2023). [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1877050924001558>
- [6] República de Colombia, "Código sustantivo del trabajo," 1950, adoptado por el Decreto Ley 2663 del 5 de agosto de 1950, "Sobre Código Sustantivo del Trabajo", publicado en el Diario Oficial No 27.407 del 9 de septiembre de 1950.
- [7] CCMA, "Estudio económico 2023," CCMA, 2023.
- [8] —, "Camara indica: Economía de antioquia 2024 trimestre iv," 2024, informe económico.
- [9] DANE, "Producto interno bruto (pib) nacional trimestral colombia," 2024, consulta de datos nacionales.

Autonomous Microalgae Control with AI and IoT for CO₂ Capture and Renewable Energy

First Author Diana Carolina Montealegre Rodriguez^[0000-0001-6745-2878], Second Author Agustín Rafael Marulanda^[0000-0001-6739-2641], and Third Author Nobel Eliecer Castellanos Bejarano^[0000-0003-1549-224X].

Abstract— This research presents the design and implementation of a smart system for the controlled cultivation of *Chlorella vulgaris*, a microalga known for its fast growth, high photosynthetic efficiency, and ability to help in environmental cleanup. The system uses Internet of Things (IoT) technology to monitor and regulate key conditions such as light, temperature, pH, and carbon dioxide (CO₂) levels. It includes a one-liter glass container illuminated by white and yellow LEDs to simulate ideal light conditions for photosynthesis, avoiding stress or damage to the microalgae. An ESP32-WROOM microcontroller manages sensors and actuators to control the environment and send real-time data to the cloud. Sensors monitor temperature, pH, and CO₂, while actuators adjust heaters, pumps, and CO₂ valves to maintain optimal conditions. The system also includes solar panels to power the components, making it energy-efficient and suitable for remote or off-grid use. This prototype makes it possible to automatically grow microalgae while capturing CO₂ from the air and reducing particulate matter (PM). At the end of each cycle, the biomass and PM collected are measured to assess system efficiency. The results show that this integrated approach can support the production of biomass and renewable energy, while also contributing to cleaner air and climate change mitigation.

Index Terms— Microalgas, Internet de las Cosas (IoT), Captura de CO₂, Energía removable, Biorremediación.

I. INTRODUCTION

In recent decades, the use of microalgae as a biotechnological resource has sparked growing interest due to their high potential for biofuel and bioproduct production, and environmental remediation processes. Among the most studied species, *Chlorella vulgaris* stands out for its high photosynthetic capacity, rapid growth, and ability to adapt to diverse environmental conditions. These characteristics make it an ideal candidate for the development of sustainable systems focused on carbon dioxide capture, renewable energy production, and wastewater treatment (Borowitzka, 1991; Safi et al., 2014; Castillo, 2017; Araujo et al., 2020).

This doctoral research describes the design and evaluation of an intelligent system for the controlled cultivation of *Chlorella vulgaris*, based on IoT technologies. This system includes a one-liter containment unit, constructed with a glass cylinder that facilitates internal light irradiation. To simulate optimal lighting conditions, white and yellow light-emitting diodes were used. White light, with a full spectrum of 400–700 nm, promotes the photosynthetic process, while yellow light, with a wavelength between 570 and 590 nm, modulates growth and can even inhibit germination in certain species (Hu et al., 2008; Kim et al., 2013).

In addition to light control, the system allows for maintaining precise temperature (between 36 and 38°C) and pH (0.7 to 0.8) conditions, critical factors for optimal microalgae development. These parameters are managed through a platform based on the ESP32-WROOM microcontroller, which integrates sensors and actuators with wireless connectivity, enabling real-time monitoring and process automation using artificial intelligence algorithms (Wang et al., 2021).

Therefore, the aim is to evaluate the technical feasibility of the developed prototype, as well as to explore the advantages of integrating automated control and artificial intelligence into microalgae cultivation systems. The aim is to demonstrate the potential of *Chlorella vulgaris* not only as a source of biomass and renewable energy, but also as an effective tool for addressing contemporary environmental challenges such as climate change and water pollution.

II. IMPLEMENTATION OF AN IOT SYSTEM FOR THE MONITORING AND DEVELOPMENT OF MICROALGAE

The microalga *Chlorella vulgaris* has established itself as a highly relevant biotechnological resource due to its high growth rate, its capacity to produce biomass rich in bioactive compounds, its potential for bioenergy generation, and its

*This paragraph of the first footnote will contain the date on which you submitted your paper for review, which is populated by IEEE. It is IEEE style to display support information, including sponsor and financial support acknowledgment, here and not in an acknowledgment section at the end of the article. For example, “This work was supported in part by the U.S. Department of Commerce under Grant BS123456.” The name of the corresponding author appears after the financial information, e.g. (Corresponding author: M. Smith). Here you may also indicate if authors contributed equally or if there are co-first authors.

effectiveness in environmental remediation processes (Becker, 2007; Kothari et al., 2010; Lee et al., 2015). In this context, this research focuses on the implementation of IoT technologies for the automated monitoring and control of *Chlorella vulgaris* crops, with the aim of optimizing their development and biological efficiency in controlled environments.

A prototype smart cultivation system was designed and developed, consisting of a one-liter water containment unit housed in a glass cylinder, enabling controlled exposure to radiation using white and yellow light-emitting diodes (LEDs). This lighting simulates UV radiation conditions necessary to stimulate photosynthesis without inducing light stress (Chisti, 2007; Brennan & Owende, 2010). The system enables precise control of critical environmental variables such as temperature, maintained between 36 and 38°C, and pH, adjusted within a range of 0.7 to 0.8. For this purpose, specialized sensors connected to an IoT platform based on the ESP32-WROOM microcontroller were integrated, enabling data acquisition, real-time analysis, and remote control of the system's actuators (Sharma et al., 2020; Restrepo et al., 2022).

This technological integration not only optimizes constant crop monitoring, but also lays the groundwork for scalability, energy efficiency and CO₂ capture, as well as automation in microalgae production systems designed for environmental and bioindustrial applications (Del Río et al., 2021; López et al., 2023).

2.1 Controlled cultivation system:

The system design prioritizes an ideal environment for microalgae development, optimizing CO₂ absorption and particulate matter (PM) capture from the air. The device avoids direct exposure to intense solar radiation, using filtered light to prevent photoinhibition and heat stress. The controlled intake of contaminated air ensures adequate oxygenation without oversaturating the medium. Additionally, pH and temperature sensors adjust the environment to maintain an optimal growth rate.

The IoT platform continuously monitors the culture, assessing population growth and allowing estimates of biomass, captured CO₂, and settled PM. At the end of each cycle, the generated biomass is quantified as a direct indicator of system efficiency.

This integrated approach of biotechnology and IoT promotes sustainable solutions for water treatment, bioenergy production, and climate change mitigation, moving toward a circular bioeconomy based on the smart use of biological resources.

2.2. Control System Features (ESP32-WROOM):

This microcontroller incorporates a 32-bit Xtensa® Dual-Core processor, enabling simultaneous execution of tasks such as data acquisition, control logic, and wireless communication via Wi-Fi. It integrates multiple peripherals: up to 34 GPIOs, 12-bit ADCs for analog sensors (temperature and pH), PWM channels for actuator control (pumps, resistors, CO₂ valves), and an 8-bit DAC for analog signal generation. The architecture enables efficient operation, even in low-power modes, making it suitable for systems powered by alternative

energy sources. The architecture enables efficient operation, even in low-power modes, ideal for systems powered by alternative energy.

The prototype includes DS18B20 temperature sensors and SEN0161 pH modules, along with actuators such as 40W at 12V resistors and peristaltic dosing pumps. For real-time monitoring of carbon dioxide levels, the system incorporates MH-Z19C or SCD30 NDIR sensors, which provide accurate CO₂ concentration measurements via UART or I2C. Based on these readings, gas solenoid valves are modulated using PWM to control CO₂ injection, ensuring optimal photosynthetic conditions for microalgae growth. All collected data is transmitted to a cloud platform for visualization, analysis, and remote control.

2.2.1 Wireless Connectivity and Peripheral:

WiFi: 802.11 b/g/n

The ESP32-WROOM microcontroller allows connection to standard Wi-Fi networks, facilitating data transmission to the cloud and receiving remote control commands. Temperature and pH data are sent to IoT platforms (such as AWS IoT, Azure IoT Hub, among others) for visualization, analysis, and storage. Furthermore, the system is also designed to integrate specific platforms for CO₂ capture and power generation. This data is managed in real time, optimizing the performance of microalgae cultivation systems based on CO₂ fixation efficiency and bioenergy production. The microcontroller can receive commands from the cloud to adjust temperature and pH setpoints, activate or deactivate actuators, and make adjustments to the CO₂ capture and power generation systems to maximize environmental and energy benefits.

Peripherals:

GPIO (General Purpose Input/Output): Up to 34 GPIO pins are available for connecting temperature and pH sensors, relays that control heaters/coolers, dosing pumps for pH adjustment, indicator LEDs, and other peripherals required for the system. These pins allow for flexible and expandable configuration for various monitoring and control devices.

ADC (Analog-to-Digital Converter): The ESP32 features a 12-bit ADC, used to read voltage signals from the temperature and pH sensors and convert them into precise digital values. These values are processed, enabling precise monitoring and control of the system's environmental conditions.

DAC (Digital-to-Analog Converter): With an 8-bit DAC, the ESP32 can control actuators that require analog signals, such as proportional valves for precise dosing of acids or bases, facilitating fine-tuning of pH parameters.

I2C Protocol: Some temperature and pH sensors use the I2C protocol for efficient communication between devices, allowing the integration of multiple sensors with minimal pinouts. Additionally, the system can include an OLED display for local data visualization, improving accessibility and on-site monitoring.

PWM (Pulse Width Modulation): The ESP32 features up to 16 PWM channels, allowing precise control of devices such as peristaltic pumps for pH adjustment, heating resistors, CO₂ solenoid valves, and the intensity of LED indicators. This

modulation capability is key to dynamically adapting the growing system's conditions based on monitored environmental parameters.

MH-Z19C or SCD30 (NDIR) Sensors: Non-dispersive infrared sensors that enable real-time measurement of carbon dioxide (CO₂) concentration in the growing environment. These devices are integrated via UART or I2C interfaces to the ESP32 microcontroller, providing continuous data that is processed for precise air quality monitoring in the system.

CO₂ Solenoid Valves + PWM Control: The system incorporates solenoid valves controlled by pulse-width modulation (PWM), enabling controlled dosing of CO₂ based on the values measured by the sensors. This integration ensures dynamic regulation of the gas supply, optimizing the photosynthetic conditions of the microalgae cultivation.

2.2.2 Environmental and Peripheral Sensors:

Temperature and pH Sensors:

Temperature: The DS18B20 digital sensor provides accurate measurements in the -55°C to 125°C range, suitable for system thermal monitoring.

pH: The SEN0161 module includes a signal-conditioned probe that delivers a stable analog output, ideal for conversion using the ESP32's 12-bit ADC.

CO₂ Sensors:

MH-Z19C or SCD30 (NDIR): Non-dispersive infrared sensors that enable real-time measurement of carbon dioxide (CO₂) concentration, a crucial parameter for optimizing photosynthesis in microalgae cultures.

Environmental Control Actuators: Heating: 40W 12V relay-controlled heating element to maintain optimal cultivation temperatures.

pH Adjustment: PWM-controlled peristaltic dosing pumps for precise addition of acidic or basic solutions.

CO₂ Injection: PWM-controlled solenoid valves regulate CO₂ delivery based on data from NDIR sensors, helping to improve photosynthetic efficiency.

2.2.3. Energy Management and Efficient Consumption:

Power and Energy Management: The microcontroller requires a regulated 3.3V power supply for basic operation, while the actuators (resistors, pumps, and valves) operate with external 5V and 12V sources. The design includes the integration of alternative power sources, such as solar panels, to ensure energy autonomy in continuous monitoring systems.

Low Power Consumption: The ESP32 offers multiple low-power modes (deep sleep and light sleep), ideal for power-limited or battery-based applications. These modes significantly reduce power consumption during periods of inactivity, optimizing system efficiency without compromising its responsiveness to critical variations in temperature, pH, or CO₂.

2.3. Design of an IoT Microalgae Cultivation System for CO₂ Capture, Particulate Matter Reduction, and Renewable Energy Generation:

The device has been designed to establish optimal conditions for microalgae growth, maximizing carbon dioxide (CO₂)

uptake and particulate matter (PM) retention. The system operates under controlled parameters of light, temperature, CO₂, pH, oxygenation, and power generation, with the goal of creating an ideal environment for microalgae cultivation in contaminated media (see Figure 1: Controlled Microalgae Cultivation System for CO₂ and Particulate Matter Capture).

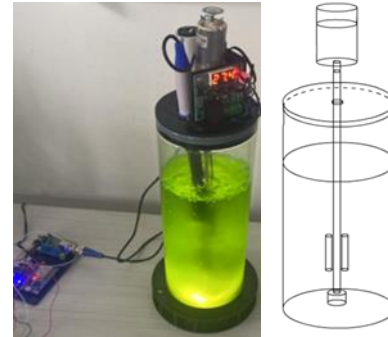


Figure 1. Montealegre, D. C., Marulanda, A.R. & Castellano, N. E. (2025). Controlled Microalgae Cultivation System for CO₂ and Particulate Matter Capture

Solar Radiation Control: The system is designed to avoid direct exposure to solar radiation, as excessive light intensity can raise the internal temperature and cause cellular damage in microalgae, negatively affecting their metabolism. To mitigate these effects, light filtering or scattering mechanisms are used to maintain adequate light levels, promoting efficient photosynthesis without causing thermal stress or photoinhibition. This regulation optimizes the microalgae's CO₂ capture process.

Polluted Air Flow and Oxygenation: The intake of polluted air into the system is precisely controlled, as it constitutes the main source of carbon dioxide (CO₂) and particulate matter (PM) to be captured. A moderate flow ensures the entry of pollutants without saturating the system, allowing for effective oxygenation through bubbling. This mechanism not only maintains adequate dissolved oxygen levels but also optimizes conditions for a high rate of CO₂ fixation by the microalgal biomass.

Temperature Control: The system's temperature is maintained within the optimal range for the cultured species using electronically controlled sensors and heating elements. An appropriate temperature maximizes the microalgae's metabolic activity, increasing their efficiency in photosynthesis and, consequently, in CO₂ capture. The system is also energy-efficient, leveraging renewable energy sources such as solar panels to power the heating elements and sensors.

pH Regulation: The culture medium is maintained within a neutral or slightly alkaline pH range, depending on the microalgae used. This is achieved through pH sensors and an automatic dosing system controlled by a peristaltic pump, which adjusts the levels in real time. A stable pH promotes cell growth and, therefore, the microalgae's ability to continuously and efficiently fix carbon dioxide.

Monitoring, growth evaluation, and CO₂ capture: The system incorporates sensors that collect data on environmental variables and microalgae development. This data allows for

estimating cell density and, using algal growth models, calculating the amount of CO₂ absorbed during cultivation. Additionally, the system integrates NDIR sensors (such as the MH-Z19C or SCD30) to measure CO₂ concentration in the air in real time before and after treatment, thus quantifying the system's capture efficiency.

Particulate Matter Assessment and Power Generation: The system's efficiency is also assessed by quantifying the particulate matter retained by sedimentation at the base of the device. This accumulation is measured at the end of each cycle as evidence of the air purification process. To ensure continuous operation in autonomous or remote environments, the system incorporates power generation modules, primarily solar panels, which power the sensors, actuators, and the ESP32 microcontroller. This ensures energy sustainability without compromising system functionality.

Acknowledgments.

V. CONCLUSION

The research demonstrated that the integration of Internet of Things (IoT) technologies with biotechnological processes applied to the cultivation of *Chlorella vulgaris* is technically and environmentally viable. The intelligent system developed enabled precise control of critical variables such as light, temperature, pH, and CO₂ concentration, ensuring optimal conditions for the microalgae's growth and strengthening its carbon and particulate matter capture capacity. The prototype demonstrated that automation, combined with the use of renewable energy, improves cultivation efficiency and biomass production, enabling its implementation in urban, industrial, or remote settings. This presents a scalable model that contributes to climate change mitigation, bioenergy generation, and the bioremediation of contaminated environments..

ACKNOWLEDGMENT

This study was made possible thanks to the financial and academic support provided by the Minuto de Dios University Corporation, in conjunction with the Julio Garavito Colombian School of Engineering, within the framework of the PhD program in Engineering. I am deeply grateful to both institutions for providing the resources, guidance, and support necessary for the development of this project. I also extend my gratitude to the professors, researchers, and collaborators who contributed their experience and knowledge during the different stages of this research.

REFERENCES

Basic format for periodicals:

- [1] Araújo, G. S., Matos, L. J. B. L., Fernandes, J. O., Cartaxo, S. J. M., Gonçalves, L. R. B., Fernandes, F. A. N., & Farias, W. R. L. (2020). *Chlorella vulgaris*: Potential applications and challenges in biofuel production and wastewater treatment. *Renewable and Sustainable Energy Reviews*, 118, 109431. <https://doi.org/10.1016/j.rser.2019.109431>
- [2] Becker, E. W. (2007). Micro-algae as a source of protein. *Biotechnology Advances*, 25(2), 207–210.
- [3] Borowitzka, M. A. (1991). Commercial production of microalgae: ponds, tanks, tubes, and fermenters. *Journal of Biotechnology*, 70(1-3), 313–321. [https://doi.org/10.1016/0168-1656\(91\)90050-D](https://doi.org/10.1016/0168-1656(91)90050-D)
- [4] Brennan, L., & Owende, P. (2010). Biofuels from microalgae—A review of technologies for production, processing, and extractions of biofuels and co-products. *Renewable and Sustainable Energy Reviews*, 14(2), 557–577.
- [5] Castillo, F. J. (2017). Microalgae as a sustainable energy alternative: Biofuel production from *Chlorella vulgaris*. *Revista Colombiana de Biotecnología*, 19(1), 45–55.
- [6] Chisti, Y. (2007). Biodiesel from microalgae. *Biotechnology Advances*, 25(3), 294–306.
- [7] Del Río, L., García, J. L., & Muñoz, R. (2021). Technological advancements in microalgae cultivation for industrial applications: A review. *Algal Research*, 58, 102404.
- [8] Hu, Q., Sommerfeld, M., Jarvis, E., Ghirardi, M., Posewitz, M., Seibert, M., & Darzins, A. (2008). Microalgal triacylglycerols as feedstocks for biofuel production: perspectives and advances. *The Plant Journal*, 54(4), 621–639. <https://doi.org/10.1111/j.1365-3113X.2008.03492.x>
- [9] Kim, J., Yoo, G., Lee, H., Lim, J., Kim, K., & Lee, Y. (2013). Effects of wavelength and intensity of LED light on the growth of green alga *Chlorella vulgaris*. *Biochemical Engineering Journal*, 56(3), 20–25. <https://doi.org/10.1016/j.bej.2013.04.012>
- [10] Kothari, R., Pathak, V. V., Kumar, V., & Singh, D. P. (2010). Experimental study for growth potential of unicellular alga *Chlorella pyrenoidosa* on domestic wastewater. *Bio-resource Technology*, 121, 479–482.
- [11] Lee, Y. K., Ding, S. Y., & Low, C. S. (2015). Microalgae for biodiesel production and CO₂ sequestration. *Journal of Environmental Science and Technology*, 49(15), 8986–8994.
- [12] López, A., Ortega, M., & Ruiz, F. (2023). Smart photobioreactors using ESP32 for scalable algal biomass production. *Sensors*, 23(4), 1738.
- [13] Restrepo, D., Pérez, H., & Martínez, L. (2022). Development of an IoT system for the automated cultivation of microalgae in the laboratory. *Revista Colombiana de Biotecnología*, 24(2), 143–154.
- [14] Safi, C., Zebib, B., Merah, O., Pontalier, P. Y., & Vaca-Garcia, C. (2014). Morphology, composition, production, processing, and applications of *Chlorella vulgaris*: A review. *Renewable and Sustainable Energy Reviews*, 35, 265–278. <https://doi.org/10.1016/j.rser.2014.04.007>
- [15] Sharma, A., Jain, R., & Raj, T. (2020). IoT-based smart system for monitoring algal growth in photobioreactors. *Sensors and Actuators B: Chemical*, 320, 128276.
- [16] Wang, Y., Yu, Y., Hu, Y., Liu, Y., & Lin, J. (2021). Smart photobioreactor system with artificial intelligence for optimizing microalgae production. *Biotechnology Advances*, 49, 107759. <https://doi.org/10.1016/j.biotechadv.2021.107759>



First Author Espinal, Tolima (Colombia), January 30, 1982. She holds a degree in Industrial Engineering from the Universidad Cooperativa de Colombia and is a doctoral student at the Julio Garavito Colombian School of Engineering. Her areas of interest focus on the research and development of alternative energies, particularly the generation, storage, and

utilization of non-conventional energies.

She currently works as a professor in the Industrial Engineering program at the Corporación Universitaria Minuto de Dios (UNIMINUTO), where she has taught courses such as Management Fundamentals, Introduction to Engineering, Cleaner Production, Industrial Processes, Drawing, Management Information Systems, Quality Management, Quality Tools, and Research Seminars, among others. She has also worked as an Industrial Engineering professor at the Fundación Universitaria del Área Andina, teaching courses related to Introduction to Engineering, Waste Management, Environmental Safety, Statistical Quality Control, Integrated Management Systems, Quality, Safety, and Occupational Health, among others. In her role as a researcher, she has contributed to the development of scientific articles in the fields of engineering and sustainability.



Second Author Venezuela, January 7, 1970. He holds a degree in Electrical Engineering from the University of Zulia, Venezuela, and a PhD from the University of Seville, Spain. His areas of interest include power system modeling and optimization, power system stability, renewable resource integration, power quality, and electric

vehicles.

Trabajó en la Escuela de Ingeniería Eléctrica de la Universidad del Zulia (1995-2016) como Profesor Asistente, Profesor Titular y Profesor titular, ocupado los cargos de Jefe del Departamento de Potencia, Director de la Escuela de Ingeniería Eléctrica y Director de la Maestría de Ingeniería Eléctrica y Coordinador Académico de la Facultad de Ingeniería de la Universidad del Zulia. De igual forma trabajó en la creación y puesta en marcha de la Maestría en Ingeniería Eléctrica de la Universidad del Zulia. Actualmente trabaja como Profesor Asociado en el Programa de Ingeniería Eléctrica de la Escuela Colombiana de Ingeniería Julio Garavito, impartiendo cursos de Pregrado y Posgrado. Ha recibido la Orden al Mérito Universitario Jesús Enrique Lossada en su Segunda y Primera Clase otorgado por la Universidad del Zulia en los años 2010 y 2014, respetivamente. También ha recibido el Premio Dr. Francisco Eugenio Bustamante como investigador otorgado por el Vicerectorado Académico de la Universidad del Zulia en el año 2008, así como reconocimientos de IEEE-Inelectra por tutorías académicas en los años en los años 2008 y 2010.



Third Author, Bogotá, Colombia, October 3, 1982. Electronic Engineer, PhD in Complex Thinking and Master's in Design and Management of Technological Projects. Areas of interest and experience focus on the research and development of innovative technologies with a strong social focus, Alternative Energy (generation, storage, and use of

non-conventional energies), Direction and management of research and engineering projects, Innovation and Prototyping (development of high-impact technological products and prototyping in pre-commercial stages). Scientific Thinking (the pursuit of new knowledge through experimentation and transdisciplinary methodology).

Academic and Professional Career. He worked at the Minuto de Dios University Corporation between 2011 and 2024, serving as a Research Associate and Professor in the Electronics Technology and Industrial Engineering programs. At this institution, he taught courses on industrial automation, programmable logic controllers (PLCs), pneumatic systems, and industrial systems simulation. He also taught in the Electronics Technology program, teaching microcontrollers, robotics, analog and digital communication systems, and digital electronics. He has held the roles of Robotics and Automation Seedbed Leader and Research Project Leader. Currently, he works as a professor of the Master's program in Climate Change at the Minuto de Dios University Corporation, teaching the Ecotechnologies and Renewable Energy Systems component. He is also an Industrial Engineering professor at the Andean Area University Foundation 2024 and 2025, where he teaches courses on Big Data, Data Mining, and Advanced Python Programming. Similarly, in his role as a researcher, he has contributed to the development of scientific articles and supported the supervision of undergraduate theses. Recognitions He has been recognized by the Minuto de Dios University Corporation for the university's first invention patent. This patent, granted under number NC2017/0004741, is entitled "Alternative Energy Generation Systems."

FPGA-Based Real-Time Heat Solver for Navier-Stokes-Inspired UWOC Modeling

Miguel Solarte-Sanchez, David Marquez-Viloria, Isnardo Arenas-Navarro, and Neil Guerrero-Gonzalez

Abstract—This paper presents a real-time FPGA implementation of the heat equation as a computational core for modeling underwater optical communication channels. By employing the explicit finite difference method with a Red-Black ordering scheme, we achieve parallelized updates that are well-suited for hardware acceleration. The solver was deployed on a Xilinx ZCU104 FPGA platform and validated with a two-dimensional diffusion scenario using 64 spatial points and 252 time steps. The implementation satisfies the Courant–Friedrichs–Lewy stability condition and exploits a memory-efficient architecture using 310 BRAMs and 415 DSP slices. The system achieved a total latency of 37,510 clock cycles while maintaining accurate spatiotemporal thermal profiles consistent with theoretical expectations. These results confirm the viability of the design for real-time physical simulations. Future work includes coupling this thermal solver with Navier–Stokes-based fluid models to enable more comprehensive underwater channel characterization.

Index Terms—FPGA implementation, Real-time simulation, Finite difference method, Heat equation solver, Red-Black scheme, Underwater optical communication

Underwater optical communication (UWOC) has gained significant attention in recent years due to its potential for high-bandwidth, low-latency data transmission in aquatic environments. Unlike acoustic communication, optical systems offer higher data rates but are severely affected by water turbidity, temperature gradients, and flow-induced refractive index variations. Accurately modeling these dynamic conditions remains an open challenge, especially in scenarios where real-time emulation is necessary for testing and validating communication systems [1], [2].

Traditionally, underwater channels have been modeled using empirical or statistical approaches, such as Monte Carlo simulations or Beer-Lambert-based atten-

uation models. While useful, these methods lack the capacity to represent the fluid dynamics and transient behavior that can influence optical signal propagation in realistic conditions [3], [2]. In contrast, physical models derived from the Navier-Stokes equations can offer a more comprehensive representation of the underwater medium, capturing the interplay between temperature, fluid motion, and light scattering [4]. However, their computational complexity makes real-time execution impractical on conventional processors [5].

To address this limitation, simplified partial differential equations (PDEs) such as the heat equation have been employed as surrogates for fluid dynamics under specific assumptions. The heat equation captures temperature diffusion over time and serves as a computationally tractable approximation of the Navier-Stokes behavior in certain regimes, particularly for modeling thermal plumes and refractive index fluctuations that affect optical channels [6].

Recent advances in hardware acceleration using Field-Programmable Gate Arrays (FPGAs) have opened new possibilities for real-time PDE solving [7]. FPGAs offer parallelism, low latency, and reconfigurability, making them ideal platforms for implementing physical models in high-speed communication testbeds [8]. Several works have demonstrated FPGA-based solvers for the heat equation and shallow-water equations, but few have targeted their use in the context of real-time channel distortion for underwater optical communications [2].

This work positions itself at the intersection of these domains by proposing an FPGA-based real-time solver of the heat equation as a core component in approximating the Navier-Stokes behavior of underwater optical channels. The proposed hardware implementation utilizes the explicit finite difference method with a Red-Black scheme to optimize memory access and parallelism, enabling efficient real-time simulation on FPGA platforms. This implementation aims to serve as the dynamic kernel of a hardware-in-the-loop system that emulates signal distortion in subaquatic environments for testing and future characterization using arbitrary modulation constellations [9], [10]. In addition, this approach opens the door to further advancements in real-time modeling for complex fluid dynamics, providing

M. Solarte-Sanchez is a Master's student in Automation and Industrial Control in the Department of Electronics and Telecommunications, Instituto Tecnológico Metropolitano, Medellín, Colombia, e-mail: miguelsolarte244621@correo.itm.edu.co.

D. Marquez-Viloria is with Instituto Tecnológico Metropolitano, Medellín, Colombia, e-mail: davidmarquez@itm.edu.co.

I. Arenas-Navarro is with the Mathematics Department, Universidad Militar Nueva Granada, Cajicá, Colombia., e-mail: isnardo.arenas@unimilitar.edu.co.

N. Guerrero-Gonzalez is with the Department of Electrical, Electronic and Computer Engineering, Universidad Nacional de Colombia, Manizales, Colombia., e-mail: nguerrero@unal.edu.co.

Manuscript received April 19, 2005; revised August 26, 2015.

a valuable tool for the development and validation of underwater optical communication systems [11], [12].

I. METHODOLOGY

The heat equation was numerically solved using the explicit finite difference method, a widely used technique for time-domain simulations due to its simplicity and ease of implementation on parallel computing platforms such as FPGAs. The one-dimensional heat equation is discretized using forward differences in time and central differences in space, leading to the following update rule:

$$u_i^{n+1} = u_i^n + \frac{\alpha \Delta t}{\Delta x^2} (u_{i+1}^n - 2u_i^n + u_{i-1}^n)$$

Here, u_i^n represents the temperature at spatial position x_i and time t_n , α is the thermal diffusivity, Δt is the time step, and Δx is the spatial discretization step. This formulation allows for an iterative computation of the solution over time, using only local neighboring values from the previous time step.

To ensure numerical stability, the Courant–Friedrichs–Lewy (CFL) condition must be satisfied:

$$\frac{\alpha \Delta t}{\Delta x^2} < \frac{1}{2}$$

This inequality constrains the choice of Δt and Δx such that the simulation remains stable and does not diverge. It reflects the requirement that information must not propagate faster than the numerical scheme can resolve.

Figure 1 illustrates the general space-time mesh layout corresponding to the explicit scheme, where temperature values are updated row-wise as time progresses.

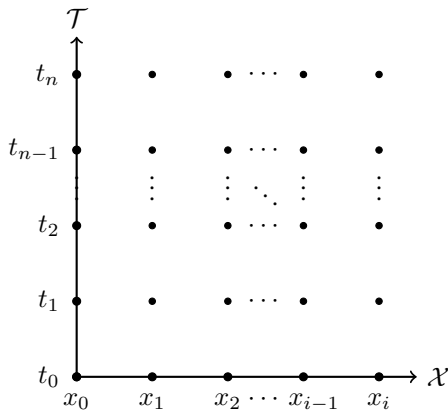


Figure 1. General space-time mesh for the explicit finite difference method.

A. Red-Black Scheme and Hardware Mapping

To exploit parallelism and improve computational throughput in hardware, a Red-Black ordering scheme was adopted. This method alternates updates between two interleaved sets of nodes (red and black), allowing concurrent updates of non-adjacent points in the computational domain without data hazards. This scheme minimizes memory access contention and improves pipeline efficiency in hardware accelerators.

Figure 2 illustrates the Red-Black update pattern. In the first half of each time step, the red nodes are updated using values from the previous time level; in the second half, the black nodes are updated based on the new values of the red nodes. This division ensures no simultaneous write conflicts during a parallel update.

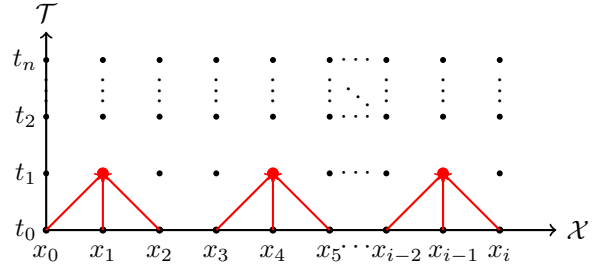


Figure 2. First iteration of Red-Black explicit method for parallel FPGA implementation.

B. FPGA Implementation Architecture

The numerical model was synthesized and deployed onto an FPGA platform using High-Level Synthesis (HLS) tools. The architecture was designed to exploit parallelism, optimize memory access, and support parameterized configurations for grid resolution and time steps.

The design, shown in Figure 3, features the following key aspects:

- Each spatial node is mapped to a dedicated Block RAM (BRAM), enabling localized memory access and parallel updates across the domain.
- Boundary conditions $u(0, t)$ and $u(L, t)$ are handled via fixed BRAM blocks or registers if they remain constant, ensuring proper enforcement of Dirichlet constraints.
- The Red-Black structure permits a three-point spacing between updated nodes, which removes the need for memory locks or arbitration mechanisms.
- The design includes a host interface for setting initial conditions and simulation parameters, providing runtime flexibility without resynthesis.
- Each BRAM stores a time history for its associated spatial point, allowing temporal indexing and enabling post-processing or verification.

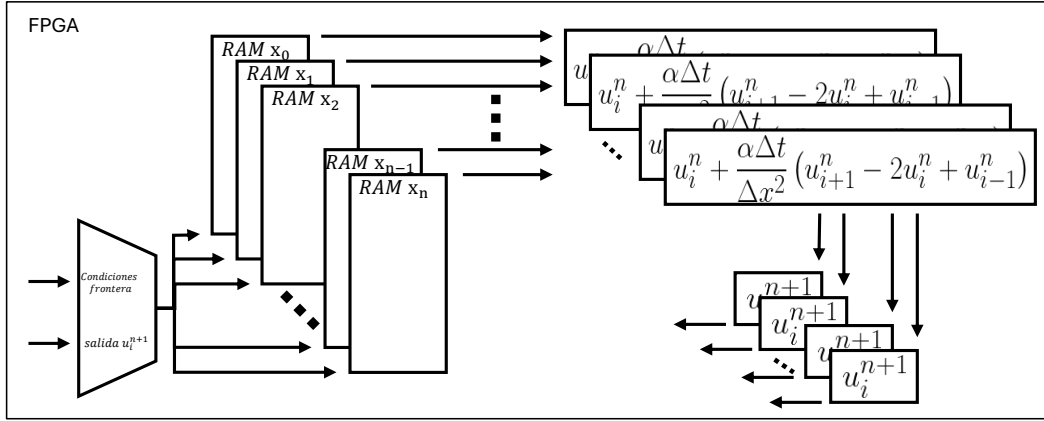


Figure 3. Design scheme for FPGA implementation of the heat equation solver.

This implementation strategy enables real-time simulation capabilities and serves as a reusable building block for more complex fluid dynamics models, such as those based on the Navier–Stokes equations.

II. RESULTS

To validate the proposed hardware implementation of the explicit finite difference method for the heat equation, a set of initial and boundary conditions was defined to simulate a two-dimensional diffusion process. The following configuration was used as the test scenario:

A. Boundary and Initial Conditions

$$\begin{aligned} N &= 64 && \text{(number of spatial points)} \\ I &= 252 && \text{(number of time steps)} \\ \text{CFL} &= 0.1 && \text{(Courant–Friedrichs–Lewy number)} \end{aligned}$$

$$u(x, 0) = \begin{cases} 2x, & 0 \leq x < 0.5 \\ 2(1 - x), & 0.5 \leq x \leq 1 \end{cases}$$

$$u(0, t) = 0, \quad u(L, t) = 0$$

This initial condition defines a symmetric triangular temperature profile along the spatial axis x , which was extended into a second spatial dimension for the two-dimensional simulation. The boundary conditions enforce zero temperature along all borders of the 2D domain, allowing the observation of pure diffusion behavior over time.

The implementation was deployed and tested on a Xilinx ZCU104 development board, which features a

Zynq UltraScale+ MPSoC. This platform provides the necessary computational resources and BRAM capacity to handle multidimensional data arrays efficiently. Furthermore, it supports high-level synthesis workflows that ease the hardware description and acceleration of finite difference computations.

Figure 4 shows the simulation results obtained on the FPGA using the described configuration. The figure represents the spatiotemporal evolution of the heat distribution in two dimensions. As time progresses, the initial temperature distribution spreads symmetrically and smooths out due to the diffusive nature of the heat equation. The central peak diminishes while the heat propagates radially outward.

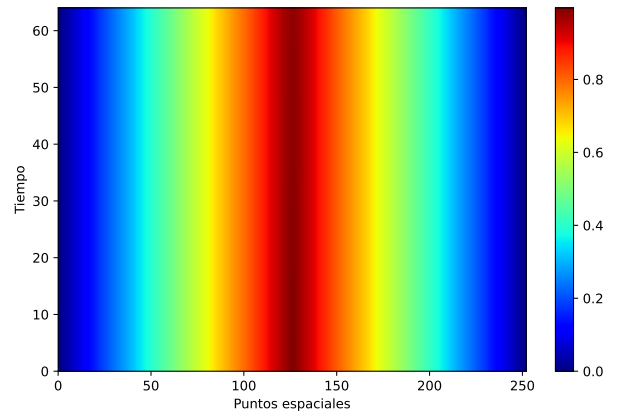


Figure 4. Simulation results using 64 spatial points and 252 time steps in two dimensions.

Thanks to the Red-Black scheme, the design was able to update non-overlapping spatial nodes concurrently, significantly reducing memory contention and maximizing parallel throughput. The achieved real-time execution

confirms that this hardware architecture is well-suited for embedded physical simulations requiring low-latency and multidimensional computation.

Table I summarizes the FPGA resource utilization and latency of the implemented heat equation module on the ZCU104 board. The design achieved a latency of 37,510 clock cycles, which demonstrates the efficiency of the proposed Red-Black structure for explicit finite difference schemes. In terms of hardware resources, the implementation required 310 BRAM blocks, 415 DSP slices, 56,444 flip-flops (FF), and 59,555 lookup tables (LUTs). These values confirm the feasibility of deploying the solver on mid-to-high-end FPGA platforms while preserving scalability for larger mesh sizes or more complex systems. The optimized memory usage and parallel computation structure allow for a significant reduction in time-to-solution, which is particularly valuable for real-time applications such as underwater optical channel characterization.

Table I
FPGA RESOURCE UTILIZATION AND LATENCY FOR THE HEAT EQUATION MODULE.

Module Name	Latency (cycles)	BRAM	DSP	FF	LUT
Heat Equation	37,510	310	415	56,444	59,555

The experiment demonstrates the practical feasibility and computational performance of the proposed FPGA-based solver as a core component for more complex physical models, including those based on Navier–Stokes equations for underwater optical channel modeling.

III. CONCLUSION

This work demonstrated the feasibility of implementing the heat equation in real-time using FPGAs, with an explicit finite difference method and Red-Black updating scheme. The design leveraged the parallelism and deterministic timing of FPGAs to achieve high-throughput computation suitable for emulating physical-layer dynamics in underwater optical communication (UWOC) channels.

The proposed architecture is scalable and resource-efficient, making it suitable for real-time testbeds or hybrid hardware-in-the-loop simulation environments. Results confirmed the solver’s ability to emulate transient thermal diffusion effects that alter the refractive index profile, providing insights into signal distortion due to temperature variations.

Future work will extend this approach to incorporate more comprehensive fluid dynamics models based on the Navier–Stokes equations, enabling the emulation of turbulence and flow-induced effects in underwater channels. This progression will support the evaluation of

advanced optical modulation formats and adaptive systems under realistic physical conditions, consolidating the role of FPGAs as core elements in reconfigurable physical emulation platforms. We will also include a sensitivity analysis of initial conditions in our future work to examine the model’s behavior.

ACKNOWLEDGMENT

This research was funded by the Office of Naval Research Global (ONRG) and the Air Force Office of Scientific Research/Southern Office of Aerospace Research and Development (AFOSR/SOARD) under grant number N62909-24-1-2088.

REFERENCES

- [1] K. Wang, Y. Gao, M. Dragone, Y. Pétillot, and X. Wang, “A deep echo state network-based novel signal processing approach for underwater wireless optical communication system with pam and ofdm signals,” *Photonics*, vol. 10, p. 763, 2023.
- [2] C. Gabriel, M. Khalighi, S. Bourennane, P. León, and V. Rigaud, “Monte-carlo-based channel characterization for underwater optical communication systems,” *Journal of Optical Communications and Networking*, vol. 5, p. 1, 2012.
- [3] M. Ju, P. Huang, Y. Li, and H. Shi, “Effect of receiver’s tilted angle on the capacity for underwater wireless optical communication,” *Electronics*, vol. 9, p. 2072, 2020.
- [4] Y. Weng, T. Matsuda, Y. Sekimori, J. Pajarinen, J. Peters, and T. Maki, “Sim-to-real transfer for underwater wireless optical communication alignment policy between auvs,” in *OCEANS 2022 - Chennai*, 2022, pp. 1–7.
- [5] R. Jiang, C. Sun, L. Zhang, X. Tang, H. Wang, and A. Zhang, “Deep learning aided signal detection for spad-based underwater optical wireless communications,” *Ieee Access*, vol. 8, pp. 20 363–20 374, 2020.
- [6] R. Tian, T. Wang, X. Shen, R. Zhu, L. Jiang, Y. Lu, H. Lu, Y. Song, and P. Zhang, “108 m underwater wireless optical communication using a 490 nm blue vecsel and an aom,” *Sensors*, vol. 24, p. 2609, 2024.
- [7] C. Li, H. Lu, W. Tsai, Z. Wang, C. Hung, C. Su, and Y. Lu, “A 5 m/25 gbps underwater wireless optical communication system,” *Ieee Photonics Journal*, vol. 10, pp. 1–9, 2018.
- [8] X. Liu, S. Yi, X. Zhou, S. Zhang, Z. Fang, Z. Qiu, L. Hu, C. Cong, L. Zheng, R. Liu, and P. Tian, “Laser-based white-light source for high-speed underwater wireless optical communication and high-efficiency underwater solid-state lighting,” *Optics Express*, vol. 26, p. 19259, 2018.
- [9] G. He, Z. Lv, C. Qiu, and Z. Liu, “21.3: the influence of nlos on underwater wireless optical communication system using an ingan-based micro-led and nrz-ook modulation,” *Sid Symposium Digest of Technical Papers*, vol. 52, pp. 286–288, 2021.
- [10] H. Huo, M. Fu, X. Liu, and B. Zheng, “Dcc-gan-based channel emulator for underwater wireless optical communication systems,” *Frontiers in Marine Science*, vol. 10, 2023.
- [11] X. Sun, W. Cai, O. Alkhazragi, E. Ooi, H. He, A. Chaaban, C. Shen, H. Oubei, M. Khan, T. Ng, M. Alouini, and B. Ooi, “375-nm ultraviolet-laser based non-line-of-sight underwater optical communication,” *Optics Express*, vol. 26, p. 12870, 2018.
- [12] T. Wu, Y. Chi, H. Wang, C. Tsai, and G. Lin, “Blue laser diode enables underwater communication at 12.4gbps,” *Scientific Reports*, vol. 7, 2017.

Una implementación de la propuesta de modelo extendido de administración de tierras (LADM-COL) para el inventario vial

A. Cruz, L. Rocha & W. Fernández

Resumen—La infraestructura vial es uno de los pilares fundamentales en el crecimiento económico de un país y en la calidad de vida de sus habitantes. Para identificar las necesidades y garantizar la conservación de la infraestructura, el gobierno y las autoridades locales deben medir, mapear e inventariar regularmente los activos viales. En Colombia, este tipo de inventario está a cargo de la administración municipal, departamental y el gobierno nacional, bajo contratos de concesión. El Modelo de Dominio de Administración de Tierras (LADM) surge como un estándar internacional (ISO 19152:2012) para describir la realidad del territorio y LADM-COL es el perfil que resulta de adaptar la norma ISO 19152:2012 a las particularidades de Colombia. Teniendo en cuenta que este modelo conceptual se puede extender a diferentes temáticas en la administración del territorio, este estudio presenta una implementación de la propuesta de modelo extendido LADM-COL para el inventario vial. El modelo de datos integra las necesidades de las diferentes entidades en un diagrama de clases UML. El flujo de implementación utiliza el lenguaje descriptivo de INTERLIS y la información proviene de inventarios de datos abiertos y de las misiones en campo realizadas con un sistema de mapeo móvil en la ciudad de Bogotá, Colombia. El resultado es una base de datos geográfica con información del inventario vial para apoyar la toma de decisiones de la gestión vial en un lenguaje interoperable para las entidades. El modelo propuesto desarrollado en la investigación tiene como finalidad estandarizar la información vial y promover el intercambio de datos entre los diferentes niveles territoriales en la temática de transporte terrestre, tomando como base el concepto de modelo extendido del núcleo LADM-COL.

Palabras clave— Red vial, Infraestructura de transporte, LADM, Modelo extendido LADM-COL, Interoperabilidad.

I. INTRODUCCIÓN

Con los años, existe una presión cada vez mayor sobre la tierra debido a factores como el cambio climático, la disponibilidad limitada de tierras, el aumento de la población, la urbanización y los problemas de transporte [1]. En la administración de tierras, el sector vial es uno de los pilares fundamentales en el crecimiento económico debido a su papel en el movimiento de bienes y personas [2]. Por lo que, el

funcionamiento de las vías en óptimas condiciones fortalece la presencia del Estado en el territorio, permite articular los centros de producción, y promueve la oferta de servicios públicos, de salud, seguridad y educación [3].

La red vial de Colombia está dividida en red primaria, secundaria y terciaria. La red primaria está a cargo del gobierno nacional bajo contratos de concesión, la red secundaria pertenece a las administraciones departamentales, y la red terciaria comparte la administración departamental, municipal, y el gobierno nacional. Este esquema no ha permitido una cobertura total del país con información precisa y actualizada. A pesar de los esfuerzos por integrar una base de datos geográfica de los diferentes administradores viales, la información que existe sobre un inventario de la red vial de Colombia, es escasa.

El Modelo de Dominio de Administración de Tierras (LADM, por sus siglas en inglés), propuesto por la Organización Internacional de Normalización (ISO, por sus siglas en inglés) 19152, proporciona una terminología común para la administración del territorio, dentro de un estándar internacional. La adaptación de la norma ISO 19152:2012 a las características propias de Colombia es el resultado de un trabajo interinstitucional que consolida el modelo núcleo LADM-COL. Los modelos extendidos y de aplicación responden a las necesidades temáticas de los objetos territoriales en conformidad con el estándar internacional y el modelo núcleo.

Este trabajo propone una implementación de la propuesta de modelo extendido basado en el modelo núcleo de administración de tierras LADM-COL que integre los datos del inventario vial, estado, rehabilitación y mantenimiento de pavimentos. El modelo es aplicado con datos abiertos dispuestos por las entidades y la detección de fallas de la ciudad de Bogotá (Colombia), adquiridas a partir de un sistema de mapeo móvil, un prototipo con componentes de hardware y software que se instala en un vehículo para la captura de información.

A. Cruz es estudiante de la Maestría en Ciencias de la Información y las Comunicaciones Énfasis Geomática de la Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. Miembro del Grupo de Investigación en Infraestructura Sostenible (e-mail: accruz@udistrital.edu.co).

L. Rocha es Doctora en Geografía, docente de la Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. Directora del semillero de investigación SDI y miembro del grupo de investigación NIDE (e-mail: lrocha@udistrital.edu.co).

W. Fernández es Post Doctor en Infraestructura del Transporte Sostenible, decano de la Facultad del Medio Ambiente y Recursos Naturales de la Universidad Distrital Francisco José de Caldas, Bogotá, Colombia. Director del Grupo de Investigación en Infraestructura Sostenible (e-mail: wfernandez@udistrital.edu.co).

II. LADM

El modelo LADM (Land Administration Domain Model) fue propuesto en 2006 con el objetivo de proporcionar una ontología integral y estandarizada para facilitar la conceptualización de la gestión de la tierra [4]. Inicialmente, se centró en el dominio catastral y en 2012, el LADM evolucionó hasta convertirse en una norma internacional ISO-19152, que abarca información básica relacionada con los componentes de la administración de la tierra (incluyendo el agua y los elementos por encima y por debajo de la superficie terrestre).

El modelo está compuesto por 3 paquetes: Interesado, Administrativo y Unidad Espacial, además de un subpaquete denominado Topografía y Representación, como se ilustra en la Fig. 1. El paquete Interesado representa a individuos o grupos de individuos asociados al paquete Administrativo que consta de dos partes esenciales, la Unidad Básica Administrativa como el objeto territorial legal en la gestión de la tierra y los Derechos, Restricciones y Responsabilidades asociados. Por su parte, el paquete de Unidad Espacial constituye la unidad fundamental que describe las características espaciales de la tierra a través de geometrías tipo punto, línea o polígono.

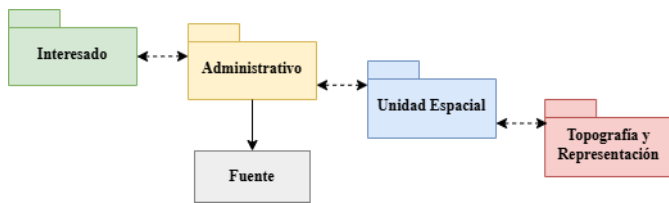


Fig. 1. Paquetes y subpaquetes del LADM.

El LADM ha sido adoptado por muchos países alrededor del mundo como modelo conceptual para la implementación efectiva de los Sistemas de Administración de la Tierra (LAS) o para la aplicación de un área específica [5], frente a la necesidad de tener una visión unificada de los objetos relevantes y las responsabilidades sobre la autoridad de estos objetos [6]. Por esa razón, Silva et al. [7] desarrolló un modelo de redes de servicios públicos de agua subterránea para el Estado de Pernambuco en Brasil bajo las clases RD_RedeAgua, RD_Ramal y RD_Hidrometra. Xu et al. [8] implementó el LADM en el sistema de tenencia de las granjas rurales de China con las subclases CN_Ownership, CN_QualificationRight y CN_UseRight, de la clase CN_Right. Radulovic et al. [9] desarrolló un perfil de país para el cálculo de impuestos de Montenegro haciendo uso de las clases MNE_VM_ValuationUnit, MNE_VM_Building, bMNE_VM_PartOfBuilding y MNE_VM_Parcel. Zamzuri et al. [10] propuso una representación tridimensional de las unidades espaciales marinas de Malasia bajo las clases MY_MarinePropertyUnit, MY_AdjacentFreeSpace y MY_Zone. Ahsan et al. [11] desarrolló un perfil de Pakistán en el contexto de los LAS con el nombre PK_LADM. Así que, el prefijo COL identifica el perfil de país para Colombia.

El modelo núcleo LADM-COL es la adaptación de la norma ISO 19152:2012, que establece la base ontológica y semántica de la administración de tierras en Colombia y un lenguaje

común bajo los principios de i) independencia legal: cada entidad es responsable de la gestión de sus datos; ii) neutralidad tecnológica: se elaboran los modelos en el Lenguaje de Modelado Unificado (UML), independientemente de la plataforma de servidores y base de datos en que se implemente; iii) normalización semántica, con base en el estándar ISO 19152:2012; y iv) un enfoque orientado a modelos en tres niveles: modelo núcleo, modelo extendido y modelo de aplicación [12].

III. SISTEMA DE MAPEO MÓVIL

Es un sistema integrado por software y hardware que permite recopilar información sobre las carreteras, intersecciones, señales de tráfico y otros elementos relacionados con la infraestructura vial. Se compone de tres módulos montados en un vehículo, como se ilustra en la Fig. 2. El módulo A determina la posición precisa mediante el Receptor del Sistema Global de Navegación por Satélite (GNSS) y la Unidad de Medición Inercial (IMU). El módulo B captura fotografías o videos que muestran las condiciones de los activos a partir de una cámara y el módulo C controla el funcionamiento de los sensores y almacena los datos a través de un computador equipado con software [13].

El sistema provee ventajas en términos de que captura datos de forma continua y en tiempo real, entrega una alta precisión de los datos, permite que el mapeo pueda realizarse de manera más frecuente, requiere un trabajo de campo mínimo, los datos se procesan en la comodidad de la oficina y representa un obstáculo mínimo para los usuarios de la vía.

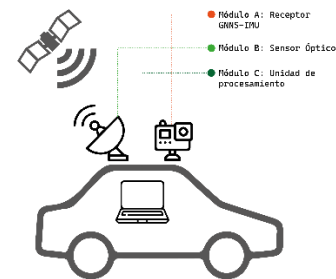


Fig. 2. Sistema de mapeo móvil.

IV. METODOLOGÍA

La metodología está dividida en tres fases, como se ilustra en la Fig. 3. En la primera fase se realiza el diagnóstico de dos temáticas principales, modelos de datos utilizados por las instituciones que producen información relacionada con el transporte terrestre y las contribuciones de LADM y LADM-COL sobre inventario, estado y mantenimiento de la malla vial. El proceso inicia con la búsqueda de referencias bibliográficas recientes en el contexto vial, la comparación de los elementos viales inventariados por las agencias viales y la depuración de la información geográfica que hará parte del modelo extendido propuesto.

En la segunda fase se realiza la conceptualización y diseño del modelo extendido LADM-COL para el inventario vial haciendo uso del lenguaje de modelado conceptual INTERLIS. El diagrama

de clases se diseña en la herramienta UML/Interlis Editor, donde se definen las clases, los atributos sin caracteres especiales, las cardinalidades de los atributos, los dominios con el rango de elementos específicos para la definición del atributo y las relaciones entre cada una de las clases. El diagrama se mejora visualmente aplicando color a las clases Interesado (verde), Administrativo (amarillo), Unidad Espacial (azul) y Fuente (gris). Finalmente, se valida la sintaxis y si no arroja errores, se exporta a un archivo de extensión ILI que contiene toda la información definida para el modelo extendido.

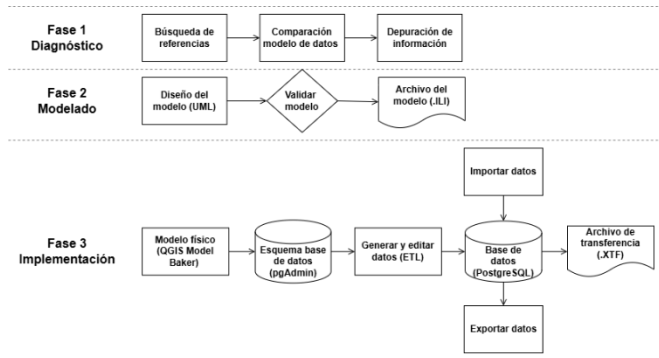


Fig. 3. Fases de la metodología.

La tercera fase abarca los procesos de implementación del modelo extendido e integración de diferentes fuentes de datos haciendo uso del sistema de información geográfica QGIS. El primer paso consiste en crear la base de datos en pgAdmin y generar el modelo físico en el motor de bases de datos PostgreSQL. Este proceso se realiza con el plugin QGIS Model Baker que traduce las clases, relaciones y tipos de datos definidos en el modelo conceptual en un esquema de base de datos. El siguiente paso se ocupa de la generación y edición de los datos geográficos y alfanuméricos mediante dos opciones. La primera consiste en la creación de geometrías y la adición de registros de forma individual para cada una de las clases y la segunda, utiliza la herramienta ETL para extraer la información, transformar los campos a los atributos definidos en el modelo y cargar los datos de forma masiva, a partir de un archivo que puede ser de tipo csv, shapefile, entre otros.

El último paso consiste en el intercambio de datos entre sistemas, dado que es posible exportar o importar información. Una vez la información ha sido cargada en la base de datos, se exporta a un archivo de transferencia de INTERLIS con extensión XTF, que almacena las clases, restricciones, relaciones, atributos, dominios y referencias al modelo núcleo LADM-COL en un formato de tipo XML, que se puede visualizar como texto plano y es independiente de cualquier software. De la misma manera, la base de datos integra información de sistemas externos importando un archivo XTF. En ambos procesos, se valida que la estructura de los datos, la cardinalidad, los valores de dominios y demás restricciones sean conformes a la definición del modelo creado en el lenguaje de INTERLIS.

V. MODELO EXTENDIDO INVENTARIO VIAL

La propuesta de modelo extendido recibe el nombre “LADM_COL_v_1_0_0_Ext_IV” de acuerdo con la

nomenclatura definida por la Infraestructura Colombiana de Datos Espaciales (ICDE) de Colombia. LADM_COL indica que se trata de un modelo extendido del perfil colombiano de LADM, v indica la versión del modelo, 1_0_0 es el número de la versión del modelo, Ext indica que se trata de un modelo extendido, IV es la abreviación de inventario vial.

A. Unidad Administrativa Básica

El modelo núcleo define la clase COL_UnidadAdministrativa Basica con los atributos de nombre y tipo, en este caso, el modelo extendido aplica en la categoría de Transporte y define cinco unidades administrativas asociadas a los elementos del inventario vial. Estas unidades son Vía, Espacio Público, Bahía, Puente y Ciclorruta, como se ilustra en la Fig. 4. La unidad Espacio Público está conformada por los elementos Alameda, Andén, Paso Peatonal, Plaza y Pompeyano que se capturan en el atributo tipo de espacio público.

Todas las unidades están asociadas a Vía mediante el atributo Código de Identificación Vial (CIV), sin embargo, cada unidad tiene una identificación individual a partir del atributo PK_ID. Los demás atributos recogen información relacionada con el tipo de suelo urbano, rural o de expansión, el tipo de material con el cual fue construido, la funcionalidad si es vehicular o peatonal y el estado de la superficie pavimentada, afirmado o en construcción.

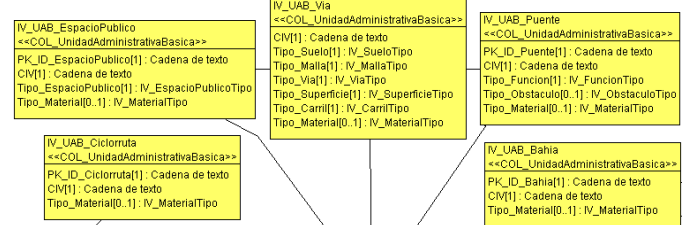


Fig. 4. Clases unidad administrativa básica.

B. Derechos, Restricciones y Responsabilidades (DRR)

Las relaciones entre la parte interesada y la unidad administrativa básica se gestionan en la clase COL_DRR a través de la identificación de los derechos, restricciones y responsabilidades como clases individuales denominadas IV_Derecho, IV_Restriccion y IV_Responsabilidad, que se ilustran en la Fig. 5. El derecho está asociado con el inventario, diagnóstico, diseño, intervención y seguimiento, la restricción es de tipo vehicular, sentido, pico y placa, cierre y velocidad, mientras que la responsabilidad consiste en el mantenimiento periódico o rutinario y la rehabilitación parcial o total.

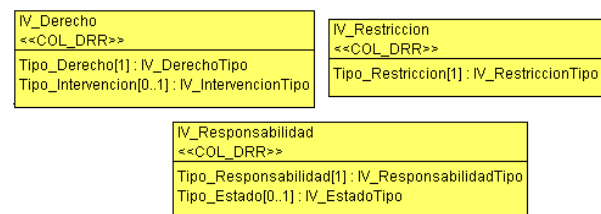


Fig. 5. Clases DRR.

C. Fuente Administrativa y Espacial

El propósito de la clase COL_Fuente es almacenar los documentos que respaldan la información de cada uno de los paquetes. La clase IV_FuenteAdministrativa que se ilustra en la Fig. 6, integra documentos como actas, contratos,

especificaciones, manuales que vinculan el interesado con los DRR, mientras que la clase IV_FuenteEspacial está asociada a los soportes en la captura de las geometrías como croquis de campo, informes técnicos, ortofotos, fotografías.

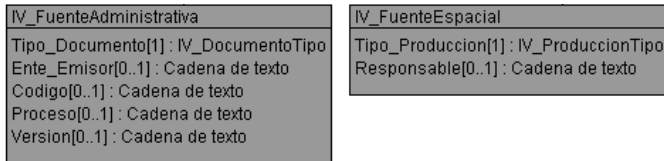


Fig. 6. Clases fuente.

D. Interesado

La clase IV_Agencia_Vial corresponde a las entidades adscritas al Ministerio de Transporte que administran la red vial nacional y están encargadas de reportar el inventario vial, ejercer obras viales y de espacio público, así como la rehabilitación y el mantenimiento periódico de la malla vial. Los atributos de la clase se aprecian en la Fig. 7 y los que hereda de la clase COL_Interesado son nombre, tipo de interesado, tipo de documento y número de documento.

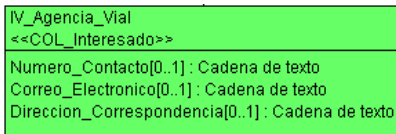


Fig. 7. Clase interesado.

E. Unidad Espacial

La clase COL_UnidadEspacial almacena la representación geométrica de la unidad administrativa básica. Las unidades espaciales del modelo de datos extendido son cinco en geometría tipo polígono y corresponden a Vía, Bahía, Puente, Ciclorruta y Espacio público, como se ilustra en la Fig. 8. La unidad Vía soporta la geometría tipo punto que se almacena en la clase IV_UE_Sitio_Critico y tipo línea en la clase IV_UE_Segmento.

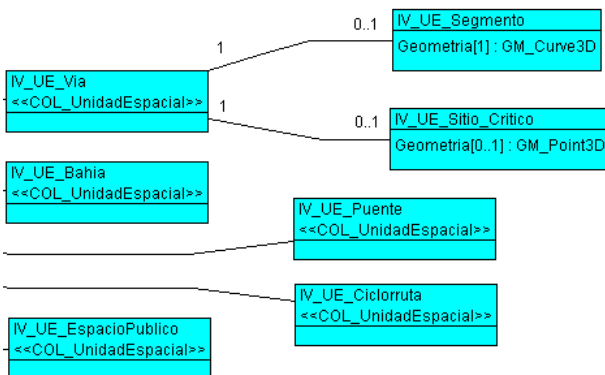


Fig. 8. Clases unidad espacial.

VI. CAPTURA DE INFORMACIÓN

La información que ingresa en el modelo proviene de dos fuentes de información. La primera es el Instituto de Desarrollo Urbano (IDU) que dispone los datos de inventario y estado de la infraestructura de los sistemas de movilidad y de espacio público construidos en la ciudad de Bogotá, Colombia. La

capital del país está dividida en 20 localidades para gestionar de forma adecuada las redes de servicios públicos como la infraestructura vial. En este trabajo, se presentan los elementos viales de la localidad Teusaquillo que arroja la siguiente cantidad de registros en la base de datos espacial: En geometría tipo línea hay 2.542 Segmentos y en geometría tipo polígono hay 3.441 Vías, 474 Bahías, 62 Puentes y 491 Ciclorrutas. En cuanto al espacio público, se tienen 4.702 Andenes, 6 Plazas, 46 Pompeyanos, 90 Pasos Peatonales, 961 Separadores y no registra Alamedas. En la Fig. 9 se resalta el elemento Vía identificado con el CIV 13001176.

La segunda fuente de información corresponde al procesamiento de los videos capturados con el sistema de mapeo móvil que selecciona el fotograma del video, encierra la falla vial en un cuadro, le asigna una categoría específica y establece un porcentaje de acuerdo con la probabilidad alcanzada en la detección. En la localidad de Teusaquillo se identificaron 6.266 fallas, representadas con geometría tipo punto de color amarillo en la Fig. 9. Las categorías más representativas son piel de cocodrilo con 2.180, seguido de otras fallas con 1.466 y desenfoque línea blanca con 917 hallazgos. Las fallas viales son incluidas en la clase IV_UE_Sitio_Critico y se asocian por su ubicación espacial con el CIV de la clase IV_UE_Via.



Fig. 9. Elemento de la clase IV_UE_Via.

VII. CONCLUSIONES

El modelo extendido propuesto permite la integración, interoperabilidad y articulación de los sistemas de información viales y constituye la base para la construcción de modelos de aplicación en temáticas asociadas al transporte por parte de las entidades adscritas al Ministerio de Transporte. Así como el modelo núcleo LADM-COL actualmente se encuentra en la versión número cuatro, el modelo extendido presentado podrá evolucionar conforma a las necesidades de información de otros elementos viales y los ajustes en la normatividad que se produzcan en adelante. Sin embargo, constituye el primer paso para que las entidades del orden municipal, departamental y nacional consideren la importancia de contar con un inventario de la infraestructura vial existente que soporte la toma de decisiones y permita una intervención y planificación eficaz en el territorio.

Los territorios más lejanos de la capital del país son los que más adolecen las necesidades de vías, sin este tipo de infraestructura, el acceso a los centros de salud, de educación de acopio y de producción es limitado. Así que las entidades que hacen parte del sistema de administración de tierras están obligadas a adoptar el estándar LADM-COL en sus procesos y procedimientos de gestión. La implementación del modelo extendido para la captura del inventario vial no implica cambiar la forma en que las entidades capturan la información en campo o en oficina, porque el modelo es independiente del origen de los datos y del software donde se edita la información, simplemente otorga una estructura común para almacenar la información geográfica y alfanumérica facilitando su consulta y análisis mediante un sistema de información interoperable entre las diferentes entidades.

REFERENCIAS

- [1] A. Burns, A. Rajabifard, y D. Shojaei, “Undertaking land administration reform: Is there a better way,” *Land Use Policy*, vol. 132, Sep. 2023, doi: 10.1016/j.landusepol.2023.106824
- [2] A. Chaudhary y A. Akhtar, “A novel approach for environmental impact assessment of road construction projects in India,” *Environmental Impact Assessment Review*, vol. 106, May. 2024, doi: 10.1016/j.eiar.2024.107477.
- [3] L. Narvaez, “Vías Terciarias: Motor del desarrollo económico rural,” *Revista de Ingeniería*, vol. 45, pp. 80-87, 2017, doi: 10.16924/revinge.45.11
- [4] C. Li et al., “Modeling the Urban Low-Altitude Traffic Space Based on the Land Administration Domain Model—Case Studies in Shenzhen, China,” *Land*, vol. 13, no. 12, Dic. 2024, doi: 10.3390/land13122062.
- [5] M. Mete, “Developing GeoAI Integrated Mass Valuation Model Based on LADM Valuation Information Great Britain Country Profile”. *Transactions in GIS*. Nov. 2024, doi: 10.1111/tgis.13273
- [6] P. Oukes, C. Okembo, J. Morales, C. Lemmen, y J. Zevenbergen, “Implementing data exchange and interoperability on LADM country profiles using the ISO framework for enterprise interoperability standard,” *Land Use Policy*, vol. 146, Ago 2024, doi: 10.1016/j.landusepol.2024.107333
- [7] W. Silva y A. Carneiro, “Subsurface utility network cadastre proposal, based on LADM (ISO / FDIS 19152),” *Bulletin of Geodetic Sciences*, vol. 26, no. 2, Mar 2020, doi: 10.1590/s1982-21702020000200006
- [8] Z. Xu, Y. Zhuo, G. Li, R. Bennett, R. Liao, y W. Cifang, “An An LADM-based model to facilitate land tenure reform of rural homesteads in China,” *Land Use Policy*, vol. 120, Sep 2022, doi: 10.1016/j.landusepol.2022.106271
- [9] A. Radulovic, D. Sladic, M. Govedarica, y D. Raicevic, “Using LADM and BIM for property and tourist taxation in the municipality of Bar, Montenegro,” *Land Use Policy*, vol. 131, Ago 2023, doi: 10.1016/j.landusepol.2023.106715
- [10] A. Zamzuri, M. Hassan, y A. Rahman, “Incorporating Adjacent Free Space (AFS) for marine spatial unit in 3D marine cadastre data model based on LADM,” *Land Use Policy*, vol. 139, Abr 2024, doi: 10.1016/j.landusepol.2024.107084
- [11] M. Ahsan et al., “Applying the Land Administration Domain Model (LADM) for Integrated, Standardized, and Sustainable Development of Cadastre Country Profile for Pakistan,” *Land*, vol. 13, no. 6, Jun 2024, doi: 10.3390/land13060883
- [12] Infraestructura Colombiana de Datos Espaciales, “Metodología para la gestión del modelo núcleo y modelos extendidos LADM (Land Administration Domain Model) en Colombia,” Versión 1.1, Dic. 14, 2023. [Online].
- [13] D. Espinel, W. Fernández, J. Moreno, D. Carranza & C. Marrugo, “A Smart Mobile Mapping Application for the Evaluation of Road Infrastructure in Urban and Rural Corridors,” *Applied Computer Sciences in Engineering*, pp. 175-185, Oct. 2024, doi: 10.1007/978-3-031-74595-9_16.

Gestión de la demanda para el aplanamiento de curva de consumo eléctrico en poblaciones densas

Francisco D. Moya Ch. PhD.

Resumen—Los sistemas eléctricos han venido enfrentando frecuentes escenarios de esfuerzo excesivo a causa del continuo crecimiento de la demanda de energía eléctrica; de ahí que conllevan a problemas de confiabilidad, baja eficiencia, pérdidas de energía, aumento de costos eléctricos asociados, entre otros. El uso de combustibles fósiles para la generación de electricidad, plantea preocupaciones ambientales a causa de las altas emisiones de gases de efecto invernadero. Ante esta situación, se plantea una alta participación en la gestión de la demanda, estableciendo cambios tecnológicos, financieros y sociales que posibiliten el suministro de energía eléctrica con mayor calidad, sostenibilidad, seguridad y confiabilidad. Esta propuesta de investigación, propone desarrollar una estrategia de solución para la gestión de la demanda de electricidad en poblaciones densas, basado la operación de cargas eléctricas controlables, manipulación de iluminación artificial, desconexión y deslastre de cargas eléctricas, junto con otros aspectos técnicos relacionados en la gestión de la demanda, enfocado en el decrecimiento de la curva de demanda en horas de alto consumo.

Palabras clave—Gestión de la Demanda.

I. INTRODUCCIÓN

EN las últimas décadas, los sistemas eléctricos han venido enfrentando frecuentes escenarios de esfuerzo excesivo a causa del continuo crecimiento de la demanda de energía eléctrica, interrupciones de las líneas de transmisión durante horas pico, envejecimiento de la infraestructura de la red, incorporación de recursos de generación distribuida, entre otros; de ahí que conllevan a problemas de confiabilidad, baja eficiencia, pérdidas de energía y aumento de costos eléctricos asociados, según lo descrito en los trabajos [1] y [2]. El uso de combustibles fósiles en la generación de electricidad, plantea preocupaciones ambientales a causa de las altas emisiones de gases de efecto invernadero, de acuerdo a lo ilustrado en los trabajos [3] y [4]. Ante esta situación, las redes eléctricas han iniciado un proceso de modernización y transformación hacia las redes inteligentes, según lo detallado en los trabajos [5] y [6], estableciendo cambios tecnológicos, financieros y sociales que posibiliten el suministro de energía eléctrica con mayor calidad, sostenibilidad, seguridad y confiabilidad. Estas redes permiten un flujo bidireccional de electricidad y datos, desde el distribuidor de energía hasta el usuario final, ya que combinan tecnologías de la información y comunicación con metodologías de automatización para el control, gestión de la

demanda, monitoreo y mantenimiento de sistemas eléctricos, tal como se ilustra en los trabajos en [6] y [7].

Este trabajo de investigación, plantea un desarrollar una estrategia de solución para la gestión de la demanda de electricidad en poblaciones densas, basado la operación de cargas eléctricas controlables, manipulación de iluminación artificial, desconexión y deslastre de cargas, y otros aspectos técnicos en la gestión de la demanda (inteligente y coordinada), de usuarios residenciales, enfocado en el decrecimiento de la curva de demanda en horas de alto consumo.

II. MARCO TEÓRICO

Afrontar los cambios mencionados en la implementación de redes inteligentes, implica el empleo de Sistemas de Gestión de Energía para Hogares y Edificios Inteligentes (HEMS, SBMS), de acuerdo a los trabajos [8] y [9]. Las redes inteligentes incorporan conceptos como: la Gestión de la Demanda (DSM: Demand-Side Management), que se encarga de planificar e implementar estrategias orientadas a modificar el horario de consumo de energía eléctrica, para reducir los picos de demanda y las emisiones de CO₂, según lo expuesto en los textos [10] y [11]; esto se logra mediante la conexión y desconexión automática de cargas eléctricas en el hogar por medio de tecnologías en domótica y modelos de optimización como lo describe el trabajo en [12] y en [13], e ilustrado también en la en la Figura 1, de manera que permite minimizar los costos de la factura de electricidad y equilibrar la curva de demanda, según los texto en [14] y [15].

El segundo concepto es la Respuesta a la Demanda (DR: Demand response), que se refiere al uso de precios dinámicos para incentivar a los usuarios finales a disminuir el consumo de electricidad durante ciertos periodos tarifarios [3]; estos precios varían según la demanda, siendo altos cuando la demanda es mayor (horas pico) y bajos cuando la demanda es menor (horas valle).

En las décadas de los 80's y 90's, se desarrollaron métodos de DSM como el recorte de picos, conservación estratégica, relleno de valles, crecimiento estratégico de la carga, desplazamiento de carga y flexibilización de la carga; sin embargo, fue a partir del año 2000 cuando se implementó la generación distribuida con tecnologías enfocadas al control, monitoreo y medición a través de equipos electrónicos

inteligentes, como es descrito en el texto [16] y [17]. Así mismo, en la DSM se emplean dos enfoques: el control indirecto y el control directo. En el control indirecto, se utilizan incentivos en los precios y la interacción social para fomentar cambios en el comportamiento de consumo mediante algoritmos de optimización. Los esquemas tarifarios utilizados incluyen el tiempo de uso, el precio crítico de pico, el precio por carga máxima y el precio dinámico o de tiempo real, tal como se describe en el texto [18] y [19]. En el control directo, la comercializadora actúa directamente sobre las cargas, lo que implica duras restricciones al usuario en la administración de sus cargas eléctricas, así como es ilustrado en el trabajo de [20] y [21].

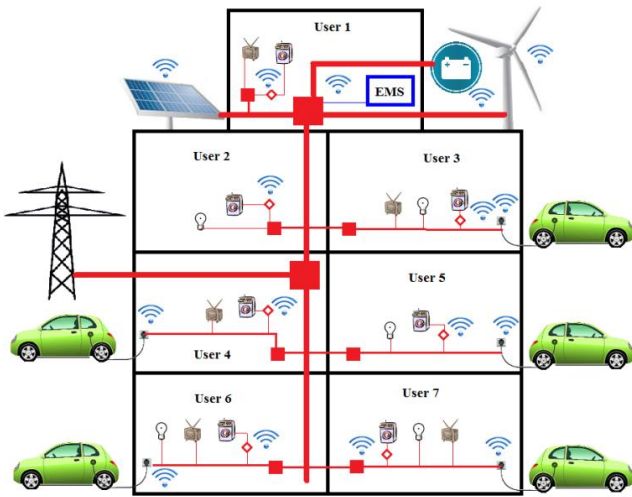


Fig. 1. Sistema de control en un edificio inteligente [14].

Hay que tener presente, que al usuario final le resulta difícil programar el encendido y apagado de los electrodomésticos manualmente debido a la falta de conocimientos técnicos, de tiempo, o no son lo suficientemente proactivos para participar en la DR en sus hogares, así como se describe en [22] y [23]. Por consiguiente, la Programación de Cargas Eléctricas Controlables (PCEC) es fundamental en un Sistema de Gestión de la Energía en los Hogares (HEMS: Home Energy Management System), siendo responsable del monitoreo y control de las cargas eléctricas según indicaciones y especificaciones por parte del usuario, de acuerdo a los textos de [24] y [25], también ilustrado en la Figura 2; además para el desarrollo del HEMS influyen variables externas como el precio, los problemas medio ambientales, el bienestar personal, la educación en eficiencia energética y la responsabilidad del usuario para ser un agente activo en la DSM, que afectan los hábitos y el comportamiento por parte del consumidor residencial a la hora de hacer uso de sus electrodomésticos, como se ilustra en [26].

Es importante mencionar, que al programar los electrodomésticos en los horarios que más les convenga a los usuarios, no garantiza el aplanamiento de la curva de carga en las redes de distribución; por ello se presenta el desafío de realizar una estrategia de mayor alcance, sobre los diferentes

usuarios, que permita la operación de sus cargas en diferentes momentos del día y, por tanto, obtener un resultado considerable sobre consumo eléctrico en horas punta. Con base en lo mencionado, esta investigación pretende contribuir al desarrollo de una estrategia la gestión de la demanda en poblaciones densas.

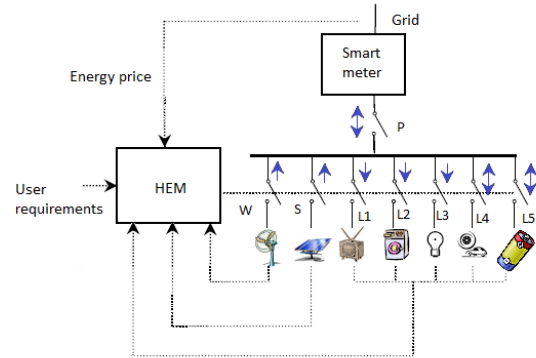


Fig. 2. Sistema de control en un hogar inteligente [13].

III. METODOLOGÍA PROPUESTA

Es necesario tener en cuenta la amplia gama de variables implicados en el problema de la gestión de la demanda para poblaciones densas; de ahí que se vincule la ingeniería de software, la informática, el modelamiento matemático y el conocimiento en ingeniería eléctrica, para proporcionar una nueva solución en la gestión de la demanda en grandes poblaciones, aprovechando el avance de las diferentes herramientas y metodologías de optimización para solución de problemas en ingeniería. El desarrollo de una metodología, debe prever el desarrollo de una herramienta computacional, incluyendo la aplicación de métodos cuantitativos y cualitativos, que cubran las condiciones relevantes del funcionamiento de redes eléctricas y cargas eléctricas.

Las redes eléctricas inteligentes son un concepto todavía en desarrollo, pero que empieza a ser una realidad en el escenario mundial, ya que combina tecnologías de la información y comunicación, junto con metodologías de automatización para el control, monitoreo y mantenimiento de redes eléctricas, como se ilustra en [27], [28] y [29]. Sin embargo, llevar este concepto a la práctica conlleva grandes cambios físicos, sociales, políticos, culturales y económicos, por lo que una forma efectiva de abordar lo anteriormente mencionado es el empleo de sistemas de gestión de energía en poblaciones densas, que permita el aplanamiento de la curva de demanda eléctrica ilustrada en la Figura 3 para el caso colombiano, la cual supera actualmente los 11000 MW en potencia demandada máxima.

Respecto a la gestión de la demanda en los hogares (HEMS), a los consumidores les resulta difícil programar manualmente los electrodomésticos, ya sea por falta de tiempo o no son lo suficientemente proactivos para realizar la RD en sus hogares; así pues, es de gran importancia estudiar la Programación de Cargas Eléctricas Controlables (PCEC), con el propósito de aprovechar no solo el bajo precio de la energía eléctrica, establecidas en diferentes periodos del día, sino también reducir las curvas de alto consumo de electricidad en horas pico, incentivando a los hogares a utilizar los electrodomésticos en horas de menor consumo. La Figura 4,

ilustra la proyección de la demanda de electricidad en Colombia para los próximos años, verificándose el aumento considerable año tras año, debido a aspectos importantes como: aumento de la población, aumento del consumo de electricidad por cada hora, uso de la electricidad en la industria, uso del transporte eléctrico como lo son los vehículos eléctricos, trenes eléctricos, entre otros.

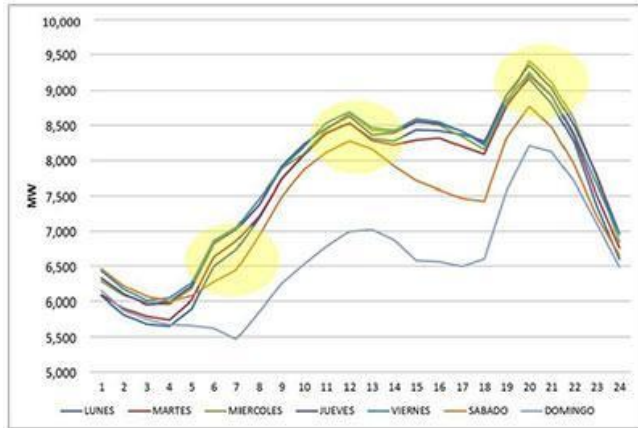


Fig. 3. Potencia máxima demandada en Colombia [30].

No obstante, en el caso colombiano existe actualmente un sistema de tarifa plana (monomía), del costo de la electricidad al usuario final, que no lo motiva a programar y administrar su consumo de energía, limitando la posibilidad de obtener los beneficios que ofrece una adecuada gestión de la demanda; por consiguiente, el país debería implementar a mediano plazo, un cambio en el sistema tarifario que facilite una mayor participación de la gestión de energía en el hogar.

Se plantea entonces desarrollar una estrategia de solución para la gestión de la demanda de electricidad en poblaciones densas, basado la operación de cargas eléctricas controlables, manipulación de iluminación artificial, desconexión y deslastre de cargas, y otros aspectos técnicos en la gestión de la demanda (inteligente y coordinada), de usuarios residenciales, enfocado en el decrecimiento de la curva de demanda en horas de alto consumo.

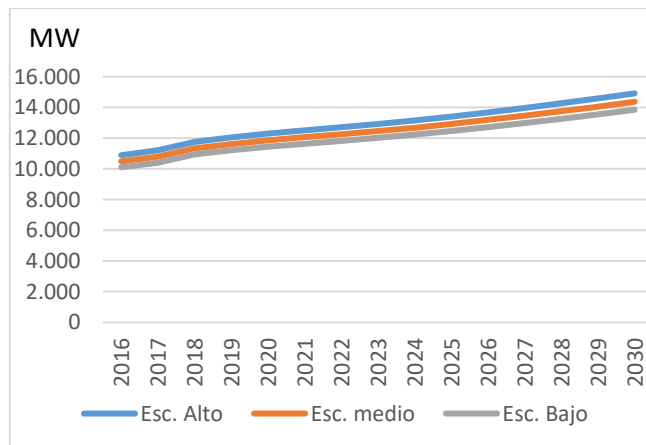


Fig. 4. Proyección de la demanda en Colombia [30].

Así pues, recae sobre el operador de red la mayor responsabilidad sobre la gestión de la demanda eléctrica en

grandes poblaciones, dado que dicho operador, no solamente es el principal propietario de las redes eléctricas, también es el encargado de dar suministro eléctrico a los diferentes usuarios, como también opera y realiza el mantenimiento de las redes públicas de iluminación, sobre las cuales se puede realizar una gestión importante de reducción de consumo, trasformando dichos sistemas para que sean inteligentes, desconectando luminarias, o atenuando su luminosidad en caso de que no sean requeridas, tal como se propone en el proyecto de investigación [tal], ilustrado en la Figura 5.

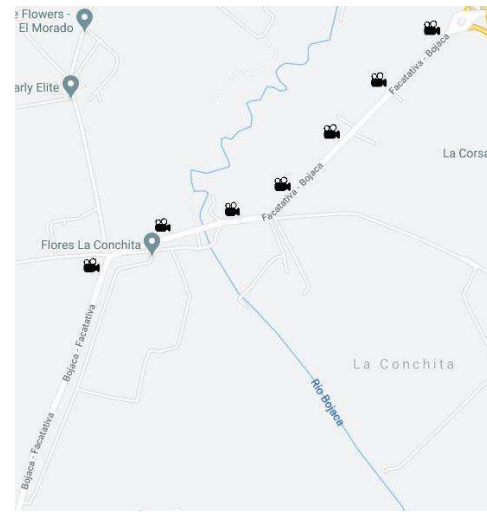


Fig. 5. Propuesta de control inteligente de iluminación pública [27].

Dada la importancia del Operador de Red eléctrica en la gestión de la demanda en grandes poblaciones, se ilustra a dicho operador como el eje central en la estrategia para el control del crecimiento de la demanda eléctrica, tal como se ilustra en la Figura 6. El Operador de Red, puede operar y manipular directamente la iluminación pública, por lo tanto, está en sus manos el reto de convertir dichos sistemas en infraestructuras inteligentes, capaces de mitigar su consumo eléctrico. Cabe al Operador de Red, monitorear los consumos de cada usuario, tanto residenciales, como comerciales e industriales, para realizar un seguimiento de aumentos excesivos de consumos de electricidad.

Los Operadores de Red, tendrán la potestad, de manipular las cargas eléctricas controlables de cada usuario, con el fin de controlar consumos excesivos en horas pico. El Operador de Red, de ser necesario, podrá desconectar cargas eléctricas controlables en usuarios residenciales, como son las neveras, lavadoras, secadoras, y demás electrodomésticos, que no causarán un impacto negativo a los usuarios durante las horas pico. El Operador de Red, también podrá operar y manipular las cargas eléctricas de sus usuarios comerciales solo en caso de ser necesario, como lo es la desconexión de escaleras eléctricas, atenuación de iluminación artificial, desconexión de refrigeradores, aires acondicionados, calefacción eléctrica, por periodos cortos, que permitan superar el tiempo de las horas de alto consumo eléctrico.

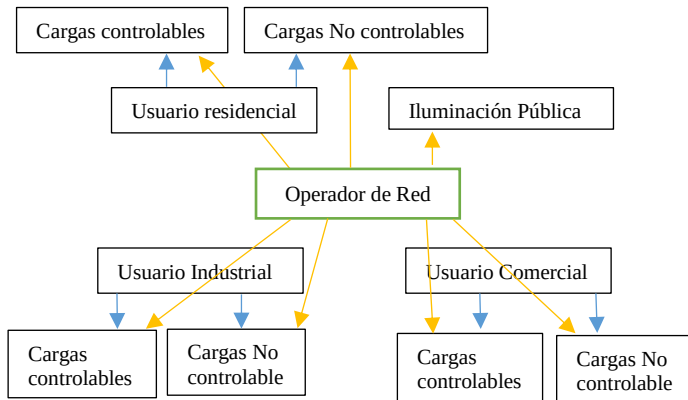


Fig. 6. Gestión de la demanda desde el Operador de Red (Elaboración del autor).

El Operador de Red, podrá operar y manipular cargas controlables en usuarios industriales solo en caso de ser necesario, esto es, con el fin de reducir los altos consumos de electricidad durante las horas pico. Sobre los usuarios industriales, podrán desconectarse sistemas de refrigeración, sistemas de calefacción, atenuación de iluminación, entre otros, por un periodo corto, buscando el menor impacto negativo sobre dichos usuarios.

V. CONCLUSIONES

Los altos consumos de la electricidad en los años venideros, no solamente se deben resolver desde el lado de la oferta; cabe realizar gestiones avanzadas del lado de la demanda, con el fin de mitigar dichos consumos.

Instalar nuevos sistemas de generación eléctricos, presenta grandes retos e inconvenientes como lo es: altos costos de instalación, impacto ambiental al instalar nuevas infraestructuras, impacto ambiental al utilizar fuentes convencionales de generación eléctricas que utilizan combustibles fósiles, entre otros.

Mitigar el alto consumo de electricidad en las horas pico, reducirá los esfuerzos sobre la red eléctrica, así como la reducción de los costos del servicio eléctrico, dado que, en las horas de mayor consumo, se encienden los generadores eléctricos con el mayor costo de generación eléctrico.

Recae sobre el operador de red la mayor responsabilidad sobre la gestión de la demanda eléctrica en grandes poblaciones, dado que dicho operador, no solamente es el principal propietario de las redes eléctricas, también es el encargado de dar suministro eléctrico a los diferentes usuarios, como también opera y realiza el mantenimiento de las redes públicas de iluminación, entre otras cargas.

REFERENCIAS

- [1] M. Pipattanasomporn, M. Kuzlu, and S. Rahman, "An Algorithm for Intelligent Home Energy Management and Demand Response Analysis," *IEEE Trans Smart Grid*, vol. 3, no. 4, pp. 2166–2173, Jun. 2012. <https://doi.org/10.1109/TSG.2012.2201182>
- [2] A. C. Duman, Ö. Güler, K. Deveci, and Ö. Gönül, "Residential load scheduling optimization for demand-side management under time-of-use

- rate," in 2018 6th International Istanbul Smart Grids and Cities Congress and Fair (ICSG), Istanbul, Turkey, 2018, pp. 193–196. <https://doi.org/10.1109/SGCF.2018.8408971>
- [3] G. Huang, J. Yang, and C. Wei, "Cost-Effective and Comfort-Aware Electricity Scheduling for Home Energy Management System," in 2016 IEEE International Conferences on Big Data and Cloud Computing (BDCloud), Social Computing and Networking (SocialCom), Sustainable Computing and Communications (SustainCom) (BDCloud-SocialCom-SustainCom), Atlanta, GA, USA, 2016, pp. 453–460.
- [4] L. Sheng-Kai, and C. Chao-Rong, "Optimal energy consumption scheduling in home energy management system," in 2016 International Conference on Machine Learning and Cybernetics (ICMLC), Jeju, Korea (South), 2016, pp. 638–643. <https://doi.org/10.1109/ICMLC.2016.7872962>
- [5] F. D. Moya Chavez, L. C. P. Da Silva, and J. C. Lopez A., "A Mathematical Model for the Optimal Scheduling of Smart Home Electrical Loads," *WSEAS Transactions on Power Systems*, Jan. 2018. <https://repositorio.usc.edu.co/items/cf5777b3-948e-4561-98a0-d01b30d47592>
- [6] R. Jovanovic, A. Bousselham, and I. S. Bayram, "Residential Demand Response Scheduling with Consideration of Consumer Preferences," *Applied Sciences*, vol. 6, no. 1, p. 16, Jan. 2016. <https://doi.org/10.3390/app6010016>
- [7] M. Amer, A. Naaman, N. K. M'Sirdi, and A. M. El-Zonkoly, "Smart home energy management systems survey," in International Conference on Renewable Energies for Developing Countries 2014, Beirut, Lebanon, 2014, pp. 167–173. <https://doi.org/10.1109/REDEC.2014.7038551>
- [8] Y. Wang, C. Hao, and T. Yoshimura, "A particle swarm optimization and branch and bound based algorithm for economical smart home scheduling," in 2017 IEEE 60th International Midwest Symposium on Circuits and Systems (MWSCAS), Boston, MA, USA, 2017, pp. 213–216. <https://doi.org/10.1109/MWSCAS.2017.8052898>
- [9] A. R. S. Vidal, L. A. A. Jacobs, and L. S. Batista, "An evolutionary approach for the demand side management optimization in smart grid," in 2014 IEEE Symposium on Computational Intelligence Applications in Smart Grid (CIASG), Orlando, FL, USA, 2014, pp. 1–7. <https://doi.org/10.1109/CIASG.2014.7011561>
- [10] E. Moreno Soto, "Análisis de la flexibilidad de la demanda en usuarios residenciales," M.S. Proyecto de Grado, Universidad Nacional de Colombia, Bogotá, Colombia, 2021. [Online]. Available: <https://repositorio.unal.edu.co/handle/unal/80431>
- [11] S. N. Bragagnolo, J. C. Vaschetti, and F. Magnago, "A Technical and Economic Criteria Comparison on Demand Side Management with Multi-Level Optimization Model," *IEEE Latin America Transactions*, vol. 19, no. 9, pp. 1494–1501, Jun. 2021. <https://doi.org/10.1109/TLA.2021.9468442>
- [12] D. H. Ruiz Amaya, "Optimización de recursos energéticos a partir de la utilización de un DER y de la integración de equipos de automatización," M.S. Proyecto de Grado, Pontificia Universidad Javeriana, Bogotá, Colombia, 2015. [Online]. Available: <https://doi.org/10.11144/Javeriana.10554.16812>
- [13] F. D. Moya Ch., L. C. P. da Silva, J. C. Lopez A., and P. Vergara B., "GRASP model for smart home electrical loads scheduling," in 2015 International Energy and Sustainability Conference (IESC), Farmingdale, NY, USA, 2015, pp. 1–6. <https://doi.org/10.1109/IESC.2015.7384384>
- [14] F. D. Moya Ch., J. C. Lopez A., and L. C. P. da Silva, "Model for smart building electrical loads scheduling," in 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), Florence, Italy, 2016, pp. 1–6. <https://doi.org/10.1109/EEEIC.2016.7555639>
- [15] A. M. Vega Escobar, "Gestión de la energía eléctrica domiciliar con base en la gestión activa de la demanda," Ph.D. tesis, Universidad Distrital Francisco José de Caldas, Bogotá, Colombia, 2018. [Online]. Available: <http://hdl.handle.net/11349/7587>
- [16] F. Y. Melhem, O. Grunder, Z. Hammoudan, and N. Moubayed, "Energy Management in Electrical Smart Grid Environment Using Robust Optimization Algorithm," *IEEE Trans Ind Appl*, vol. 54, no. 3, pp. 2714–2726, Feb. 2018. <https://doi.org/10.1109/TIA.2018.2803728>
- [17] S. Janocha B., S. S. Baum, and I. Stadler, "Cost minimization by optimization of electricity generation and demand side management," in 2016 International Energy and Sustainability Conference (IESC), Cologne, Germany, 2016, pp. 1–7. <https://doi.org/10.1109/IESC.2016.7569489>

- [18] W. Guzman, S. Osorio, and S. Rivera, "Modelado de cargas controlables en el despacho de sistemas con fuentes renovables y vehículos eléctricos," *Ingeniería y Región*, vol. 17, p. 49, Jun. 2017. <https://doi.org/10.25054/22161325.1535>
- [19] I. O. Essiet, Y. Sun, and Z. Wang, "Optimized energy consumption model for smart home using improved differential evolution algorithm," *Energy*, vol. 172, pp. 354–365, Apr. 2019.
- [20] H. T. Dinh, and D. Kim, "An Optimal Energy-Saving Home Energy Management Supporting User Comfort and Electricity Selling with Different Prices," *IEEE Access*, vol. 9, pp. 9235–9249, Jan. 2021. <https://doi.org/10.1109/ACCESS.2021.3050757>
- [21] A. Ouammi, "Optimal Power Scheduling for a Cooperative Network of Smart Residential Buildings," *IEEE Trans Sustain Energy*, vol. 7, no. 3, pp. 1317–1326, Mar. 2016. <https://doi.org/10.1109/TSTE.2016.2525728>
- [22] N. D. Rahate, and N. Kinhekar, "Demand side management for household equipment's," in *2017 International Conference on Information, Communication, Instrumentation and Control (ICICIC)*, Indore, India, 2017, pp. 1-5. <https://doi.org/10.1109/ICOMICON.2017.8279108>
- [23] N. M. Bejarano, F. D. Moya Chaves, and O. D. Montoya, "Optimization Model for Collective Energy Demand Management in Smart Homes," *Tecnológicas*, vol. 27, no. 60, e3014, 2024. <https://doi.org/10.22430/22565337.3014>
- [24] Á. R. Restrepo, S. E. Nope, and D. E. Enríquez, "Beneficios Económicos de la Gestión de la Demanda y la Energía Autogenerada en el Contexto de la Regulación Colombiana," *CIT Información tecnológica*, vol. 29, no. 1, pp. 105–116, Feb. 2018. <https://doi.org/10.4067/S0718-07642018000100105>
- [25] J. S. Gómez Marín, S. X. Carvajal, y A. Arango Manrique, "Programas de gestión de demanda de electricidad para el sector residencial en Colombia: Enfoque Sistémico," *Energética*, Jul. 2015. <https://repositorio.unal.edu.co/handle/unal/64269>
- [26] M. J. Kochenderfer, and T. A. Wheeler, "Linear Constrained Optimization" in *Algorithms for Optimization*. London, England: The MIT Press, 2019, pp. 189-206. <https://mitpress.mit.edu/9780262039420/algorithms-for-optimization/>
- [27] M. Martínez, R. Reyes, y F. Moya, "State of art and Proposal Development of a Street Lighting System Using Artificial Vision" en *6th International Conference on Smart and Sustainable Technologies 2021: SpliTech 2021*, Split and Bol, Croatia, September 8-11, 2021
- [28] A. Grandas, "Prediseño de micro-red eléctrica inteligente para la sede Edificio Administrativo y Sabio Caldas de la Universidad Distrital Francisco José de Caldas", Proyecto de Grado, Octubre 2022, Bogotá Colombia.
- [29] V. J. Gutierrez-Martinez, C. A. Moreno-Bautista, J. M. Lozano-Garcia, A. Pizano-Martinez, E. A. Zamora-Cardenas, and M. A. Gomez-Martinez, "A Heuristic Home Electric Energy Management System Considering Renewable Energy Availability," *Energies*, vol. 12, no. 4, p. 671, Feb. 2019. <https://doi.org/10.3390/en12040671>.
- [30] XM Colombia, datos publicados en página WEB. <http://www.xm.com.co>

Techno-Economic Analysis of Afro descendent Energy Communities to Empower Rural Entrepreneurships.

Y. Ulianov L, *Senior IEEE*, Sofia Rojas-Rueda, Alberto Coronado Mendoza, José María Yusta, Maximilianne Acevedo, Jhon-Esteban Viveros

Abstract— This study presents a technical analysis of the solar energy potential in three rural communities located in the Valle del Cauca region of Colombia. Utilizing HOMER Pro simulation software, the research assesses solar resource availability in relation to annual environmental variations characteristic of each community. The assessment incorporates demographic estimates and assumes a standardized household energy consumption of 1000 W, representing the typical use of a refrigerator, television, lighting, and washing machine basics. The requisite installed capacity in kilowatts (kW) is thus determined for each locality, based on these parameters. The analysis of simulation results is facilitated by the use of comparative graphics, a method that enables the identification of trends and the assessment of the feasibility conditions for the implementation of photovoltaic systems. This work makes a significant contribution to the field of decentralized energy solutions in underserved rural contexts. Its findings have implications for sustainable development and energy access policy.

Index Terms—Energy community, Renewable Energy Sources, Optimization, HOMER.

I. INTRODUCTION

ACCESS to reliable and sustainable energy remains a critical challenge in many rural areas of developing countries. In Colombia, a significant portion of the rural population lives in Non-Interconnected Zones, where access to electricity is limited or nonexistent. However, even in areas connected to the National Interconnected System, many households face energy poverty due to insufficient economic resources to afford conventional utility bills [1]. This highlights the need to explore diverse and decentralized energy solutions that are not only technically feasible but also economically accessible. Renewable energy, particularly solar photovoltaic (PV) systems, offers a viable pathway for addressing these disparities. The decreasing costs of PV technology and the high solar resource availability in Colombia make it a promising solution, especially when implemented through Small-Scale Photovoltaic Systems [2]. Energy communities is a term to identify initiatives led by citizens, allowing them to take control of their energy production and consumption. The communities aim to empower and promote productive community processes, where the benefits of utilizing local and renewable energy

resources are reflected and shared within the energy community [3], [4].

The majority of research scientists have developed models for renewable energy systems, optimizing their size and cost for various configurations to support rural electrification [5], [6], [7], [8]. The system's performance is examined and compared using different combinations of the Hybrid Renewable Energy Systems (HRES), to determine the optimal configuration with the minimum value of Net Present Cost (NPC) and Levelized Cost of Energy (LCOE)[9], [10]. Finally, the optimized system is economically feasible, with an attractive payback period and fewer emissions.

This study focuses on three rural communities in the Valle del Cauca region, analyzing their solar potential using HOMER Pro simulation software. It evaluates solar irradiance based on annual temperature profiles, estimates energy demands using a standardized household load of 1000 W, and determines the necessary installed capacity. Through simulations, this research proposes sustainable PV system configurations that reduce costs and improve energy access, supporting both electrification and socio-economic development in underserved areas.

This paper is organized as follows. In section II, Methodology is detailed, indicating how a multicriteria criteria analysis help and designers and developers a better and accurate renewable energy design. Then, in section III presents in a very detailed form, the simulation process for TIMBA, one of the rural villages studied. There, the energy resource and components selection as HOMER inputs are explained. Finally, the optimization is carried out and results presented.

II. METHODOLOGY. CASE STUDY

Small villages in the south of Valle Del Cauca region in Colombia, work with sugar cane. Majority of the population in this area is afrodenscent and live in part of a valley and hilly terrain, where sugar cane is mainsource of labour. Sugar cane is cultivated in the vicinity of numerous small villages, with the majority of the population engaged in the harvesting of sugar cane or related occupations. A second group of people is engaged in agricultural activities, cultivating a variety of crops,

Yuri Ulianov López (ORCID 0000-0001-5411-2786) is with the Universidad Autónoma de Occidente, Cali Colombia (ylopez@uao.edu.co).

Sofía Rojas Rueda is with the Universidad Autónoma de Occidente, Cali Colombia (sofia.rojas@uao.edu.co).

Maximilianne Acevedo Azuero is with the Universidad Autónoma de Occidente, Cali Colombia (maximilianne.acevedo@uao.edu.co).

Jhon Esteban Viveros Rueda is with the Universidad Autónoma de Occidente, Cali Colombia (jhon_esteban.viveros@uao.edu.co).

including pineapple, guava, lemon, sapote, and cocoa, as well as medicinal herbs for diverse applications. In addition to these crops, some people own domestic animals, such as chickens and pigs, and others have small farms. For energy load per village or small town, in this study 4 people per family are considered [3].

This paper employs a structured theoretical-experimental methodology comprising seven stages designed to evaluate the technical and economic feasibility of designing, simulating, and sizing energy communities. Each stage was developed based on best practices in rural energy communities on-grid [11], [12]. This included a detailed study of Colombian policies such as Law 2294 of 2023, Decree 2236 of 2023, and Resolution CREG 101 072 of 2025, which provide the legal and technical foundation for implementing decentralized energy models [13].

Next, a socio-economic and energy demand characterization of the community was carried out. Surveys and local interviews helped determine consumption patterns, economic activities, and energy access limitations. This stage provided the load profiles required for simulation input. Socio-spatial characterization is critical for the accurate sizing of distributed energy systems in rural contexts [14]. The third stage consisted of a resource assessment, identifying locally available renewable energy sources: solar irradiance, hydraulic potential, and biomass availability from agro-industrial waste. These parameters were evaluated using climatic data-bases and local measurements, which were essential for realistic system modeling [15], [16], [17].

The fourth and fifth stages involved system design and optimization using HOMER Pro, a simulation tool widely used for configuring hybrid renewable systems. Multiple scenarios were simulated to compare the technical performance and economic viability of different topologies (PV-only, PV-diesel, PV-hydro, PV-biomass, and hybrid PV-grid) [18], [19].

HOMER, developed by the U.S. National Renewable Energy Laboratory, is used to design and simulate hybrid renewable energy systems, including those connected to the grid. It performs three main functions: simulation, optimization, and sensitivity analysis. In the simulation phase, HOMER evaluates the technical performance and life-cycle cost of a system on an hourly basis throughout the year. During optimization, it compares various configurations to find the most efficient and cost-effective one. Sensitivity analysis examines how changes in uncertain variables, such as fuel prices or irradiance, impact system performance, making HOMER a valuable tool for planning sustainable energy systems under defined constraints and optimal conditions.[18], [19], [20]

The sixth stage focused on the economic analysis of the optimal configuration, calculating metrics such as Net Present Cost (NPC), Levelized Cost of Energy (LCOE), and user tariff. Sensitivity analysis was conducted to evaluate the robustness of the solution under variable fuel prices, demand growth, and solar availability. Finally, the seventh stage involved proposing a scalable energy community model based on the simulation results. Technical recommendations and a roadmap for implementation were drafted, taking into account legal integration to the national grid and the financial sustainability

of the system [11], [21]

III. SIMULATION AND RESULTS

A. TIMBA.

A small town with 2000 inhabitants (2010) four people for family in average. Assuming 4 people per family. The plot of geological location is shown in Fig. 1. Location: 3.12 N, 76.6W. Several software online and installed tools could help designers to have accurate data. Economic data is from 2025 and energy tariff (kWh/m) is from July 2023. Global solar irradiance is showed in figure 2.

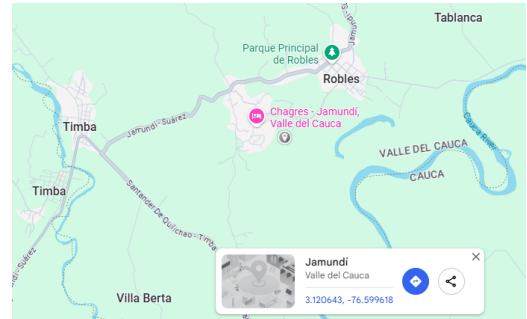


Fig. 1. Research area location (Source: Googlemaps. Mapdata 2025)

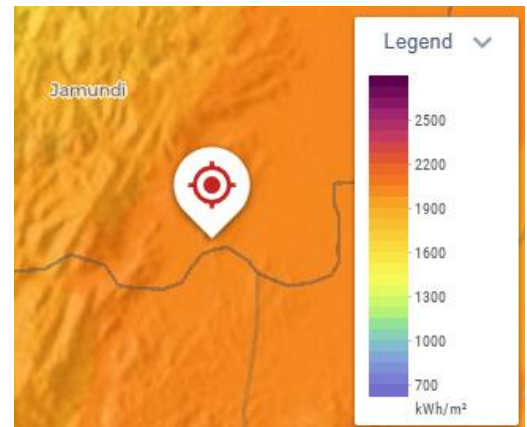


Fig. 2. Global Solar Irradiance (kWh/m^2) Source: Global Solar Atlas. Mapdata 2025)

Solar irradiance in this area is high, making solar PV systems valuable for renewable projects and energy communities developments. In figure 3, a bar figure shows average values. The PVWatts energy estimate is based on an hourly performance simulation using a typical-year weather file that represents a multi-year historical period for JAMUNDI for a Fixed (roof mount) photovoltaic system.

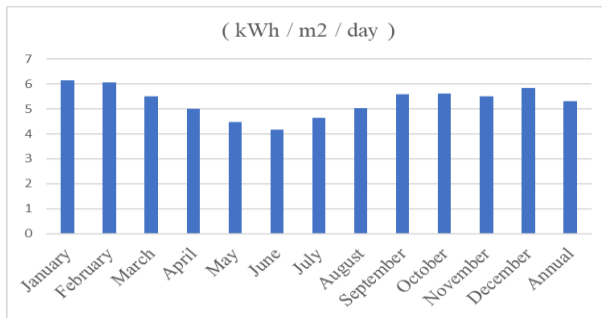


Fig. 3. Monthly Irradiance (kWh/m²/d) Source: Made by author with PVWATTs data.

Once the average solar irradiance data and energy demand or consumption per family and the entire community are available, the simulation process start. The components of the HOMER program database are selected, and the costs of local marketers are obtained to perform a more accurate and localized sizing as show in Figure 4.

A. Design

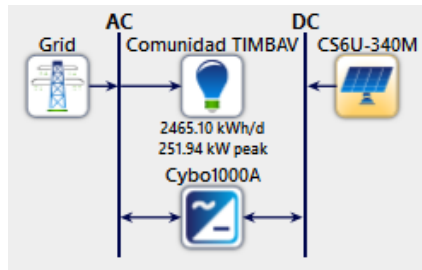


Fig. 4. Schematic.

With the values of the most common and necessary appliances, the average power that the energy community may have been entered on an hourly basis. The habits of the rural population in Valle del Cauca create a load profile, as shown in Figure 5, with the highest consumption between 7 and 9 o'clock, at midday, and 21 and 22 o'clock. With the help of the program, an hourly load profile is obtained as presented in Figure 5.

B. Demand.

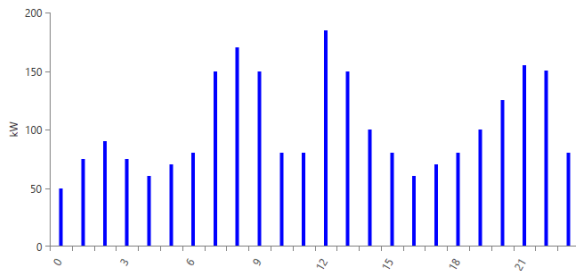


Fig. 5. Hourly demand (kW)

These towns are connected to the distribution power grid; therefore, for the homer program simulation, the power generation strategy is selected in alternating current. Average power and energy for this community is showed in Table 1.

TABLE I
POWER AND ENERGY METRICS

Metric	Baseline	Scaled
Average (kWh/d)	2465,1	2465,1
Average (kW)	102,71	102,71
Peak (kW)	251,94	251.94
Load Factor	0,41	0,41

C. Optimization Results

Once the equipment and components are selected from the Homer database, the local costs of each of these technologies are entered. Thus, the Homer program estimates the cost and feasibility of the community's energy system after iterations as presented in Figure 6.

Architecture				Cost			
CS6U-340M (kW)	Grid (kW)	Cybo1000A (kW)	Dispatch	COE (\$)	NPC (\$)	Operating cost (\$/yr)	Initial capital (\$)
694	100	201	LF	\$487.43	\$6.61B	\$436M	\$724M
694	100	201	CC	\$487.43	\$6.61B	\$436M	\$724M
694	100	197	LF	\$490.36	\$6.61B	\$437M	\$721M
694	100	197	CC	\$490.36	\$6.61B	\$437M	\$721M
681	100	199	LF	\$490.10	\$6.61B	\$437M	\$712M

Fig. 6. Technology comparison results.

D. Electrical Performance Of Optimal Results

The Homer program performs optimization for the lowest net present value and the lowest levelized cost of energy. In this case, the total net present value is \$6608,595,000 COP (\$1'612,092.28 USD). The levelized energy cost is \$487.43 (\$0.12 USD), and the operating cost is \$436,354,700 (\$106,443.8 USD). This energy performance and its costs are the result of the optimal configuration which, according to the software, is as follows:

Canadian Solar PV system	820.938 kWh/y	64.3%
Grid purchases	456.075 kWh/y	35.7%

Thus, the solar PV system is the man energy source as is presented in Figure 7. However, there is an unmet electric load of 61877 kWh/y which represents 6.88 %.

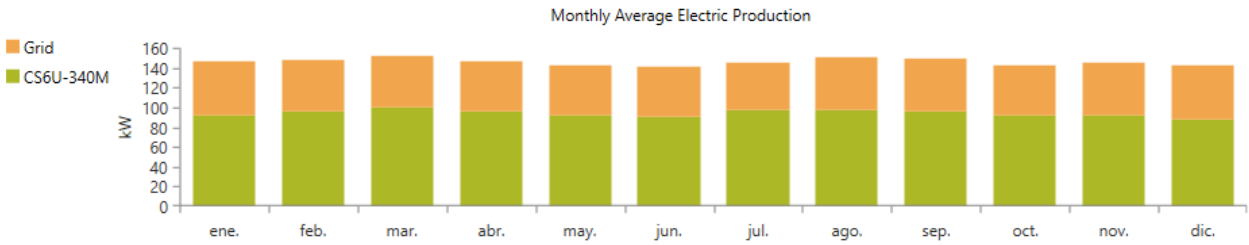


Fig. 7. Hourly demand (kW)

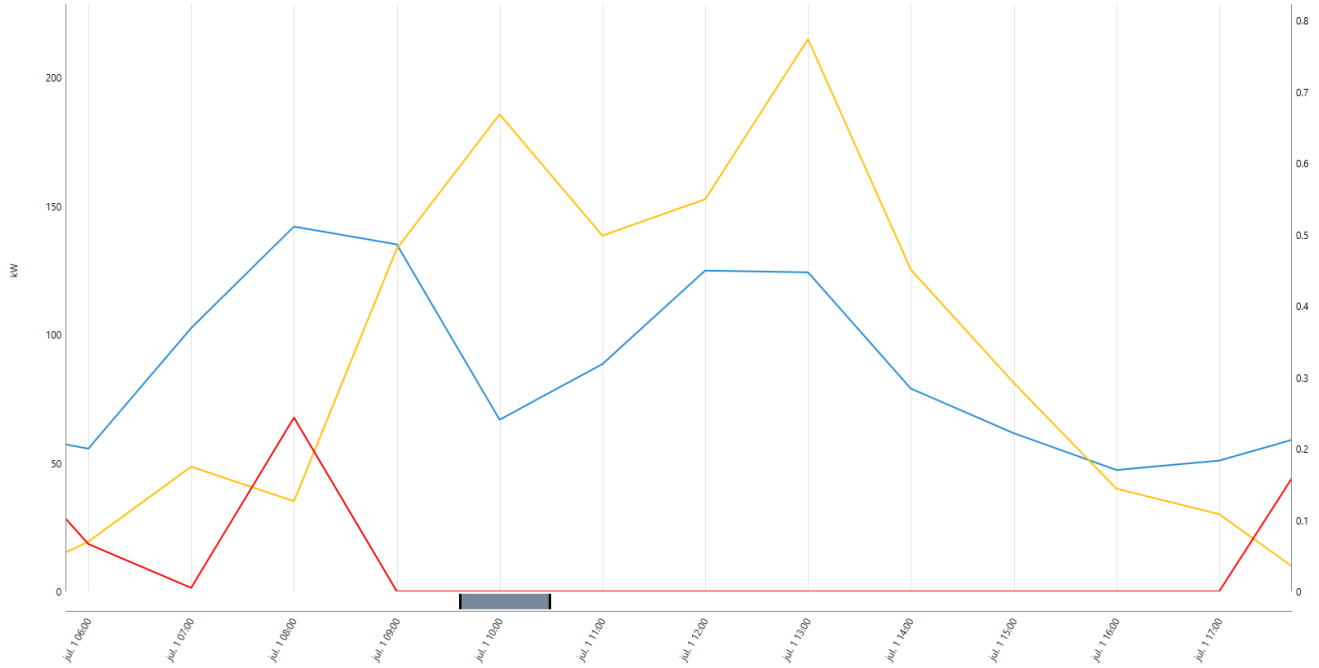


Fig. 8. Hourly demand (kW)

A more detailed analysis of the system's energy performance can be performed using the daily behavior curves shown in Figure 8. In this figure, a typical day is illustrated, where the most significant curve (yellow) represents the solar energy generated during the day. The central curve (blue) is the average energy demand of the community for that day. The slight curve (red) represents the grid and shows that for most of the day, no power purchase from the grid is required.

Similarly, two other villages in the same region have been studied: Chagres with almost 1000 population and Quinamayó with around 1500. Although the results presented are lower than those for TIMBA, the same methodology was used to ensure accurate sizing of these other two energy communities.

E. Quinamayó Design - Demand

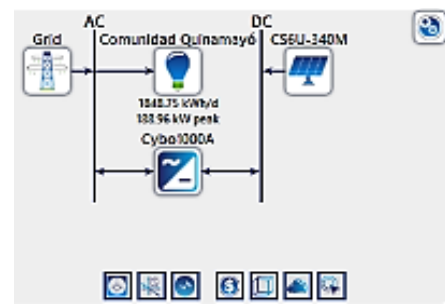


TABLE II
POWER AND ENERGY METRICS

Metric	Baseline	Scaled
Average (kWh/d)	1848,7	1848,7
Average (kW)	77,03	77,03
Peak (kW)	188,95	188,95
Load Factor	0,41	0,41

For the case of the QUINAMAYÓ the report indicates that 98 were infeasible due to restrictions of the simulation by the possible loss of energy these restrictions are adapted in the simulation process to make the simulation more accurate. Additionally, the net present value (NPV) is similarly close to \$3,100, with operating costs of approximately \$ 315 million. In this way, the software delivers several results from which the designer can select the best solution for this electrification.

F. Chagres Design

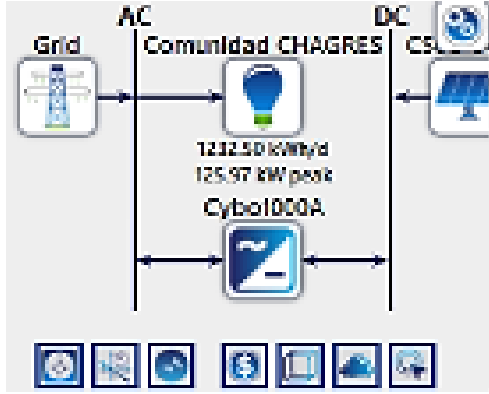


TABLE III
POWER AND ENERGY METRICS

Metric	Baseline	Scaled
Average (kWh/d)	1232,5	1232,5
Average (kW)	51,35	51,35
Peak (kW)	125,97	125,97
Load Factor	0,41	0,41

In the case of the CHAGRES population, the optimization results show that only the final 12 out of 344 simulations were feasible. The report notes that 112 simulations were deemed infeasible due to constraints related to the installed renewable energy capacity or potential energy losses. These limitations are incorporated into the simulation process to enhance its accuracy. Among the feasible outcomes, the results table indicates that the cost of energy remains consistent across the last four simulations, all averaging around 495. Similarly, the net present value (NPV) is approximately \$3,100, with operating costs near \$198 million. This approach allows the software to generate multiple viable options, enabling the designer to choose the most suitable solution for electrification.

G. Comparison results

The Table IV compares the simulation results for three communities: Chagres, Timba, and Quinamayó. Although these small rural populations differ in size and solar resource availability, these factors cause some variation in the results. For instance, communities with better solar conditions tend to experience slightly lower energy costs and more efficient performance. Population size also influences energy demand and system design. However, despite these differences, the overall pricing for the Standalone Solar Photovoltaic System

(SSFV) remains accessible and relatively similar across all three communities.

TABLE IV
METRICS OF THREE DIFFERENT VILLAGES

Village	Panel [kW]	Average [kWh/d]	Operating cost [COP\$/yr]
Timba	694	2465,1	436M
Quinamayó	500	1848,7	311M
Chagres	400	1232,5	199M

IV. CONCLUSIONS

This study highlights the significance of designing and simulating systems that utilize solar energy to establish energy communities in the southern region of Valle del Cauca, Colombia. It is among the first studies of its kind to demonstrate the application of computer tools, such as the HOMER program, which facilitates the efficient design of these energy communities. The study takes into account technical criteria, environmental regulations, and current laws in Colombia.

The article also demonstrates the benefits of using various platforms and databases. The solar resource or irradiance database, sourced from PVWatts and the Global Solar Atlas, was used to corroborate the high radiation levels at the site, with NASA data being used to make the simulation more accurate and precise. In this case, the study area receives around 4 kW/m² of solar radiation.

The irradiance of the site, the selection of optimal components, and the sizing of appropriate solar technology indicate that the solar photovoltaic system generates approximately 60% of the Timba population's average monthly electricity production, reducing this energy community's utility payment costs.

V. REFERENCES

- [1] "Plan Indicativo de Expansión de Cobertura de Energía Eléctrica - UPME." Accessed: May 28, 2025. [Online]. Available: <https://www.upme.gov.co/simec/planeacion-energetica/plan-de-expansion-de-cobertura/>
- [2] UPME and IDEAM, "Mapas de índice UV para Colombia," *Atlas de radiación solar para Colombia*, 2005.
- [3] "DANE - Censo general 2005." Accessed: May 28, 2025. [Online]. Available: <https://www.dane.gov.co/index.php/estadisticas-por-tema/demografia-y-poblacion/centro-general-2005-1>
- [4] "World Energy Outlook 2020 – Analysis - IEA." Accessed: May 28, 2025. [Online]. Available: <https://www.iea.org/reports/world-energy-outlook-2020>
- [5] H. Al-Masri, O. Dawaghreh, S. M.-J. of cleaner, and undefined 2023, "Optimal configuration of a large scale on-grid renewable energy systems with different design strategies," *Elsevier*, Accessed: May 27, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S0959652623017304>
- [6] S. Memon, R. P.-R. E. Focus, and undefined 2021, "An overview of optimization techniques used for sizing of hybrid renewable energy systems," *Elsevier*, Accessed: May 27, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1755008421000429>
- [7] C. Ammari, D. Belatrache, ... B. T.-E. and B., and undefined 2022, "Sizing, optimization, control and energy management of hybrid renewable energy system—A review," *Elsevier*, Accessed: May 27, 2025. [Online]. Available:

- <https://www.sciencedirect.com/science/article/pii/S2666123321000301>
- [8] J. Rojas-Zerpa, J. Y.-R. and sustainable energy reviews, and undefined 2015, "Application of multicriteria decision methods for electric supply planning in rural and remote areas," *Elsevier*, Accessed: Apr. 02, 2025. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S1364032115007868>
 - [9] M. I.-W. Engineering and undefined 2024, "Energy management strategies of hybrid renewable energy systems: A review," *journals.sagepub.com*, vol. 48, no. 1, pp. 133–161, Feb. 2024, doi: 10.1177/0309524X231200010.
 - [10] A. M. Samatar, S. Mekhilef, H. Mokhlis, M. Kermadi, and O. Alshammari, "Optimal design of a hybrid energy system considering techno-economic factors for off-grid electrification in remote areas," *Springer*, vol. 27, pp. 2073–2101, May 2024, doi: 10.1007/S10098-024-02939-3.
 - [11] H. Pujari, ... M. R.-T. on E. E., and undefined 2021, "Optimal design and techno-economic analysis of a hybrid grid-independent renewable energy system for a rural community," *Wiley Online Library*, vol. 31, no. 9, Sep. 2021, doi: 10.1002/2050-7038.13007.
 - [12] S. Vendoti, M. Muralidhar, R. K.- Environment, D. and, and undefined 2021, "Techno-economic analysis of off-grid solar/wind/biogas/biomass/fuel cell/battery system for electrification in a cluster of villages by HOMER software," *Springer*, vol. 23, no. 1, pp. 351–372, Jan. 2021, doi: 10.1007/S10668-019-00583-2.
 - [13] "Alejandria - Resolución 101_72 de 2025 CREG." Accessed: May 28, 2025. [Online]. Available: https://gestornormativo.creg.gov.co/gestor/entorno/docs/resolucion_creg_101-72_2025.htm
 - [14] I. Siksnelyte-Butkiene and D. Streimikiene, "Multi-Criteria Decision Making for Sustainable Energy Development in Households," *Sustainable Energy Development*, pp. 133–164, Jun. 2023, doi: 10.1201/9781003327196-5.
 - [15] S. Abdullah and A. Markandya, "Rural electrification programmes in Kenya: Policy conclusions from a valuation study," *Energy for Sustainable Development*, vol. 16, no. 1, pp. 103–110, 2012, doi: 10.1016/j.esd.2011.10.007.
 - [16] F. Fuso Nerini, M. Howells, M. Bazilian, and M. F. Gomez, "Rural electrification options in the Brazilian Amazon: A multi-criteria analysis," *Energy for Sustainable Development*, vol. 20, pp. 36–48, Jun. 2014, doi: 10.1016/J.ESD.2014.02.005.
 - [17] M. M. Kamal, A. Mohammad, I. Ashraf, and E. Fernandez, "Rural electrification using renewable energy resources and its environmental impact assessment," *Environmental Science and Pollution Research*, vol. 29, no. 57, pp. 86562–86579, Dec. 2022, doi: 10.1007/S11356-022-22001-3/FIGURES/6.
 - [18] T. Lambert, P. Gilman, and P. Lilienthal, "Micropower System Modeling with Homer," in *Integration of Alternative Sources of Energy*, John Wiley & Sons, Inc., 2006, pp. 379–418. doi: 10.1002/0471755621.ch15.
 - [19] S. Yadav, P. Kumar, and A. Kumar, "Techno-economic assessment of hybrid renewable energy system with multi energy storage system using HOMER," *Energy*, vol. 297, p. 131231, Jun. 2024, doi: 10.1016/J.ENERGY.2024.131231.
 - [20] M. K. Shahzad, A. Zahid, T. ur Rashid, M. A. Rehan, M. Ali, and M. Ahmad, "Techno-economic feasibility analysis of a solar-biomass off grid system for the electrification of remote rural areas in Pakistan using HOMER software," *Renew Energy*, vol. 106, pp. 264–273, Jun. 2017, doi: 10.1016/j.renene.2017.01.033.
 - [21] A. O. Rousis, D. Tzelepis, I. Konstantelos, C. Booth, and G. Strbac, "Design of a Hybrid AC/DC Microgrid Using HOMER Pro: Case Study on an Islanded Residential Application," *Inventions 2018*, Vol. 3, Page 55, vol. 3, no. 3, p. 55, Aug. 2018, doi: 10.3390/INVENTIONS3030055.

Propuesta Metodológica para Evaluación Técnico-Económica en la Producción de Amoníaco Verde Multipropósito en Colombia

Carlos A. Nieto Ortiz, *Estudiante*, Johann A. Hernández Mora, y Diego J. Rodríguez Patarroyo, *Docentes, Maestría en ingeniería, Facultad de ingeniería, Universidad Distrital Francisco José de Caldas Bogotá D.C., Colombia*
canietoo@udistrital.edu.co jahernandezm@udistrital.edu.co djrodriguezp@udistrital.edu.co

Resumen—Este documento presenta una propuesta metodológica para el desarrollo de investigaciones enfocadas en la viabilidad técnico-económica para la producción de amoníaco multipropósito (principalmente para fertilizantes), a partir de hidrógeno verde generado a través de Fuentes No Convencionales de Energía Renovable (FNCER), todo esto en el contexto colombiano. La metodología se basa en cinco (5) etapas principales: (1) recopilación de información científica, (2) parámetros de evaluación y ejecución, (3) viabilidad técnico económica, (4) política pública y marco regulatorio y (5) documentación y divulgación del proyecto. Dentro de cada una de estas etapas se abordan los aspectos claves que son necesarios para obtener un resultado tanto claro como conciso en la investigación, la ejecución de cada etapa es secuencial, su elaboración sirve como datos de entrada para la siguiente, con relación a las etapas (3) y (4) teniendo en cuenta la naturaleza de la investigación que es la implementación de FNCER se aplican directamente en el modelo económico los incentivos y beneficios según la normativa colombiana vigente, en la etapa (4) se expone de manera más detallada toda la reglamentación requerida y utilizada. Se considera que esta propuesta contiene todos los componentes necesarios para presentar un análisis técnico económico preciso y confiable en la ejecución de este tipo de proyectos llevándolos a una fase de ingeniería conceptual.

Palabras clave—Amoníaco, electricidad, energía, fertilizantes, FNCER, hidrógeno, metodología, técnico económica.

I. INTRODUCCIÓN

LA producción de amoníaco verde toma más fuerza en la actualidad. Investigaciones como la de [1], enfocada en dos países de Sudamérica, evidencian que “Colombia y Venezuela tienen un gran potencial para abastecer la demanda de fertilizantes nitrogenados con producción de amoníaco verde, aplicar hidrógeno verde en la fabricación y utilizar los excedentes para sustitución energética de Gas Licuado de Petróleo – GLP”. Es importante mencionar, que un elemento que contribuye a la producción de amoníaco a partir de energías

verdes en Colombia es la ley 1715 de 2014 “Por medio de la cual se regula la integración de las energías renovables no convencionales al Sistema Energético Nacional” [2], por lo cual es de vital importancia hacer uso de este tipo de incentivos para la realización de estos proyectos que finalmente aportan para activar el proceso de transición energética, y a nivel económico impactar de manera positiva en costos de producción y precios de los alimentos, por dar un ejemplo de sus aplicaciones.

Según lo indicado por [3], a la fecha se están ejecutando más de nueve (9) proyectos en toda Latinoamérica que se relacionan con la producción de amoníaco verde. Sumado a esto, en [4] indica que “el amoníaco renovable se produce a partir de hidrógeno renovable, que a su vez se produce a través de la electrólisis del agua utilizando electricidad renovable”, lo que se traduce en un proceso altamente amigable con el medio ambiente en todas sus fases. Por último, cabe mencionar que Colombia se ha propuesto desde el año 2021 iniciar el proceso de transición energética, con metas como llegar a un nivel superior al 12% en la participación de fuentes no convencionales de energías renovables en la matriz eléctrica, según se menciona en [5].

Asimismo, con la expedición de la ley 2099 de 2021 “Por medio de la cual se dictan disposiciones para la transición energética, la dinamización del mercado energético, la reactivación económica del país y se dictan otras disposiciones” [6], se creó una ruta de implementación del hidrógeno como fuente de energía renovable, lo cual facilita y fortalece la puesta en marcha de energías renovables para la obtención de amoníaco verde.

En la Figura 1 se observa un diagrama de flujo que muestra la metodología propuesta en este trabajo, la cual está compuesta por cinco (5) etapas definidas, las partes que la componen y como se relacionan. En los siguientes capítulos se presenta como está organizada cada etapa, y finalmente se presentan

Este documento es desarrollado y entregado en el mes de julio de 2025, tiene como objetivo cumplir con la producción y/o divulgación científica de la investigación realizada para atender los requerimientos necesarios de grado del programa de Maestría en Ingeniería de la facultad de ingeniería de la Universidad Distrital Francisco José de Caldas ubicada en la ciudad de Bogotá – Colombia. Los recursos utilizados en la investigación provienen del autor principal. (Autor correspondiente: Carlos A. Nieto Ortiz). El autor principal desarrollo todo el contenido presentado en este documento, el segundo y tercer autor desarrollaron actividades de revisión y aportes técnicos puntuales en la construcción de la investigación.

Carlos A. Nieto Ortiz, es estudiante de Maestría en Ingeniería de la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas, Bogotá D.C., Colombia (correo electrónico: canietoo@udistrital.edu.co).

PhD. Johann A. Hernández Mora, es docente de planta de la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas, Bogotá D.C., Colombia (correo electrónico: jahernandezm@udistrital.edu.co).

PhD. Diego J. Rodríguez Patarroyo, es docente de planta de la Facultad de Ingeniería de la Universidad Distrital Francisco José de Caldas, Bogotá D.C., Colombia (correo electrónico: djrodriguezp@udistrital.edu.co).

unas conclusiones del trabajo realizado.

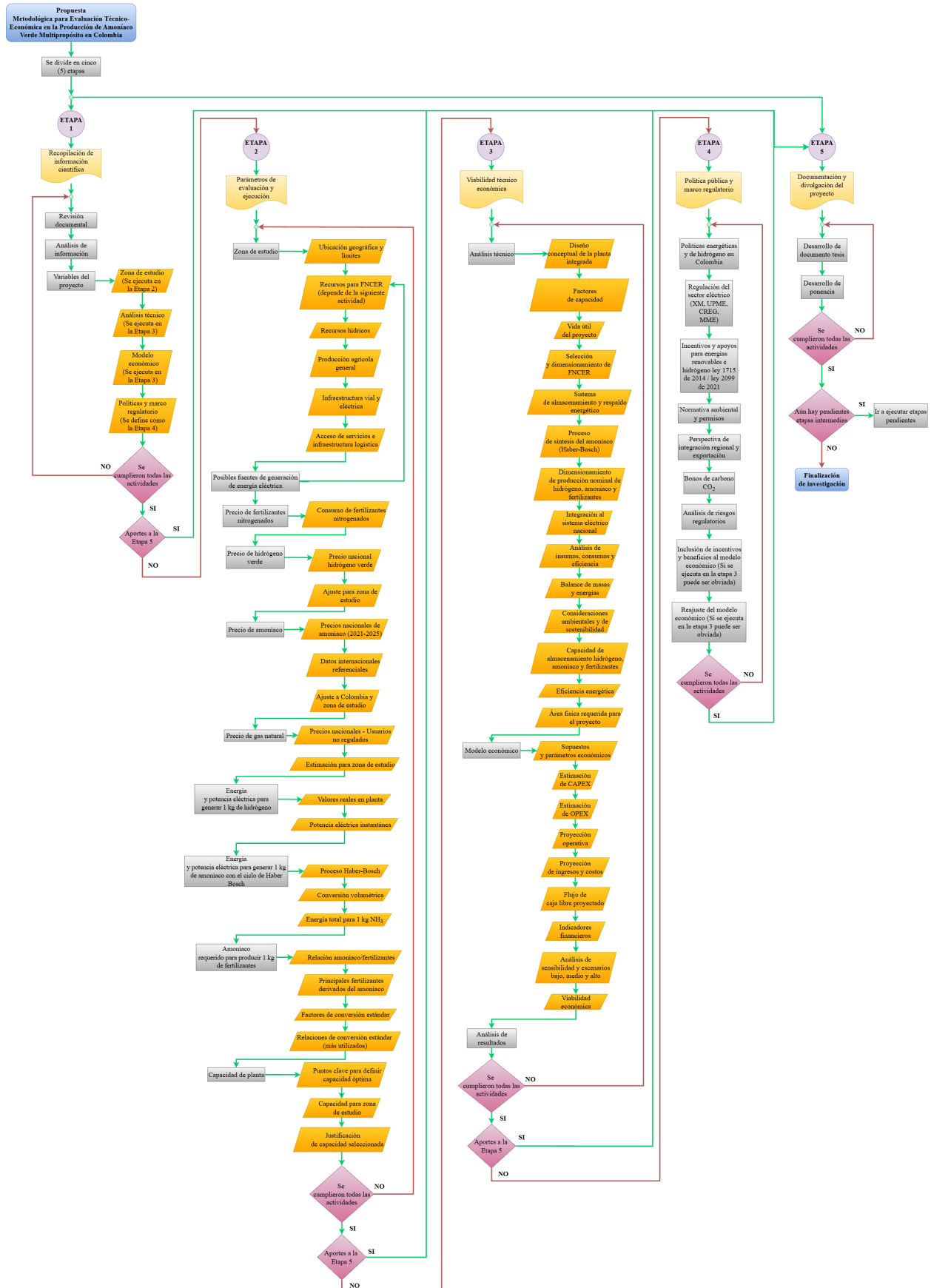


Fig. 1. Diagrama de flujo propuesta metodológica.

II. RECOPIACIÓN DE INFORMACIÓN CIENTÍFICA

Se compone de tres actividades: revisión documental, análisis de información y variables del proyecto. Las dos primeras se enfocan en la búsqueda, recopilación y estudio de todos los datos necesarios para desarrollar la investigación. La tercera tiene como objetivo definir las variables fundamentales que deben ser exploradas. En este caso se establecen cuatro: zona de estudio, análisis técnico, modelo económico y políticas y marco regulatorio, las cuales son ejecutadas en las siguientes etapas.

III. PARÁMETROS DE EVALUACIÓN Y EJECUCIÓN

Habiendo obtenido e indagado la información requerida, así como identificado las variables principales, se procede a establecer los parámetros de evaluación y ejecución. Está compuesta por diez (10) actividades: zona de estudio (que es una variable fundamental), posibles fuentes de generación de energía eléctrica, precio de fertilizantes nitrogenados, precio de hidrógeno verde, precio de amoníaco, precio de gas natural, energía y potencia eléctrica para generar 1 kg de hidrógeno, energía y potencia eléctrica para generar 1 kg de amoníaco con el ciclo de Haber Bosch, amoníaco requerido para producir 1 kg de fertilizantes y capacidad de la planta.

A. Zona de estudio

- 1) **Ubicación geográfica y límites**
- 2) **Recursos para FNCER (depende de la actividad B)**
- 3) **Recursos hídricos**
- 4) **Producción agrícola general**
- 5) **Infraestructura vial y eléctrica**
- 6) **Acceso a servicios e infraestructura logística**

B. Posibles fuentes de generación de energía eléctrica

Definida la zona de estudio se evalúa que FNCER es la más apropiada y con mayor potencial de uso, es importante contrastar al menos tres (3) fuentes de energía que puedan ser desarrolladas en el área.

C. Precio de fertilizantes nitrogenados

- 1) **Consumo de fertilizantes nitrogenados**

D. Precio de hidrógeno verde

- 1) **Precio nacional del hidrógeno verde**
- 2) **Ajuste para zona de estudio**

E. Precio de amoníaco

- 1) **Precios nacionales de amoníaco (2021-2025)**
- 2) **Datos internacionales referenciales**
- 3) **Ajuste a Colombia y zona de estudio**

F. Precio de gas natural

- 1) **Precios nacionales – Usuarios no regulados**

2) Estimación para zona de estudio

G. Energía y potencia eléctrica para generar 1 kg de hidrógeno

- 1) **Valores reales en planta**
- 2) **Potencia eléctrica instantánea**

H. Energía y potencia eléctrica para generar 1 kg de amoníaco con el ciclo de Haber Bosch

- 1) **Proceso Haber-Bosch**
- 2) **Conversión volumétrica**
- 3) **Energía total para 1 kg de NH_3**

I. Amoníaco requerido para producir 1 kg de fertilizantes

- 1) **Relación amoníaco/fertilizantes**
- 2) **Principales fertilizantes derivados del amoníaco**
- 3) **Factores de conversión estándar**
- 4) **Promedio para fertilizantes nitrogenados en Colombia**
- 5) **Relaciones de conversión estándar (más utilizados)**

J. Capacidad de la planta

- 1) **Puntos clave para definir capacidad óptima**
- 2) **Capacidad para zona de estudio**
- 3) **Justificación de capacidad seleccionada**

IV. VIABILIDAD TÉCNICO ECONÓMICA

Esta etapa de la investigación arroja resultados clave que deben ser procesados, analizados y evaluados al detalle. Está compuesta por tres (3) actividades: análisis técnico (variable fundamental), modelo económico (variable fundamental) y análisis de resultados.

A. Análisis técnico

- 1) **Diseño conceptual de la planta integrada**
- 2) **Factores de capacidad**
- 3) **Vida útil del proyecto**
- 4) **Selección y dimensionamiento de FNCER**
- 5) **Sistema de almacenamiento y respaldo energético**
- 6) **Proceso de síntesis del amoníaco (Haber-Bosch)**
- 7) **Dimensionamiento de producción nominal de hidrógeno, amoníaco y fertilizantes**
- 8) **Integración al sistema eléctrico nacional**
- 9) **Análisis de insumos, consumos y eficiencia**
- 10) **Balance de masas y energías**
- 11) **Consideraciones ambientales y de sostenibilidad**
- 12) **Capacidad de almacenamiento hidrógeno, amoníaco y fertilizantes**
- 13) **Eficiencia energética**
- 14) **Área física requerida para el proyecto**

B. Modelo económico

- 1) **Supuestos y parámetros económicos**
- 2) **Estimación de CAPEX**

- 3) **Estimación de OPEX**
- 4) **Proyección operativa**
- 5) **Proyección de ingresos y costos**
- 6) **Flujo de caja libre proyectado**
- 7) **Indicadores financieros**
- 8) **Análisis de sensibilidad y escenarios bajo, medio y alto**
- 9) **Viabilidad económica**

C. Análisis de resultados

Allí se reúnen todos los datos terminales a lo largo de la investigación tanto técnicos como económicos y se generan las conclusiones, es posible plantear y evaluar alternativas de mejora en este punto.

V. POLÍTICA PÚBLICA Y MARCO REGULATORIO

Esta etapa expone e identifica las políticas, regulaciones, leyes e incentivos que rigen y pueden ser aplicadas a este tipo de proyectos. Son nueve (9) las actividades efectuadas: políticas energéticas y de hidrógeno en Colombia, regulación del sector eléctrico (XM, UPME, CREG, MME), incentivos y apoyos para energías renovables e hidrógeno ley 1715 de 2014 / ley 2099 de 2021, normativa ambiental y permisos, perspectiva de integración regional y exportación, bonos de carbono CO₂, análisis de riesgos regulatorios, inclusión de incentivos y beneficios al modelo económico y reajuste del modelo económico. Las dos (2) últimas actividades pueden ser obviadas si en la etapa (3) se aplican los incentivos directamente los incentivos al modelo económico.

VI. CONCLUSIONES

Esta propuesta metodológica permite realizar un análisis de profundidad de la temática objetivo de la investigación, su estructura facilita evaluar paso a paso los aspectos técnicos, económicos y normativos requeridos. Además, mantiene una línea base de trabajo mientras a la vez brinda alternativas de ejecución de las etapas establecidas. Es importante aclarar que recae sobre el investigador efectuar cada una de las actividades en el flujo diseñado, así como garantizar toda la información necesaria para que el resultado final logrado sea de buena calidad.

Esta propuesta metodológica tiene un alcance a nivel de ingeniería conceptual, para que pueda ser utilizada en una fase superior se considera necesario la conformación de un grupo interdisciplinario para reestructurar y hacer la adhesión de nuevas actividades en cada etapa.

Actualmente, esta metodología está siendo aplicada en un proyecto en el departamento de Arauca, su implementación ha arrojado resultados positivos permitiendo definir que, para la producción de hidrógeno en la zona, la FNCER más adecuada es la energía solar fotovoltaica, obtener los precios base de los productos principales como lo son el amoníaco, el hidrógeno y los fertilizantes, además, se estableció la capacidad de la planta teniendo en cuenta dos opciones: una conservadora y otra óptima. Todos los datos mencionados anteriormente son relevantes en la investigación y

precisarlos adecuadamente aumenta la probabilidad de éxito.

AGRADECIMIENTOS

Se agradece el apoyo técnico al grupo de investigación Laboratorio de Investigación de Fuentes Alternativas de Energía (LIFAE) de la Universidad Distrital Francisco José de Caldas.

REFERENCIAS

- [1] W. Nadaleti, E. Gomes de Souza, y V. Lourenc, "Green hydrogen-based pathways and alternatives: Towards the renewable energy transition in South America's regions-Part B," *International Journal of Hydrogen Energy*, pp. 1-15, 2022.
- [2] Congreso de la República de Colombia, "Ley 1715: Por medio de la cual se regula la integración de las energías renovables no convencionales al sistema energético nacional," Bogotá D.C., 13 de mayo de 2014.
- [3] IEA, "Hydrogen in Latin America. From near-term opportunities to large-scale deployment," 2021.
- [4] IRENA & AEA, "Innovation Outlook: Renewable Ammonia," *International Renewable Energy*, 2022.
- [5] Ministerio de Minas y Energía, "Transición energética: un legado para el presente y el futuro de Colombia," 2021.
- [6] Congreso de la República de Colombia, "Ley 2099: Por medio de la cual se dictan disposiciones para la transición energética, la dinamización del mercado energético, la reactivación económica del país y se dictan otras disposiciones," Bogotá D.C., 10 de julio de 2021.
- [7] INVIAS, "Inventario de carreteras nacionales," 2024.
- [8] UPME, "Plan de expansión eléctrica," 2023.
- [9] XM, "Informe de operación SIN," 2023.
- [10] UPRA, "Frontera agrícola y aptitud agropecuaria," 2023.
- [11] Agronet, "Estadísticas agrícolas regionales," 2023.
- [12] MME, "Hoja de Ruta para la Transición Energética," 2021.
- [13] XM, "Estadísticas de generación renovable en Colombia," 2024.
- [14] MME, Hoja de Ruta para la Transición Energética. 2021.
- [15] MME, UPME, "Hoja de Ruta del Hidrógeno en Colombia," 2021, 2024.
- [16] IEA, "Global Hydrogen Review," 2021–2024.
- [17] IRENA, "Hydrogen: A Renewable Energy Perspective," 2023.
- [18] UPME, "Actualización hidrógeno Colombia," 2024.
- [19] CREG, "Reporte de precios de gas natural usuarios no regulados," 2024.

Análisis de la dinámica de los embarazos no deseados y los abortos inseguros en Colombia mediante dinámica de sistemas

Chacón-Tibaduiza, Natalia; Franco Franco Carlos Alberto; Guzmán-Cortés, Diana Carolina*

Abstract—Este trabajo presenta un análisis de la problemática de embarazos no deseados y el aborto inseguro en Colombia, mediante dinámica de sistemas. Se identifican las variables que tienen relación directa o indirecta con la problemática con el fin de entender la relación causal entre éstas, y luego se construye y parametriza el modelo de simulación de flujos y niveles o modelo de Forrester para analizar el comportamiento. El modelo de simulación desarrollado logra representar el comportamiento de las principales variables, el cual a futuro permitirá la implementación de distintos escenarios.

Index Terms— dinámica de sistemas, embarazos no deseados, abortos; simulación.

I. INTRODUCCIÓN

Cifras del *Guttmacher Institute* [1] evidencian que el embarazo no planificado continúa siendo un desafío relevante en Colombia, con una tasa anual estimada de 89 casos por cada 1.000 mujeres en edad reproductiva. Estas cifras presentan marcadas diferencias regionales: mientras en las zonas Central y Oriental la tasa se sitúa en 67 por 1.000, en Bogotá alcanza los 113 por 1.000, lo que refleja profundas disparidades territoriales. Según [1], el embarazo no planeado conduce a nacimientos no deseados, de manera particular, en Colombia la incidencia de nacimientos no deseados pasó de 36% en 1990 a 51% en 2010, cifra que representa un reto para las políticas de salud pública y educación sexual. Así mismo, en el contexto colombiano, aproximadamente el 44% de los embarazos no deseados concluyen en una interrupción inducida del embarazo, es decir, un aborto inducido. lo que representa aproximadamente 400.400 abortos anuales.

De acuerdo con el *Ministerio de Salud y Protección Social* [2], el embarazo no planeado y el aborto inducido constituyen problemas de salud pública estrechamente vinculados, cuya ocurrencia responde a factores individuales, culturales, educativos y socioeconómicos. [2] menciona que se ha identificado que las tasas más elevadas de fecundidad coinciden con contextos de pobreza, baja escolaridad y deterioro de

indicadores de salud como la mortalidad materna e infantil. A esto se suma la demanda insatisfecha de métodos anticonceptivos, especialmente en poblaciones vulnerables con mayores barreras de acceso a los servicios de salud [2].

Este mismo informe [2], reporta que, a nivel mundial ocurren cerca de 46 millones de abortos al año, de los cuales alrededor de 22 millones se practican en condiciones inseguras, principalmente en países en desarrollo. Según la Organización Mundial de la Salud, una de cada ocho muertes maternas está asociada a complicaciones derivadas de abortos inseguros, lo que representa aproximadamente 47 mil fallecimientos anuales y múltiples secuelas permanentes, incluida la infertilidad. En el caso colombiano, se estima que cada año podrían registrarse hasta 400.000 abortos, de los cuales el 99% se realizan en la clandestinidad. Los datos oficiales del año 2016 reportan entre 3.100 y 3.600 procedimientos médicos y un promedio de 15.000 abortos espontáneos anuales en los últimos cinco años.

Por lo tanto, este estudio aborda la problemática de los embarazos no deseados y los abortos inseguros en Colombia mediante un enfoque de dinámica de sistemas, con el propósito de identificar las variables que explican este comportamiento, analizar las relaciones causales que las vinculan y construir un modelo de simulación de Forrester para su representación.

Este documento se organiza así: la sección II presenta la metodología, la sección III los resultados y la sección IV las conclusiones.

II. METODOLOGÍA

Para el desarrollo de este trabajo se siguió la metodología de Sterman [3]. El proceso de modelado en dinámica de sistemas comprende cinco pasos, se inicia con la selección y articulación del problema, posteriormente se plantea la hipótesis dinámica, luego se realiza la formulación del modelo de flujos y niveles o de Forrester, en cuarto lugar, realizar el proceso de

*Corresponding author: Diana Carolina Guzmán-Cortés

Natalia Chacón-Tibaduiza is with Universidad del Rosario (e-mail: [giseth.chacon@urosario.edu.co](mailto:giset.chacon@urosario.edu.co))

Diana Carolina Guzmán-Cortés is with Universidad Distrital Francisco José de Caldas (e-mail: dcguzmanc@udistrital.edu.co)

Carlos Alberto Franco-Franco is with Universidad Distrital Francisco José de Caldas (e-mail: cafrancof@udistrital.edu.co)

validación/verificación, y finalmente realizar la evaluación de escenarios o políticas. En este trabajo realizamos los primeros cuatro pasos.

Siguiendo esta metodología, en primer lugar, se identificó la problemática de salud pública de los embarazos no deseados y su relación con los abortos inseguros en Colombia, y se realizó un proceso de levantamiento de información para tener los modos de referencia de variables como índice de fecundidad, embarazos, total de nacimientos, nacimientos no planeados, niños dados en adopción y abortos clandestinos, los cuales representan el comportamiento de algunas de las variables relacionadas con la problemática.

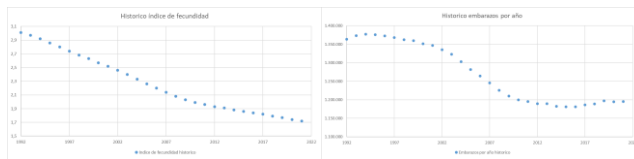


Fig. 1. Histórico de índice fecundidad y embarazos.

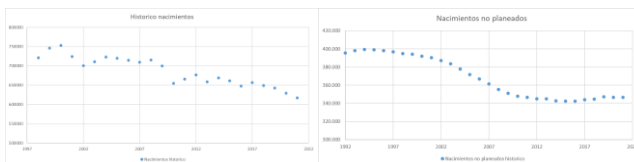


Fig. 2. Histórico de nacimientos y nacimientos no planeados.

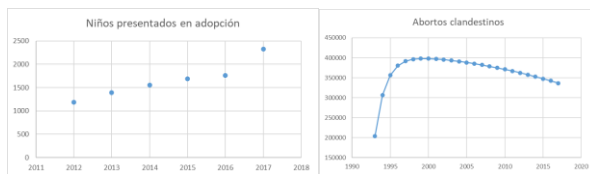


Fig. 3. Histórico de niños de adopción y abortos clandestinos

En segundo lugar, identificaron las variables más importantes dentro de la problemática, así como sus relaciones causales con el fin de plantear la hipótesis dinámica. Una vez planteada la hipótesis dinámica, se procedió a la construcción del modelo de simulación y finalizamos con algunas pruebas de validación como verificación de unidades y pruebas de comportamiento extremo.

III. RESULTADOS

El modelo de simulación construido se presenta en la Figura 4, allí se pueden observar distintas tasas utilizadas como índice de fecundidad, tasa de mujeres fértiles, tasa de mujeres que se embarazan, población de mujeres en edad fértil, tasa de embarazos no deseados, tasa de embarazos deseados, tasas de nacimientos planeados y no planeados, tasas de abortos legales e ilegales, incidencia de la pobreza, tasas de intervención social con relación a los niños dados en adopción en intervención

social, entre otras. Se pueden apreciar también los diferentes flujos de entrada y salida utilizados para sus respectivos niveles. Algunos de los niveles utilizados corresponden a embarazos totales, nacimientos planeados y no planeados, abortos legales e ilegales y nivel de niños adoptados.



Fig. 4. Modelo de simulación de Forrester o de flujos y niveles

La figura 5 presenta la prueba de verificación de unidades del modelo propuesto.

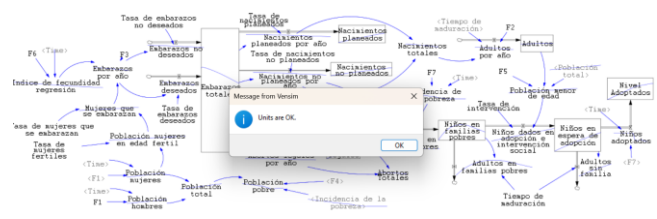


Fig. 5. Verificación de unidades

Algunos resultados del modelo de simulación se presentan en las figuras 6 y 7. En estas figuras podemos observar que el modelo de simulación construido logra representar el comportamiento histórico de las variables

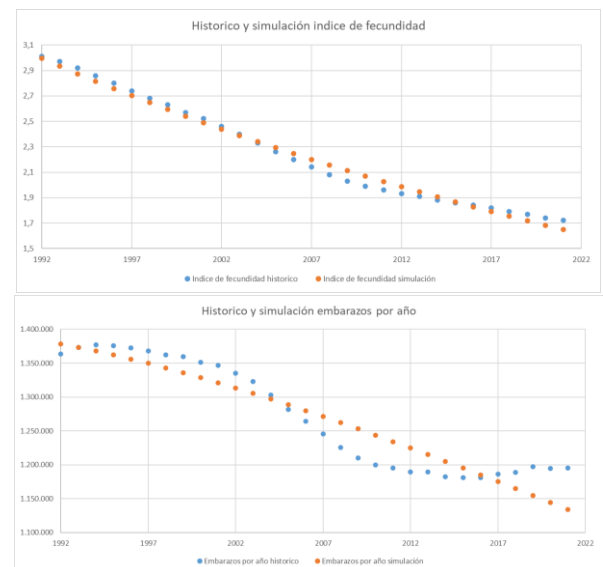


Fig. 6. Resultados simulados versus histórico de datos de índice de fecundidad y embarazos por año

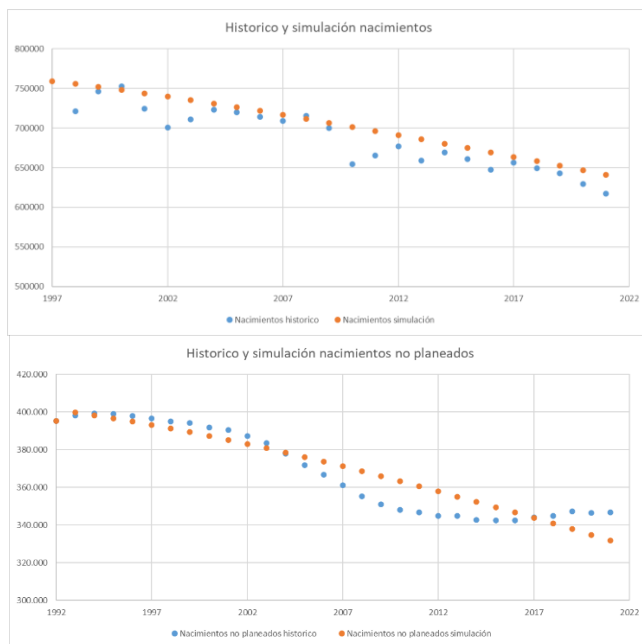


Fig. 7. Resultados simulados versus histórico de datos de total de nacimientos y nacimientos no planeados

IV. CONCLUSIONES

La dinámica de sistemas es una herramienta útil para abordar problemáticas sociales. En este caso, se abordó una problemática de salud pública como lo son los embarazos no deseados que pueden conllevar a abortos inseguros.

Se observa que el modelo propuesto lograr representar el comportamiento histórico de las variables relacionadas, lo cual es útil para poder proponer escenarios o reglas de decisión y evaluar el impacto de éstos para la problemática analizada.

REFERENCIAS

- [1] GUTTMACHER INSTITUTE, EMBARAZO NO DESEADO EN COLOMBIA. [ONLINE]. AVAILABLE: [HTTPS://WWW.GUTTMACHER.ORG/SITES/DEFAULT/FILES/PDFS/PUBS/EMBARAZO-NO-DESEADO-COLOMBIA.PDF](https://www.guttmacher.org/sites/default/files/pdfs/pubs/EMBARAZO-NO-DESEADO-COLOMBIA.PDF)
- [2] MINISTERIO DE SALUD Y PROTECCIÓN SOCIAL, ABC DE LA MATERNIDAD ELEGIDA, BOGOTÁ D.C., COLOMBIA. [ONLINE]. AVAILABLE: [HTTPS://WWW.MINSALUD.GOV.CO/SITES/RID/LISTS/BIBLIOTECADIGITAL/RIDE/VS/PP/ABC-MATERNIDAD-ELEGIDA.PDF](https://www.minsalud.gov.co/sites/rid/lists/bibliotecaDigital/RIDE/VS/PP/ABC-MATERNIDAD-ELEGIDA.PDF)
- [3] J. D. STERMAN, BUSINESS DYNAMICS: SYSTEMS THINKING AND MODELING FOR A COMPLEX WORLD. BOSTON, MA, USA: IRWIN/MCGRAW-HILL, 2000.

Comprendiendo la dinámica del cultivo de coca en Colombia: un enfoque desde la dinámica de sistemas

S. García & L. Cárdenas

Abstract— Este estudio analiza la persistencia y expansión del cultivo de coca en Colombia mediante un modelo de simulación basado en dinámica de sistemas. El modelo integra variables económicas, sociales y territoriales, como la rentabilidad del cultivo, la presión de la demanda internacional y la presencia de Grupos Armados Ilegales, para identificar los mecanismos de realimentación que sostienen el fenómeno. Los resultados evidencian que las políticas centradas exclusivamente en la oferta, como la erradicación manual forzada, resultan limitadas o incluso contraproducentes si no se acompañan de transformaciones estructurales. Se propone avanzar hacia enfoques integrales que incorporen tanto las dinámicas locales como los factores globales asociados a la demanda. La simulación permite explorar escenarios alternativos y evaluar el impacto de distintas decisiones de política, brindando insumos valiosos para el diseño de estrategias que transformen las condiciones que favorecen la continuidad del cultivo de coca en el país.

Palabras claves: Cultivo de coca, Colombia, Dinámica de Sistemas, Políticas, Simulación.

I. INTRODUCTION

La problemática asociada al cultivo de coca constituye uno de los desafíos más importantes para Colombia [1]. Más allá de su relación con economías ilegales, este fenómeno refleja condiciones estructurales de desigualdad social, baja presencia institucional en áreas rurales y debilidad de las políticas de desarrollo territorial [2]. La permanencia del cultivo en el tiempo y su expansión evidencian tanto la magnitud del problema como las limitaciones de las estrategias adoptadas para enfrentarlo. Durante más de cinco décadas, el Estado colombiano, en articulación con actores internacionales, han adoptado diversas estrategias para mitigar su crecimiento. Entre ellas destacan, la estrategia de fumigación aérea con glifosato, Erradicación Manual Forzada y la Erradicación Manual Voluntaria. No obstante, los resultados obtenidos han sido insuficientes, y en algunos casos han generado efectos no deseados, lo cual pone de manifiesto la necesidad de replantear los enfoques tradicionales de intervención [3].

A pesar de los esfuerzos, las cifras recientes muestran que el problema no solo persiste, sino que se ha intensificado. Colombia continúa siendo el país con la mayor extensión de

cultivos de coca a nivel mundial [4]. De acuerdo con los datos más recientes de la Oficina de las Naciones Unidas contra la Droga y el Delito [5], el área sembrada ha alcanzado un récord histórico de 253,000 hectáreas para el año 2023. Esta tendencia evidencia las limitaciones de las estrategias aplicadas, las cuales no han logrado generar cambios ni romper con los mecanismos que reproducen el fenómeno. En consecuencia, la expansión y persistencia del cultivo de coca ha despertado un interés creciente en el ámbito académico [6,7,8,9]. Gran parte de la literatura se ha orientado a evaluar las respuestas institucionales al fenómeno, con especial atención a las políticas de control y erradicación. Estas investigaciones han identificado una serie de efectos no deseados derivados de dichas intervenciones, entre los que destacan el desplazamiento de los cultivos hacia nuevas regiones, el fortalecimiento de redes criminales más sofisticadas, y el deterioro del tejido social en comunidades rurales expuestas a estas dinámicas [10,11]. Sin embargo, dada la naturaleza persistente y multifactorial del fenómeno, se hace cada vez más evidente la necesidad de incorporar un nuevo enfoque que permita reconocer la complejidad del sistema en su conjunto.

Uno de los principales aportes de esta investigación consiste en ofrecer una nueva perspectiva para comprender y abordar el problema del cultivo de coca, a partir del uso de herramientas metodológicas provenientes del pensamiento sistémico. Se propone, en particular, la utilización de simulación en dinámica de sistemas, el cual permite representar y analizar la evolución de fenómenos complejos a partir de la identificación de sus componentes, las relaciones causales que los vinculan y los mecanismos de realimentación que determinan su comportamiento en el tiempo [12]. Esta herramienta permite capturar la interacción entre múltiples variables, así como explorar patrones de comportamiento. Esta aproximación pretende contribuir a un entendimiento de la situación, que permita explorar de manera informada distintas estrategias de intervención.

II. METODOLOGÍA

A. Dinámica de sistemas

La metodología de dinámica de sistemas fue escogida para el

S García. Es estudiante de Ingeniería Industrial en la Universidad de Antioquia, es miembro del semillero Ingeniería y Sociedad SIS (e-mail: sara.garciac1udea.edu.co)

L Cárdenas. Es profesora en la Universidad de Antioquia. Tiene un Doctorado en Ingeniería de la Universidad Nacional de Colombia, es coordinadora del Semillero Ingeniería y Sociedad SIS (e-mail: lauram.cardenas@udea.edu.co)

desarrollo de este estudio por su utilidad en la representación de comportamientos sociales complejos, su capacidad para proyectar escenarios futuros y su potencial para evaluar en entornos simulados, los efectos de distintas estrategias y políticas. La dinámica de sistemas se distingue de otras metodologías al capturar elementos como realimentación, la complejidad estructural del sistema y las relaciones no lineales entre variables [12,13].

B. Diagrama causal

La figura 1. presenta las relaciones presentes entre las variables más representativas de la problemática alrededor de los cultivos de coca. Estas variables resultan del proceso de consulta y análisis de la documentación científica y gubernamental de la problemática. Las letras B indican la presencia de ciclos de balance y las letras R indican la presencia de ciclos de refuerzo en la hipótesis.

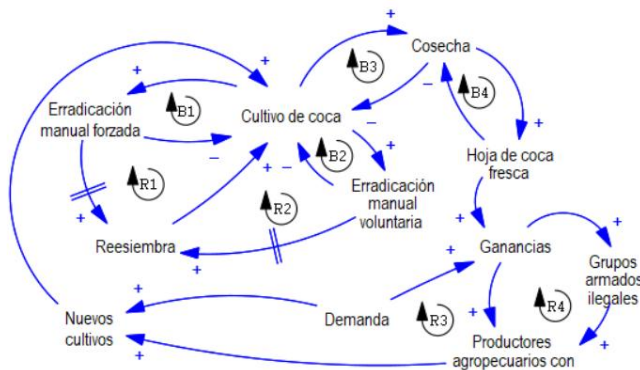


Fig. 1. Diagrama causal.

La cantidad de hectáreas de coca cultivada es una variable clave para entender la dinámica del narcotráfico en Colombia [2]. Para reducir este fenómeno, el gobierno ha implementado estrategias como la Erradicación manual forzada (B1), cuyo objetivo es eliminar directamente los cultivos ilícitos. Sin embargo, su efecto puede ser ambiguo, en muchos casos incentiva la resiembra (R1) limitando su efectividad a largo plazo [15]. Como alternativa, el Programa Nacional Integral de Sustitución de Cultivos de Uso Ilícito (PNIS) promovió la Erradicación manual voluntaria (B2) con participación comunitaria, asistencia técnica y desarrollo alternativo [2]. No obstante, también se han reportado efectos de rebote donde se incentiva la resiembra (R2) [8].

Como es común en los procesos agrícolas, la cosecha (B3) es una etapa importante, pues implica la recolección de la materia prima, en este caso hoja de coca (B4) [3]. Cada ciclo de cosecha genera ingresos para los productores, lo que les permite reinvertir en nuevas siembras para asegurar el siguiente ciclo productivo. Este proceso refuerza la capacidad de los productores para expandir el área cultivada. De este modo, la alta rentabilidad del cultivo (R3) se convierte en el principal incentivo económico para los Productores Agropecuarios con Coca (PAC), quienes amplían las hectáreas sembradas buscando mayores ganancias. A su vez, la presencia y el control territorial de los Grupos Armados Ilegales (GAI) (R4) refuerzan

esta dinámica, al ofrecer protección y condiciones para la producción y comercialización, consolidando así el crecimiento del sistema [10,14].

C. Estructura del modelo

Con el fin de proporcionar una visión general de la estructura del modelo de simulación desarrollado, la figura 1 presenta una representación modular del modelo, destacando sus principales componentes.

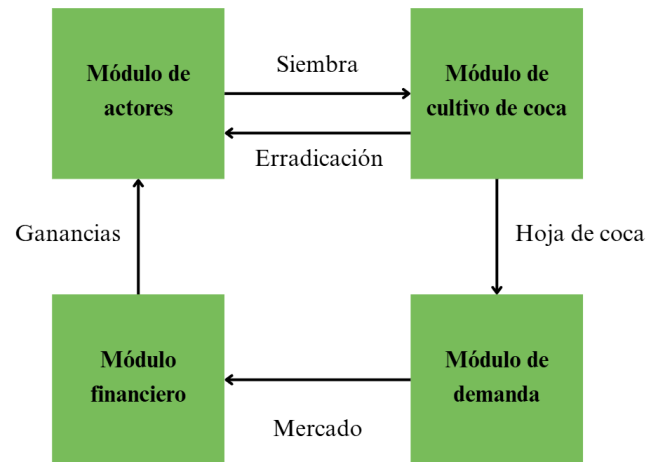


Fig. 2. Módulos representativos del modelo.

El módulo de cultivo de coca representa las dinámicas relacionadas con las hectáreas sembradas en Colombia, incorporando procesos como la siembra, la cosecha y la erradicación, así como los efectos derivados de estas estrategias. Este módulo permite capturar los cambios en la superficie cultivada a lo largo del tiempo. Por su parte, el módulo de demanda considera la presión ejercida por los mercados internacionales, que condicionan el comportamiento del sistema [2]. Este componente es fundamental, dado que la hoja de coca constituye la materia prima para la transformación en cocaína, lo cual mantiene activa la cadena de producción.

El módulo financiero se centra en los aspectos económicos asociados a la producción, incluyendo los costos de insumos, mano de obra y transporte, así como los ingresos y utilidades obtenidas por los actores involucrados. Este componente permite analizar cómo las variaciones en los márgenes de rentabilidad influyen en las decisiones de vinculación al negocio. Finalmente, el módulo de actores aborda el rol que desempeñan los principales agentes involucrados: GAI [14] y los PAC [2]. Sin embargo, su participación no ocurre de forma aislada. Los GAI ejercen control territorial en múltiples zonas del país, y tienen una fuerte presencia en áreas donde se concentra la actividad agrícola asociada a la coca [15]. Estos grupos no sólo garantizan la protección de los cultivos, sino que también intervienen activamente en la organización y supervisión de la producción, ofreciendo insumos y asistencia técnica a los PAC [2].

Debido al enfoque sistémico de la metodología, estos módulos están interrelacionados. El vínculo entre los actores y el cultivo es bidireccional: porque los PAC son quienes realizan las labores de siembra y cosecha. En muchas regiones, esta población depende casi exclusivamente del cultivo como fuente de ingresos [2]. Además, las políticas de erradicación tanto forzada como voluntaria, influyen sobre la vinculación o desvinculación de los GAI y PAC al sistema. La relación entre el cultivo y la demanda es igualmente directa. La existencia de una demanda a nivel internacional genera presión constante sobre la producción, incentivando la siembra para garantizar el abastecimiento del mercado. Esta demanda, a su vez, está vinculada con el módulo financiero: genera ingresos por la comercialización de la hoja, lo que incrementa la rentabilidad y refuerza la decisión de los actores de mantenerse en el sistema. En consecuencia, las ganancias económicas derivadas del cultivo son un factor clave que explica su persistencia, al constituir uno de los principales incentivos para la participación tanto de los PAC como de los GAI. Esta estructura interdependiente permite entender el cultivo de coca como parte de un sistema complejo, donde las decisiones individuales están influenciadas por múltiples factores económicos, institucionales y sociales que interactúan entre sí. El modelo busca capturar estas relaciones para ofrecer una representación más realista de las dinámicas que sostienen el fenómeno.

D. Validación

Como parte del proceso de validación del modelo, se realizaron varias pruebas para validar la estructura y los parámetros del modelo. Para confirmar la estructura teórica del modelo propuesto, se realizó una revisión de literatura para establecer relaciones causales. Durante la fase de construcción, se revisaron y ajustaron los valores de los parámetros con base en referencias académicas pertinentes. La Tabla 1 resume los principales parámetros utilizados, incluyendo sus valores, unidades de medida y respectivas fuentes de información

TABLA I
PARÁMETROS PRINCIPALES DEL MODELO

Parámetro	Valor	Unidades	Fuente
Tasa de Erradicación Manual Forzada	0,01	Adimensional	Determinado a partir de [2]
Tasa de Erradicación Manual Voluntaria	0,07	Adimensional	Determinado a partir de [2]
Tasa de cosecha	0,2	Adimensional	[7]
Rendimiento hoja de coca	8,5	Ton/Ha	[2]

También se realizaron pruebas de condiciones extremas, en las cuales se sometió al modelo a escenarios límite, como una demanda inexistente. En este caso, el modelo reaccionó de forma coherente con la lógica del sistema: al eliminarse completamente la demanda, se produjo una disminución

progresiva en la rentabilidad del cultivo, lo que desencadenó una reducción en la cantidad de productores y una disminución en los cultivos, ya que desaparecieron los incentivos económicos que sostenían la actividad y su continuidad dejó de ser viable.

E. Escenarios

Se implementaron diferentes escenarios de simulación para explorar posibles trayectorias del sistema ante modificaciones en variables críticas, como un aumento de la demanda y una disminución en la intensidad de las estrategias de erradicación. Este último escenario adquiere especial relevancia en el contexto actual, en el cual el gobierno del presidente Gustavo Petro ha reducido significativamente la aplicación de la Erradicación Manual Forzada como medida de control, mientras que la demanda internacional permanece alta, configurando un entorno particularmente desafiante para el cultivo [16].

III. RESULTADOS

El modelo fue implementado en Powersim Studio 2005, con un horizonte de tiempo 20 años y un paso de tiempo de un año. Los resultados del modelo base indican que el cultivo de coca continuará expandiéndose dentro de los 20 años simulados como se muestra en la figura 3, mientras la rentabilidad se mantenga como principal incentivo para los PAC y no se modifiquen las condiciones estructurales que permiten la operación de los GAI en los territorios. Asimismo, las estrategias de erradicación, si no se complementa con intervenciones que transformen los mecanismos de realimentación asociados al atractivo económico del cultivo, podría tener un impacto limitado o incluso efectos no deseados. El análisis también confirma que la demanda cumple un rol central en la configuración del sistema: mientras persista una presión constante desde los mercados de consumo, los incentivos para la producción seguirán activos, independientemente de los esfuerzos centrados en restringir la oferta.

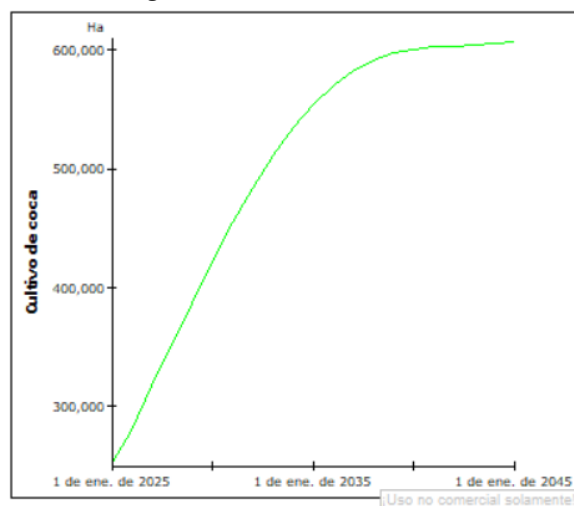


Fig. 3. Cultivo de coca en el tiempo.

En la figura 4 se presenta una comparación entre el escenario base y un escenario alternativo caracterizado por un aumento

significativo en la demanda y una disminución en la intensidad de la Erradicación manual forzada. Los resultados muestran una diferencia marcada entre ambas trayectorias: mientras que en el escenario base el sistema tiende a estabilizar el área cultivada en torno al nivel que satisface la demanda, en el escenario alternativo el cultivo de coca experimenta un crecimiento acelerado, alcanzando cifras elevadas de hectáreas cultivadas. Este comportamiento evidencia cómo los incentivos económicos derivados de un mercado en expansión, combinados con una menor intervención estatal en términos de erradicación, fortalecen los mecanismos de realimentación que favorecen la persistencia y expansión del negocio de cultivo de coca. A mayor rentabilidad y menor estrategias de intervención, se incrementa tanto el número de productores como la producción, consolidando el cultivo como una actividad económicamente para amplios sectores rurales. Por tanto, se vuelve necesario incorporar estrategias que actúen también sobre la demanda, considerando marcos de intervención más amplios, que contemplen tanto las dinámicas territoriales como los factores globales que sostienen el sistema.

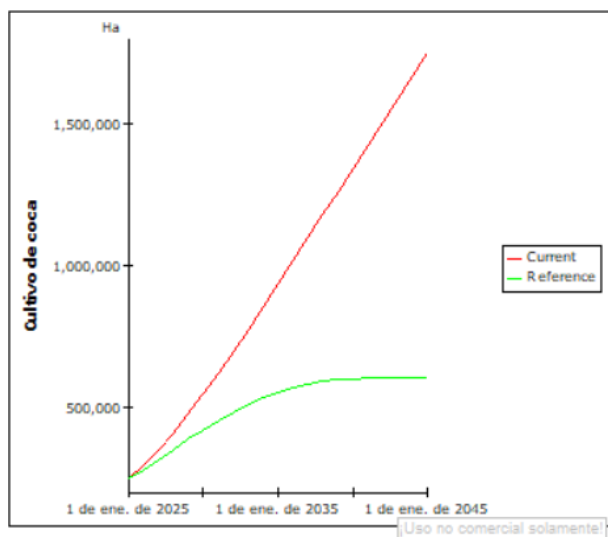


Fig. 4. Comparación de escenarios cultivo de coca.

IV. CONCLUSIÓN

Esta investigación constituye una aproximación al análisis del fenómeno del cultivo de coca en Colombia a partir de la metodología de dinámica de sistemas. La construcción del diagrama causal y el modelo de simulación permitieron identificar los ciclos de refuerzo y balance que interactúan entre sí para explicar la persistencia del cultivo. Los resultados obtenidos permiten comprender de manera articulada los factores que favorecen su expansión, los actores que intervienen en su reproducción y las posibles trayectorias del sistema bajo distintas condiciones. Lejos de ofrecer una visión simplificada, el modelo busca hacer visible la complejidad estructural del problema y proveer una herramienta analítica que contribuya al diseño de estrategias de intervención más coherentes con la naturaleza del sistema.

En este sentido, los enfoques orientados exclusivamente al control de la producción han mostrado resultados insuficientes,

ya que no logran alterar los fundamentos económicos y sociales que sostienen el cultivo, ni desarticular las redes de actores que lo reproducen. La persistencia del fenómeno evidencia que es necesario superar las intervenciones fragmentadas y avanzar hacia estrategias integrales que combinen control, desarrollo rural, fortalecimiento institucional y transformación de las condiciones que hacen del cultivo de coca una alternativa viable para muchas comunidades. Como línea futura de trabajo, se plantea la ampliación del modelo mediante la incorporación de nuevas variables relevantes con información territorial desagregada y su utilización como insumo para procesos participativos de construcción de política pública, en coordinación con comunidades locales, instituciones gubernamentales y organizaciones sociales. Asimismo, es importante considerar que el modelo actual no incorpora explícitamente restricciones físicas, como la disponibilidad máxima de tierra. Por ello, una mejora futura será la inclusión de parámetros que representen la saturación de recursos y los límites productivos reales, con el fin de reflejar de forma más precisa la dinámica propia de los sistemas agrícolas.

REFERENCIAS

- [1] D. Mejía y P. Restrepo, "The economics of the war on illegal drug production and trafficking", *Journal Of Economic Behavior & Organization*, vol. 126, pp. 255-275, abr. 2016, doi: 10.1016/j.jebo.2015.11.003.
- [2] UNODC-SIMCI, *Monitoreo de territorios afectados por cultivos ilícitos 2022*. Bogotá: Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC) - Sistema Integrado de Monitoreo de Cultivos Ilícitos, 2023.
- [3] F. I. Bastos, W. Caiaffa, D. Rossi, M. Vila, y M. Malta, "The children of mama coca: Coca, cocaine and the fate of harm reduction in South America", *International Journal Of Drug Policy*, vol. 18, n.o 2, pp. 99-106, ene. 2007, doi: 10.1016/j.drugpo.2006.11.017.
- [4] J. P. Cote, "The Effect of Interdiction on Coca Cultivation in Colombia," *SSRN Electronic Journal*, 2019. doi: 10.2139/ssrn.3449268.
- [5] UNODC-SIMCI, *Monitoreo de territorios con presencia de cultivos de coca 2023*. Bogotá: Oficina de las Naciones Unidas contra la Droga y el Delito (UNODC) - Sistema Integrado de Monitoreo de Cultivos Ilícitos, 2024.
- [6] S. Jaén and I. Dyner, "A system dynamics approach to the study of Colombian coca cultivation and the counter-intuitive consequence of law enforcement," *International Journal of Drug Policy*, vol. 25, no. 2, pp. 226-234, 2014. doi: 10.1016/j.drugpo.2014.01.010.
- [7] A. Córdoba, *Modelamiento de la relación entre la población humana y el recurso natural Erythroxylum Coca*, 2015.
- [8] D. C. Barrios Rodríguez, C. E. Olaya Nieto, and I. d. J. Beltrán Pacheco, *Innovación estratégica para el control del transporte de coca en Colombia desde la perspectiva de dinámica de sistemas*, Bachelor's thesis, Universidad de los Andes, 2011.
- [9] C. Olaya, J. Gómez-Quintero, A. M. Rueda, and P. Méndez, *Control de cultivos de Coca. Una perspectiva operacional de ingeniería*. Bogotá: Centro de Estudios sobre Seguridad y Drogas (CESED), 2021.
- [10] O. Dube and S. Naidu, "Bases, Bullets, and Ballots: The Effect of US Military Aid on Political Conflict in Colombia," *The Journal of Politics*, vol. 77, no. 1, pp. 249-267, ene. 2015. doi: 10.1086/679021.
- [11] J. D. Angrist and A. D. Kugler, *Rural windfall or a new resource curse? Coca, income, and civil conflict in Colombia*.
- [12] S. Bayer, *Systems Thinking and Modeling for a Complex World*, vol. 34. Chicago, IL, USA: Irwin/McGraw-Hill, 2004. ISBN: 007238915X.
- [13] J. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World*. Chicago, IL, USA: McGraw-Hill, 2000.
- [14] D. Mejía and D. Rico, *La microeconomía de la producción y tráfico de cocaína en Colombia*. Bogotá, Colombia: Universidad de los Andes, Facultad de Economía, CEDE, 2010.
- [15] F. Londoño Niño, *¿A dónde fue la cocaína?: oferta y demanda de cocaína*.
- [16] Euronews, «Colombia | Gustavo Petro suspende la erradicación forzosa de los cultivos de coca», *Euronews*, 24 de agosto de 2022. [En línea]. Disponible en: <https://es.euronews.com/2022/08/24/colombia-gustavo-petro-suspende-la-erradicacion-forzosa-de-los-cultivos-de-coca>

Dynamic Modeling of the Natural Gas Market in Colombia in the Framework of a Sustainable Energy Transition

Derlyn Franco Castaño, Juan Carlos Osorio Gómez, and Diego Fernando Manotas Duque

Abstract—This research proposal aims to design a dynamic model of the supply and demand of natural gas in Colombia, considering socioeconomic variables such as GDP, urbanization rate and behavior of new energies consumption, and determine the variables that affect them within the framework of a sustainable energy transition. First the demand model will be presented, then the supply model and finally the integration of both models under twelve scenarios. It is concluded that whatever the scenario, it is essential to think about strategies that provide a rapid response within the framework of the energy transition because the available natural gas is not enough to meet Colombia's internal demand from 2027.

Index Terms— Energy Transition, Natural Gas, System Dynamics.

I. INTRODUCTION

Colombia, like many countries in Latin America and the Global South, faces a multitude of crises and political challenges, including social, ecological, economic, public finance, security, and internal law and order issues. Climate change, with its impacts and adaptation and mitigation needs, adds to these problems. All these crises, which are common to many countries in the Global South, converge in what is perhaps the greatest global challenge in human history: the energy transition as a response to the climate crisis resulting from the historical project of modernity/coloniality [1].

Colombia, participate to the 2015 Paris Agreement and subsequent Conferences of the Parties (COP), committed to reducing 20% of its Greenhouse Gas (GHG) emissions by 2030. However, this is not the only reason to undertake an energy transition, beyond complying with international climate agreements, reducing significant emissions or reducing adverse effects despite the environmental and social impacts of fossil fuel extraction, an energy transition is linked, among other things, to the fact that Colombia in the near future take the risk of becoming an importer of fossil fuels and losing export revenues and losing competitiveness of its industry if it does not decarbonize [2].

In the energy transition process, Colombia has been promoting natural gas production and its use in electricity generation as cleaner sources. The consumption of energy currently comes from oil, coal, and, increasingly, natural gas. These hydrocarbons supply electricity, heat, and cooling to homes and workplaces, and fuel transportation systems [3].

At the end of 2023, the Agencia Nacional de Hidrocarburos (ANH) in Colombia reported that the Proven Reserves/Production (R/P) ratio of natural gas is 6.1 years, indicating that hydrocarbon reserves could be depleted by 2030 if no changes are made in the exploration of new reserves.

System Dynamics will be a strategy that will help to better understand future changes, due to the different factors involved in the behavior of natural gas supply and consumption and its variability. The System Dynamics approach will allow to explain endogenously how the system evolves over time through cause-and-effect feedback [4], [5] and its systemic perspective is particularly relevant when considering alternative future scenarios.

In this context, this research seeks to model the supply and demand of natural gas in Colombia under social and economic aspects within the framework of an energy transition policy, reviewing scenarios with the introduction of new clean energy sources to replace natural gas and the introduction of natural gas from new wells.

II. NATURAL GAS DEMAND MODEL IN COLOMBIA

Natural gas plays a crucial role in Colombia, reaching over 11 million users across industries, businesses, and homes [6]. Natural gas consumption in Colombia is classified into six sectors: Industrial, Thermoelectric, Residential, Oil, Vehicle and Petrochemical.

Eighteen (18) variables were defined that are considered determinants in the dynamics of the natural gas demand in Colombia, the influence diagram is shown in Fig.1.

This work was supported by the Universidad del Valle under Internal Grant.

Derlyn Franco Castaño is student of Master in the Escuela de Ingeniería Industrial, Universidad del Valle, Cali 760000, Colombia (e-mail: Derlyn.franco@correounivalle.edu.co).

Juan Carlos Osorio Gomez is researcher in the Logística y Producción research group, and professor of the Escuela de Ingeniería Industrial, Universidad del Valle, Cali 760000, Colombia (e-mail: juan.osorio@correounivalle.edu.co).

Diego Fernando Manotas Duque is researcher in the Finanzas Cuantitativas research group, and professor of the Escuela de Ingeniería Industrial, Universidad del Valle, Cali 760000, Colombia (e-mail: diego.mannotas@correounivalle.edu.co).

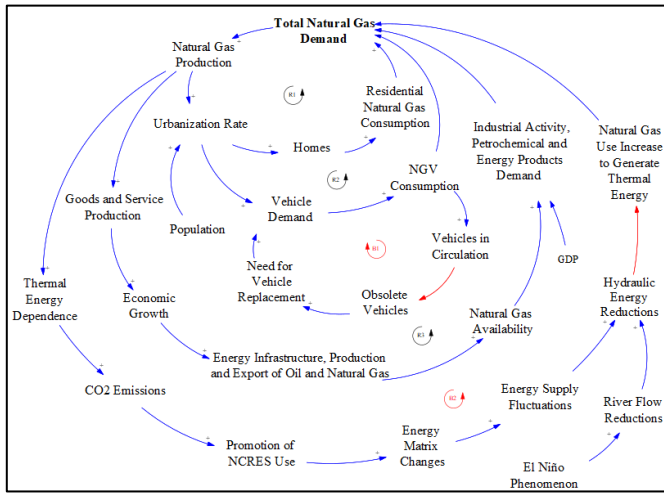


Fig. 1. Natural Gas Demand Influence Diagram in Colombia.

The causal diagram above consists of 3 reinforcing loops and 2 balancing loops, which are detailed below:

Reinforcing Loop Residential Sector (R1): a larger Population, higher Urbanization Rate (percentage of habitants who reside only in the municipal capitals), and the number of households will increase, which generates greater Natural Gas Consumption in the Residential Sector, this generates an increasing Demand for Natural Gas and as a consequence more Natural Gas must be Produced, which in turn can promote more people to move to the city due to the availability of this resource, driving a growth in the Urbanization Rate.

Reinforcing Loop in the Vehicle Sector (R2): As the population grows, the Urbanization Rate increases, generating greater demand for vehicles. The greater the demand for vehicles, the greater the number of vehicles that consume natural gas. This generates an increase in the natural gas demand, and consequently, more natural gas must be produced.

Reinforcing Loop Oil, Industrial and Petrochemical Sector (R3): GDP growth generates an increase in industrial activity, demand for petrochemical and energy products, this causes the natural gas demand to be greater, which requires greater production of natural gas, by having greater availability of natural gas as fuel, there is more fuel to increase the production of goods and services, which generates economic growth, and economic growth means that there is greater energy infrastructure, production and export of oil and natural gas, which generates greater availability of natural gas so that with the increase in GDP a greater increase in industrial activity, oil and petrochemicals is generated.

For this research, the industrial and petrochemical sectors consumption has been taken the percentage of growth projected by [7]. For the oil sector, we have taken the information provided by Ecopetrol on its prospective consumption of natural gas associated with the production of crude oil, natural gas and refined products in the country.

Balancing Loop Vehicle Sector (B1): Greater Vehicle Demand generates greater Natural Gas Vehicle NGV

Consumption by having more Vehicles in Circulation. The more Vehicles in Circulation, the more Vehicles become Obsolete, where the Need for Replacement will be greater, which causes greater Demand for Vehicles. This model was taken from [8].

Balancing Loop Thermoelectric Sector (B2): When the El Niño Phenomenon occurs, river flows decrease, which causes that, by having less water, the Capacity to Generate Hydraulic Energy will be lower, by having less Hydraulic Energy Generation, Thermal Energy must be generated with Natural Gas to compensate for the Electric Energy Internal Demand, this causes an increase in the Natural Gas Demand, which requires greater Production of Natural Gas, by having greater Availability of Natural Gas, the Dependence on Thermal Energy increases, which generates greater CO2 Emissions to the environment, this encourages the use of other Non-Conventional Renewable Energy Sources NCRES such as wind and solar, among others. This leads to changes in the Energy Matrix and fluctuations in energy supply, which, combined with the El Niño phenomenon, causes a decrease in hydropower production.

Once the main variables of the system and their interrelationships, the Forrester Diagram presented in Fig. 2 was constructed.

III. NATURAL GAS SUPPLY MODEL IN COLOMBIA

The natural gas supply is determined by available reserves and the capacity of the logistics associated with them, so it is important to define this. Reserves are those quantities of hydrocarbons anticipated to be commercially recoverable through the implementation of development projects in known accumulations, from a given date onward, under defined conditions.

For the natural gas supply model, twelve (12) variables were defined that are considered determinants in the dynamics of the Natural Gas Supply in Colombia, as shown in Fig. 3:

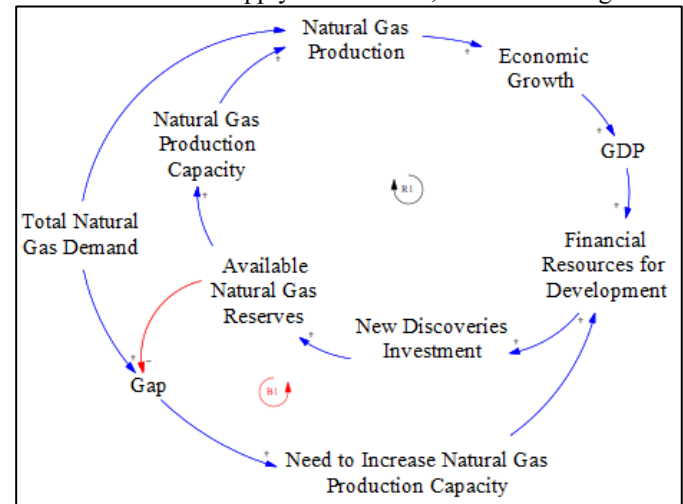


Fig. 3. Supply of Natural Gas Causal Diagram

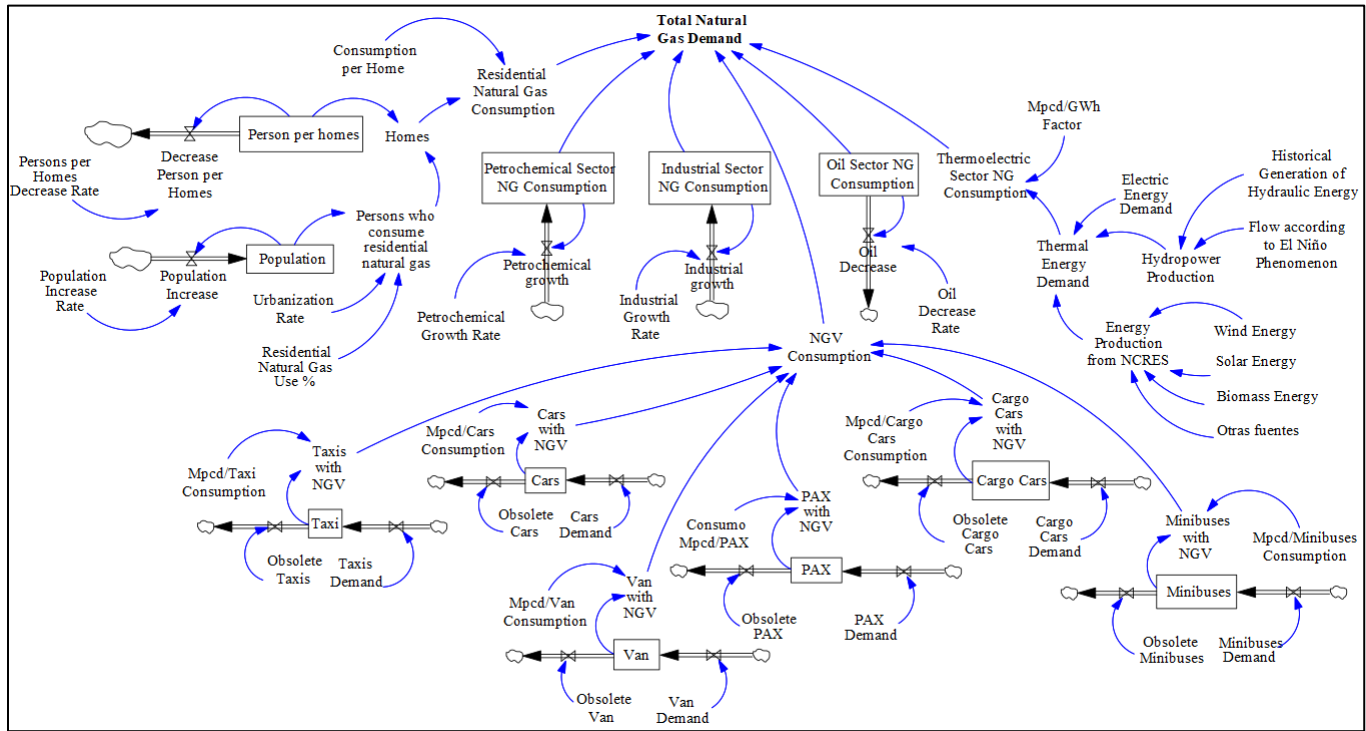


Fig. 2. Natural Gas Demand Forrester Diagram in Colombia.

The causal diagram in Fig. 3 consists of 1 Reinforcing loop and 1 Balancing loop, which are detailed below:

Reinforcing Loop Natural Gas Production (R1): As natural gas demand grows, increased natural gas production is required, which represents greater economic growth and a greater contribution to GDP growth. With higher GDP, more financial resources are available for investment in new discoveries and production expansion, which generates growth in available natural gas reserves. With more available natural gas reserves, there is a greater need to expand natural gas production capacity to achieve higher natural gas production.

Balancing Loop Demand/Availability Gap (B1): The increase in Natural Gas Demand increases the Gap between demand and availability of Natural Gas as demand grows, but availability continues to decrease, as a greater Gap occurs, the Need to increase Natural Gas Production Capacity grows, which requires an increase in Financial Resources for the Development and production of natural gas. Having more Financial Resources implies an increase in Investment in New Discoveries of Natural Gas fields. Investing more in New Discoveries generates greater Natural Gas Reserves to be used and by having more Available Reserves, a decrease in the Gap between demand and availability of Natural Gas is generated.

It is important to highlight that the Produced Natural Gas (PNG) corresponds to the sum of the Extracted Natural Gas (ENG) plus the Natural Gas Inventory (NGI) less the Total Demand (TD):

$$PNG = ENG + NGI - TD. \quad (1)$$

After defining the main variables of the system and their interrelations, the Forrester diagram is presented in Fig. 4.

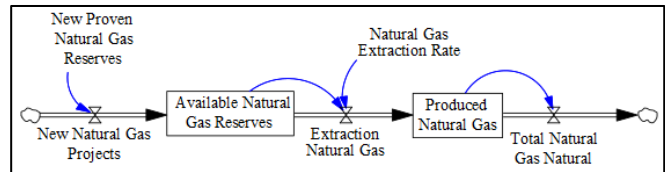


Fig. 4. Natural Gas Supply Forrester Diagram in Colombia.

IV. INTEGRATION OF DEMAND AND SUPPLY NATURAL GAS MODELS IN COLOMBIA

The demand and supply natural gas models have been integrated and the projected results are shown in Fig. 5:

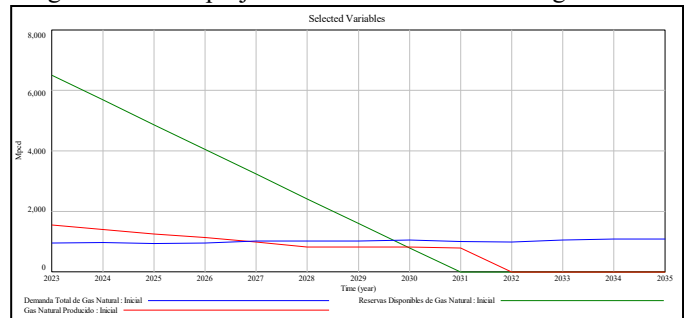


Fig. 5. Total Natural Gas Demand, Available Natural Gas Reserves and Produced Natural Gas Projection 2023-2035

Fig. 5. in the second semester of 2026 there is a crossover between the Produced Natural Gas (red line) and the Total Natural Gas Demand (blue line), in other words until 2026 the extracted natural gas meets the projected consumption, from that year onwards a deficit appears because consumption exceeds the available natural gas, this situation suggests thinking about different strategies to meet demand. By the second semester of 2030 the Available Natural Gas Reserves

(green line) reach zero, that is to say, that year the internal natural gas would be completely exhausted. Produced Natural Gas (red line) reaches zero by the second semester of 2031, meaning that even though Natural Gas Reserves are expected to run out by 2030, there will be enough inventory to cover part of Total Demand through 2031. In conclusion, starting in 2027, other alternatives must be considered to meet projected demand.

Below, different scenarios are defined to analyze how the initial situation changes in each of them:

- Scenario 1: Include hydrogen demand projections in the natural gas consumption model to cover the natural gas shortage.
- Scenario 2: Include wind energy growth projections in the natural gas consumption model, which will help cover natural gas requirements in the thermoelectric sector and leave more natural gas to cover demand in other sectors.
- Scenario 3: Include solar energy growth projections in the natural gas consumption model, which will help cover natural gas requirements in the thermoelectric sector and leave more natural gas to cover demand in other sectors.
- Scenario 4: Include new natural gas reserves that are being announced to cover domestic natural gas demand.
- Scenario 5: Integrate scenarios 1 and 2.
- Scenario 6: Integrate scenarios 1 and 3.
- Scenario 7: Integrate scenarios 1 and 4.
- Scenario 8: Integrate scenarios 2 and 3.
- Scenario 9: Integrate scenarios 2 and 4.
- Scenario 10: Integrate scenarios 3 and 4.
- Scenario 11: Integrate scenarios 1, 2, and 3.
- Scenario 12: Integrate scenarios 4 and 11.

The results presented below consolidate the analysis of each of the scenarios:

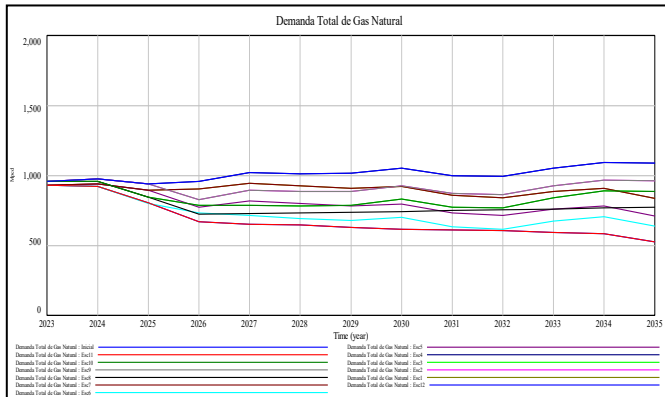


Fig. 6. Comparison of Total Natural Gas Demand in the reviewed scenarios.

Fig. 6 Total Natural Gas Demand for presents the greatest reduction in scenario 11 (red line), corresponding to the introduction of Hydrogen as a replacement for natural gas and the increase in the use of wind and solar energy, here there is an increase in the use of NCRES and, consequently, a replacement of the use of natural gas in the thermoelectric sector, which is reflected in the decrease in the Total Natural Gas Demand. Scenarios 6, 8 and 10 are the next respectively that present reduced Total Natural Gas Demand, it is worth noting that they

are precisely the ones that contain the use of capacity expansion in solar energy, which has high growth projections and represents a great contribution to the energy transition in Colombia.

It is important to clarify that the reduction in Total Natural Gas Demand is mainly reflected by the change in the thermoelectric sector, also in the vehicle and industrial sectors in lesser proportion, the remaining natural gas consumption sectors are not altered, in this way it is shown in Fig. 7 how natural gas consumption varies in each of the reviewed scenarios in the thermoelectric sector. In line with Fig. 6, Fig. 7 shows the significant reduction in natural gas use in the thermoelectric sector in scenario 11, where greater use is made of NCRES such as Hydrogen, wind and solar energy.

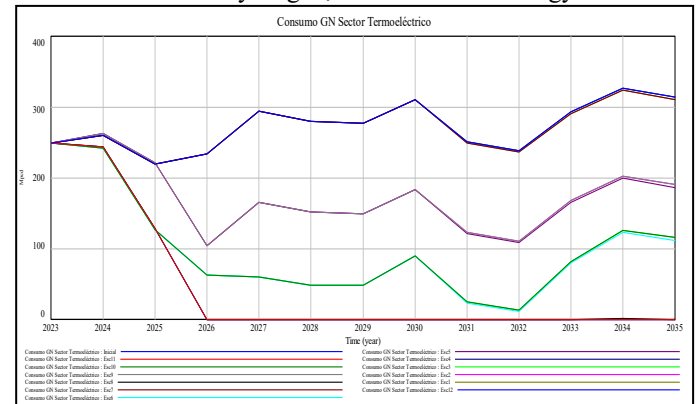


Fig. 7. Comparison of Natural Gas Demand in the Thermoelectric Sector in the Proposed Scenarios

Fig. 8 shows that scenario 12, the highest value of Produced Natural Gas is balanced, where the downward trend changes to upward. The decrease in natural gas use in the thermoelectric sector is reflected in the increase in Produced Natural Gas. This occurs because Total Natural Gas Demand decreases and the extraction rate remains constant, generating a natural gas inventory that accumulates, obtaining greater natural gas availability.

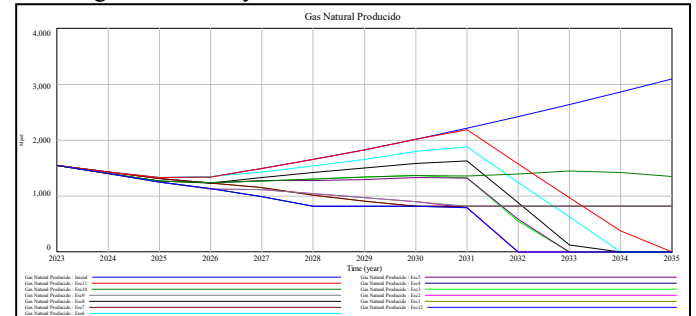


Fig. 8. Comparison of Produced Natural Gas in the different scenarios proposed.

Regarding the Available Natural Gas Reserves, Fig. 9 shows that with a constant extraction rate of 816.75 MMcf/d of natural gas, the reserve will be available until the second semester of 2030 according to the initial, 1, 2, 3, 5, 6, 8, and 11 scenarios. The 4, 7, 9, 10, and 12 Scenarios, which consider the availability of a new well, show a significant increase in 2027.

Regarding the availability of natural gas, the behavior for each scenario can be observed in Table 1, with the constant extraction rate of 816.75 MMcf/d.

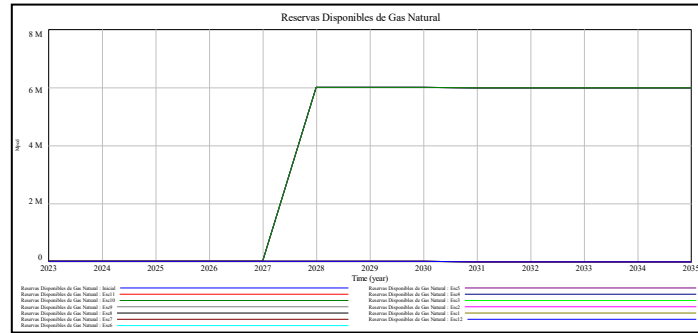


Fig. 9. Comparison of Available Natural Gas Reserves in the different scenarios proposed.

TABLE 1
COMPARISON OF VARIABLES FOR EACH SCENARIO

Concept	Maximum date for 100% coverage of Natural Gas Demand	Reserve depletion date	Produced Natural Gas depletion date
Initial	2026-II	2030-II	2031-I
Scenario 1	2029-I	2030-II	2031-II
Scenario 2	2029-II	2030-II	2031-II
Scenario 3	2031-II	2030-II	2032-II
Scenario 4	2026-II	-	-
Scenario 5	2031-II	2030-II	2032-II
Scenario 6	2032-II	2030-II	2033-II
Scenario 7	2029-I	-	2029-I
Scenario 8	2032-I	2030-II	2033-II
Scenario 9	2029-II	-	-
Scenario 10	-	-	-
Scenario 11	2033-II	2030-II	2034-II
Scenario 12	-	-	-

According to Table 1, in 10 and 12 scenarios, between 2023 and 2035, Produced Natural Gas always meets projected demand. These two scenarios coincide in that they consider the arrival of natural gas from the new well and the expansion of solar energy capacity projected by UPME, representing ideal scenarios under the country's current conditions. It is important to highlight that 4, 7, and 9 scenarios, which consider the arrival of natural gas from the new well, present problems covering demand, not due to availability, but due to the extraction rate, which has been constant since before in this model. This can be solved by increasing that rate and thus meeting projected consumption.

Without considering the arrival of natural gas from a new well, scenario 11 turns out to be the most convenient since it covers 100% of the demand until the second semester of 2033, despite the fact that Natural Gas Reserves are depleted in the second semester of 2030, and this may occur because the projection of natural gas consumption becomes increasingly

lower because in the thermoelectric sector the consumption of natural gas stops due to having greater capacity in the generation of wind and solar energy, and by using Hydrogen, and present a decrease in consumption in the industrial and vehicular sectors due to the use of Hydrogen consumption projections as a replacement for natural gas. By maintaining the extraction rate, and as its use decreases, natural gas reserves are accumulated, presenting a prolonged duration, which allows meeting the projected demand for a longer time. It is also important to note that a combination of expanding wind power generation capacity with the use of hydrogen as a replacement for natural gas presents the same conditions as in scenario 3, covering 100% of the projected demand until second semester 2031, that is, 5 years more than in the initial model.

Respect comparisons between the scenarios, it can be concluded that the one that brings the greatest benefit to the country would be scenario 10, where solar energy generation capacity grows and the arrival of natural gas from the new well in 2027. Scenario 12 is also beneficial, in addition to what is included in scenario 10, which includes the growth in wind energy generation capacity and the projections of hydrogen use as a replacement for natural gas. Although the use of other alternative sources to natural gas in the different consumption sectors helps extend the time to continue meeting 100% of the projected demand, it does not solve the problem of the availability of natural gas reserves, which suggests that the best thing that can happen to the country is the incorporation of the new Sirius 2 well by 2027.

REFERENCES

- [1]. Escobar, A. Reinterpretando las civilizaciones: De la crítica a las transiciones. ARQ (Santiago), 111, 24–41. <https://doi.org/10.4067/S0717-69962022000200024> (2022).
- [2]. Thema, J., & Roa García, M. C. The energy transition in Colombia: current situation, projections, challenges, narratives and public policies in relation to the energy transition in Germany (Vol. 200). Wuppertal Institut für Klima, Umwelt, Energie. (2023).
- [3]. Shell Homepage, Transición Energética en Colombia y el mundo. Transición Energética. <https://www.shell.com.co/sostenibilidad/transicion-energetica.html#:~:text=Colombia%2C%20en%201%C3%ADnea%20con%20esta,vez%20m%C3%A1s%2C%20del%20gas%20natural>, last accessed 2023/11/21.
- [4]. Sterman, JD. Business dynamics systems thinking and modeling for a complex world. first ed. USA: McGraw-Hill. (2000).
- [5]. Forrester JW. Industrial dynamics. J Oper Res Soc; 48:1037e41. <https://doi.org/10.1057/palgrave.jors.2600946> (1997).
- [6]. Ministerio de Minas y Energía Homepage, Cobertura del servicio de gas natural, segundo trimestre 2023. <https://naturgas.com.co/wp-content/uploads/2023/11/INFORME-COMPLETO-2.pdf>, last accessed 2023/11/21
- [7]. UPME Homepage. Proyección de la Demanda de Energía Eléctrica y Potencia Máxima 2023-2037. https://www1.upme.gov.co/DemandayEficiencia/Documents/UPME_Pr_oyeccion_demanda_2023-2037_VF2.pdf, last accessed 2023/11/21.
- [8]. Martínez-Escobar, N., Osorio-Gómez, J. C., & Sierra-Castillo, J. A. Dynamic Evaluation of Environmental Effects Associated with the Inclusion of Hydrogen-powered Vehicles in Colombia. International Journal of Energy Economics and Policy, 14(3), 339–348. <https://doi.org/10.32479/ijeep.15694>. (2024).

Identificación de Técnicas de Analítica para la mejora de la precisión en modelos de Simulación, una revisión crítica de literatura

Alejandro Álvarez. Yony Fernando Ceballos.
Departamento de Ingeniería Industrial
Universidad de Antioquia

Abstract— La simulación discreta es una herramienta que, en la actualidad es ampliamente utilizada para analizar y representar sistemas complejos de una forma clara y precisa a través de modelos, los cuales dependen en gran medida de utilizar datos de entrada precisos para obtener un buen desempeño. En los últimos años, algunas técnicas de analítica de datos, como análisis estadístico, modelos predictivos, entre otros, han sido utilizadas como alternativas para mejorar el desempeño de los parámetros mencionados para los modelos de simulación. Con base en lo anterior, en este artículo se desarrollará una revisión sistemática de literatura con el objetivo de identificar y analizar cómo algunas técnicas de analítica de datos han sido aplicadas con el fin de mejorar la estimación de parámetros de entrada, y si éstas han tenido un impacto positivo en la precisión y calidad de los modelos de simulación. A partir de un análisis estructurado de artículos e investigaciones desarrolladas en la última década, se identifican enfoques, herramientas y tendencias en la integración de métodos analíticos con la simulación. Los resultados permiten establecer un panorama general sobre las técnicas más empleadas, su efectividad y los principales desafíos asociados en este tema, ofreciendo una base para futuras investigaciones relacionadas.

Index Terms— Analítica de datos, Datos de entrada, Estimación de parámetros, Modelado, Precisión, Simulación discreta

I. INTRODUCCIÓN

A analítica de datos ha sido una herramienta que en los últimos años ha generado grandes avances en términos tecnológicos, sociales, económicos y empresariales [10], [3], [13]. En este último sentido se identifica que al utilizar diferentes técnicas de representación de entornos al interior de las empresas es posible integrar procesos de análisis de datos o de evaluación de resultados asociados a los modelos que se emplean como base en la toma de decisiones [4], [14]. Estos modelos pueden ser de diversa índole, lo cual presenta un reto para los decisores ya que dicha forma de evaluación debe ser ágil, oportuna y clara. Por tal razón, los datos que se emplean en tales modelos deben cumplir con criterios de oportunidad y utilidad.

La simulación aparece junto con otras herramientas de toma de decisiones al final de la década de los años cincuenta y con el

auge de los computadores personales toma gran fuerza siendo comúnmente utilizada en las empresas a partir de la década del 2000. En la actualidad esta herramienta se debe integrar con tecnologías emergentes en analítica que permitan que sea flexible y rápida al momento de ser empleada y que permita unas decisiones fundamentadas en un análisis estadístico cuidadoso [11], [16].

Por tal razón, el análisis de datos se vuelve una necesidad en estas herramientas de toma de decisiones. Este convierte datos sin procesar en información útil, e incluye diferentes herramientas, tecnologías y procesos para encontrar comportamientos. El análisis de datos también ayuda a las empresas a tener un conocimiento más profundo de sus procesos y servicios [7]. Al cambiar el paradigma más allá de los datos para conectar los conocimientos con la acción, las empresas pueden optimizar las operaciones y aumentar la productividad de los empleados.

Con base en lo anterior, en este trabajo de investigación se realizará un análisis exploratorio (revisión de literatura) en términos de revisión bibliográfica acerca de cuál es el estado de la integración de diversas técnicas de analítica con software de simulación en la interpretación y utilización de datos de entrada.

Los modelos empleados tradicionalmente para hacer análisis del comportamiento de la empresa se fundamentan en la simulación [5], [20]. La simulación es una herramienta muy útil en la industria, ya que permite modelar y analizar el comportamiento de sistemas complejos en un entorno controlado y seguro. La simulación se utiliza en muchos campos de la industria, desde la manufactura hasta la logística y la planificación de la producción. Esta herramienta necesita una gran cantidad de datos para lograr unos ajustes satisfactorios y después de un proceso de ejecución genera una cantidad significativa de información que debe analizarse de manera rápida y objetiva para lograr realizar cambios al interior de la organización que impacten de manera inmediata y puedan generar valor agregado [19].

La analítica de datos se ha convertido en una herramienta

fundamental para las organizaciones que buscan mejorar sus procesos y tomar decisiones más informadas. En particular, en el ámbito de la simulación discreta, la analítica de datos se utiliza para analizar los datos de entrada y salida de los modelos de simulación. La simulación discreta es una técnica que se utiliza para modelar sistemas complejos, donde los eventos ocurren en momentos discretos en el tiempo. En estos modelos, los datos de entrada se utilizan para simular el comportamiento del sistema, mientras que los datos de salida proporcionan información sobre cómo el sistema responde a diferentes situaciones.

La identificación de patrones y tendencias que puedan ayudar a mejorar el rendimiento del sistema, para la evaluación de cuellos de botella, la optimización de los recursos utilizados o la identificación de oportunidades de mejora en el proceso de simulación.

En la presente investigación se pretende responder a la pregunta: ¿Qué técnicas de analítica de datos son más efectivas para mejorar la precisión de los modelos de simulación discreta en líneas de servicio al aplicarlas en los datos de entrada, según la literatura científica actual? Como medio para realizar esta investigación se realizará un proceso de revisión de literatura en la cual se identifiquen las formas aplicadas recientemente de los diferentes métodos de analítica de datos en modelos de simulación.

II. OBJETIVOS

Objetivo General:

Realizar una revisión de la literatura sobre la aplicación de la analítica de datos en modelos de simulación discreta, realizando un comparativo desde los procesos necesarios para realizar una simulación discreta.

Objetivos específicos:

- Diseñar estrategia de búsqueda a través de operadores (*and, or, not, etc*) sobre las técnicas de analítica de datos utilizadas en la preparación y limpieza de datos de entrada para modelos de simulación discreta.
- Seleccionar estudios relevantes en las bases de datos especializadas (*Scopus, Science Direct, Emerald*) para identificar las principales oportunidades y limitaciones asociadas con la aplicación de la analítica de datos para modelos de simulación discreta.
- Identificar si hay software especializado que integre elementos de analítica de datos en modelos de simulación discreta.
- Sintetizar los hallazgos recopilados sobre la búsqueda realizada en el informe final de investigación.

III. MARCO TEÓRICO

A. Analítica de datos

En la era digital, la analítica se ha convertido en un aspecto primordial para la toma de decisiones en las empresas. Su capacidad para procesar grandes volúmenes de datos, identificar patrones y generar información accionable ha

transformado la manera en que las organizaciones operan y compiten en el mercado [3]. Desde su origen como un conjunto de técnicas estadísticas hasta su integración con tecnologías avanzadas como la inteligencia artificial y el aprendizaje automático, la analítica de datos ha evolucionado significativamente, impactando sectores como la manufactura, la salud, el comercio y los servicios financieros [2].

El uso de la analítica de datos en un ámbito empresarial ha logrado hazaas que hace años eran impensables: optimización de procesos, reducción de tiempos y costos, y mejoras en la experiencia del cliente, siendo útil en diferentes sectores. Por ejemplo, en la industria del retail, los algoritmos de analítica permiten identificar patrones de compra y ajustar los inventarios en tiempo real, mientras que, en el sector financiero, se utilizan modelos predictivos para la detección de fraudes y la evaluación de riesgos crediticios [8].

El desarrollo de la analítica de datos ha seguido una evolución progresiva, pasando de modelos descriptivos y estadísticos básicos a técnicas avanzadas de predicción y prescripción. Inicialmente, las organizaciones utilizaban herramientas simples de análisis basadas en hojas de cálculo y reportes tradicionales. Sin embargo, con el avance de la tecnología y la llegada del big data, surgieron nuevas metodologías que permiten analizar grandes volúmenes de información en tiempo real [4].

La relevancia de la analítica de datos radica en su capacidad para transformar datos sin procesar en conocimiento útil, permitiendo a las empresas anticiparse a cambios en el mercado y mejorar su competitividad. En un entorno donde la cantidad de información crece exponencialmente, las organizaciones que no implementan estrategias de análisis de datos corren el riesgo de quedar rezagadas frente a sus competidores [1].

B. Simulación

La simulación, de forma general, es una técnica numérica para realizar experimentos en una computadora. Estos experimentos involucran ciertos tipos de modelo matemático y lógico que describen el comportamiento de sistemas de negocios, económicos, sociales, biológicos, físicos o químicos a través de largos periodos de tiempo [15].

Podemos encontrar varios tipos de software de simulación, agrupados en cuatro categorías [14]:

- **Software de simulación de circuitos:** permite verificar diseños electrónicos sin necesidad de implementarlos físicamente.
- **Software para simulación de sistemas:** simula situaciones de la vida real con el objetivo de probar el funcionamiento de sistemas.
- **Software para análisis de riesgo y simulación:** pronostica resultados de escenarios peligrosos para desarrollar planes de contingencia.
- **Software de simulación acústica:** analiza la calidad y rendimiento del sonido.

Este trabajo se enfoca en la revisión de literatura que utiliza software de simulación para sistemas, con el objetivo de analizar patrones que permitan identificar y proyectar soluciones a eventos reales. Esto se logra a partir de los datos generados por el software de simulación, en conjunto con herramientas de analítica y big data, las cuales permiten

procesar grandes volúmenes de información y obtener resultados simulados más precisos y de alta calidad.

El big data y la analítica de datos se pueden definir como el crecimiento exponencial en la generación de grandes volúmenes de información provenientes de diversas fuentes, y la necesidad de su captura, almacenamiento y análisis para maximizar su valor en la toma de decisiones organizacionales. Según Aguilar [9], el big data representa una nueva generación de tecnologías, arquitecturas y estrategias diseñadas para capturar y analizar grandes volúmenes de datos heterogéneos a alta velocidad, con el fin de generar valor económico y operativo.

Si bien es posible realizar simulaciones infinitas con datos obtenidos a partir de variables parametrizadas, estas deben llevarse a procesos confiables que generen resultados veraces y aplicables en contextos reales. En particular, en el ámbito de las líneas de servicio (presenciales o virtuales), la simulación permite caracterizar y analizar datos de entrada para mejorar la atención tanto a personas como a productos.

Cuando un sistema de servicio está en funcionamiento, es posible realizar un análisis inicial basado en datos reales, identificar patrones de comportamiento a lo largo del tiempo y, mediante procesos de simulación, proyectar nuevos escenarios que optimicen la calidad del servicio. Así, se pueden desarrollar estrategias más eficientes, mejorando los resultados actuales y anticipando posibles desafíos operacionales.

Para lograr estos objetivos, se pueden emplear diversas técnicas de analítica de datos que faciliten la interpretación de la información y la toma de decisiones basada en simulaciones.

Algunas de estas técnicas incluyen:

- **Análisis descriptivo:** Comprensión del comportamiento de los datos históricos y su distribución.
- **Modelos predictivos:** Algoritmos de aprendizaje automático y estadística para anticipar tendencias.
- **Clustering o segmentación:** Agrupa datos similares para identificar patrones.
- **Análisis de series temporales:** Evalúa la evolución de los datos en el tiempo.
- **Optimización matemática:** Toma de decisiones mediante modelos que maximizan la eficiencia de los recursos.

El uso de estas técnicas, en conjunto con la simulación y el big data, permite generar análisis precisos y decisiones fundamentadas para mejorar los sistemas de servicio y otros ámbitos de aplicación.

IV. METODOLOGÍA

Teniendo en cuenta que el proyecto se basa en una revisión de literatura y en analizar técnicas de simulación y analítica, se empleará un enfoque cualitativo-exploratorio.

Lo principal es el tema de la recolección de información, donde para garantizar una cobertura amplia y confiable se proponen los siguientes pasos:

- Se consultarán y buscarán artículos en bases de datos científicas como Scopus, ScienceDirect, Emerald, entre otras, para obtener información relevante.
- Para incluir una publicación en nuestro pull de información se buscarán principalmente artículos o publicaciones que estén publicadas en revistas indexadas en los últimos 10 años (Debido a la evolución de la información). Se buscarán publicaciones que traten sobre el uso de software de simulación en sistemas y/o que aborden la aplicación de técnicas de analítica y Big Data en simulaciones. En función de los hallazgos puede redefinirse esta restricción.

Teniendo una recolección significativa de publicaciones que cumplan con los estándares mencionados previamente, se procede a realizar un análisis y clasificación de la información. La información se podría clasificar de acuerdo con los aspectos claves como por ejemplo información sobre tipos de software de simulación, información sobre técnicas de analítica usadas junto a la simulación, beneficios y desafíos, aplicaciones prácticas, entre otros.

Teniendo la información clara y analizada, se propone realizar y sintetizar los datos en una especie de cuadro comparativo, o algún esquema con enfoque similar, donde se puedan resumir las principales características y hallazgos encontrados en la literatura, con el fin de identificar patrones y oportunidades respecto al uso de las técnicas de analítica para la simulación.

Luego de tener claridad sobre lo mencionado anteriormente, se diseñará una simulación basada en datos reales, donde se validará la aplicabilidad de las técnicas encontradas en la literatura, con esto se podría demostrar la precisión y efectividad de las técnicas de analítica de datos integradas con la simulación, aplicados en un contexto práctico, además de que se puede realizar una comparación con los resultados teóricos obtenidos en la revisión de literatura. Para concluir, se sintetizarán los principales hallazgos del estudio, identificando oportunidades de mejora y recomendaciones para futuras investigaciones que estén asociadas al tema.

V. RESULTADOS ESPERADOS

- **Identificación de técnicas analíticas relevantes:** Obtener un marco conceptual que describa qué técnicas de machine learning e inteligencia artificial son más utilizadas en la obtención de datos para modelos de simulación.
- **Clasificación de metodologías:** Categorizar diferentes enfoques analíticos según su aplicabilidad.
- **Identificación de estrategias óptimas:** Identificar estrategias para integrar técnicas de analítica en la obtención y preparación de datos para simulaciones.
- **Evaluación del impacto en la calidad de modelos:** Analizar cómo la implementación de métodos analíticos podría mejorar la precisión y utilidad de los modelos de simulación.

- **Propuesta de oportunidades para investigaciones futuras:** Identificar vacíos u oportunidades de mejora en los artículos ya publicados de cara a futuras investigaciones.

VI.REFERENCIAS

- [1] L. J. Aguilar, “Big Data: Análisis de grandes volúmenes de datos en organizaciones,” *Journal of Chemical Information and Modeling*, vol. 53, no. 9, 2013.
- [2] M. Ahmed and A.-S. K. Pathan, *Data Analytics Concepts, Techniques, and Applications*. Taylor and Francis, vol. 91, no. 5, 2012. doi: 10.1017/CBO9781107415324.004.
- [3] N. Alias, W. M. Al-Rahmi, N. Yahaya, and Q. Al-Maatouk, “Big data, modeling, simulation, computational platform and holistic approaches for the fourth industrial revolution,” *Int. J. Eng. Technol. (UAE)*, vol. 7, no. 4, pp. 3722–3725, 2018. doi: 10.14419/ijet.v7i4.21244.
- [4] D. Bourguignon, *Closing the loop: New circular economy package*. European Parliamentary Research Service, Jan. 2016.
- [5] A. Borshchev and A. Filippov, “From system dynamics and discrete event to practical agent based modeling: reasons, techniques, tools,” *Proc. Conf. System Dynamics*, 2004. [Online]. Available: <http://scholar.google.com/...>
- [6] E. Brynjolfsson and K. McElheran, “The rapid adoption of data-driven decision-making,” *Am. Econ. Rev.*, vol. 106, no. 5, pp. 133–139, 2016. doi: 10.1257/aer.p20161016.
- [7] B. A. Cerinza, *Análisis de sistemas de producción mediante el uso de herramientas digitales orientadas a la simulación de procesos productivos que integren modelado, análisis y visualización*. Univ. de Pamplona, 2022. [Online]. Available: <http://repositoriodspace.unipamplona.edu.co/jspui/handle/20.500.12744/5310>
- [8] H. Chen, R. H. Chiang, and V. C. Storey, “Business Intelligence and Analytics: From Big Data to Big Impact,” *MIS Q.*, vol. 36, no. 4, pp. 1165–1188, 2012. doi: 10.2307/41703503.
- [9] R. Coss Bu, *Simulación: un enfoque práctico*, 2nd ed. Limusa Noriega, 1996.
- [10] T. H. Davenport and J. G. Harris, *Competing on Analytics: The New Science of Winning*. Harvard Business Review Press, 2007.
- [11] P. Fernández, P. Livacic-Rojas, and G. Vallejo, “How to elect the best statistical analysis for to analyze a repeated measures design,” *Int. J. Clin. Health Psychol.*, vol. 7, no. 1, pp. 153–175, 2007.
- [12] A. Gandomi and M. Haider, “Beyond the hype: Big data concepts, methods, and analytics,” *Int. J. Inf. Manag.*, vol. 35, no. 2, pp. 137–144, 2015. doi: 10.1016/j.ijinfomgt.2014.10.007.
- [13] P. J. Giabbanelli and M. Baniukiewicz, “Visual analytics to identify temporal patterns and variability in simulations from cellular automata,” *ACM Trans. Model. Comput. Simul.*, vol. 29, no. 1, pp. 1–26, 2019. doi: 10.1145/3265748.
- [14] A. L. Medaglia, D. Hueth, J. C. Mendieta, and J. A. Sefair, “A multiobjective model for the selection and timing of public enterprise projects,” *Socio-Econ. Plan. Sci.*, vol. 42, no. 1, pp. 31–45, 2008. doi: 10.1016/j.seps.2006.06.009.
- [15] A. McAfee and E. Brynjolfsson, “Big Data: The Management Revolution,” *Harv. Bus. Rev.*, vol. 90, no. 10, pp. 60–68, 2012.
- [16] E. R. Pozo-Safla, S. M. Aquino-Arroba, and M. A. Ordoñez-Viñan, “Análisis estadístico para validar la simulación por elemento finito en el diseño a deformación de una viga en voladizo,” *Polo Del Conocimiento*, vol. 6, no. 6, pp. 586–611, 2021. doi: 10.23857/pc.v6i6.2771.
- [17] F. Provost and T. Fawcett, *Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking*. O’Reilly Media, 2013.
- [18] W. Raghupathi and V. Raghupathi, “Big data analytics in healthcare: Promise and potential,” *Health Inf. Sci. Syst.*, vol. 2, no. 1, p. 3, 2014. doi: 10.1186/2047-2501-2-3.
- [19] I. Sitova and J. Pecerska, “Data Mining Techniques in Simulation Results Analysis,” *Proc. 59th Int. Sci. Conf. Inf. Technol. Manag. Sci. Riga Tech. Univ.*, 2018. doi: 10.1109/ITMS.2018.8552972.
- [20] A. A. Tako and S. Robinson, “The application of discrete event simulation and system dynamics in the logistics and supply chain context,” *Decis. Support Syst.*, vol. 52, no. 4, pp. 802–815, 2012. doi: 10.1016/j.dss.2011.11.015.
- [21] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed. Morgan Kaufmann, 2016.

Desafíos en el Modelado de Comportamientos Colectivos: Revisión de la Literatura sobre Metodologías de Simulación Basada en Agentes

Danna Gabriela Espitia Cano²; Lindsay Álvarez Pomar²

Resumen—La simulación basada en agentes se ha consolidado como una herramienta clave para representar sistemas complejos, como las multitudes, donde surgen dinámicas colectivas, decisiones distribuidas y comportamientos emergentes. Sin embargo, muchas metodologías actuales para el modelado de agentes se centran en aspectos estructurales y funcionales, dejando de lado dimensiones fundamentales como las intenciones, motivaciones internas de los agentes y mecanismos detallados de interacción.

Este artículo presenta una revisión sistemática de metodologías orientadas a la ingeniería de sistemas multiagente (AOSE), con énfasis en Prometheus, evaluando su capacidad para capturar el comportamiento autónomo y distribuido de los agentes en contextos dinámicos. A través de un análisis comparativo, se identifican vacíos metodológicos relacionados con la falta de estandarización, la baja cobertura de procesos cognitivos a nivel micro y la limitada aplicabilidad práctica en simulaciones complejas.

Palabras Claves - Simulación Basada en Agentes, Modelado, Metodologías.

I. INTRODUCTION

La simulación basada en agentes (ABM) se ha posicionado como una herramienta eficaz para modelar sistemas complejos con interacciones descentralizadas entre entidades autónomas, lo que permite observar comportamientos emergentes [1,2]. Ha sido utilizada con éxito en campos como evacuación en emergencias [3,4], logística humanitaria [5,6], desarrollo urbano [7] y gestión del riesgo [8,9].

El grado de detalle en la parametrización de los agentes —incluyendo atributos, reglas, valores, creencias y mecanismos adaptativos— es clave para capturar el vínculo entre comportamientos individuales y patrones colectivos [10–12]. Para guiar este proceso, existen metodologías de ingeniería orientada a agentes (AOSE) como GAIA [13], Tropos y MaSE [14], INGENIAS [15] y Prometheus [16]. No obstante, diversas revisiones indican que muchas de ellas no abordan con

suficiente profundidad aspectos como emociones, intenciones o capacidades de aprendizaje [18–21].

Esta limitación ha motivado el desarrollo de metodologías más flexibles para modelar agentes con motivaciones internas, sensibilidad contextual y capacidad evolutiva [22, 3, 23]. Asimismo, se han incorporado enfoques como el diseño sensible a valores [5], la simulación híbrida [24] y modelos multiescala [25], promoviendo una integración más sólida entre la computación y las teorías del comportamiento humano [26,27].

II. METODOLOGÍAS ORIENTADAS A AGENTES EN LAS SIMULACIONES DE MULTITUDES

En décadas recientes, el desarrollo de sistemas multiagente ha buscado reflejar la complejidad de entornos sociales y organizacionales mediante la representación de interacciones autónomas y comportamientos emergentes [1,28]. Sin embargo, muchas metodologías se enfocan en aspectos estructurales, descuidando componentes motivacionales y colaborativos [20], lo cual limita la representación de fenómenos como la toma de decisiones colectiva o el aprendizaje grupal [29].

Adicionalmente, se ha observado una adopción limitada y no sistemática de metodologías formales (AOSE) en proyectos ABM, predominando enfoques híbridos que no siempre aseguran coherencia ni trazabilidad [30,31]. Ante esta situación, este estudio propone un análisis comparativo de metodologías orientadas a agentes aplicadas a la simulación de multitudes, con el objetivo de identificar fortalezas, debilidades y posibilidades de integración para una representación más robusta de sistemas dinámicos y complejos.

A. Revisión de literatura en Scopos Fase 1

Para abordar el problema relacionado con la diversidad y formalización de metodologías en la simulación basada en agentes para modelar multitudes, se formuló la ecuación de búsqueda:

¹ Ingeniera Industrial, Universidad Distrital Francisco José de Caldas, Grupo de Investigación: ARCOSES. Especialista de Calidad. Lean Solutions Group, Correo Electrónico: dgespitia@udistrital.edu.co

² Docente Académico de Ingeniería Industrial, Universidad Distrital Francisco José de Caldas. Doctorado en Ingeniería de la Universidad Distrital Francisco José de Caldas, Maestría en Ingeniería Industrial, Universidad de los Andes, Ingeniería Industrial, Universidad Distrital Francisco José de Caldas, Correo Electrónico: lavarez@udistrital.edu.co

TITLE-ABS-KEY ("multi-agent simulation" OR "agent-based model" OR "multiagent") AND ("methodology" OR "modeling approach") AND ("crowd").

Esta ecuación permitió identificar estudios relevantes sobre enfoques metodológicos aplicados al modelado colectivo. La búsqueda en Scopus arrojó 62 documentos, los cuales fueron analizados bibliométricamente para identificar tendencias temporales y temáticas predominantes, los cuales fueron analizados bibliométricamente para detectar tendencias temporales y temáticas. Se usó VOSviewer para visualizar redes de palabras clave y refinar el análisis.

El análisis temporal de publicaciones en Scopus muestra un interés variable en metodologías para la simulación de multitudes basada en agentes, con picos en 2012, 2017 y 2024. Resulta llamativo que durante la pandemia por COVID-19 no se evidencie un aumento significativo en este tipo de estudios, lo cual podría deberse a cómo está formulada la ecuación de búsqueda, especialmente al incluir el término crowd. No obstante, al ajustar la búsqueda a agent-based simulation sin crowd, los años 2020-2021 reflejan un auge, indicando que el campo de estudio tuvo desarrollo.

La red de co-ocurrencia de palabras clave (Figura 1) ubica a "collective intelligence" como nodo central, conectado a términos como "agent-based model" y "self-organization". Se observa un cambio de temas consolidados (2017–2019) a emergentes (2020–2021), reflejando un giro hacia evaluaciones dinámicas del comportamiento colectivo. No obstante, la baja presencia de términos como "methodology" evidencia que el desarrollo metodológico formal no es prioritario, dejando una brecha en la sistematización de enfoques para modelar multitudes.

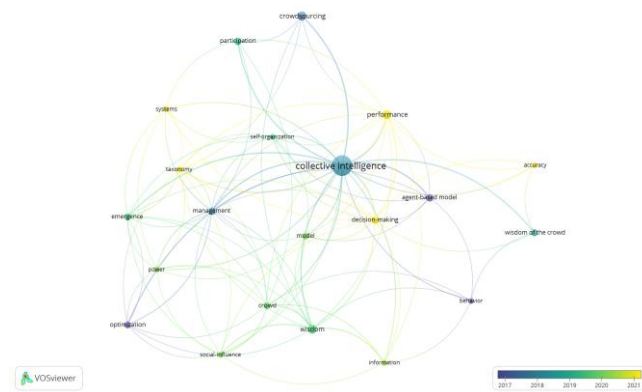


Fig. 1. Red de Co-ocurrencia temporal. Fuente VOSviewer - Ecuación de búsqueda.

B. Revisión de literatura en Scopus Fase 2

Este hallazgo condujo a una segunda fase de revisión, centrada en evaluar cómo se estructuran y justifican las metodologías empleadas en estudios recientes de simulación basada en agentes. Se seleccionó una muestra intencionada de 15 artículos representativos de diversos dominios, con el fin de analizar si utilizan marcos metodológicos explícitos o enfoques aislados.

La muestra abarca aplicaciones en logística humanitaria, salud pública, transporte, gestión de multitudes, IoT, biología, agricultura y migración forzada, lo que permitió una visión transversal sobre el uso metodológico del modelado basado en agentes en distintos contextos.

La Tabla I presenta la clasificación de los artículos según su dominio y tipo de modelado empleado.

TABLA I
CLASIFICACIÓN GENERAL DE LOS ARTÍCULOS SEGÚN SU APLICACIÓN Y TIPO DE MODELADO

Autor(es)	Año	Aplicación	Tipo de Modelado
Christine Boshuijzen-van Burken et al.	2020	Logística humanitaria	ABM
Juan Camilo López-Vargas et al.	2025	Logística humanitaria	ABM
Shatrudhan Pandey et al.	2025	Industria	ABM
Ludmil B. Alexandrov et al.	2009	Biología	EBAM
Samreen Laghari, Muaz A. Niazi	2016	Tecnología/IoT	EBAM
Makoto Mizuno et al.	2015	Negocios/Finanzas	EBAM
Daniel O. Pessoa et al.	2025	Transporte	ABM + MAS
Mahdi Keykhaei et al.	2024	Logística Humanitaria	ABM + FIS
R. A. Williams	2025	Gestión empresarial	ABM
Moon Gi Seok, Daejin Park	2024	Tecnología/IoT	ML-AMB
Huu-Tu Dang et al.	2024	Gestión de multitudes	ML-AMB
Favour Nerrise	2021	Logística Humanitaria	ML-AMB + ABM
Jo-An Atkinson et al.	2018	Salud pública	P-AMB
Will, M. et al.	2024	Agricultura	Theory-Driven ABM
Arika Ligmann-Zielinska	2009	Urbanismo	Theory-Driven ABM

En los últimos años, la simulación basada en agentes (ABM) ha crecido y se ha diversificado notablemente [28]. Este desarrollo ha integrado enfoques como machine learning, lógica difusa, técnicas estadísticas y metodologías participativas, dando lugar a variantes como ML-ABM, Hybrid ABM y P-ABM.

Los estudios revisados aplican ABM en campos como logística humanitaria, industria, transporte y salud. Por ejemplo, Boshuijzen-van Burken et al. modelaron valores en contextos de refugiados [5], López-Vargas et al. simulaban coordinación ante desastres [6], y Pandey et al. aplicaron ABM para mitigar riesgos industriales [32]. En transporte, Pessoa et al. integraron ABM + MAS [24], y Keykhaei et al. usaron lógica difusa para evacuaciones [33].

Algunos modelos integran ABM con inteligencia artificial, como ML-ABM aplicado en IoT por Seok y Park [34], arquitecturas con aprendizaje automático para simular multitudes de Dang et al. [25], y el uso combinado de ML-ABM con Random Forest para desastres naturales por Nerrise [8]. Atkinson et al. emplearon P-ABM para políticas públicas [26]. También se destacan enfoques evolutivos bajo EBAM en biología celular [13], IoT adaptativo [35] y servicios financieros

[23], que fomentan adaptabilidad, pero carecen de una estructura metodológica clara.

Por otro lado, se observa diversidad en el uso de técnicas como ML, ODD y algoritmos evolutivos, pero estas se aplican de forma aislada, lo que limita la trazabilidad y replicabilidad de los modelos [36,37]. Aunque se infiere el empleo de ODD en estudios como Boshuijzen-van Burken [5], López-Vargas [6], Pandey [32], Dang [25] y Nerrise [8], y la presencia de arquitectura BDI en Laghari y Niazi [35] o algoritmos evolutivos en Alexandrov [13], ninguno integra un marco metodológico integral que articule desde el nivel micro de los agentes. Atkinson [26] utilizó POM y modelado participativo, pero también sin un enfoque sistemático completo.

Aunque existen muchas técnicas aplicadas para el modelado basado en agentes, sigue faltando metodologías formales que aseguren coherencia, trazabilidad y detalle, especialmente en contextos colectivos y dinámicos [37][38]. Esta ausencia dificulta la estandarización y comparación objetiva entre enfoques. Por ello, se avanza hacia una tercera fase de revisión, enfocada en analizar y comparar metodologías especializadas según su aplicabilidad, cobertura del modelado y capacidad para parametrizar agentes a nivel micro.

C. Revisión de literatura en Scopus Fase 3

Esta fase se enfoca en metodologías de simulación basada en agentes, esenciales para diseñar y validar sistemas complejos de forma estructurada, desde la conceptualización hasta la implementación [11], [39]. Estas metodologías operan en distintos niveles de abstracción: el nivel micro, que modela el comportamiento individual, y el nivel meso, que estructura interacciones colectivas [40].

Metodologías como MAS-CommonKADS [41], GAIA [14], CASSIOPEIA [42], ZEUS [43] y Vowel Engineering [44] se centran en el nivel meso, representando roles, jerarquías y dinámicas organizativas [18], [19]. Si bien eficaces para modelar estructuras colectivas, presentan debilidades en la representación del comportamiento interno de los agentes [45], además de carecer de guías claras de evaluación y validación [21] y de criterios estandarizados [20].

En contraste, metodologías como BDI [46] y TROPOS [14] abordan el nivel micro desde la lógica de creencias, deseos e intenciones, aunque BDI es solo una arquitectura sin soporte metodológico completo, y TROPOS se centra en requisitos iniciales, con poca cobertura en fases posteriores [19,29]. Por otro lado, modelos híbridos como MaSE [15], INGENIAS [16] y Prometheus [47] integran estructuras organizativas y modelado interno, abarcando más fases del desarrollo. Entre ellos, Prometheus destaca por su claridad funcional, soporte de herramientas y modelado visual [19], aunque aún debe mejorar para representar dinámicas emergentes complejas. En general, ninguna metodología logra integrar totalmente los niveles funcional y organizacional, pero Prometheus se presenta como una base sólida para incorporar principios de inteligencia colectiva. Sin embargo, persisten limitaciones comunes: falta de estandarización [48], baja validación [5], escasa integración de fases [20], herramientas fragmentadas [21] y alta complejidad técnica en metodologías como INGENIAS y TROPOS [4,49].

En este contexto, Promethues, desarrollada por Padgham y Winikoff, se basa en arquitecturas BDI y destaca por su enfoque práctico, claridad y accesibilidad [18]. Su estructura se compone de tres fases: (1) la especificación del sistema, que define objetivos, funcionalidades y escenarios de uso; (2) el diseño de la arquitectura, donde se asignan funcionalidades a agentes mediante diagramas y protocolos; y (3) el diseño detallado, que modela la lógica interna de los agentes (eventos, planes y capacidades).

Gracias a esta estructura y su alineación con BDI, Prometheus permite representar con precisión tanto el comportamiento individual como el colaborativo. Su claridad metodológica la convierte en una base sólida para extenderse hacia propuestas que integren principios de inteligencia colectiva [17].

III. DESCRIPCIÓN DE LA REVISIÓN DE LA LITERATURA

La simulación basada en agentes (ABM) es una herramienta eficaz para modelar sistemas dinámicos con múltiples entidades autónomas [50], destacando en logística para analizar comportamientos colectivos de vehículos, personas o drones con fines de optimización y eficiencia [51]. Además, el modelado de comportamientos emergentes se ha mejorado mediante técnicas como simulación paralela para aumentar el rendimiento [3], métodos de enjambre y ruido aleatorio en la toma de decisiones [4], y estimación de parámetros para calibrar modelos con datos reales [51].

A. Consideraciones sobre Metodologías Aplicadas a Modelos Basados en Agentes

El desarrollo de sistemas multiagente ha promovido metodologías específicas para guiar su análisis, diseño e implementación [52]. Estas buscan estructurar comportamientos e interacciones entre agentes, permitiendo modelar también fenómenos colectivos como la autoorganización [53,20,54].

No obstante, persisten limitaciones: muchas metodologías carecen de mecanismos para representar dinámicas complejas, validar modelos formalmente o mantener trazabilidad entre fases [35]. Además, estudios coinciden en que metodologías como Prometheus, GAIA o MaSE tienden a centrarse en roles y estructuras, sin profundizar en las intenciones colectivas o aspectos internos de los agentes [49,55].

B. Limitaciones y brechas en las metodologías de desarrollo de sistemas basado en agente

Aunque la ingeniería orientada a agentes ha avanzado, aún existen brechas importantes que dificultan la estandarización, comparación y aplicación práctica de las metodologías. Principalmente, falta un marco común y criterios formales para evaluar y comparar enfoques, lo que complica la selección adecuada [2, 20, 21].

Muchas metodologías no cubren todo el ciclo de vida del sistema, enfocándose solo en requisitos o careciendo de soporte para validación y mantenimiento [1, 19, 29]. También hay ambigüedades en notación y terminología que afectan la comunicación y reutilización, además de incompatibilidades entre metamodelos que limitan la interoperabilidad [19, 56]. Así mismo, muchas propuestas son poco maduras, con poca

validación en entornos reales y dificultades para integrar coherentemente las fases del desarrollo [2, 20, 21]. En síntesis, se requiere avanzar hacia propuestas más integrales que cierren estas brechas, estableciendo estándares claros de evaluación, ampliando su cobertura funcional y reforzando su aplicabilidad en escenarios dinámicos y complejos.

C. Consideraciones sobre el modelado de agentes en las metodologías aplicadas

Las metodologías actuales para el modelado de agentes en sistemas multiagente presentan limitaciones, pues suelen omitir componentes internos clave como creencias, metas y planes, lo que restringe la expresividad y adaptabilidad del comportamiento [49]. Además, la interacción e interoperabilidad entre agentes en entornos heterogéneos no se aborda sistemáticamente, dificultando la cooperación y la toma de decisiones colectiva [49, 57].

Muchas metodologías usan modelos estáticos sin aprendizaje ni adaptación, limitando su validez en entornos dinámicos [48, 58], y carecen de estructuras formales para roles y colaboración organizacional, afectando su aplicación en logística y crisis [49, 56]. Por ello, es crucial mejorar el modelado micro para reflejar motivaciones y comportamientos colectivos más precisos.

Aunque metodologías como Prometheus, GAIA e INGENIAS incluyen metas compartidas y razonamiento distribuido, ninguna integra completamente las dimensiones cognitivas, sociales y motivacionales del comportamiento colectivo [55], lo que restringe la capacidad de modelar sistemas complejos de manera integral. Por ello, es importante plantear una metodología estructurada con un enfoque que incorpore estas dimensiones puede mejorar la representación de agentes explicativos y adaptativos, cerrando vacíos en el modelado de comportamientos emergentes en sistemas [29, 3].

IV. DISCUSIÓN Y CONCLUSIONES

Este análisis evidenció avances importantes en la aplicación de la simulación basada en agentes y su vinculación con la inteligencia colectiva. No obstante, persisten limitaciones metodológicas, especialmente en el diseño interno de los agentes y su capacidad para representar dinámicas colectivas y decisiones distribuidas.

Aunque se utilizan técnicas como ODD, BDI y enfoques participativos, su integración suele ser fragmentada, dificultando la construcción de modelos explicativos robustos a nivel micro. La revisión de metodologías como GAIA, MaSE, Tropos, INGENIAS y Prometheus permitió identificar vacíos como la falta de estandarización, escasa validación empírica y limitada representación cognitiva de los agentes [18, 51, 59]. Se destaca la necesidad de avanzar hacia modelos que integren estructuras formales con una comprensión profunda de agentes autónomos, dotados de objetivos, creencias e intenciones. Esto permitiría desarrollar simulaciones más interpretables y adaptativas, alineadas con la complejidad de los sistemas distribuidos.

En suma, aunque se reconocen avances, la revisión revela brechas clave que deben abordarse para fortalecer el valor

explicativo y operativo de los modelos en contextos colectivos complejos.

REFERENCES

- [1] C. M. Macal and M. J. North, "Tutorial on agent-based modeling and simulation," in *Proceedings of the Winter Simulation Conference*, 2005, pp. 14–pp.
- [2] C. M. Macal and M. J. North, "A multidimensional framework for characterizing the use of agent-based models in practice," in *Proceedings of the Winter Simulation Conference*, 2010, pp. 1–10.
- [3] M. F. Mohd Ibrahim, N. Mustaffa, S. Z. M. Hashim, and I. E. Khairuddin, "Agent-based simulation model for emergency evacuation in a smart building," *Sustainable Cities and Society*, vol. 77, p. 103532, 2022.
- [4] A. Malinowski, F. Osorio, and R. Torres, "Parallel agent-based simulation for large-scale evacuation scenarios," *Simulation Modelling Practice and Theory*, vol. 77, pp. 1–15, 2017.
- [5] C. Boshuijzen-van Burken, A. L. I. M. van de Walle, S. Cunningham, and L. M. Coolen, "Agent-Based Modelling of Values: The Case of Value Sensitive Design for Refugee Logistics," *Journal of Artificial Societies and Social Simulation*, vol. 23, no. 4, 2020.
- [6] J. C. López-Vargas, D. Castaño-Cano, and R. F. Gómez-Acosta, "Coordination Mechanisms Applied to Logistical Systems for Local Disaster Preparedness: A Latin American Case," *Journal of Humanitarian Logistics and Supply Chain Management*, vol. 15, no. 2, 2025.
- [7] A. Ligmann-Zielinska, "The Impact of Risk-Taking Attitudes on a Land Use Pattern: An Agent-Based Model of Residential Development," *Journal of Land Use Science*, vol. 4, no. 1–2, 2009.
- [8] F. Nerrise, "Predictive Agent-Based Modeling of Natural Disasters Using Machine Learning," Ph.D. dissertation, University of Maryland-College Park, 2021.
- [9] V. Padilla and J. Dávila, "Simulación multi-agente para gestión de desastres y reducción de riesgo," *Revista Ciencia e Ingeniería*, edición especial: "Jornada de Modelado y Simulación," pp. 21–30, 2012.
- [10] M. Wooldridge, *An Introduction to MultiAgent Systems*, 2nd ed. Wiley, 2009.
- [11] S. Chen and Y. Wang, "Towards a methodology of modeling emergent behaviors in multi-agent systems," *Simulation Modelling Practice and Theory*, vol. 19, no. 5, pp. 1020–1036, 2011.
- [12] L. B. Alexandrov, S. Nik-Zainal, D. C. Wedge, J. P. W. Edwards, S. A. Roberts, and M. R. Stratton, "Computer Modeling Describes Gravity-Related Adaptation in Cell Cultures," *PLoS ONE*, vol. 4, no. 11, 2009.
- [13] M. Wooldridge, N. R. Jennings, and D. Kinny, "The Gaia Methodology for Agent-Oriented Analysis and Design," *Autonomous Agents and Multi-Agent Systems*, vol. 15, 2000.
- [14] P. Bresciani, A. Perini, P. Giorgini, F. Giunchiglia, and J. Mylopoulos, "Tropos: An Agent-Oriented Software Development Methodology," *Autonomous Agents and Multi-Agent Systems*, vol. 8, no. 3, pp. 203–236, 2004.
- [15] S. DeLoach, "Analysis and Design using MaSE and agentTool," in *Proc. 12th Midwest Artificial Intelligence and Cognitive Science Conf. (MAICS)*, 2001.
- [16] J. Gomez Sanz, *Modelado de Sistemas Multi-Agente*, Ph.D. dissertation, Facultad de Informática, Univ. Complutense, Spain, 2002.
- [17] L. Padgham and M. Winikoff, *Developing Intelligent Agent Systems: A Practical Guide*. John Wiley & Sons, 2004.
- [18] S. A. DeLoach, M. Wood, and C. Sparkman, "Multiagent systems engineering," *International Journal of Software Engineering and Knowledge Engineering*, vol. 11, no. 3, pp. 231–258, 2001.
- [19] K. H. Dam and M. Winikoff, "Comparing agent-oriented methodologies," in *Agent-Oriented Software Engineering III*, Springer, 2003, pp. 78–93.
- [20] J. J. Gómez-Sanz, "Metodologías para el desarrollo de sistemas multi-agente," *Inteligencia Artificial. Revista Iberoamericana de IA*, vol. 7, no. 18, pp. 51–64, 2003.
- [21] M. Moreno Espino, A. Rosete Suárez, A. Simón Cuevas, et al., "Ingeniería de software orientada a agentes: roles y metodologías," *Revista Ingeniería Industrial*, vol. XXVII, no. 2–3, pp. 35–44, 2006.
- [22] R. A. Williams, "Towards an Agent-Based Model Using a Hybrid Conceptual Modelling Approach: A Case Study of Relationship Conflict Within Large Enterprise System Implementations," *Journal of Simulation*, vol. 19, no. 1, 2025.

- [23] M. Mizuno, T. Nakayama, and S. Takada, "Simulating Customer-to-Customer Interaction in a B2B Financial Service Business by Empirical Agent-Based Modeling," in BICT Conference, 2015.
- [24] D. O. Pessoa, P. Gomes, F. L. Ribeiro, and R. Rossetti, "Developing Routes Using Hybrid Simulations: Public Transport vs. Bike Sharing," *International Journal of Simulation Modelling*, vol. 24, no. 1, 2025.
- [25] H.-T. Dang, J. Jeong, and H. Lim, "HyPedSim: A Multi-Level Crowd-Simulation Framework—Methodology, Calibration, and Validation," *Sensors*, vol. 24, no. 4, 2024.
- [26] J.-A. Atkinson, A. Page, A. Wells, and M. Milat, "Harnessing Advances in Computer Simulation to Inform Policy and Planning to Reduce Alcohol-Related Harms," *International Journal of Public Health*, vol. 63, no. 5, 2018.
- [27] M. Will, A. Bruns, S. Buerkert, and K. Schmid, "From Primary Data to Formalized Decision-Making: Open Challenges and Ways Forward to Inform Representations of Farmers' Behavior in Agent-Based Models," *Ecology and Society*, vol. 29, no. 1, 2024.
- [28] C. Macal and M. North, "Tutorial on agent-based modelling and simulation," *Journal of Simulation*, vol. 4, no. 3, pp. 151–162, 2010.
- [29] J. Aguilar, I. Bessembel, M. Cerrada, et al., "Una metodología para el modelado de sistemas de ingeniería orientado a agentes," *Inteligencia Artificial. Revista Iberoamericana de IA*, vol. 12, no. 38, pp. 85–97, 2008.
- [30] G. Madey, V. Freeh, and R. Tynan, "Agent-based modeling and simulation of collaborative networks," *ACM Computing Surveys*, vol. 35, no. 1, pp. 137–169, 2003.
- [31] M. Cossentino, S. Galland, and N. Gaud, "Agent-Oriented Software Engineering," in *Multi-Agent Systems*, J. P. Müller, F. Klügl, and J. Thangarajah, Eds. Springer, 2015, pp. 107–152.
- [32] S. Pandey, R. Shukla, A. Sinha, and A. Ghosh, "Towards Safer Steel Operations with a Multi-Model Framework for Accident Prediction and Risk Assessment Simulation," *Scientific Reports*, vol. 15, 2025.
- [33] M. Keykhaei, M. Khakpour, and A. Rezaeian, "A Situation-Aware Emergency Evacuation (SAEE) Model Using Multi-Agent-Based Simulation for Crisis Management After Earthquake Warning," *Geospatial Information Science*, vol. 27, no. 1, 2024.
- [34] M. G. Seok and D. Park, "A Novel Multi-Level Evaluation Approach for Human-Coupled IoT Applications," *Journal of Ambient Intelligence and Humanized Computing*, vol. 15, no. 3, 2024.
- [35] S. Laghari and M. A. Niazi, "Modeling the Internet of Things, Self-Organizing and Other Complex Adaptive Communication Networks," *PLoS ONE*, vol. 11, no. 6, 2016.
- [36] C. García-Díaz and C. Olaya, "A survey on agent-based modeling and simulation using formal methodologies," *Simulation Modelling Practice and Theory*, vol. 92, pp. 232–254, 2019.
- [37] C. BERNON, V. Camps, M.-P. Gleizes, and G. Picard, "Engineering adaptive multi-agent systems: The ADELFE methodology," in *Agent-Oriented Methodologies*. IGI Global, 2005, pp. 172–202.
- [38] V. Dignum and F. Dignum, "Designing agent systems: state of the practice," *International Journal of Agent-Oriented Software Engineering*, vol. 1, no. 1, pp. 1–20, 2007.
- [39] L. Sterling and K. Taveter, *The Art of Agent-Oriented Modeling*. MIT Press, 2009.
- [40] T. Juan, A. Pearce, and L. Sterling, "ROADMAP: Extending the Gaia methodology for complex open systems," in *Proc. 1st International Joint Conference on Autonomous Agents and Multiagent Systems*, 2002, vol. 3, pp. 3–10.
- [41] C. Iglesias, *Definición de una metodología para el desarrollo de Sistemas MultiAgente*, Ph.D. dissertation, Dept. Ing. de Sistemas Telemáticos, Univ. Politécnica de Madrid, Spain, 1998.
- [42] A. Collinot, P. Carle, and K. Zeghal, "Cassiopeia: A Method for Designing Computational Organizations," in *Proc. 1st Int. Workshop on Decentralized Multi-Agent Systems (DIMAS'95)*, Cracow, Poland, Nov. 1995, pp. 124–131.
- [43] H. S. Nwana, D. T. Ndumu, L. C. Lee, and J. C. Collis, "ZEUS: A Toolkit for Building Distributed Multi-Agent Systems," *Applied Artificial Intelligence Journal*, vol. 13, no. 1, pp. 129–185, 1999.
- [44] P. M. Ricordel, *Programmation Orientée Multi-Agents: Développement et Déploiement de Systèmes Multi-Agents Voyelles*, Ph.D. dissertation, INP de Grenoble, France, 2001.
- [45] L. Rodríguez, "Análisis comparativo de métodos de elicitación de requisitos para sistemas basados en agentes," *Academia.edu*, 2010.
- [46] D. Kinny, M. Georgeff, and A. Rao, "A Methodology and Modelling Technique for Systems of BDI Agents," *Technical Report*, 1997.
- [47] L. Padgham and M. Winikoff, "Prometheus: A Methodology for Developing Intelligent Agents," in *Agent-Oriented Software Engineering III*, P. Giunchiglia, J. Odell, and G. Weiss, Eds., Springer, LNCS 2585, 2003, pp. 174–185.
- [48] R. Junges and F. Klügl, "How to Design Agent-Based Simulation Models Using Agent Learning (MABLE)," in *Proceedings of the Winter Simulation Conference*, 2012.
- [49] Q. N. N. Tran and G. Low, "MOBMA: A Methodology for Ontology-Based Multi-Agent Systems Development," *Information and Software Technology*, 2008.
- [50] L. R. Izquierdo, J. M. Galán, J. I. Santos, and R. Del Olmo, "Modelado de sistemas complejos mediante simulación basada en agentes y mediante dinámica de sistemas," *EMPIRIA. Revista de Metodología de las Ciencias Sociales*, no. 16, pp. 85–112, jul.–dic. 2008.
- [51] J. Wojtusiak, A. Jain, and R. S. Michalski, "Agent-based decision support system for emergency management," in *Proceedings of the 2012 Winter Simulation Conference*, 2012, pp. 1–12.
- [52] Y. Hadfeg Fernández, M. Moreno Espino, and M. D. Delgado Dapena, "Propuesta de actividades de pruebas para Ingenierías," *Ingeniería Informática*, vol. III, no. 3, pp. 47–56, sep.–dic. 2012.
- [53] V. Padilla and J. Dávila, "Simulación multi-agente para gestión de desastres y reducción de riesgo," *Revista Ciencia e Ingeniería*, edición especial: "Jornada de Modelado y Simulación," pp. 21–30, 2012.
- [54] A. Sabas, M. Badri, and S. Delisle, "A multidimensional framework for the evaluation of multiagent system methodologies," *Université Laval and Université du Québec à Trois-Rivières*, 2002.
- [55] C. A. Iglesias, *Definición de una metodología para el desarrollo de Sistemas MultiAgente*, Ph.D. dissertation, Universidad Politécnica de Madrid, 1998.
- [56] G. Beydoun, G. Low, B. Henderson-Sellers, H. Mouratidis, J. J. Gómez-Sanz, J. Pavón, and C. Gonzalez-Perez, "FAML: A Generic Metamodel for MAS Development," *IEEE Transactions on Software Engineering*, 2009.
- [57] T. De Wolf and T. Holvoet, "Towards a Methodology for Engineering Self-Organising Emergent Systems," in *Self-Organization and Autonomic Informatics*, IOS Press, 2005.
- [58] F. Klügl and A. L. C. Bazzan, "Agent-based modeling and simulation: introduction and review," *Journal of Artificial Societies and Social Simulation*, vol. 15, no. 2, 2012.
- [59] J. Surowiecki, *The Wisdom of Crowds*. New York: Anchor Books, 2005.
- [60] D. S. W. Tansley and C. Hayball, *Knowledge Based Systems Analysis and Design: A KADS Developer's Handbook*. Prentice Hall, 1993.

Dinámica de la cadena de suministro Laboral sujeta a Disrupciones. Caso Cacao

Carlos Arango L., Dany Guerra A., Andres Hualpa Z.

Abstract— Las cadenas de suministro han mostrado su vulnerabilidad a diversas disrupciones, incluyen entre otras, guerras arancelarias, inestabilidad geopolítica, pandemias y cambios tecnológicos. Dado su papel fundamental en el desarrollo económico de las regiones, los estudios en cadena de suministro laborales deben abordarse desde una perspectiva sistémica y dinámica que aporten a la comprensión de la influencia del impacto de las disrupciones no solo en aspectos operacionales sino en aspectos comportamentales, aportando evidencia de la incidencia de los procesos de toma de decisiones por parte de los agentes en la estabilidad/ inestabilidad de la cadena.

Este estudio tiene como objetivo caracterizar la dinámica de la gestión de inventarios en la cadena de suministro laboral mediante el modelado de disrupciones en el tiempo de aprovisionamiento. Se simularán disrupciones para evaluar la propagación de las perturbaciones y la aparición de respuestas comportamentales oportunas por parte de los tomadores de decisiones usando el heurístico de anclaje y ajuste dando cuenta de la racionalidad acotada de dichos decisores mientras gestionan inventarios en la cadena de suministro laboral tomando como referencia la cadena del cacao en Colombia.

Hasta el momento se encuentra que el flujo de capital humano se ve afectado por las respuestas comportamentales de los agentes cuando controlan los inventarios de producto en proceso y producto terminado en la cadena de suministro laboral. Sin embargo, a diferencia de los contextos estáticos donde se decide de una vez por todas, dicha afectación suele moderarse en tanto sus reacciones comportamentales son parte de una sucesión de decisiones en el horizonte temporal considerado.

Palabras Claves—Disrupciones, Efecto Ripple, Cadena de Suministro, Política de Inventarios.

I. INTRODUCTION

Una disrupción es un evento inesperado con una probabilidad baja de ocurrencia y efectos que se caracterizan por la interrupción del flujo de materiales y recursos dentro de una cadena de suministro causando impactos significativos en la eficiencia de la misma [1].

Las disrupciones pueden ser frecuentes o raras, leves o severas, pero todas tienen el potencial de desestabilizar las operaciones en la cadena de suministro. Comprender su dinámica es esencial para diseñar estrategias de gestión de riesgos y promover la resiliencia en la cadena de suministro [2, 3, 4].

Por su parte, la gestión de inventarios es un aspecto clave

dentro de la cadena de suministro laboral por lo cual constituye un área de investigación activa en la gestión de operaciones, particularmente cuando se examina la interacción entre los flujos de capital humano y los niveles de inventario en proceso (WIP) y productos terminados. En este sentido, la cadena de suministro laboral depende fundamentalmente de la disponibilidad de inventario para garantizar la continuidad de la producción debido a los procesos interconectados que rigen la asignación, el desarrollo y la distribución de los recursos humanos.

Las disrupciones en los tiempos de adquisición generan retrasos en el flujo de materiales y/o variabilidad en los plazos de entrega, afectando profundamente la eficiencia en la utilización de la mano de obra y el rendimiento general de la cadena de suministro. Estas disrupciones derivan incertidumbre que se propaga a lo largo de la cadena de suministro laboral, afectando a la planificación de la fuerza laboral, la programación y las estrategias de almacenamiento de inventario. A partir de lo anterior, este estudio se esfuerza por delinear estas dinámicas modelando las disrupciones del tiempo tomando como estudio de caso el proceso de cosecha en la gestión de la cadena de suministro laboral del cacao, enfatizando las interdependencias entre el capital humano y las existencias de inventario [5, 6].

La gestión de inventarios dentro de las cadenas de suministro se ha centrado tradicionalmente en la supervisión de los bienes físicos. Sin embargo, la integración del capital humano como componente dinámico presenta desafíos. La disponibilidad de mano de obra, los niveles de cualificación y la flexibilidad de la asignación tienen un impacto directo en el procesamiento de cultivos en curso (WIP) y en la tasa de producción de Cacao [7]. Las disrupciones en el tiempo de adquisición pueden provocar desequilibrios en el inventario, lo que resulta en escasez que hace que la mano de obra permanezca inactiva o en un exceso de inventario que ejerce presión sobre la capacidad de almacenamiento. Ambos escenarios interrumpen la productividad laboral y elevan los costos operativos. Estos escenarios subrayan la necesidad de incluir modelos robustos que permitan evaluar adecuadamente los bucles de retroalimentación entre la variabilidad de las adquisiciones, los niveles de inventario y la dinámica laboral.

Para abordar este problema, el estudio toma como punto de

Date submitted for review: 24/05/2025.

Arango, C. es docente del programa de Ingeniería Industrial de la Universidad de la Salle, Bogotá Colombia (e-mail: cararango@unisalle.edu.co).

Guerra, D. es docente del programa de Ingeniería Industrial de la Universidad de la Salle, Bogotá Colombia (e-mail: ferguerave@unisalle.edu.co).

Hualpa, A., es docente del programa de Ingeniería Industrial de la Universidad de la Salle, Bogotá Colombia (e-mail: amhualpa@unisalle.edu.co).

partida la heurística de anclaje y ajuste propuesta por Tverski & Kahneman [8], aplicada por Sterman [9] a cadenas de suministro laboral en el marco de dinámica de sistemas. La idea es simular el flujo de capital humano junto con los cultivos en curso (WIP) y los inventarios de productos terminados en presencia de disrupciones que impactan la duración del cultivo de cacao. el modelo propuesto permite identificar umbrales críticos en los que las interrupciones comprometen la estabilidad del sistema.

La metodología considerada está soportada en investigaciones destacando el papel de las políticas de gestión de inventarios como amortiguadoras en presencia de incertidumbre en la cadena de suministro [10]. A través de una primera indagación, el análisis presentado busca enriquecer la literatura al proporcionar información sobre la caracterización de la respuesta de la cadena de suministro laboral permitiendo generar estrategias de gestión resiliente de la cadena de suministro laboral.

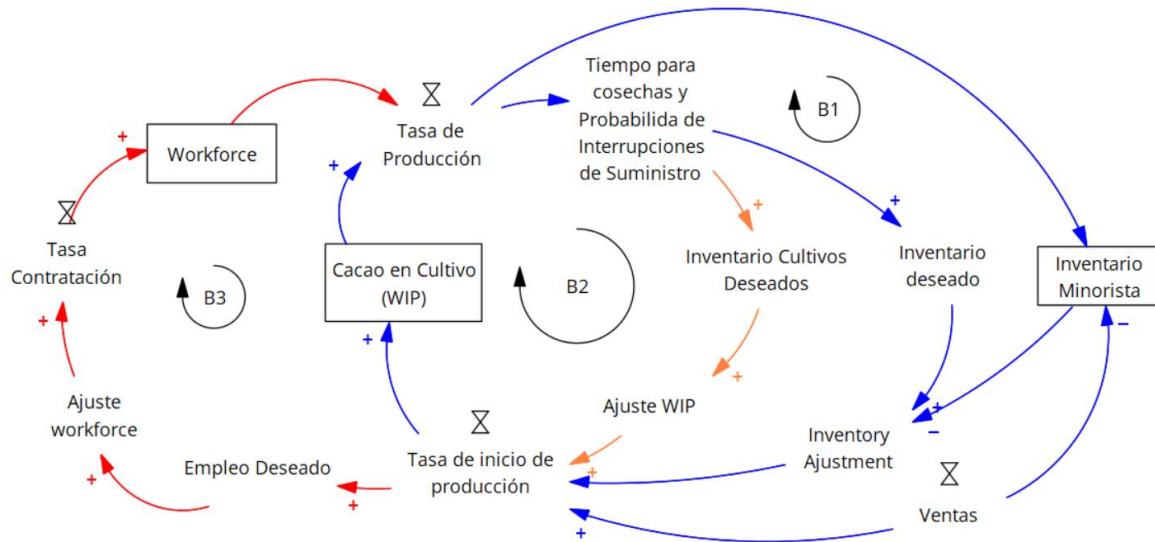


Fig. 1. Diagrama Causal Cadena de suministro Laboral Caso - Cacao

II. DEFINICIÓN DEL SISTEMA Y VARIABLES

La cadena de suministro del cacao en el sector agrícola se define como un sistema dinámico compuesto por múltiples actores y procesos interrelacionados, que van desde la siembra hasta la exportación del producto. Este sistema incluye tanto la producción agrícola como el acopio del cacao, además de incorporar la dimensión de la cadena laboral que sostiene los cultivos antes relacionados. Los actores clave que intervienen en esta red son los productores agrícolas y los intermediarios, quienes articulan la oferta con la demanda tanto en el mercado local como internacional.

Los límites del sistema se han definido para incluir desde el inicio del cultivo de cacao hasta su venta en mercados a distribuidores, permitiendo un análisis que abarque las transformaciones físicas del producto y los flujos logísticos asociados. Este análisis se realiza a mediano y largo plazo, dado que muchos de los procesos agrícolas y logísticos presentan comportamientos acumulativos y retardos temporales significativos.

En cuanto a las variables de estado del sistema se identifican el cacao en cultivo, el inventario minorista y la fuerza laboral. Asimismo, las variables de flujo incluyen la tasa de producción, la tasa de inicio de producción, las ventas, y la tasa de contratación de mano de obra. Las variables auxiliares que influyen sobre los flujos son el inventario deseado, el tiempo promedio de permanencia del inventario, el cubrimiento del

inventario y el inventario de cultivos deseado (Ver figura 1).

Sterman [11] formuló el primer modelo de cadena de suministro dinámica con la cobertura de inventario endógena deseada y las líneas de suministro de inventario bajo pedido. La figura 1 muestra la interacción entre un minorista, el cacao en cultivo y el empleo necesario para el proceso productivo. Para mantener niveles de producción adecuados el sistema mantiene inventario representado por el cacao en cultivo WIP. El canal también mantiene una línea de suministro de pedidos con la fuerza laboral. Los Minoristas realizan pedidos a los cultivadores para (1) reponer los kg de cacao vendido a los clientes, (2) cerrar cualquier brecha entre su nivel de inventario disponible deseado y real, y (3) cerrar cualquier brecha entre el nivel de cacao en cultivo deseado y real. Estas acciones forman varias retroalimentaciones negativas.

En primer lugar, si el inventario del minorista cae por debajo del nivel objetivo, los gestores piden más cacao a sus proveedores. Los cultivadores aumentan los envíos, aunque tardan en hacerlo. Las entregas a los minoristas elevan los inventarios disponibles, cerrando el ciclo negativo (de equilibrio) de Control de Inventario (B1). En segundo lugar, los cultivadores ajustan el nivel de sus cultivos para mantener la línea de suministro de pedidos no completados a un nivel suficiente para satisfacer la demanda, cerrando el ciclo de control del cacao en cultivo (B2). En tercer lugar, un aumento en los cultivos lleva a los administradores a elevar la cantidad de personal necesario para atender los cultivos cerrando el ciclo

negativo de disponibilidad (B3).

El sistema presenta relaciones causales con retroalimentaciones negativas, por ejemplo: un mayor tiempo de entrega desde las fincas genera una necesidad de mayor inventario deseado, lo que incrementa la discrepancia entre inventario actual y deseado, provocando así más pedidos, y consecuentemente, un aumento en los pedidos atrasados. Estas interacciones permiten observar cómo pequeñas demoras pueden amplificarse en la cadena.

III. FORMULACIÓN DEL MODELO

Las cadenas de suministro del cacao integra múltiples actores, las cuales, a partir de las expectativas del mercado, ajustan los cultivos y la capacidad de producción para satisfacer los cambios en la demanda. Cada eslabón mantiene y controla los inventarios de materiales y cultivo en proceso. El comportamiento de una cadena de suministro está regido por la realimentación negativa. Esto quiere decir que el retraso de fase causa su comportamiento inherente de oscilación y amplificación.

Inicialmente se considera la regla de decisión usada por Sterman [12] basada en la heurística de anclaje y ajuste de Tversky & Kahneman [8]. En esta, los minoristas asumen como ancla la cantidad que esperan que el distribuidor les ordene (Ventas Esperadas) y luego van ajustando esta cantidad de modo que se minimice la discrepancia entre el inventario actual y el deseado. Adicionalmente, se reduce la discrepancia entre el estado actual y deseado de los cultivos de cacao. Se parte de la tasa de inicio de producción (TIP) la cual no puede ser negativa.

$$TIP_t = \max(0, V_t^e + IA_t + AWIPC_t) \quad (1)$$

El ajuste del inventario del minorista IA_t se asume lineal y con discrepancia entre los niveles deseados y el inventario neto. Esta diferencia se divide en el tiempo para el ajuste del inventario TAI .

$$IA_t = \frac{(S_t^* - S_t)}{TAI} \quad (2)$$

Además, el ajuste de la línea de suministro $AWIPC_t$ incluye el tiempo de ajuste del inventario de cultivos TAW , en este caso de cacao.

$$AWIPC_t = \frac{(WIPC_t^* - WIPC_t)}{TAW} \quad (3)$$

Se asume que tanto S^* como $WIPC^*$ tienen dinámica en función del cambio en las ventas,

$$S^* = Inv. Cov. \times V_t = CI \times \frac{S}{AVL} \quad (4)$$

$$WIPC_t^* = \left(\frac{(S_t^* - S_t)}{TAI} + \frac{S}{AVL} \right) T_{Cosecha} \quad (5)$$

En particular para $WIPC_t^*$, se asume que el decisor realiza su estimación contemplando además de las ventas, el ajuste del Inventario del minorista y el tiempo de cosecha. De este modo, la tasa de inicio de producción (TIP) puede reescribirse.

$$TIP_t = \max \left\{ 0, CI \times \frac{S_t}{AVL} + \frac{(S_t^* - S_t)}{TAI} + \frac{(WIPC_t^* - WIPC_t)}{TAW} \right\} \quad (6)$$

En cuanto al componente laboral, se modela la tasa de contratación siguiendo el heurístico de anclaje y ajuste tomando como ancla la sustitución de las renuncias de empleados y adicionando el ajuste de la fuerza laboral de acuerdo a la discrepancia entre la fuerza laboral esperada E^* , la fuerza laboral actual E y su respectivo tiempo de ajuste del empleo TAE (medido en unidades de tiempo)

$$\frac{dE}{dt} = \frac{E}{DPE} + \frac{E^* - E}{TAE} \quad (7)$$

La formulación asume que el ajuste es lineal en la discrepancia entre el nivel de empleo deseado E^* y el nivel de empleo actual E . A su vez, E^* tiene su propia dinámica derivada de la tasa de inicio de producción y la productividad PD por empleado

$$E^* = \frac{TIP_t}{PD} \quad (8)$$

de tal forma que la tasa de contratación se resume en la expresión:

$$\frac{dE}{dt} = \frac{E}{DPE} + \frac{\frac{TIP_t}{PD} - E}{TAE} \quad (9)$$

El ajuste para la fuerza laboral crea el bucle de retroalimentación negativo de Control Laboral. El ajuste laboral forma un proceso lineal de retroalimentación negativa. TAE es el tiempo de ajuste del empleo (medido en unidades de tiempo). El ajuste laboral forma un proceso lineal de retroalimentación negativa.

La inclusión de la tasa de deserción esperada de empleados TDE en la formulación de la tasa de Contratación deseada TCD es una suposición conductual importante. Las pérdidas esperadas se incluyen por dos razones: en primer lugar, omitir el reemplazo de las pérdidas esperadas conduce a un error de estado estacionario en el que el trabajo difiere de su valor deseado incluso en equilibrio. El error de estado estacionario significa que una brecha entre los estados deseado y real persiste incluso después de que el sistema ha tenido tiempo de establecerse en su estado estacionario (es decir, incluso después de que las relaciones entre las variables de estado dejan de cambiar). El error de estado estacionario puede surgir a pesar de la existencia de un bucle de retroalimentación negativa, como el bucle de ajuste de empleo, que se esfuerza por eliminar las discrepancias entre el estado deseado y el real del sistema. Finalmente, la tasa de contratación deseada:

$$TCD^* = Tasa\ deserción + ALE \quad (10)$$

El escenario base utiliza los siguientes valores: Tiempo promedio del inventario de cacao de minoristas $AVL=4$ meses, Tiempo para cosecha $=19.5$ meses, para este se toman en cuenta que en las fincas cacaoteras existen cultivos en distintas etapas productivas, [13, 14, 15], Tiempo de ajuste del inventario y tiempo para ajuste de inventarios en cultivo $= 12$ meses.

IV. SIMULACIÓN Y ANÁLISIS DE ESCENARIOS

El modelo, desarrollado en Vensim PLE, tiene como objetivo analizar el comportamiento dinámico de la cadena de suministro del cacao desde la siembra hasta el distribuidor incorporando elementos como la producción agrícola, el acopio, los inventarios y la estructura laboral. El sistema cuenta con tres variables de estado: el inventario minorista y el inventario de cacao en cultivo. Este último representa los cultivos de cacao entre la producción primaria y la llegada al punto inicial de acopio.

4.1 Escenario Base: condiciones estables

El escenario base representa un sistema estable en el que no se presentan perturbaciones ni alteraciones en los parámetros clave. Los valores del AVL y $T_{Cosecha}$ se mantienen constantes y en equilibrio, de modo que las tasas de producción y de ventas son iguales. En estas condiciones, la discrepancia entre el inventario minorista y el inventario deseado es nula, al igual que la discrepancia del inventario en proceso. Por lo tanto, los ajustes que se aplican a los flujos son mínimos o inexistentes, y la tasa de inicio de producción coincide con las ventas reales.

El sistema se mantiene balanceado, sin acumulación de inventarios ni oscilaciones, lo que refleja una operación eficiente y predecible a lo largo del tiempo de simulación.

$$\frac{dS}{dt} = 0 ; \frac{dWIP}{dt} = 0 ; \frac{dE}{dt} = 0$$

Para (S)

$$\frac{dS}{dt} = \frac{WIP}{T_{Cosecha}} - \frac{S}{AVL} = 0$$

$$\frac{WIP}{T_{Cosecha}} = \frac{S}{AVL}$$

Dado que el flujo de Cacao es:

$$\frac{S}{AVL} = 6.140 \frac{\text{ton.}}{\text{mes}}; \quad AVL = 4 \text{ meses}$$

$$S = 6.140 \times 4 = 24.560 \text{ toneladas}$$

Con $T_{Cosecha}=19.5$ antes de $t=10$

$$\frac{WIP}{T_{Cosecha}} = \frac{S}{AVL} = 6.140 \text{ ton./mes}$$

$WIP = 6.140 \times T_{Cosecha} = 119.730 \text{ toneladas}$ las cuales representan la cantidad de cacao sembrado para Colombia. En el caso del empleo, la tasa de inicio de producción la cual en equilibrio es igual a 6.400 ton/mes y con una productividad para Colombia de 450 kg-año/hectárea [16, 17, 18]

$$\frac{dE}{dt} = \frac{\frac{TIP_t}{PD} - Empleo}{TAE}; \quad Empleo = \frac{6.140 \frac{\text{Ton}}{\text{mes}} \times 0.5 \text{ em} - ha/}{0,037 \text{ ton} - ha/me}$$

El Empleo= 81.866 lo cual corresponde a los empleos directos creados por los cultivos de cacao actuales en Colombia.

Inmediatamente después que $t=10$, con $WIP=117.930$ (ya que los inventarios no cambian instantáneamente) se cambia el tiempo de cosecha, para simular esta situación se utilizan Retrasos materiales (Delays). Se modelan como stocks (acumulación) y flujos con una función de retraso en este caso se utiliza la función Pulse la cual introduce un cambio temporal y breve en el tiempo de cosecha, como un "pico" o "choque". La variable aumenta (en $t=10$ llevando el tiempo de cosecha a 30 meses) durante tres periodos y luego regresa a su valor original (Ver figura 2). Posteriormente en $t=60$ se crea otra disrupción, en este caso el tiempo de cosecha pasa a 9,5 meses, para las simulaciones se utiliza un Delay Fixed de seis periodos.

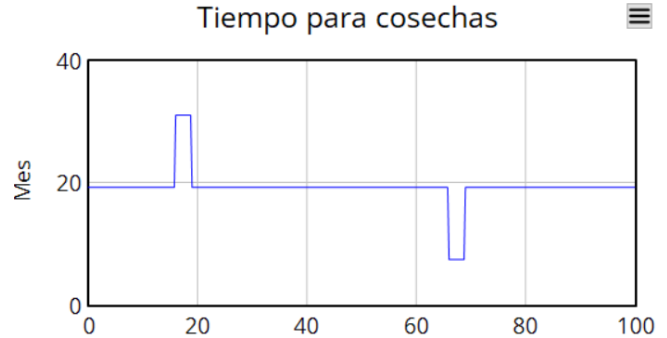


Figura 2. Retrasos en tiempos de cosecha utilizando función Pulse

En circunstancias normales, la demanda de cacao es creciente por lo que los cultivadores venden toda su producción y en general son capaces de cumplir con las expectativas en su totalidad dentro del plazo de entrega normal. Los minoristas cuentan con pequeños stocks de seguridad (las existencias de seguridad solo serían necesarias para amortiguar la incertidumbre en las ventas).

La dinámica del sistema está dominada por las tres retroalimentaciones negativas descritas anteriormente. Ahora, sin embargo, en el tiempo 10 la capacidad de los cultivadores para realizar envíos no se cumple. La causa podría ser un cambio climático o un desastre naturales. A medida que el tiempo de cultivos aumenta, la tasa de producción cae por debajo de los niveles deseados. Para garantizar un suministro constante de entregas, los Minoristas ahora deben realizar pedidos con más anticipación. En el caso mencionado, si el plazo de entrega de los cultivadores aumenta de 19,5 a 30 meses, los minoristas (en equilibrio) deben mantener 6.400 toneladas de cacao en pedido. Los pedidos aumentan aún más, mientras que las asignaciones disminuyen y la fiabilidad de la entrega disminuye, cerrando la retroalimentación positiva (reforzante) de Ordenes Fantasma, R1. Además, al aumentar los tiempos de cosecha, los cultivadores aumentan sus existencias de seguridad objetivo. Para ello, deben pedir aumentar los cultivos, lo que desencadena existencias de seguridad objetivo aún mayores y cierra la retroalimentación positiva del acaparamiento, R2. Contrariamente al funcionamiento normal de los mercados, la escasez aumenta la demanda (Ver figura 3). Adicional podemos observar cómo los factores comportamentales afectan a las variables de estado inventario en cultivo y Numero de Empleados (Workforce), cuando ocurre la disrupción en el tiempo de cosecha, el número de cultivos en Cacao se reduce, sin embargo los decisores a

pesar de realizar una reducción inicial de fuerza laboral, al darse cuenta del aumento en cultivos reaccionan de forma más abrupta, además el número de operarios sigue aumentando por cinco periodos antes de reconocer el impacto en la reducción de cultivos.

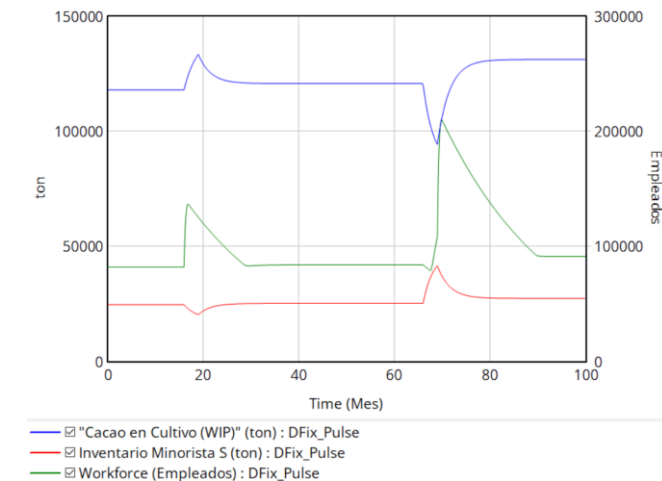


Figura 3. Comportamiento variables de estado

La Figura 4 muestra la dinámica del modelo. El tiempo de cosecha sigue un ciclo que depende de la edad de cultivo. El aumento en el tiempo de cosecha reduce la tasa de producción de los cultivadores reduciendo el inventario del minorista, y las ventas realizadas por ellos, los minoristas deben aumentar los pedidos realizados a los cultivadores por encima de las ventas realizadas. La capacidad de los cultivadores es limitada debido a los aumentos de tiempo para realizar la cosecha. Los cultivadores responden realizando pedidos fantasmas y aumentando el cacao en cultivo, impulsando los cultivos por encima de las ventas finales, adicional la tasa de contratación responde de manera brusca pasando de 3.411 operarios mensuales a contratar 164.416 operarios en el periodo 16, inmediatamente sucede la disrupción, podemos observar el impacto de los factores comportamentales en esta decisión y en la que ocurre inmediatamente en periodos posteriores, significando que no se contraten personas por cerca de 20 periodos. Las retroalimentaciones positivas dominan la dinámica.

Ahora en $t=60$ se revierten los efectos a causa de una disrupción en el tiempo de cosecha. Con el aumento de la tasa de producción y la disponibilidad más fácil de cacao, los minoristas cancelan sus pedidos fantasmas. De esta manera, con entregas más rápidas y fiables no es necesario mantener niveles de existencia grande para los cultivos existentes. La tasa de inicio de producción se desploma, ya que el tiempo de cosecha también se redujo. Los cultivadores luchan por reducir la tasa a la que crece el inventario en cultivo, pero el retraso en el tiempo de cultivo conduce a la acumulación de inventarios excesivos.

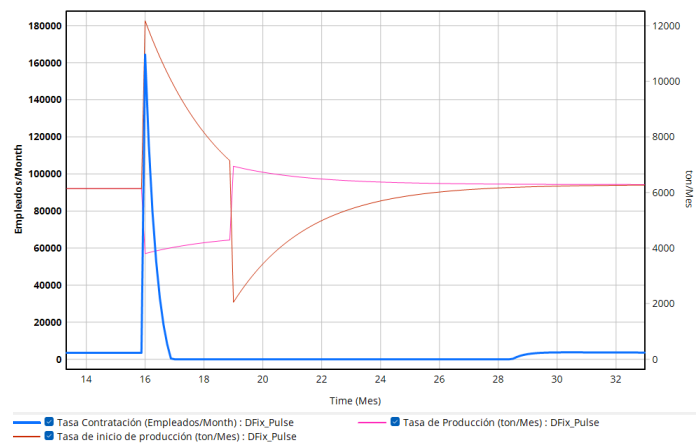


Figura 4 Comportamiento variables de flujo

El sistema encuentra su equilibrio hacia el mes 51, sin embargo, las ordenes en este punto son de 6.283,25 ton indicando un aumento de 163,18 kg (2.3%).

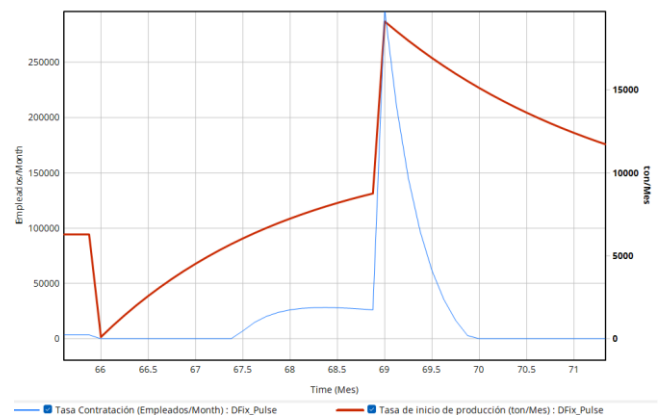


Figura 5. Tasa de Contratacion vs Tasa de Inicio de Produccion

La figura 5 muestra fluctuaciones significativas entre los 60 y 70 meses, que pueden atribuirse a la presencia de órdenes fantasmas. En el gráfico, los picos en empleados, producción y ventas sugieren una sobreproducción frente a una demanda no plenamente realizada. Estos patrones son coherentes con las características de las órdenes fantasmas, que introducen distorsiones temporales en las dinámicas del sistema. Las órdenes fantasmas se modelan como un choques exógenos o variables de ruido. El gráfico 5 pone de manifiesto cómo, bajo presión de demandas impredecibles, los agentes tienden a aplicar atajos mentales que provocan amplificaciones sucesivas en la cadena de decisiones —ya sea en la contratación de personal o en la tasa de inicio de producción—, desembocando en patrones oscilatorios. La comprensión de estos fenómenos comportamentales resulta clave para diseñar mecanismos de realimentación más objetivos (por ejemplo, sistemas de señalamiento de inventarios o indicadores de eficiencia de entrenamiento) que mitiguen tanto la sobrecontratación reactiva como la subutilización de capacidad. Su impacto se amplificó mediante los bucles de retroalimentación, como ajustes de inventario o políticas de contratación.

Después del mes 60 tenemos una nueva oscilación generada por un pulso. Sin embargo, en este caso el pico brusco a los 65

meses indica una perturbación ocasionada por un cambio repentino en S^* y en el WIP.

Si el inventario deseado de cacao en cultivo incrementa de forma repentina, el ajuste realizado por el inventario de producto en proceso (Ajuste WIP) crece sustancialmente, llevando el inventario de producto en proceso a crecer de forma rápida.

El sistema se estabiliza alrededor de 6806,64 kg/mes lo que indica un nuevo equilibrio después de la perturbación. Esto podría ser el resultado de: Un cambio permanente en S^* o WIP*, cambiando el equilibrio.

V. CONCLUSIONES

Se muestra un sistema con retardo de material de segundo orden en el que el inventario de cacao en cosecha alimenta a los inventarios del minorista. Antes de $t=10$, el sistema está en equilibrio con flujos equilibrados.

Se simulan disrupciones utilizando la función pulse, esta simula la situación en la cual tenemos un evento temporal y breve como puede ser un desastre natural. En $t = 10$, el aumento del tiempo de cosecha de 19.5 a 30 reduce el flujo de salida del inventario del minorista causando que el cacao en cultivo se acumule y el inventario del minorista disminuya debido a un menor flujo de entrada. Los términos de la retroalimentación ajustan las ordenes de producción, conduciendo el sistema a un nuevo equilibrio. Las pequeñas ganancias de retroalimentación dan como resultado una respuesta sobre amortiguada, con transiciones suaves al nuevo equilibrio sin oscilaciones.

En el análisis conjunto de las variables de flujo —la tasa de contratación de empleados y la tasa de inicio de producción— se observan dinámicas claramente no lineales y marcadas por respuestas comportamentales que trascienden la simple lógica de equilibrio. Ante una desaceleración de la producción, los responsables de los cultivos incrementan de manera abrupta la contratación, este comportamiento ilustra un sesgo de disponibilidad y heurística de precaución: ante la amenaza de no satisfacer la demanda, se recurre a “pedir más personal por si acaso”, generando un sobreajuste que luego debe corregirse. El retraso entre el pico de contrataciones y el máximo productivo refleja tiempos de inducción y aprendizaje, propios de cualquier sistema productivo, pero también pone de manifiesto la tendencia humana a subestimar dichos retardos (optimism bias) y a compensar tarde los desequilibrios. Tras alcanzar un punto máximo de producción, se produce una fuerte reducción en la tasa de contratación conforme los indicadores internos confirman un exceso de capacidad. Este vaivén exhibe el fenómeno de órdenes fantasmas en la contratación de personal.

Las variables de estado confirman que la aplicación de la heurística de anclaje y ajuste conduce a ciclos de sobre aprovechamiento de capacidad y acumulación de inventarios. Para mitigar estas oscilaciones, es imprescindible diseñar políticas que incorporen indicadores de rotación de stock y tiempos de ciclo, así como mecanismos de ajuste rápido basados en información de demanda en tiempo real. Solo así podrá alcanzarse una gestión más estable y eficiente de los inventarios distributivos y de producción.

Desde la óptica de las heurísticas adaptativas [19] [20], estas decisiones no son necesariamente irracionales: responden a reglas simples (“contrato rápido—produce rápido”) diseñadas para entornos de alta incertidumbre. No obstante, al ignorar información sobre el tiempo de cosecha y la verdadera capacidad instalada de los cultivos WIP, dichas heurísticas introducen distorsiones que redundan en ciclos de sobre aprovisionamiento y subutilización. Los resultados obtenidos demuestran que tanto el AVL como el Tiempo de cosecha son parámetros críticos que afectan directamente la estabilidad del sistema. Un aumento del AVL deteriora la capacidad del sistema para estimar adecuadamente la demanda, mientras que un incremento del Tiempo de cosecha reduce la capacidad de respuesta de la oferta. En ambos casos, se generan ciclos de retroalimentación negativa que afectan el equilibrio de los inventarios, provocando acumulaciones o desabastecimientos según el caso.

REFERENCIAS

- [1] D. Ivanov, Introduction to supply chain resilience: Management, modelling, technology, Springer Nature, 2021.
- [2] M. S. Golan, L. H. Jernegan and I. Linkov, "Trends and applications of resilience analytics in supply chain modeling: systematic literature review in the context of the COVID-19 pandemic," *Environment Systems and Decisions*, vol. 40, no. 2, pp. 222-243, 2020.
- [3] M. M. Queiroz, D. Ivanov, A. Dolgui and S. Fosso Wamba, "Impacts of epidemic outbreaks on supply chains: mapping a research agenda amid the COVID-19 pandemic through a structured literature review," *Annals of operations research*, vol. 319, no. 1, pp. 1159-1196, 2022.
- [4] J. El Baz and S. Ruel, "Can supply chain risk management practices mitigate the disruption impacts on supply chains' resilience and robustness? Evidence from an empirical survey in a COVID-19 outbreak era," *International journal of production economics*, p. 2021, 2021.
- [5] D. Ivanov, B. Sokolov and A. Dolgui, "The Ripple effect in supply chains: trade-off 'efficiency-flexibility-resilience' in disruption management," *International Journal of Production Research*, vol. 52, no. 7, pp. 2154-2172, 2014.
- [6] A. Dolgui, D. Ivanov and B. Sokolov, "Ripple effect in the supply chain: an analysis and recent literature," *International journal of production research*, vol. 56, no. 1-2, pp. 414-430, 2018.
- [7] W. J. Hopp and M. L. Spearman, Basic factory dynamics. Factory Physics, 3rd ed., New York: McGraw-Hill, 2008.
- [8] A. Tversky and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases: Biases in judgments reveal some heuristics of thinking under uncertainty," *Science*, vol. 185, no. 4157, pp. 1124-1131, 1974.

- [9] J. Sterman, *Business Dynamics: Systems Thinking and Modeling for a Complex World*, McGraw Hill, 2000.
- [10] S. Chopra and P. Meindl, *Supply Chain Management—Strategy, Planning, and Operation 6 th Edition.*, 2016.
- [11] J. W. Forrester, *Industry dynamics.*, Cambridge, Massachusetts, 1961.
- [12] J. D. Sterman, "Modeling managerial behavior: Misperceptions of feedback in a dynamic decision making experiment," *Management science*, vol. 35, no. 3, pp. 321-339, 1989.
- [13] L. L. Nalley, B. L. Dixon and J. S. H. Popp, "Increasing Profitability of Small Scale Orchard Producers through Optimizing Replacement Rate: The Case Study of Ghana.," in *Southern Agricultural Economics Association Annual meeting*, Birmingham, 2012.
- [14] F. Aneani, V. M. Anchirinah, F. Owusu-Ansah and M. Asamoah, "Adoption of some cocoa production technologies by cocoa farmers in Ghana," *Sustainable Agriculture Research*, vol. 1, no. 1, 2012.
- [15] F. M. Aneani, V. M. Anchirinah, F. Owusu-Ansah and M. Asamoah, "An analysis of the extent and determinants of crop diversification by cocoa (Theobroma cacao) farmers in Ghana.," *African Journal of Agricultural Research*, vol. 6, no. 18, pp. 4277-4287, 2011.
- [16] E. Antolinez, P. Almanza, A. Barahona, E. Polanco and P. Serrano, "Estado actual de la cacaocultura: una revisión de sus principales limitantes," *Ciencia y Agricultura*, vol. 17, no. 2, 2020.
- [17] P. Abbott, T. Benjamin, G. Burniske, M. Croft, M. Fenton, C. Kelly, M. Lundy and F. & W. M. Rodriguez-Camayo, *Análisis de la Cadena Productiva del Cacao en Colombia*, Purdue University , 2019.
- [18] E. E. Escobar-Quíñonez, M. Ramírez Mutis and L. F. Pérez Ángel, *Análisis de la competitividad de los pequeños productores de cacao con fairtrade (Working papers N.º1).*, Cali: Ediciones Universidad Cooperativa de Colombia, 2022.
- [19] G. Gigerenzer and W. Gaissmaier, "Heuristic Decision Making," *Annual Review of Psychology*, vol. 62, pp. <https://doi.org/10.1146/annurev-psych-120709-145346>, 2011.
- [20] K. Katsikopoulos, "Psychological heuristics for making inferences: Definition, performance, and the emerging theory and practice," *Decision Analysis*, vol. 8, pp. 10-29. <http://dx.doi.org/10.1287/deca.1100.0191>, 2011.

Persistence Bias and the Bullwhip Effect: An Agent-Based Model of a Behavioral Driver in Supply Chains

Edgar Cerón, *PhD Student, Tecnológico de Monterrey*, Claudia P. Estévez-Mujica, *Assistant Professor, Pontificia Universidad Javeriana*, and Eva S. Hernández-Gress, *Professor, Tecnológico de Monterrey*.

Abstract—Cognitive biases significantly influence decision-making in supply chains, yet the specific impact of persistence bias, understood as the tendency to commit with previous decisions despite evolving conditions, remains insufficiently examined. This study introduces an agent-based model of a four-echelon supply chain (retailer, wholesaler, distributor, and manufacturer) to examine how individual resistance to change (RTC) interacts with performance feedback to shape replenishment behavior, ultimately leading to the Bullwhip Effect (BWE). The model reflects the decentralized and dynamic characteristics of real-world supply chains, embedding behavioral rules modulated by agents' RTC levels. Agents revise their ordering decisions based on outcome feedback: low-RTC agents are more responsive to negative signals (e.g., stockouts), while high-RTC agents are prone to maintaining previous decisions. The analysis investigates to what extent varying RTC levels influence demand amplification as orders propagate upstream. The model emphasizes the behavioral roots of supply chain volatility and the relevance of integrating psychological factors into simulation-based supply chain research.

Index Terms—Agent-based modeling, Beer Game, Bullwhip effect, Persistence bias, Resistance to change.

I. INTRODUCTION

In multi-echelon supply chains, replenishment decisions are made by individual actors that do not always follow purely rational principles. Although traditional models assume that decision-makers make optimal decisions based on complete information and logical processing, empirical evidence shows that individuals are influenced by cognitive limitations [1], heuristics [2], and biases [3]. Among these, persistence bias, defined as the tendency to continue prior decisions regardless of changing circumstances [4], has received comparatively little attention, despite its potential to explain inertial ordering behavior that contributes to demand amplification across the supply chain.

This tendency is consistent with the observation by Samuelson and Zeckhauser [5] that, when faced with new options, decision-makers often prefer doing nothing or maintaining previous decisions. Such inertia is not merely situational but can also reflect a deeper psychological predisposition to resist change, even when it can be beneficial.

The Resistance to Change (RTC) scale, developed by Oreg [6], captures this predisposition by measuring the extent to which individuals avoid or devalue change. Thus, RTC can be viewed as a behavioral moderator that influences the strength of persistence bias in dynamic decision-making contexts. Recent research further suggests that feedback plays a crucial role in modulating the influence of RTC. Crivelli et al. [7] shows that decision-makers with high RTC are more likely to persist with previous choices when feedback is interpreted as positive or neutral, whereas negative feedback can encourage them to reassess and adjust their decisions. Within supply chains, this dynamic implies that the interaction between RTC and environmental feedback may substantially affect replenishment behaviors.

Decision-makers may interpret positive feedback, such as high service levels or low inventory costs, as an incentive to maintain their established ordering practices. Conversely, they may view negative feedback, such as stockouts or high inventory costs, as a signal to revise their ordering strategy, although their RTC level may still prevent them from doing so. Ultimately, the decision to update (or not update) the ordering policy can lead to systematic deviations from the optimal order level, thereby contributing to the bullwhip effect (BWE).

The BWE has been traditionally studied from an operational perspective, focusing on the structural characteristics of a system that leads rational decision-makers to amplify demand. However, as Croson et al. [8] stated, even when operational causes (order batching, demand signaling, and price fluctuations, among others) are removed, the BWE remains due to behavioral factors related to human judgment and decision-making.

This study presents an agent-based model of a four-echelon supply chain (retailer, wholesaler, distributor, and manufacturer), in which each agent makes periodic replenishment decisions influenced by their RTC level and by feedback received from prior performance. The model is designed to implement a decision rule in which the RTC level moderates the probability to maintain or revise decisions. How this behavior influences inventory dynamics, costs, and demand amplification is also examined.

Edgar Cerón (Corresponding author) is with the Tecnológico de Monterrey, Pachuca, HGO 42080 MEX (e-mail: eceron@tec.mx).

Claudia P. Estévez-Mujica is with Pontificia Universidad Javeriana, Bogotá, 110231 COL (e-mail: estevez_cp@javeriana.edu.co).

Eva S. Hernández-Gress is with the Tecnológico de Monterrey, Pachuca, HGO 42080 MEX (e-mail: evahgress@tec.mx).

Supplementary materials, including the model's ODD (Overview, Design Concepts, and Details) protocol and its NetLogo implementation, are available upon request by email.

II. REPLENISHMENT DECISIONS IN A FOUR-ECHELON SUPPLY CHAIN.

This study examines replenishment decisions in a four-echelon supply chain that replicates the roles and structure of the widely used Beer Distribution Game (BDG). Originally developed at MIT, the game simulates the flow of supply and demand across four supply chain members: retailer, wholesaler, distributor, and factory [2] (see Fig. 1). During the game, each participant is responsible for fulfilling incoming orders while placing replenishment orders with the upstream member. The primary objective is to satisfy customer demand efficiently, keeping inventory levels as low as possible while avoiding stockouts and backorders.

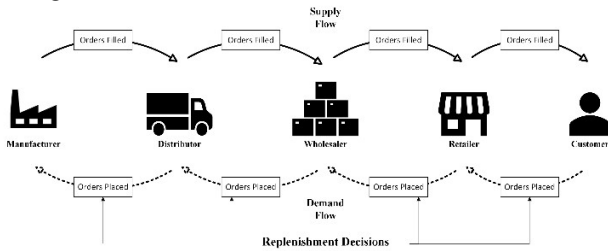


Fig. 1. Replenishment Decisions in a Four-Echelon Supply Chain.

The BDG has become a widely adopted framework in supply chain research due to its ability to capture core dynamics such as inventory management challenges and decentralized decision-making [8]. These characteristics make it a powerful tool for evaluating behavioral and systemic phenomena in supply chains. Researchers have successfully implemented BDG models using various computational approaches, including system dynamics [9], agent-based modeling [10], and optimization techniques [11], often obtaining comparable and insightful results about the behavioral causes of the bullwhip effect [12].

The replenishment process during the game follows a simple yet structured logic. Shipments from upstream members are received and added to inventory, transferring goods from their point of origin to the next echelon in the supply chain (for example, from manufacturer to distributor). At the same time, orders are placed by the immediate downstream member. If sufficient inventory is available, these orders are fulfilled; otherwise, they are registered as backorders. Unlike physical shipments, which flow downstream, orders travel upstream through the supply chain. While fulfilled orders exit the system once completed, unfulfilled ones persist as backorders and accumulate at their corresponding stage. Finally, a new order is placed and transmitted to upstream members.

III. MODEL DESCRIPTION.

The model description follows the ODD (Overview, Design concepts, Details) protocol for describing individual- and agent-based models [13], as updated by Grimm et al. [14].

A. Purpose.

The overall purpose of our model is to illustrate how different levels of individual RTC shape inventory replenishment

dynamics, and how such behavior ultimately contributes to the emergence of the bullwhip effect.

Specifically, we are addressing the following question: *To what extent does individual RTC influence the BWE in a decentralized supply chain environment?* Consequently, the model should be able to replicate the following pattern: variation in the magnitude of the BWE as the level of RTC changes.

Following [10], we evaluate the BWE by comparing the variance of orders across successive echelons as in (1). Our model is considered valid if it reproduces the expected amplification of orders (i.e., when the standard deviation of a supply chain member's orders exceeds that of its immediate customer). The bullwhip effect's magnitude (BWE_M) is defined as the ratio between variance of orders of the last echelon (σ_N^2) and the variance of the final customer's demand (σ^2).

$$BWE_M = \sigma_N^2 / \sigma^2. \quad (1)$$

B. Entities, state variables, and scales.

The model includes the following entities: one final customer, four supply chain members (Retailer, Wholesaler, Distributor, Manufacturer), four demand links, four supply links, and an observer. The state variables characterizing these entities are listed in Table I. The temporal resolution is one week, and the model runs for 36 weeks, consistent with the standard game settings of the BDG [2].

C. Process overview and scheduling.

The most important processes of the model are: (1) model initialization, (2) shipping of available inventory, (3) receipt of deliveries, (4) placement of replenishment orders, and (5) performance reporting. During each simulated week, all supply chain members sequentially update their internal state variables based on a common event structure.

Fig. 2 shows the decision rule for replenishment decisions updated in response to feedback from previous decisions. Positive feedback (i.e., low inventory and no backorders) leads agents to maintain their previous ordering strategy (status quo). In contrast, negative feedback triggers a probabilistic decision: agents may update their order based on a behavioral replenishment rule, or maintain the status quo, depending on their RTC value.

D. Design concepts.

The most important design concepts of the model are the behavioral foundations that drive inventory decisions and the emergent properties resulting from individual-level rules. These are listed as following:

1) Basic principles

The model analyzes the BWE as a consequence of persistence bias. Depending on their RTC level, agents rely more-or-less on a heuristic-based decision rule that anchors on past orders and adjusts based on inventory imbalance, simulating the anchoring-and-adjustment process.

TABLE I
SUMMARY OF ENTITIES AND STATE VARIABLES

Entity	State Variable	Dynamic/Static	Range	Units	Description
Members	role	Static	{Retailer, Wholesaler, Distributor, Manufacturer}	-	Defines the function of each agent in the supply chain
	cost	Dynamic	≥ 0	\$	Total inventory and backlog cost
	inventory	Dynamic	≥ 0	units	On-hand inventory
	backorders	Dynamic	≥ 0	units	Unfulfilled demand
	supply-line	Dynamic	≥ 0	units	Units ordered but not yet received
	status-quo	Dynamic	≥ 0	units	Previous ordering decision retained unless updated
	feedback	Dynamic	{Positive, Negative}	-	Performance evaluation of past order
	ordered	Dynamic	≥ 0	units	Current order placed to the upstream member
	orders-placed	Dynamic	≥ 0	units/week	Order quantity transmitted upstream
	orders-filled	Dynamic	≥ 0	units/week	Quantity delivered downstream
Customer	clock	Dynamic	[1, 36]	week	Simulation time tracker

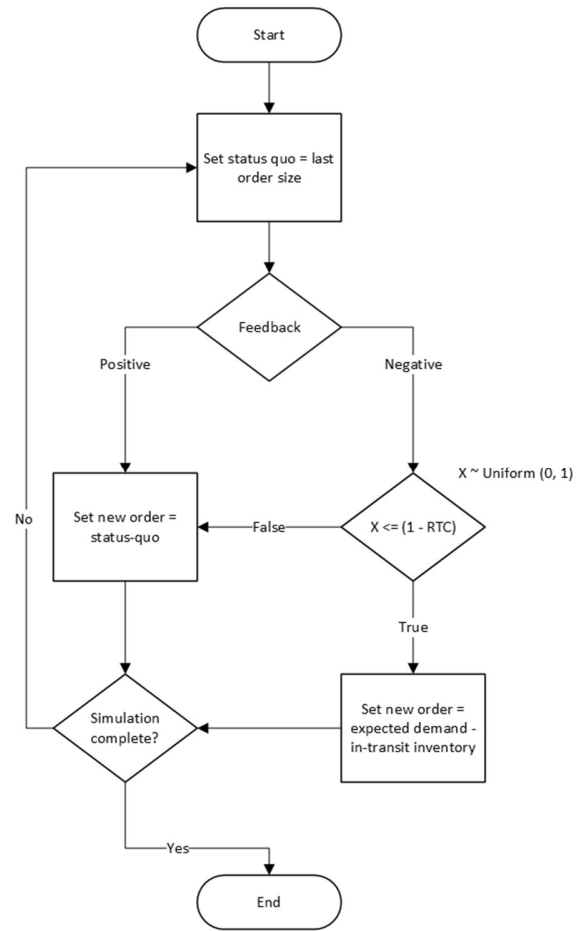


Fig. 2. Replenishment Decision Rule.

2) Emergence

The bullwhip effect emerges endogenously from the interactions of decentralized agents, driven primarily by their RTC levels and feedback evaluations. Order variability accumulates across echelons due to these adaptive behaviors.

3) Adaptation

Agents adapt their ordering strategies based on qualitative feedback from prior decisions. Positive feedback reinforces the status quo. Negative feedback triggers a probabilistic decision: with probability $1 - \text{RTC}$, the agent updates its order using the behavioral rule proposed by [10]; otherwise, it repeats the prior order.

4) Objectives

Each agent seeks to maintain low inventory levels while avoiding backorders. The decision variable ordered is updated weekly using feedback and the RTC parameter. The decision rule balances stability (via status quo) and responsiveness (via adaptive adjustment), capturing bounded rationality in decentralized settings.

5) Learning

Learning is not implemented. Agents do not modify behavior based on long-term trends or accumulated experience.

6) Prediction

Agents assume that future demand will mirror recent demand. This tendency has been empirically studied in previous studies [15][16].

7) **Sensing**

Agents have access only to local information: their own inventory and backlog levels, incoming orders from downstream partners, and received shipments from upstream partners. No global system-wide information is available.

8) **Interaction**

Interactions are limited to direct upstream and downstream connections. Orders flow upstream; shipments flow downstream. No global communication or coordination exists among agents.

9) **Stochasticity**

Stochasticity enters the model through a uniformly distributed random variable that determines whether an agent changes its ordering strategy after negative feedback. This captures observed variability in how individuals respond to performance evaluations.

10) **Collectives**

No collectives are modeled. Each agent operates independently.

11) **Observation**

Model outputs include plots of inventory levels, order quantities, and inventory costs for each agent. These are monitored to assess the emergence and magnitude of the bullwhip effect under different RTC levels.

E. Simulation experiments.

We will conduct a parameter variation experiment with 1,000 replications of the RTC parameter for each of the four agents (retailer, wholesaler, distributor, and manufacturer). The RTC value ranges from 0.1 to 0.9 in increments of 0.1, resulting in nine levels per agent and a total of 6,561 simulation runs. The experiment will be executed using NetLogo's BehaviorSpace tool. The response variable is the magnitude of the bullwhip effect, evaluated through order variance amplification across echelons. Results will first be analyzed descriptively (mean, variance, and distribution of bullwhip effect outcomes) and then complemented with sensitivity analysis to identify which agents' RTC levels exert the strongest influence on demand amplification.

IV. CONCLUSION

This work contributes by embedding psychological theory into a supply chain decision-making task, offering a model that explicitly analyzes how different levels of individual resistance to change (RTC) influence the generation of the bullwhip effect. The results aim to provide twofold contributions. For scholars, the model will advance the field of Behavioral Operations Research by addressing a largely unexplored bias in the supply chain literature, persistence bias, while grounding the analysis in well-established psychological theory. For practitioners, the model will show how decision-makers with different levels of RTC can modify replenishment dynamics in firms and affect overall supply chain

performance. This suggests the importance of considering the RTC level of those appointed to manage inventory.

REFERENCES

- [1] A. Narayanan and B. B. Moritz, "Decision Making and Cognition in Multi-Echelon Supply Chains: An Experimental Study," *Prod Oper Manag*, vol. 24, no. 8, pp. 1216–1234, Aug. 2015, doi: 10.1111/poms.12343.
- [2] J. D. Sterman, "Misperceptions of feedback in dynamic decision making," *Organ Behav Hum Decis Process*, vol. 43, no. 3, pp. 301–335, Jun. 1989, doi: 10.1016/0749-5978(89)90041-1.
- [3] M. Pournader, A. Narayanan, M. F. Kebliis, and D. Ivanov, "Decision Bias and Bullwhip Effect in Multiechelon Supply Chains: Risk Preference Models," *IEEE Trans Eng Manag*, 2023, doi: 10.1109/TEM.2023.3292348.
- [4] C. R. Carter, L. Kaufmann, and A. Michel, "Behavioral supply management: a taxonomy of judgment and decision-making biases," *International Journal of Physical Distribution & Logistics Management*, vol. 37, no. 8, pp. 631–669, Sep. 2007, doi: 10.1108/09600030710825694.
- [5] W. Samuelson and R. Zeckhauser, "Status quo bias in decision making," *J Risk Uncertain*, vol. 1, no. 1, pp. 7–59, Mar. 1988, doi: 10.1007/BF00055564.
- [6] S. Oreg, "Resistance to change: Developing an individual differences measure," *Journal of Applied Psychology*, vol. 88, no. 4, pp. 680–693, 2003, doi: 10.1037/0021-9010.88.4.680.
- [7] D. Crivelli, R. A. Allegratta, and M. Balconi, "The 'status quo bias' in Response to External Feedback in Decision-Makers," *Adaptive Human Behavior and Physiology*, vol. 9, no. 4, pp. 426–441, Nov. 2023, doi: 10.1007/s40750-023-00230-1.
- [8] R. Croson, K. Donohue, E. Katok, and J. Sterman, "Order Stability in Supply Chains: Coordination Risk and the Role of Coordination Stock," *Prod Oper Manag*, vol. 23, no. 2, pp. 176–196, Feb. 2014, doi: 10.1111/j.1937-5956.2012.01422.x.
- [9] J. D. Sterman and G. Dogan, "'I'm not hoarding, I'm just stocking up before the hoarders get here.': Behavioral causes of phantom ordering in supply chains," *Journal of Operations Management*, vol. 39–40, pp. 6–22, Nov. 2015, doi: 10.1016/j.jom.2015.07.002.
- [10] S. Cannella, C. Di Mauro, R. Dominguez, A. Ancarani, and F. Schupp, "An exploratory study of risk aversion in supply chain dynamics via human experiment and agent-based simulation," *Int J Prod Res*, vol. 57, no. 4, pp. 985–999, Feb. 2019, doi: 10.1080/00207543.2018.1497817.
- [11] M. H. Khan, S. Ahmed, and D. Hussain, "Analysis of Bullwhip effect: A Behavioral Approach," *Supply Chain Forum: An International Journal*, vol. 20, no. 4, pp. 310–331, Oct. 2019, doi: 10.1080/16258312.2019.1661756.

- [12] B. B. Moritz, A. Narayanan, and C. Parker, "Unraveling Behavioral Ordering: Relative Costs and the Bullwhip Effect," *Manufacturing & Service Operations Management*, vol. 24, no. 3, pp. 1733–1750, May 2022, doi: 10.1287/msom.2021.1030.
- [13] V. Grimm *et al.*, "A standard protocol for describing individual-based and agent-based models," *Ecol Modell*, vol. 198, no. 1–2, pp. 115–126, Sep. 2006, doi: 10.1016/j.ecolmodel.2006.04.023.
- [14] V. Grimm *et al.*, "The ODD Protocol for Describing Agent-Based and Other Simulation Models: A Second Update to Improve Clarity, Replication, and Structural Realism," *Journal of Artificial Societies and Social Simulation*, vol. 23, no. 2, 2020, doi: 10.18564/jasss.4259.
- [15] T. T. Niranjana, N. K. Ghosalya, R. R. Menon, K. Rotaru, and S. Gavirneni, "Unpacking the cognitive processes of the boundedly rational newsvendor," *Prod Oper Manag*, vol. 32, no. 10, pp. 3138–3157, Oct. 2023, doi: 10.1111/poms.14027.
- [16] S. Paul, N. Craig, and E. Bendoly, "Ordering behavior in a supply chain with customers that respond to changes in service level," *Decision Sciences*, vol. 55, no. 1, pp. 88–103, Feb. 2024, doi: 10.1111/deci.12558.



Edgar Cerón (PhD Student, Tecnológico de Monterrey) is currently pursuing the Ph.D. degree in Engineering Sciences. He is a part-time lecturer at the School of Engineering and Sciences, Tecnológico de Monterrey. His research interests include behavioral operations, supply chain management, operations research, system dynamics, and agent-based modeling. His current doctoral work examines the impact of cognitive biases on inventory replenishment decisions and their role in generating the bullwhip effect.

Claudia P. Estévez-Mujica (Assistant Professor, Pontificia Universidad Javeriana), photography and biography not available at the time of publication.

Eva S. Hernández-Gress (Professor, Tecnológico de Monterrey), photograph and biography not available at the time of publication.

Immersive Virtual Reality vs. Scale-Model Simulation in Emergency Medical Response: A Study on Stress and Training Performance

Manuela Gonzalez-Arango, Daniel Morales-Corredor, Edilberto Mejia-Ruda, and Oscar I. Caldas

Abstract— Adequate emergency medicine training is crucial for developing the essential skills to all front-line healthcare providers and thus guarantee quality service and safety to patients. Modern training programs rely on emergency-based simulations that intend to approach trainees to in-site situations and the inevitable emotional states that might affect timely reaction and decision-making, such as stress or anxiety. However, immersive technologies included in medical contexts, e.g. Virtual Reality education, have proven to provide a safe and multisensory environment that improves training performance, along with a positively perceived user experience. Therefore, we propose an experimental design to explore the influence of immersive virtual environments in stress expression by medical trainees compared to conventional simulations and how it affects training performance. A case study scenario was selected (hazardous chemical spill road transportation accident) to be presented to volunteers that will be randomly assigned to either a physical scale-model simulation or an immersive virtual serious game. Simulation performance and objective (psychophysiological signals) indicators of stress will be measured to find significant statistical differences between experimental groups and association between both behavioral and affective variables.

Index Terms— Human-Computer Interaction, Emergency Management, Training, Stress, Virtual Reality

I. INTRODUCTION

AQUIRE, modifying, and improving skills and behaviors implicate internal (e.g., motivation, attention, induced emotions) and external (e.g., methodologies and available resources) factors that might promote or hinder the learning process [1]. Appearance and influence of internal and external factors depend on the learning skills involved, i.e., whether cognitive (remembering, understanding, applying, analyzing, evaluating, and creating), affective (emotions and feelings), or psychomotor (repetitions for motor memory), however, it is believed that practical training involves all three domains and requires a multidimensional assessment [2-4].

One notable factor negatively interfering with learning engagement and effort allocation is stress. Trainees perceiving over-demanding or uncontrollable challenges might experience

academic stress and evolve to anxiety, emotional distress and difficulty to adapt [6]. That affective phenomenon is also followed by non-desired physical issues, such as fatigue, insomnia, muscle tension, which harm learning interest, adherence, and performance [6], [7]. Several studies have reported high stress levels in learning tasks associated with increased mental load, elevated information processing during short lapses, and higher professional expectations, such as in health sciences [8], [9].

Regarding medical training, there is a special study case that reflects the need for multiple and simultaneous skills pursuing (cognitive, affective, and psychomotor): Emergency Medical Services (EMS), since it demands early intervention at the scene of an accident or medical emergency, from basic first aid to advanced medical care [5]. EMS trainees must face stressful challenges mostly related to the need for critical and effective time-sensitive decision making [10], [11].

It is widely accepted that simulations allow us to recreate complex pre-clinical and clinical scenes by safe and controlled environments, where trainees are asked to make under-pressure decisions to assess their responses and subsequent effects, with no real physical or psychological risk. However, previous studies have highlighted conventional simulations, e.g., physical scale-modeling or role-playing, might elicit higher anxiety and stress when perceived as a means for assessment rather than learning opportunities [12], [13].

To reduce symptoms and effects related to stress and anxiety, it is recommended to exploit well-proven technological approaches, such as gamification and virtual reality (VR) [14]. By incorporating game-based dynamics and mechanics (e.g., rewards, levels, and timely feedback) in learning contexts, has been reported to improve motivation and active participation [15], [16]. Moreover, the enhanced sensory interaction and realistic immersion provided by VR has shown efficient learning with emotion regulation [17-19]. Gamification and VR have been jointly used in educational technology to strengthen the learning process in highly demanding scenarios.[19].

Consequently, this study aims to explore whether VR-based serious gaming is likely to reduce learning stress and increase

(Emails: {est.manuela.gonzalez, est.daniele.morales, edilberto.mejia, oscar.caldas} @unimilitar.edu.co.

Study protocol received ethical approval from the UMNG's review board.

Color versions of one or more of the figures in this article are available online at <http://ieeexplore.ieee.org>

This work was financially supported by the Research Vice rectorate in Universidad Militar Nueva Granada under Grants PIC-ING-4231 and IMP-ING-4081." (Corresponding author: M. Gonzalez-Arango).

M. González-Arango, D. Morales-Corredor, and E. Mejia-Ruda are with the Biomedical Engineering program and O.I. Caldas is with the Mechatronics Engineering program, Universidad Militar Nueva Granada, 250247 Colombia.

task performance during EMS training simulations. Subjects will be randomly assigned to one of the experimental groups (conventional or VR simulation) while performance and stress indicators are compared. The following experimental proposal is intended to test two research questions:

RQ1: Do an immersive VR simulation elicits lower self-reported and psychophysiological stress in an EMS simulation?

RQ2: Do VR-driven lower levels of stress contribute to a better task performance in an EMS simulation?

II. METHODS

The current proposal is a comparative experimental study with key indicator measurements pre and post intervention. The main purpose is to ask participants to perform an EMS simulation and measure stress elicitation and task performance at two different immersion levels to analyze significant differences and associations between study variables.

A. Emergency Response Background

The proposed procedure was designed to train healthcare providers on emergencies and disaster management. Each participant is exposed to a sequence of events that will encourage them to make decisions about the three groups required skills: a) dispatching information: ability to identify the type of disaster/accident and the required resources (e.g., human, logistics, or facilities); b) recognition and communication: the ability to effectively in-site acting and communicating with the authorities to describe the accident, size of damage, risk elements, number of patients and the need for further resources; c) risk management: triage performing, patient transfer, and coordination among entities.

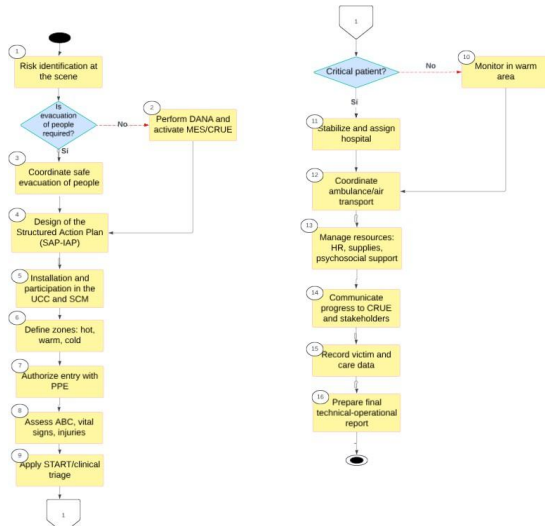


Fig. 1. EMS flowchart, considering Operative (1-7) and Clinical (8-16) stages.

The flowchart proposed in Figure 1 integrates the operative and clinical phases of emergency response for EMS trainees, according to the guidelines provided by the Colombian National Strategy for Emergency Attention and National

System for Disaster Risk Management [20]. The operative stage was organized into nine steps to perform event notification (type of event and general safety conditions), risk identification (structures collapse, landslide, electrical hazard, or chemical leaks), Damage Assessment and Needs Analysis (DANA), Medical Emergency System (MES) activation along with the Urgency and Emergency Regulation Center (CRUE), evacuation management (e.g., muster points and safe pathways), the unified command center (UCC) deployment to gather the local government, the law enforcement authorities and healthcare entities, the intervention areas definition, and the application of preliminary medical procedures.

The clinical stage includes the initial patient assessment, triage performance, hospital location and allocation based on complexity level, safe transportation, and resources management. Finally, status communication, victims and actions registry, and technical report writing.

This action sequence aims to guarantee coherence with the official protocols and objective assessment of EMS trainees technical, ethical and operative proficiency. In addition, the flowchart integrates a specific decision-making sequence for hazardous materials incidents, including ammonia leaks, which emphasizes early detection, area delimitation (hot, warm, and cold zones), HAZMAT team activation, and the correct use of personal protective equipment (PPE). It also incorporates coordinated actions with the fire department, police, and civil defense to ensure perimeter security and risk containment. Patient care follows the START triage model, distinguishing between contaminated and non-contaminated victims, applying decontamination procedures when necessary, and guaranteeing adequate stabilization before evacuation.

This action sequence aims to guarantee coherence with official protocols while ensuring an objective assessment of EMS trainees' technical, ethical, and operative proficiency, reinforcing their ability to respond effectively to complex emergencies that involve both clinical care and inter-institutional coordination.

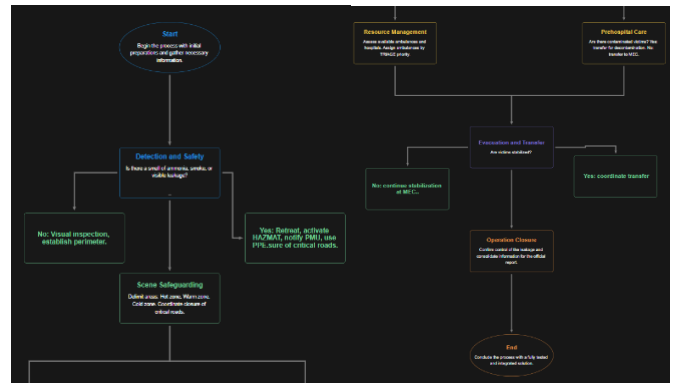


Fig. 2. Decision-making sequence for hazardous materials incidents.

In addition, there is a flowchart that integrates a specific decision-making sequence for hazardous materials incidents, including ammonia leaks, which emphasizes early detection,

area delimitation (hot, warm, and cold zones), HAZMAT team activation, and the correct use of personal protective equipment (PPE). It also incorporates coordinated actions with the fire department, police, and civil defense to ensure perimeter security and risk containment. Patient care follows the START triage model, distinguishing between contaminated and non-contaminated victims, applying decontamination procedures when necessary, and guaranteeing adequate stabilization before evacuation.

This action sequence aims to guarantee coherence with official protocols while ensuring an objective assessment of EMS trainees' technical, ethical, and operative proficiency, reinforcing their ability to respond effectively to complex emergencies that involve both clinical care and inter-institutional coordination.

B. Study case

As shown in Figures 2 and 3, the simulation reproduces the accident between a public transport vehicle (city bus) and a chemical transport truck containing ammonia. Some considerations need to be made to better set the scene: the simulation contains the common buildings and distribution of a city during a light traffic period and moderate damage (about 10 am and 12 patients only). Each participant must decide the type of ambulance to be called, building to be evacuated, patient trauma scoring, among others, while being stimulated with radio messages from the Emergency Regulation and Disaster Coordination Center providing updated information and asking for decision-making.



Fig. 3. Case study: hazardous chemical spill road transportation accident. Physical scale model (left) and VR (right) for EMS.

The accident was recreated in a physical scale model, as it is usually presented to trainees, and via a VR environment which seeks recreation streets, buildings, people, vehicles, and public areas near the simulated accident. It is expected that the VR simulation provides more immersion and realism by keeping the concept of a simplified aerial view. To also maintain similarities in terms of interaction, the VR environment relies on first-person view and hand tracking to promote embodiment with virtual hands.

Figure 3 illustrate the simulation of a collision between a city bus and a truck carrying ammonia at an urban intersection. The virtual environment was developed in Unity and deployed on the Meta Quest 2 headset, recreating a city setting with streets, buildings, hospitals, a fire station, and light traffic of vehicles and pedestrians to enhance realism.

Participants interact in first-person view using hand tracking, which allows them to manipulate objects directly with their hands, encouraging a more natural and intuitive decision-making process. The exercise begins with audio instructions from the Emergency and Disaster Response Center (CRUE), which announces the accident location and requests the immediate dispatch of an ambulance. Additional messages are provided throughout the activity to update the situation and guide further actions.



Fig. 4. Functionality of the ambulance interaction

Ambulance management is fully interactive: users can pick up ambulances positioned at hospitals and place them on the road (Figure 4), automatically triggering a virtual dispatch toward the accident site. Once the ambulance arrives, an information panel appears, confirming the resource's arrival and presenting initial on-site assessment data (Figure 5).



Fig. 5. Information provided upon the arrival of the ambulance at the accident

To define risk areas, participants can deploy interactive cones that generate a visible perimeter line, designating exclusion or evacuation zones depending on the hazards involved. Certain buildings may also be evacuated, requiring participants to decide which areas to clear based on the incident's risks. The simulation also includes a zoom function that switches between an aerial overview of the city and a detailed view of the crash scene, supporting both global and local situational awareness.



Fig. 6. Traffic cone perimeter setup

For consistency, the same accident scenario was also reproduced with a physical scale model, as typically presented to students. While the virtual reality version enhances immersion and realism, the physical model preserves a simplified aerial perspective, ensuring methodological comparability between the two experimental groups.

C. Participants

At least 75 participants per group should be recruited to expect large effect size (Cohen's $d = 0.08$) in testing the hypotheses by statistical power analysis. Random assignments might need to be counterbalanced by gender, experience in VR, and professional background.

Inclusion criteria: Medical trainees with no significant experience in EMS-training simulations and no cognitive conditions that limit them from reading and signing the informed consent and following instructions.

Exclusion criteria: Participants that report proclivity to developing VR-induced symptoms and effects (VRISE), such as vertigo, seizure, or dizziness.

D. Experimental protocol



Fig. 7. Flow chart of the study protocol

The experiment starts with a preparation and introduction stage, where informed consent is presented and sensor placement is done after voluntarily accepting to participate. A 1-minute psychophysiological measure is applied at a seated position to get the resting baseline (described below), and then VR devices were given with basic usage instructions if the participant reported non-specific experience. Figure 2 gives an overview of the entire procedure, including experimental trials and resting periods.

The first resting period allows participants to fill in a demographic survey to gather potential confounding variables that could be counterbalanced during experimental groups assignment, such as age, gender, previous experience in video games and VR, and professional background.

Fig. 3. Flow chart of the study protocol.

Participants are randomly assigned to any of two groups, each exploring a different level of immersion to the EMS training: Group A (non-immersive scale-model simulation) or Group B (immersive VR simulation).

Trial 1 is a training stage to mitigate the risk of learning bias by means of simple and short tasks that introduce the participants to the use of the simulator (tutorial). Participants in Group A will only listen to a brief description of the elements in the scale model shown in Figure 1. Participants in group B will be engaged in a demonstration and brief exploration of the VR interface. This stage helps that any changes in physiological signals or subjective perception of the simulation and the ease of use do not occur during the

actual experimental trial.

After the VR tutorial, all participants must fill in the VRISE subscale of the Virtual Reality Neuroscience Questionnaire (VRNQ) to evaluate the five items via 7-point Likert-scale questions: Absence or insignificant presence of nausea, disorientation, dizziness, fatigue, and instability [21]. Despite not manifesting any VRISE, if a participant scores less than the minimum cut-off in the scale (25 points), they will be kindly asked to withdraw. This resting period allows mitigation or elimination of any moderate VRISE and avoids any accumulation during the experimental trial.

Trial 2 is the actual experiment. Regarding the group assigned, each participant is introduced to the accident simulated and will be asked to make decisions based on a sequence of four-choices questions. After each decision, the participant is scored according to their choice (3=best; 2=good; 1=bad; 0=worst) and it will deploy the subsequent events leading to the next question. The final score will be provided only after the simulation ends. During Trial 2, psychophysiological signals are measured, processed as described below and stored for offline analysis.

E. Hardware and software

VR environment is being developed in Unity 3D version 2022.3.56f1. Immersion is provided via a Meta Quest 3 headset. Psychophysiological signals were used to assess stress—related indicator: Heart Rate (HR) and Heart-Rate Variability were measured by electrocardiography (ECG) and Skin Conductance (SC) was measured by Galvanic Skin Response (GSR), by means of a BiosignalsPlux device. ECG electrodes are placed in V2 and V6 positions of the 12-Lead scheme and the sternum for ground, whereas GSR electrodes were placed on the intermediate phalanges of the index and middle fingers in the non-dominant hand, similar to previous studies [22].

Signals processing and statistical analysis were performed in Matlab 2024b.

III. EXPECTED RESULTS AND DISCUSSION

It is expected to find significant differences between the study groups (non-immersive and immersive simulation), in terms of task performance (score) and the stress-related indicators (HR and SC features) as result to the training task.

Association is also expected between task performance and stress, as result of the correlation and interaction analysis. Discussion can be made about stress increasing or decreasing by the level of immersion in the training task, as well as its interactive or independent effect on task performance. Preferences and guidelines could be proposed to improve the training method while using simulations for EMS education.

C. Limitations and future work

There is still limited generalizability since the study is expected to be conducted with a specific population (medical trainees) and simulation (chemical spill road transportation accident). Future work can drive efforts to create or adapt different VR scenarios and be applied to a wider population.

In addition, despite including time limit for decision-making, trials duration is still affected by the subjective human-computer interaction is not possible to be normalized. Further effort must be made to reduce trial variability.

A future study could be conducted to quantitatively determine the level of immersion in VR-based learning tasks and to examine whether such immersion levels modulate participants' stress responses. The expected outcome is to identify an optimal degree of immersion that maximizes performance, enabling personalized learning experiences. This would be achieved by understanding and controlling immersion throughout task execution, through real-time feedback based on participants' psychophysiological signals.

ACKNOWLEDGMENT

The authors would like to thank Laidy Quiñones Castro and Sandra Mondragon from the Prehospital Attention program in UMNG to help on providing fruitful insights and support regarding the case study.

REFERENCES

- [1] M. Schultes, D. Graf, J. Holzer, B. Schober, C. Spiel, "Implementation and evaluation of service learning at higher education institutions", *Evaluation and Program Planning*, vol. 112, 2025, 102622, doi: 10.1016/j.evalprogplan.2025.102622.
- [2] J. West, "Utilizing Bloom's taxonomy and authentic learning principles to promote preservice teachers' pedagogical content knowledge", *Social Sciences & Humanities Open*, vol. 8, Issue 1, 2023, 100620, doi: 10.1016/j.ssaho.2023.100620.
- [3] S. R. Toala Ponce, L. Y. Gómez Pinillo, R. N. Guevara Heredia, and E. C. Quiñonez Ortiz, "Application of Bloom's taxonomy to improve teaching-learning", *Sapientia*, vol. 3, no. 6, pp. 176–189, 2022, doi: 10.51798/sijis.v3i6.507.
- [4] Y.K Kuo, S. Batool, S. devi, T. Tahir, J. Yu, "Exploring the impact of emotionalized learning experiences on the affective domain: A comprehensive analysis", *Heliyon*, vol. 10, no. 1, 2024, e23263, doi: 10.1016/j.heliyon.2023.e23263.
- [5] Institute of Medicine. "Crisis Standards of Care: A Systems Framework for Catastrophic Disaster Response: Volume 3: Prehospital Care Emergency Medical Services (EMS)". Washington, DC: The National Academies Press; 2012. doi: 10.17226/13351.
- [6] P. Frazier, A. Gabriel, A. Merians, and K. Lust, "Understanding stress as an impediment to academic performance," *Journal of American College Health*, vol. 67, pp. 562–570, 2019, doi: 10.1080/07448481.2018.1499649.
- [7] K. Kalinkov, V. Markova and T. Ganchev, "The Effect of Acute Stress on the Performance of Students in Engineering Education," in *2021 56th International Scientific Conference on Information, Communication and Energy Systems and Technologies (ICEST)*, Sozopol, Bulgaria, 2021, pp. 235–238, doi: 10.1109/ICEST52640.2021.9483572.
- [8] A. F. Bastos et al., "Academic-related stressors predict depressive symptoms in graduate students: A machine learning study", *Behavioural Brain Research*, vol. 478, Feb. 2025, doi: 10.1016/j.bbr.2024.115329.
- [9] Y. Guibin and L. Jinyang, "Research on the influence of Student Role Stress on Learning Satisfaction in Blended Learning Based on SPSS21.0 and Amos24.0," in *2nd International Conference on Education, Knowledge and Information Management (ICEKIM)*, Xiamen, China, 2021, pp. 167–174, doi: 10.1109/ICEKIM52309.2021.00044.
- [10] L. Carmona, C. Costa, S. Gascón, G. Ribeiro, M.J Chambel, "Prevalence and risk factors for anxiety, stress and depression among higher education students in Portugal and Brazil", *Journal of Affective Disorders Reports*, vol. 17, 2024, 100825, doi: 10.1016/j.jadr.2024.100825.
- [11] A. Alzahra, J. Johnson, C. Keyworth, "Prevalence of Anxiety, Depression, and Post-Traumatic Stress Disorder among Paramedic Students: A Systematic Review and Meta-Analysis", *Resuscitation*, vol. 203, Supplement 1, 2024, S232, doi: 10.1016/S0300-9572(24)00778-0.
- [12] Q. Xu, T. L. Nwe, and C. Guan, "Cluster-Based Analysis for Personalized Stress," *IEEE J Biomed Health Inform*, vol. 19, no. 1, pp. 275–281, 2015, doi: 10.1109/JBHI.2014.2311044.
- [13] B. Wheeler and E. Dippenaar, "The use of simulation as a teaching modality for paramedic education: a scoping review," *Br Paramed J*, vol. 5, no. 3, pp. 31–43, 2020, doi: 10.29045/14784726.2020.12.5.3.31.
- [14] C. K. Javvaji, H. Reddy, J. D. Vagha, A. Taksande, A. Kommareddy, and N. sudheesh Reddy, "Immersive Innovations: Exploring the Diverse Applications of Virtual Reality (VR) in Healthcare", *Cureus*, vol. 16, no. 3, pp. 1–10, 2024, doi: 10.7759/cureus.56137.
- [15] B. Thanasekhar, N. Gomathy, A. Kiruthika and S. Swarnalaxmi, "Machine Learning Based Academic Stress Management System," in *11th International Conference on Advanced Computing (ICoAC)*, Chennai, India, 2019, pp. 147–151, doi: 10.1109/ICoAC48765.2019.246831.
- [16] M. S. Bracq, E. Michinov, and P. Jannin, "Virtual Reality Simulation in Nontechnical Skills Training for Healthcare Professionals: A Systematic Review," *Simulation in Healthcare*, vol. 14, pp. 188–194, Jun. 2019, doi: 10.1097/SIH.0000000000000347.
- [17] J. Qian, M. Yancong, P. Zhigeng, Y. Xubo "Effects of Virtual-real fusion on immersion, presence, and learning performance in laboratory education", *Virtual Reality & Intelligent Hardware*, vol. 2, no. 6, 2020, pp 569–584, doi: 10.1016/j.vrih.2020.07.010.
- [18] Q. Xu, T. L. Nwe, and C. Guan, "Cluster-based analysis for personalized stress evaluation using physiological signals," *IEEE J Biomed Health Inform*, vol. 19, no. 1, pp. 275–281, 2015, doi: 10.1109/JBHI.2014.2311044.
- [19] L. JianBang, M.C Ang, J.K Chaw, K. Ah-Lian, N. Kok, L.C Meng, "Assessing the impact and development of immersive VR technology in education: Insights from telepresence, emotion, and cognition", *Technological Forecasting and Social Change*, vol. 213, 2025, 12402, doi: 10.1016/j.techfore.2025.124024.
- [20] UNGRD, "Manual de Logística para la Atención de Emergencias Unidad Nacional para la Gestión del Riesgo de Desastres," p. 39, 2013
- [21] P. Kourtesis, S. Collina, L.A. Doumas, S.E. MacPherson, "Validation of the virtual reality neuroscience questionnaire: maximum duration of immersive virtual reality sessions without the presence of pertinent adverse symptomatology", *Front. Human Neurosci.* vol. 13, 2019, 417, doi: 10.3389/FNHUM.2019.00417/BIBTEX.
- [22] O. I. Caldas, M. Mauleudoux, O. F. Aviles, C. Rodriguez-Guerrero, "Breaking presence in Immersive Virtual Reality toward behavioral and emotional engagement". *Computer Methods and Programs in Biomedicine*, vol. 248, 2024, 108124, doi: 10.1016/j.cmpb.2024.108124.
- [23] G. Giannakakis, D. Grigoriadis, K. Giannakaki, O. Simantiraki, A. Roniotis, and M. Tsiknakis, "Review on Psychological Stress Detection Using Biosignals," *IEEE Trans Affect Comput*, vol. 13, pp. 440–460, 2022, doi: 10.1109/TAFFC.2019.2927337.
- [24] X. Navarro, "Physiology of the autonomic nervous system," *Rev Neurol*, vol. 35, no. 6, pp. 553–562, 2002, doi: 10.33588/rn.3506.2002013.
- [25] V.C. Goessl, J.E. Curtiss, S.G. Hofmann, "The effect of heart rate variability biofeedback training on stress and anxiety: a meta-analysis", *Psychol. Med.* vol. 47, 2017, pp. 2578–2586, doi: 10.1017/S0033291717001003.
- [26] A. Arsalan, M. Majid, S. M. Anwar, and U. Bagci, "Classification of Perceived Human Stress using Physiological Signals," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, pp. 1247–1250, 2019, doi: 10.1109/EMBC.2019.8856377.