UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Revista Ingeniería

Four-monthly scientific journal

2025

Volume 30 - Issue 1 ISSN 0121-750X E-ISSN 23448393

REVISTA Ingeniería

Volume 30 · Issue 1 · Year 2025 · ISSN 0121-750X · E-ISSN 2344-8393

Four-monthly Scientific Journal

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Carrera 7 No. 40-53
Edificio Administrativo
Piso 7 - Facultad de Ingeniería
Bogotá, Colombia
Teléfono: + 57 (1) 323 93 00 ext. 2413
Correo revista:
revista_ing@udistrital.edu.co

http://revistas.udistrital.edu.co/ojs/index.php/reving

Focus and Scope

The *Ingeniería* journal is an open-access academic-scientific online publication specialized in knowledge related to the fields of Engineering and Technology, according to the classification of scientific areas established by the Organization for Economic Cooperation and Development (OECD). Its main goal is to disseminate and promote debates about advances in research and development in the diverse areas of Engineering and Technology. We focus on disseminating original and unpublished articles in English that are relevant at both the national and international level. We seek to provide a platform to share knowledge and foster collaboration between researchers and professionals in the field. The journal does not request any Article Processing Charges (APC), as it is directly funded by Universidad Distrital Francisco José de Caldas. Moreover, the journal operates following a double-blind review process, thus ensuring the impartiality and quality of the published works. The journal is published in a continuous fashion and follows a quarterly numbering.

Open Access Policy

The *Ingeniería* journal provides free access to its content. This free access is granted under the principle of making research freely available to the public, which en-courages a greater exchange of global knowledge.

Article Processing Charges

No publication fees are charged to the authors or their institutions, nor are any payments made to expert peer reviewers or associate or adjunct editors. The *Ingeniería* journal is funded by Universidad Distrital Francisco José de Caldas, its Faculty of Engineering, and its Central Research Office.

Indexed in

Scopus
SciELO
Publindex Colombia Categoría B
Redalyc
Directorio Latindex
Fuente Académica Premier (EBSCO)
Applied Science & Technology Source Ultimate (EBSCO)
Emerging Sources Citation Index
Bibliographic Bases (SC)
Fuente Académica Plus (EBSCO)
OpenAlex
Catálogo 2.0 Latindex
DOAJ
Dialnet
ROAD
Google Schola Metrics
PKP Index
MIAR

Peer-reviewers in this issue

Carlos Fabián Ávila
Universidad UTE, Ecuador

Makoena Sebatjane
Stellenbosch University, South Africa

Seyed Reza Mousavi-Aghdam
University of Mohaghegh Ardabili, Iran

Surendran Rajendran
Saveetha School of Engineering, India

Shankha Shubhra Goswami
Abacus Institute of Engineering and Management, India

Walter Julián Gil González
Universidad Tecnológica de Pereira, Colombia

Scheila Mânica
University of Dundee, United Kingdom

Ahmed S.A. Al-Gharbawi
University of Technology, Iraq

Kah Chan Teh
School of Electrical and Electronic Engineering, Singapore

Zhen Liu
Southeast University, China

Vivek Vivek
National Institute of Technology Srinagar, India

David Luong
Carleton University, India

Blas Manuel Rodríguez-Lara
Tecnológico de Monterrey, Mexico

Vassilis E. Fotopoulos
Hellenic Open University, Greece

Michael Yu Rachkov
Moscow Polytechnic University, Russia

Oscar Danilo Montoya Giraldo
Universidad Distrital Francisco José de Caldas, Colombia

Renato P. Cunha
Universidade de Brasília, Brazil

Amit Joshi
The ICFAI University, Dehradun, India

Paulo Mendes
Universidade do Porto, Portugal

Table of contents

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.23566

Editorial

RIBIERSE-Cyted: Network for the Large-Scale Integration of Renewable Energies in Electrical Systems) – Results for 2024

A. Medina[1], O. D. Montoya[2], C. R. Baier[3], W. Gil-González[4], and J. C. Hernández[1]*

[1]University of Jaén, Spain
[2]Universidad Distrital Francisco José de Caldas, Colombia
[3]Universidad de Talca, Curicó, Chile
[4]Universidad Tecnológica de Pereira

1. Description of the Scientific-Technical Network

The RIBIERSE-CYTED network, *i.e.*, the Network for the Large-Scale Integration of Renewable Energies in Electrical Systems (723RT0150; 2023-2026), is a hub for researchers and technologists attached to Ibero-American universities, companies, and local administrations. From a scientific and technical perspective, this network contributes to the decarbonization of the electricity sector by favoring the large-scale integration of renewable sources into electric power systems. It promotes and articulates a framework for joint university-business cooperation and scientific research with a shared Ibero-American vision, and it encourages knowledge of the renewable context from service providers to end users.

To this effect, the proposed lines of action, which are derived from the specific objectives (SOs) of this network, are grouped into three areas. The first line includes the assessment of energy planning models and tools, as well as the prediction, evaluation, control, and management of the renewable sources currently in use in Ibero-American countries. All this, with the aim of extending, developing, and systematizing their application. The second line deals with analyzing the effective application of methodologies in real cases. Finally, the third line is dedicated to the technical training of researchers and suppliers of sustainable systems, as well as to the electrical eco-literacy of end users.

Editorial

* **Correspondence:** jcasa@ujaen.es

The RIBIERSE-CYTED network is made up of 48 groups and 272 researchers from 22 Ibero-American countries: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Cuba, Ecuador, El Salvador, Spain, Guatemala, Honduras, Mexico, Nicaragua, Panama, Paraguay, Peru, Portugal, Puerto Rico, Dominican Republic, Uruguay, and Venezuela. 29 of the participating partners are research groups from different universities (Fig. **1**) (Appendix 1) and the rest belong to companies and local administrations.

Figure 1. Geographical distribution of the Ibero-American universities in RIBIERSE-CYTED

2. Results

The significant dynamism of each of the partners and their strengths in relation to the network's objectives has allowed for a noteworthy number of achievements in 2024. Thereupon, the following actions have been taken as initially planned:

- Holding the annual coordination meeting
- Holding Workshop #1
- Setting up and monitoring multiple pilot projects

- Analyzing opportunities for external funding for the CYTED
- Holding Workshop #2
- Planning technical reports

As a consequence of this strong dynamism, all actions, distributed among 13 SOs, have led to the following activities:

- **23 research projects**, most of which are mainly aligned with SO04 (*Development of technical and economic tools for optimal electrical management in the hybridization of different renewable technologies, storage types, and 'E-Mobility'), SO05 (Identification of technological, regulatory, and socio-economic benefits and barriers for the massive integration and hybridization of renewables in Ibero-American countries), and SO06 (Study of the impact on current/future electrical systems of this integration, with an emphasis on temporal variability; promotion of mitigation measures)*.
- One infrastructure work, *i.e.*, the construction of a new microgrid at the Automatic Control Laboratory of Universidad Nacional de San Luis, Argentina.
- Two technological contracts, one for technical and scientific contract consultancy for a local company named *Candela Solar* (Argentina) and another one involving computational methods for the planning and operation of distribution networks while considering renewable energies (Colombia).
- 38 Master/Bachelor final projects from network member universities in Spain, Colombia, Brazil, and Cuba. In addition, four doctoral theses from partner universities in Spain, Colombia, and the Dominican Republic, one of which is collaborative. Out of these works, a total of 40 is aligned with SO04.
- Four internships carried out by PhD students. Two of them were carried out at the University of Jaén (Spain), another one at Oak Ridge National Laboratory (USA), and the last one at Universidad de Talca (Chile).
- 23 short mobilities, with hosting centers in Spain, Chile, Guatemala, Panama, Mexico, Colombia, Argentina, Puerto Rico, Brazil, and Portugal. Most of these initiatives received external funding for their implementation.
- Five national conferences organized by member universities of the network (Argentina, Brazil, and Panama).
- Four international conferences: the VII International Ibero-American Congress on Smart Cities, ICSC-CITIES 2024, co-organized by the University of Valladolid (Spain), Universidad de la República (Uruguay), Universidad Autónoma del Estado de Morelos (Mexico), Instituto Tecnológico de Costa Rica, and the University of Jaen (Spain); the 3rd International Congress on Electrical Engineering, Electronics, and Automation and the 8th International Scientific Convention (CCIUTM-2024), co-organized by Universidad Técnica de Manabí (Ecuador) and the University of Jaén (Spain); and the 18th Brazilian Power Electronics Conference (COBEP), co-organized by two member universities from Brazil, *i.e.*, (Universidade Federal Fluminense and Universidade Federal do Juiz de Fora).
- Two postgraduate courses in Argentina: i) *Methodology for the dissemination of engineering sciences and ii) Distributed generation systems*.
- Four technical courses in Spain, Mexico, and Colombia.

- One diploma course in Panama.
- One Technical Day, organized by a Spanish university (Miguel Hernández University of Elche).
- The annual coordination meeting of the RIBIERSE-CYTED in the city of Mérida, Mexico, co-organized by Universidad Autónoma de Yucatán (Mexico) and the University of Jaén (Spain).
- Four events in other categories, organized by member universities in Spain, Honduras, and Ecuador.
- One Engineering Week organized by Argentina's partner university.
- One seminar organized by Universidad de Puerto Rico.
- One symposium organized by the Cuban university of the network, the First International Symposium on Electromobility and Energy Transition (Electromovilidad 2024).
- Two webinars organized by universities in Spain and Colombia.
- One workshop organized by the University of Jaén: I International Workshop. Hardware and computational tools applied to the management/optimization of networks and microgrids with renewable energies.
- Three technical workshops organized by the Cuban member university.
- One cycle of postgraduate conferences with speakers from members from Spain and Mexico, which was published on the CYTED website.
- 143 articles published in scientific journals, 63 of which were collaborative works.
- Three scientific books, two of which were collaborative works.
- 39 articles related to participation in scientific events (conferences), 31 of which were collaborative works.
- Two deliverables with the participation of all RIBIERSE members.
- 18 internal deliverables.
- 33 news pieces available on the CYTED website, reporting on the achievements made by the network. In addition, 41 CYTED pieces have been disseminated through other channels and are related to congresses, courses, conferences, engineering weeks, workshop seminars, doctoral theses, short mobilities, and attendance to scientific events.

All these activities, categorized as *Achievements*, *Events*, and *Publications*, are aligned with the RIBIERSE-CYTED network's SOs and have been materialized in a total of 417 deliverables.

3. Partner interrelation activity in the network

Coordination work for the RIBIERSE-CYTED network was carried out through different meetings, held in person and/or online with different actors and objectives, *i.e.*,

- An annual coordination meeting
- Partner coordination meetings
- Meetings at international conferences

- Events (conferences, congresses, postgraduate courses, diploma courses, courses, seminars, other events, thematic weeks, seminar, symposium, workshop, webinar, technical days)
- Meetings on the lines of action of the RIBIERSE-CYTED network.

In 2024, each of the activities was carried out individually by a partner or collaboratively. The total number of activities in which each partner has participated is shown in Fig. 2. In addition, Table I and Fig. 3 show the individual and collaborative activities conducted by each partner.

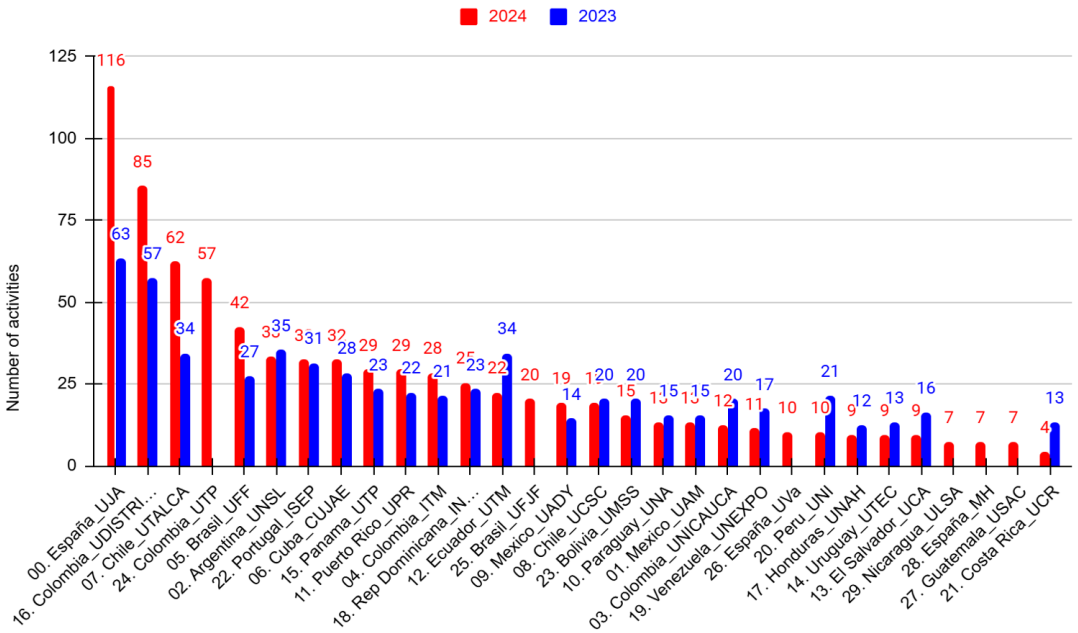


Figure 2. Total activities conducted by each partner

The second year of development of the RIBIERSE-CYTED network (2024) has been characterized by significant growth in terms of collaboration between partners, which is evidenced by the significant number of collaborative activities conducted during this period. In all categories, the percentage of collaborative activities has grown substantially, as shown in Fig. 4.

Fig. 5 presents a circular diagram of collaborations. Here, line thickness reflects the number of collaborative activities, and circle size denotes the number of individual activities. This figure highlights the change in the interrelation between the partners involved in collaborative activities. The graph for 2023 faithfully reflects what happened in the first year of the network's journey, as collaboration between member partners was still difficult, sometimes due to lack of knowledge regarding each research group's lines of work or the personal relationships that could benefit professionals in the medium and long term. However, in 2024, more than 50 % of the partners had already established solid collaboration with a significant number of their counterparts.

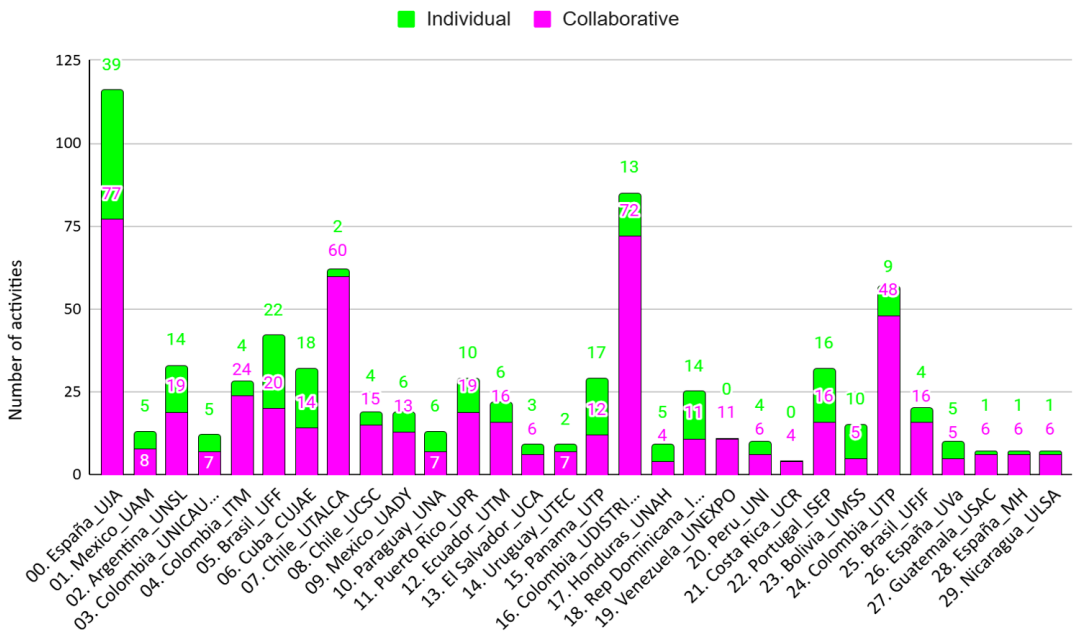


Figure 3. Individual and collaborative activities carried out in 2024

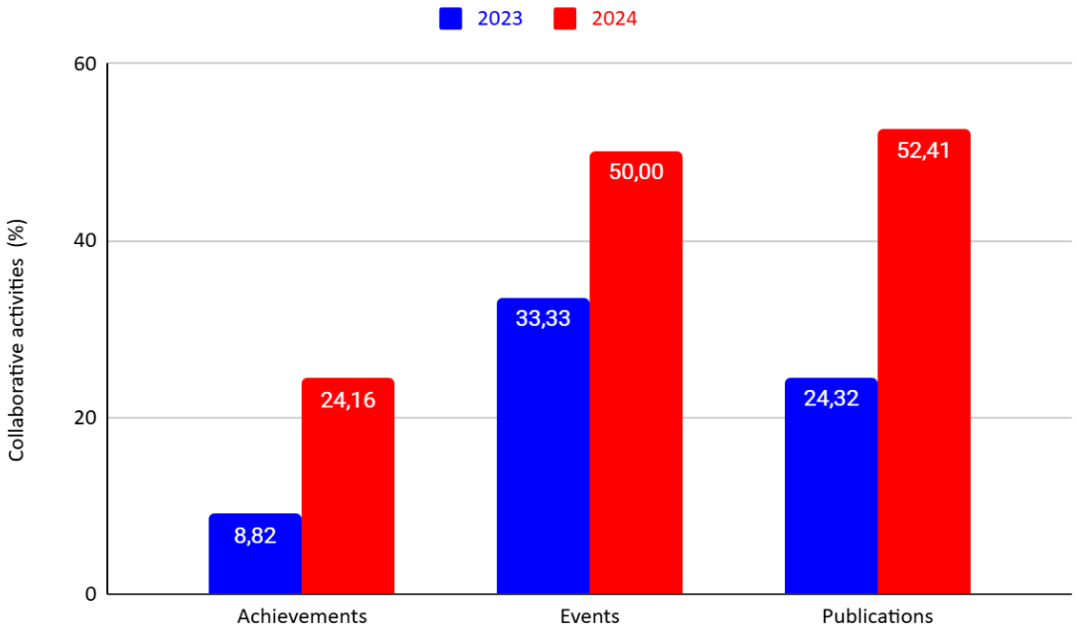


Figure 4. Percentage of collaborative activities by category 2023 *vs.* 2024

Table I. Individual and collaborative activities conducted in 2024

	00. España_UJA	01. México_UAM	02. Argentina_UNSL	03. Colombia_UNICAUCA	04. Colombia_ITM	05. Brasil_UFF	06. Cuba_CUJAE	07. Chile_UTALCA	08. Chile_USSC	09. México_UADY	10. Paraguay_UNA	11. Puerto Rico_UPR	12. Ecuador_UTM	13. El Salvador_UCA	14. Uruguay_UTEC	15. Panamá_UTP	16. Colombia_UDISTRITAL	17. Honduras_UNAH	18. Rep. Dominicana_INTEC	19. Venezuela_UNEXPO	20. Perú_UNI	21. Costa Rica_UCR	22. Portugal_ISEP	23. Bolivia_UMSS	24. Colombia_UTP	24. Brasil_UFJF	26. España_UVa	27. Guatemala_USAC	28. España_MH	29. Nicaragua_ULSA
00. España_UJA	39	6	12	6	8	11	9	22	10	9	4	14	11	6	7	12	20	4	10	11	5	4	10	5	12	5	5	6	6	4
01. México_UAM		5	4	4	4	4	4	4	4	6	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
02. Argentina_UNSL			14	4	4	8	5	14	7	4	4	7	5	4	4	4	13	4	5	4	4	4	6	4	7	4	4	4	4	4
03. Colombia_UNICAUCA				5	4	4	4	5	4	6	4	6	4	4	4	6	5	4	4	4	4	4	4	4	5	4	4	4	4	4
04. Colombia_ITM					4	4	4	21	4	4	4	6	4	4	4	4	21	4	4	4	4	4	4	4	6	4	4	4	4	4
05. Brasil_UFF						22	5	10	4	4	4	7	5	4	4	4	7	4	5	4	4	4	8	4	4	12	4	4	4	4
06. Cuba_CUJAE							18	5	6	4	4	5	10	6	4	6	5	4	5	6	4	4	4	4	4	4	4	4	4	4
07. Chile_UTALCA								2	11	4	5	9	5	4	4	4	37	4	7	4	4	4	8	4	20	5	4	4	4	4
08. Chile_USSC									4	4	4	4	4	5	4	6	4	4	5	6	4	4	4	4	4	4	4	4	4	4
09. México_UADY										6	6	6	4	4	4	6	4	4	4	5	4	4	5	4	4	4	4	4	4	4
10. Paraguay_UNA											6	6	4	4	4	6	4	4	4	5	4	4	5	4	4	4	4	4	4	4
11. Puerto Rico_UPR												10	5	4	4	6	10	4	8	4	4	4	6	4	5	4	4	4	4	4
12. Ecuador_UTM													6	4	4	4	4	5	4	5	4	4	4	4	4	4	4	4	4	4
13. El Salvador_UCA														3	4	6	4	4	4	6	4	4	4	4	4	4	4	4	4	4
14. Uruguay_UTEC															2	4	4	4	4	4	4	4	4	4	4	4	4	4	4	4
15. Panamá_UTP																17	4	4	4	4	4	4	4	4	4	4	4	4	4	4
16. Colombia_UDISTRITAL																	13	4	5	4	4	6	4	41	4	4	4	4	4	4
17. Honduras_UNAH																		5	4	4	4	4	4	4	4	4	4	4	4	4
18. Rep. Dominicana_INTEC																			14	4	4	4	4	4	4	4	4	4	4	4
19. Venezuela_UNEXPO																				0	4	4	5	4	4	4	4	4	4	4
20. Perú_UNI																					4	4	4	4	4	4	4	4	4	4
21. Costa Rica_UCR																						0	4	4	4	4	4	4	4	4
22. Portugal_ISEP																							16	4	5	9	4	4	4	4
23. Bolivia_UMSS																								10	4	4	4	4	4	4
24. Colombia_UTP																									9	4	4	4	4	4
24. Brasil_UFJF																										4	4	4	4	4
26. España_UVa																											5	4	4	4
27. Guatemala_USAC																												1	4	4
28. España_MH																													1	4
29. Nicaragua_ULSA																														1

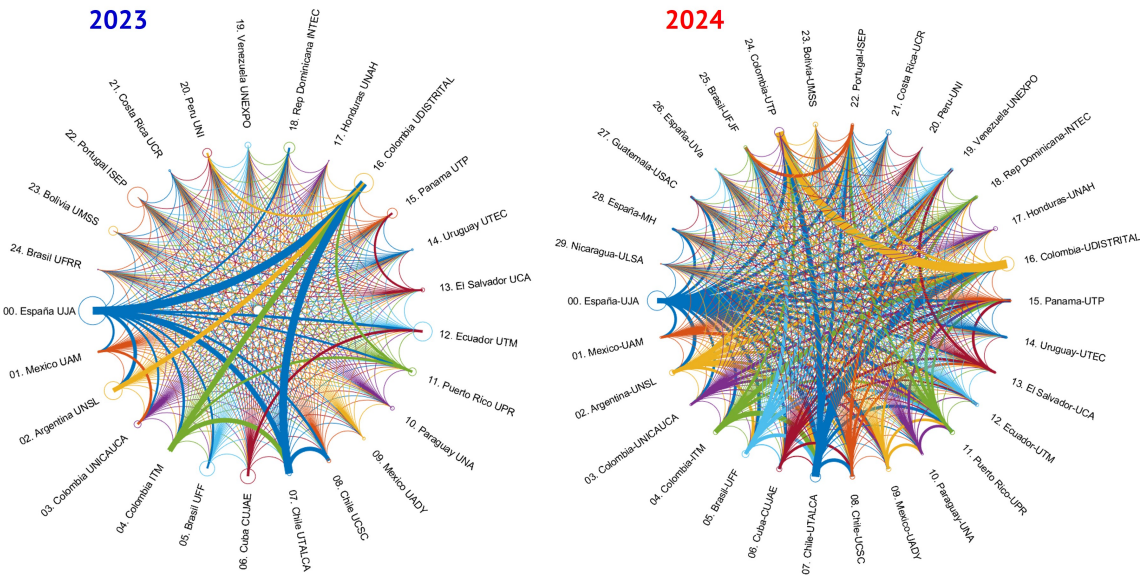


Figure 5. Interrelation between partners in collaborative activities (2023 *vs.* 2024)

4. Conclusion

The RIBIERSE-CYTED network includes partners from 22 countries, involving a total of 272 people. In this context, administrative management and expense forecasting has become a complex issue, especially regarding the management and organization of the II Annual Meeting of the RIBIERSE-CYTED network, held in the city of Mérida, Mexico, as well as regarding the mobility of several partners.

The different time zones of the 22 countries in the network make it difficult to agree on work schedules for events and collective meetings. To overcome this challenge, the proposals for new activities have been divided into subtasks, linking the partners in work nodes, an approach that was strengthened after the II Annual Meeting.

There has been great synergy between partners during the network's second year thanks to the initial dynamism, which has enabled the extension of our work modes to other groups, as well as their involvement, completing the activities scheduled for 2024 and exceeding the expectations in many cases. It is worth noting that all these achievements demonstrate the great commitment of the members, who are grateful for the support offered by the Ibero-American Program of Science and Technology for Development (CYTED)[**]. Nonetheless, the funding and support provided to the RIBIERSE-CYTED network by the University of Jaén, Universidad Autónoma de Yucatán (UADY), and all the participating universities and companies external to the network have been crucial for achieving so many activities.

[*]*The CYTED program was created by the governments of the Ibero-American countries to foster cooperation in science, technology, and innovation while aiming for harmonious development.

Appendix 1. Research group leaders

Name	Affiliation	ORCID	Email
María de los Angeles Medina Quesada	Department of Electrical Engineering, University of Jaen, Jaen 23071, Spain;	https://orcid.org/0000-0003-3924-7986	aquesada@ujaen.es
Oscar Danilo Montoya Giraldo	Facultad de Ingeniería, Universidad Distrital Francisco José de Caldas, Bogotá 110231, Colombia;	https://orcid.org/0000-0001-6051-4925	o.d.montoyagiraldo@ieee.org
Carlos Rodrigo Baier Fuentes	Department of Electrical Engineering, Universidad de Talca, Curicó, Chile;	https://orcid.org/0000-0002-1752-1625	cbaier@utalca.cl
Walter Julián Gil González	Department of Electrical Engineering, Universidad Tecnologica de Pereira, Pereira 660003, Colombia;	https://orcid.org/0000-0001-7609-1197	wjgil@utp.edu.co
Rubén Iván Bolaños	Instituto Tecnologico Metropolitano, Medellın, Antioquia 050034, Colombia;	https://orcid.org/0000-0002-0910-6579	rubenbolanos@itm.edu.co
Bruno Wanderley França	Departamento de Engenharia Elétrica, Núcleo de Inovação Tecnológica em Engenharia Elétrica - NITEE, Universidade Federal Fluminense, Niterói, RJ, Brazil;	https://orcid.org/0000-0003-0788-7997	bwfranca@id.uff.br
Federico Martin Serra	Laboratorio de Control Automático, Facultad de Ingeniería y Ciencias Agropecuarias, Universidad Nacional de San Luis, San Luis, Argentina;	https://orcid.org/0000-0002-4467-7836	serrafederico@gmail.com
Fabio Andrade Rengifo	Department of Electrical and Computer Engineering, University of Puerto Rico, Mayaguez, Mayaguez campus, PR 00681, USA;	https://orcid.org/0000-0002-8859-7336	fabio.andrade@upr.edu
Joao André Pinto Soares	GECAD - Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development, LASI - Intelligent Systems Associate Laboratory, Polytechnic of Porto, Porto, Portugal;	https://orcid.org/0000-0002-4172-4502	jan@isep.ipp.pt
Ney Raul Balderramo Velez	Department of Electrical Engineering, Technical University of Manabí, Manabí, Ecuador;	https://orcid.org/0000-0001-8502-4332	ney.balderramo@utm.edu.ec
Bruno Henriques Dias	Electrical Energy Department, Federal University of Juiz de Fora, Juiz de Fora, MG, 36036-900, Brazil;	https://orcid.org/0000-0002-7663-235X	bhdias@gmail.com
Eduardo Enrique Espinosa Neira	Department of Electrical Engineering, Faculty of Engineering, Universidad Católica de la Santísima Concepción, Talca 3467769, Chile;	https://orcid.org/0000-0002-8896-0852	eespinosa@ucsc.cl
Miguel Castro Fernández	Centro de Investigaciones y Pruebas Electroenergéticas Centre for Electroenergetic Research and Testing (CIPEL), Faculty of Electrical Engineering, Technological University of Havana José Antonio Echeverría (CUJAE), La Habana 19390, Cuba;	https://orcid.org/0000-0002-3983-469X	mcastro1956@gmail.com
Roberto Eduardo Quintal Palomo	Facultad de Ingeniería, Universidad Autónoma de Yucatán, Cordemex, Mérida, Yucatán 97310, México;	https://orcid.org/0000-0001-8265-3977	roberto.quintal@correo.uady.mx
Félix Ramón Henríquez Espinosa	Faculty of Mechanical Engineering, Technological University of Panama Panama, Republic of Panama	https://orcid.org/0000-0002-9009-2599	felix.henriquez@utp.ac.pa
Miguel Euclides Aybar Mejía	Área de Ingeniería, Instituto Tecnológico de Santo Domingo, Santo Domingo 10602, Dominican Republic;	https://orcid.org/0000-0002-4715-3499	miguel.aybar@intec.edu.do
Carmen Luisa Vásquez Stanescu	Universidad Nacional Experimental Politécnica Antonio José de Sucre, Parque Tecnológico, Barquisimeto, Venezuela;	https://orcid.org/0000-0002-0657-3470	cvasquez@unexpo.edu.ve
Francisco Beltrán Carbajal	Departamento de Energía, Unidad Azcapotzalco, Universidad Autónoma Metropolitana, Azcapotzalco, Ciudad de Mexico 02200, Mexico;	https://orcid.org/0000-0001-5244-5587	fbeltran@azc.uam.mx
Osvaldo Julián González Barrios	Laboratory of Power Systems and Control, Facultad de Ingeniería, Universidad Nacional de Asunción, Luque, Paraguay;	https://orcid.org/0000-0002-2191-9636	ogonzalez@ing.una.py
Félix Rustan Roca Subirana	Department of Mechanical Engineering, Universidad Mayor de San Simón, Cochabamba, Cercado 0301, Bolivia;	https://orcid.org/0009-0001-1302-3311	rustanroca.s@fcyt.umss.edu.bo
Jorge Luis Mírez Tarrillo	Group of Mathematical Modeling and Numerical Simulation, Faculty of Oil, Natural Gas and Petrochemical Engineering, National University of Engineering, Lima, Peru	https://orcid.org/0000-0002-5614-5853	jmirez@uni.edu.pe
Luis Aarón Martínez Figueroa	Department of Energy and Fluid Sciences, Universidad Centroamericana José Simeón Cañas, El Salvador	https://orcid.org/0000-0002-6553-8819	lamartinez@uca.edu.sv
Franciele Weschenfelder	Offshore Wind Energy Research Group, Universidad Tecnológica del Uruguay, Durazno, Uruguay	https://orcid.org/0009-0003-1525-0766	franciele.weschenfelder@utec.edu.uy
Jesús de la Casa Hernández	Department of Electrical Engineering, University of Jaen, Jaen 23071, Spain;	https://orcid.org/0000-0001-9117-1689	jcasa@ujaen.es

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.22383

Research

Development and Validation of a Computer Vision-Based Tool for Automated 3D Dental Model Arch Prediction

Desarrollo y validación de una herramienta basada en visión por computadora para la predicción automatizada de la forma del arco dental en modelos 3D

J.D. Tamayo-Quintero[*1], J.B. Gómez-Mendoza[2], and S.V. Guevara-Pérez[3]

[1]Tecnológico de Antioquia - Institución Universitaria, Medellín, Colombia
[2]Universidad Nacional de Colombia, Manizales, Colombia
[3]Universidad Nacional de Colombia, Bogotá, Colombia

Abstract

Context: Accurate dental arch shape prediction is crucial for orthodontic treatment and personalized dental appliance creation. This study introduces a computer vision-based tool for predicting arch shapes in 3D dental models.
Objective: To automate the selection of dental arch shapes through mathematical model analysis.
Method: A dataset of 484 digital dental models was narrowed to 50 through specific criteria. Experts classified these into ovoid, square, and tapered shapes using 3M templates. An automated 3D dental arch shape prediction tool was developed, incorporating automatic alignment, cusp detection, curve fitting with a sixth-order polynomial, and model comparison. Our validations employed attribute agreement analysis, the root mean squared error, the sum of squared errors, and a Gage R&R study.
Results: This study achieved a 90% agreement rate in the evaluator *vs.* standard comparison for the lower jaw, as well as 78% for that of the upper jaw. The Gage R&R study confirmed measurement reliability, and the sixth-order polynomial model was identified as optimal for arch shape description. The tool's predictive accuracy was validated through comparative analysis.
Conclusion: This research introduces an effective automated method for selecting dental arch shapes. The tool demonstrated substantial accuracy, with the potential to significantly enhance orthodontic diagnostic and treatment planning processes. Future research could further refine this methodology by exploring advanced mathematical models and incorporating machine learning techniques to optimize the selection process.

Keywords: 3D dental models, arch form analysis, dental morphometrics, computer vision, orthodontic automation

Article history

Received:
17th/June/2024

Modified:
18th/March/2024

Accepted:
16th/April/2025

Ing, vol. 30, no. 1, 2025. e22383

[*] **Correspondence:** jdtamayoq@unal.edu.co

Resumen

Contexto: La predicción precisa de la forma del arco dental es crucial para el tratamiento ortodóncico y la creación de aparatos dentales personalizados. Este estudio introduce una herramienta basada en visión por computadora para predecir las formas de arcos en modelos dentales digitales 3D.
Objetivo: Automatizar la selección de formas del arco dental mediante el análisis matemático de modelos.
Método: De 484 modelos dentales digitales, se seleccionaron 50 mediante criterios específicos, que fueron clasificados por expertos en formas ovoides, cuadradas y cónicas usando plantillas 3M. Se desarrolló una herramienta de predicción automatizada de la forma del arco dental 3D, incorporando alineación automática, detección de cúspides, ajuste de curvas con un polinomio de sexto orden y comparación de modelos. Nuestras validaciones emplearon el análisis de concordancia de atributos, la raíz del error cuadrático medio, la suma de errores cuadráticos y un estudio Gage R&R.
Resultados: El estudio logró una tasa de concordancia del 90 % en la comparación de evaluador *vs.* estándar para la mandíbula inferior, además de 78 % para la mandíbula superior. El estudio Gage R&R afirmó la fiabilidad de la medición, y el modelo de polinomio de sexto orden fue identificado como óptimo para la descripción de la forma del arco. La precisión de la herramienta fue validada mediante análisis comparativo.
Conclusión: Esta investigación introduce un método automatizado para seleccionar formas de arcos dentales. La herramienta demostró una precisión sustancial, con potencial para mejorar significativamente los procesos de diagnóstico y planificación de tratamientos ortodóncicos. Futuras investigaciones podrían refinar esta metodología explorando otros modelos matemáticos e incorporando técnicas de aprendizaje automático.

Palabras clave: modelos dentales 3D, análisis de la forma del arco, morfometría dental, visión por computadora, automatización en ortodoncia

Table of contents

1 Introduction

The analysis and classification of dental arch shapes play a pivotal role in orthodontics and dentistry. Dental arch shapes are not only indicative of individual oral health but also serve as fundamental elements in the diagnosis and treatment planning of various dental conditions (1,2). Understanding the variations in dental arch forms is essential for effective diagnosis, treatment planning, and orthodontic interventions (3). Traditionally, the classification of dental arch shapes has heavily relied on manual observation and subjective interpretation by orthodontic professionals (4). However, with the rapid advancements in digital imaging technology and computational methods, there are unprecedented opportunities to enhance the accuracy and objectivity of arch shape classification (5–7).

The integration of digital imaging technology has revolutionized the field of dentistry, providing clinicians with high-resolution, three-dimensional representations of dental structures (8). These digital models offer detailed insights into the morphology and spatial relationships of dental arches, allowing for more precise analysis and classification (9). Moreover, computational methods, such as machine learning algorithms and geometric modeling techniques, have emerged as powerful tools for automating the classification process and reducing the inherent subjectivity associated with manual assessments (10, 11). Recent developments in these methods have shown particular promise in reconstructing complete jaw morphology from partial fragments – a capability with significant implications for both clinical dentistry and forensic investigations (12).

Beyond clinical applications, dental arch analysis holds significant value in forensic odontology. The unique morphological characteristics of dental arches serve as reliable biometric markers for human identification in mass disasters, criminal investigations, and anthropological studies (13). Unlike other biological features that may degrade *postmortem*, dental structures maintain their integrity, which makes them crucial for identification when traditional methods fail. Automated classification systems can significantly enhance the efficiency and accuracy of matching *antemortem* and *postmortem* dental records, providing valuable tools for forensic experts (14). The potential to predict complete jaw morphology from fragmentary remains could revolutionize disaster victim identification, particularly in cases involving traumatic injuries or skeletal fragmentation.

In addition to these technological advancements, recent studies have explored the efficacy of various machine learning models in accurately classifying dental arch shapes based on specific geometric features and landmarks. For instance, neural networks and support vector machines have shown promising results in identifying unique dental arch characteristics with high levels of

precision (15). These computational models benefit from large datasets of digital dental impressions, enhancing their learning capabilities and accuracy over time (16).

By leveraging digital imaging technology and computational methods, orthodontic professionals can now conduct comprehensive analyses of dental arch shapes with greater efficiency and precision (17). Automated classification algorithms can rapidly process large volumes of digital data, identifying subtle variations and patterns that may go unnoticed during manual assessments (18). Furthermore, these algorithms can incorporate objective criteria and standardized metrics, leading to more consistent and reproducible classification outcomes across different practitioners (20).

This study aims to develop a novel methodology for the automated prediction and classification of dental arch shapes using digital dental models and computational techniques. By integrating expertise from orthodontics, mathematics, and computer science, it seeks to address the limitations of traditional classification methods and provide a robust framework for the objective analysis of arch shapes. This methodology is grounded in the application of geometric modeling techniques to identify patterns and distinctive features in digital models of dental arches. Furthermore, the goal is to develop a classification system that can adapt to various dental morphologies and provide accurate and consistent results across different clinical settings.

This document is structured to first introduce the innovative methodology for the automated prediction and classification of dental arch shapes, which leverages digital models and computational techniques. It then presents the findings from the application of this novel approach, followed by a discussion on its efficiency, accuracy, and potential implications for clinical practice. Finally, the conclusions section summarizes the study's contributions to the field of dental arch shape classification and outlines suggestions for future research and development.

2 Methodology

2.1 Data collection

The dataset used included the Colombian population, covering various sexes and ethnic origins, represented through original dental models made from plaster. These models were digitized using the AutoScan-DS-EX scanner, a key step in capturing the intricate details necessary for analyzing dental arch shapes. This study utilized a database of 484 digital models of dental pairs from patients at the Faculty of Dentistry of Universidad Nacional de Colombia, which were collected between 2016 and 2019. To this extensive collection, a detailed selection process was applied, guided by the following inclusion criteria.

This allowed refining the dataset to 50 dental digital models. A significant number of models from the initial group were excluded due to factors such as poor scanning quality or missing teeth in the plaster models.

Table I. Inclusion and exclusion criteria

Inclusion criteria	Exclusion criteria
☑Fully erupted and complete dentitions from first molar to first molar ☑ Digital plaster models of good quality, ensuring visibility of all teeth ☑ Individuals must be legal adults, aged 18 years and older	☑At least one missing tooth in maxilla or mandible ☑Incomplete dental models ☑Altered digital dental models (improperly scanned, few points of view and other errors) ☑Direct and indirect restoration work that affects the normal size and shape of the teeth ☑Congenital defects or deformed teeth

2.2 Arch form classification

This study involved the participation of three evaluators with experience in orthodontics:

- **Evaluator 1:** Resident from the Orthodontics and Maxillofacial Orthopedics specialty at Universidad Nacional de Colombia
- **Evaluator 2:** Resident from the Orthodontics and Maxillofacial Orthopedics specialty at Universidad Nacional de Colombia
- **Evaluator 3:** General dentist from Universidad Autónoma de Manizales

Three evaluators, experts in traditional arch shape classification methods such as observation and the use of morphometric templates, were tasked with classifying the arch shapes. For this analysis and diagnosis, an approach based on the majority vote (21) among the evaluators was used, in order to minimize the subjectivity of individual assessment, thereby improving the objectivity of the evaluation. In total, 50 dental models were classified, and each model was evaluated twice with a 15-day interval to assess the reproducibility and repeatability of the evaluators' classifications.

The arch shapes were classified into three main types: ovoid, square, and tapered. This classification aligns with common clinical practice in orthodontics (22).

The study protocol received approval from the Ethics Committee of the Faculty of Dentistry at Universidad Nacional de Colombia. This took place on September 28, 2020, under minute number 20-20 B.CIEFO-154-2020. This approval demonstrates the ethical and methodological rigor with which the study was approached.

2.2.1 Utilization of 3M morphometric arch form templates

In this study, we used 3M arch form templates designed to classify dental arches into three main shapes: ovoid, tapered, and square (Fig. 1).

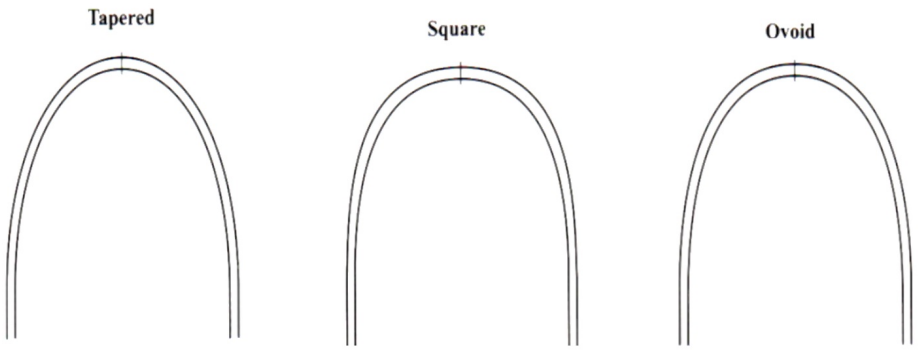


Figure 1. 3M morphometric arch form templates: standardized shapes for orthodontic assessment, showcasing tapered, square, and ovoid forms for classifying patient arch types

2.2.2 Majority voting method

The majority voting method is an approach for aggregating assessments by multiple evaluators into a consensual decision. This method excels in scenarios with evaluators of varied experience levels and mitigates subjectivity in decision-making. By employing this technique, the collective judgment of the evaluators was harnessed for a more precise classification of dental arch shapes (ovoid, tapered, and square) as commonly carried out in orthodontics.

Majority vote consensus for dental arch classification. During the dental arch shape classification, the evaluators independently assessed each model and cast their vote for one of the three shapes. The classification receiving the majority of votes was designated as the winner for that particular model. Mathematically, the consensual decision C for a sample can be represented as follows:

$$C = \underset{k \in \{\text{ovoid, tapered, square}\}}{\arg\max} \sum_{i=1}^{n} \mathbb{1}(E_i = k) \tag{1}$$

where E_i denotes the classification chosen by evaluator i, and $\mathbb{1}$ is an indicator function equal to 1 if the internal condition is true (*i.e.*, if evaluator i chooses classification k), and 0 otherwise. The classification k accumulating the highest total sum of indicator function values among all evaluators is selected as the consensus classification.

Weighted majority vote. To further refine the technique and account for the varying levels of experience among evaluators, a weighted majority vote was implemented. Here, the votes of each evaluator were assigned a weight reflecting their level of expertise, thereby enhancing the reliability of the decision-making process based on experience. The weighted consensus decision C_w was determined by

$$C_w = \underset{k \in \{\text{ovoid, tapered, square}\}}{\arg\max} \sum_{i=1}^{n} w_i \, \mathbb{1}(E_i = k) \tag{2}$$

where w_i represents the weight given for experience, which is attributed to the vote of evaluator i, wherein more experience carries more weight. This approach ensures that the final classification decision is fair by the evaluators and thus yields a more reliable result.

The use of majority vote consensus and weighted majority voting methods provides reliability to the classification of dental arch shapes, integrating the collective expertise of the evaluators.

2.3 Tool development

This study introduces a comprehensive four-phase methodology for analyzing dental arch shapes, which leverages advanced imaging and mathematical modeling techniques. The initial phase focuses on automatic dental alignment through 3D to 2D image transformation, simplifying the complexity of spatial structures for a more accessible analysis. The subsequent phases involve detailed cusp detection, precise curve fitting using a sixth-order polynomial model based on 21 reference points, and the optimization of arch shape selection through model comparison.

2.3.1 First phase: automatic dental alignment

This phase is based on the 3D to 2D projection and allows applying 2D image processing techniques. This not only simplifies the spatial complexity inherent in 3D shapes but also aligns the models to the center of the image in order to ensure a uniform orientation across all dental models.

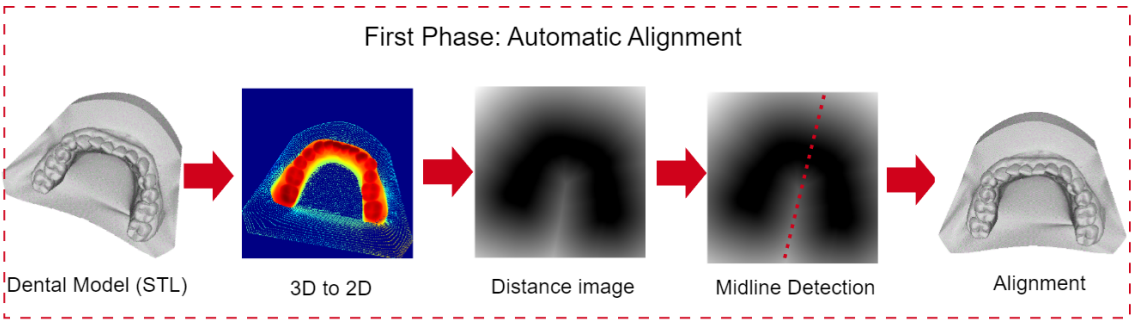


Figure 2. Automatic alignment overview: transformation and alignment of 3D dental models to 2.5D for consistent analysis orientation

Transformation of images from 3D to 2D. The transformation of 3D point clouds into 2D images facilitates the use of established image processing techniques by encoding spatial information into a two-dimensional format. This is achieved through a projection mechanism that maps 3D coordinates onto a 2D plane, thus simplifying the representation of spatial structures.

The projection is executed using a virtual camera system, which is conceptually placed on a plane parallel to the XY plane of the 3D space. The orientation of the camera is defined such that its viewing direction is perpendicular to this plane, aligned along the negative Z-axis in order to capture the depth of the points as variations in pixel intensity.

The mathematical expression for this is given by

$$P_{2D}(x, y) = \text{Project}(P_{3D}(x, y, z)) \tag{3}$$

where $P_{3D}(x, y, z)$ represents the coordinates of a point within the three-dimensional space, while $P_{2D}(x, y)$ corresponds to the resulting coordinates on the two-dimensional flat image. The projection function, denoted as Project, maps each point in the 3D space (P_{3D}) to a pixel on the 2D image (P_{2D}). This mapping process encodes the depth information of the Z-axis of each 3D point into the grayscale intensity of the 2D image pixels. By adopting this approach, we preserve the spatial relationships of the original 3D structure, allowing for further analysis and visualizations within a computationally more accessible two-dimensional framework.

When two or more points from the 3D space are projected onto the same point in the 2D space, their centroid is calculated. The centroid represents the average position of all the points, effectively summarizing their collective position in the two-dimensional representation. This approach helps to maintain the integrity of the spatial information despite the projection from 3D to 2D, ensuring that overlapping points are represented by a single, meaningful location on the 2D image.

Distance transformation. The concept of the *distance transform* (DT) fundamentally deals with calculating the distance from each point on a plane to a specified subset within it. Consider a map $I : \Omega \subset \mathbb{Z}^2 \to [0, 1]$ where the domain Ω is discretely defined as $\{1, ..., n\} \times \{1, ..., n\}$. This mapping I is analogous to a digital binary image, with 0 denoting black and 1 representing white.

The foreground set $Ow \in \Omega$ is defined as the collection of all pixel locations in Ω where the image value equals 1 (white pixels):

$$Ow = \{\mathbf{p} \in \Omega \mid I(\mathbf{p}) = 1\} \tag{4}$$

In contrast, the background set $Ow^C \in \Omega$ constitutes the complement of Ow, encompassing all pixels not in the foreground.

The DT then generates a map $DT : Ow \to \mathbb{R}$, where each pixel $\mathbf{p} \in Ow$ is assigned the shortest distance to the set Ow^C. This is formally expressed as follows:

$$DT(\mathbf{p}) = \min_{\mathbf{q} \in Ow^C} [d(\mathbf{p}, \mathbf{q})] \tag{5}$$

Here, $d(\mathbf{p}, \mathbf{q})$ signifies a distance metric between points $\mathbf{p}$ and $\mathbf{q}$, with the goal of identifying the nearest point to $\mathbf{p}$ in Ow^C. To this effect, the Euclidean distance is employed for d, thereby designating DT as the Euclidean distance transform (EDT).

To enhance the utility of the EDT, we incorporate three subsequent refinement steps: *i)* identifying the maximum EDT value along the Y-axis, *ii)* conducting a simple linear regression analysis, and *iii)* aligning the model with the centerline of the Cartesian plane, which serves as a standardized reference in our methodology.

Midline detection. The final step in our methodology involves aligning the dental model with the Cartesian plane's centerline. This is achieved by calculating the angle between the detected dental midline and a standard 90º reference line, utilizing the slope β_1 derived from our regression analysis. The angle θ, representing the deviation from the vertical, is determined as follows:

$$\theta = \begin{cases} 90 - \left(\arctan\left(\frac{y_i}{x_i}\right) \cdot \frac{180}{\pi}\right), & \text{if } \beta_1 > 0 \\ 0, & \text{if } \beta_1 = 0 \\ 90 + \left(\arctan\left(\frac{y_i}{x_i}\right) \cdot \frac{180}{\pi}\right), & \text{if } \beta_1 < 0 \end{cases} \tag{6}$$

This equation accommodates the orientation of the gradient β_1, ensuring that the angle θ finely tunes the model's alignment to the vertical axis of the Cartesian plane by standardizing the reference alignment for all models.

2.3.2 Second phase: cusp detection

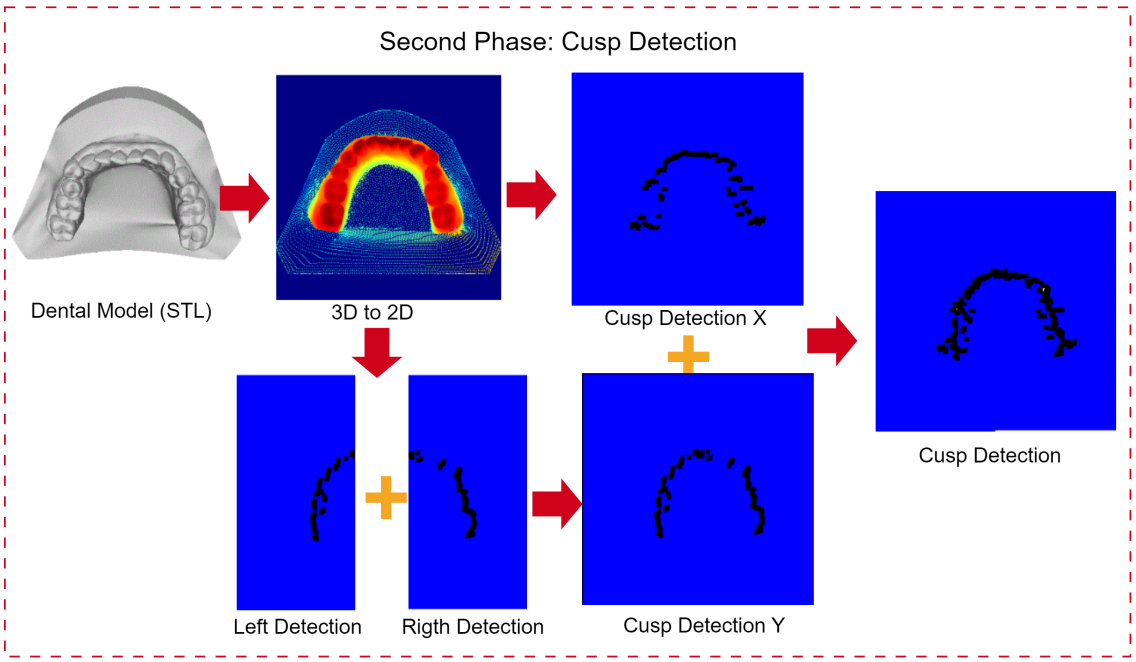


Figure 3. Cusp detection. This approach transforms a 3D dental model into a 2D representation, detects cusps separately on each half along the Y-axis, integrates these detections, and further applies cusp detection along the X-axis, resulting in a detailed cusp image.

This phase begins by transforming the projections from 3D to 2D and then duplicating the 2D image to identify the maximum intensity values along both the X- and Y-axes. It is important to note that this 2D representation is not a simple flat image; it serves as a false-color depth map, where the depth information is encoded in the image's colors

For the Y-axis, an additional process is carried out: the image is bisected to locate the maximum values or cusps. Once these maximum values are identified, the images are merged to produce a complete representation of the cusps in the digital dental model.

During this phase, the following steps were taken:

1. **Image halving along the Y-axis**. The projected image is divided along the X-axis, enabling a targeted search for maximum pixel values, which enhances cusp detection precision.

$$\text{Imagehalf} = \text{Split}(P2D(x, y)) \tag{7}$$

2. **Maximum value search along the Y-axis.** Within each half, an exhaustive search is conducted to find the maximum values along the Y-axis. The corresponding equation is given by:

$$\text{MaxVal}_{\text{right/left}}(y) = \max x \in \text{Half}_{\text{right/left}}(P2D(x, y)) \tag{8}$$

3. **Maximum value search along the X-axis.** Unlike that for the Y-axis, the search along the X-axis spans the entire width of the image, so it is not necessary to divide the image to search for the maximum intensity pixels on this axis. The corresponding equation is given by:

$$\text{MaxVal}x = \max y(P_{2D}(x, y)) \tag{9}$$

4. **Image integration.** The previously found maxima are merged into an image that contains all the cusps or maximum intensities along the X- and Y-axes. The process is mathematically represented as:

$$P_{\text{Integrated}} = \text{Merge}(\text{MaxVal}_{\text{right/left}}(y), \text{MaxVal}x) \tag{10}$$

This formula illustrates the merging of the maximum intensity values found in the left and right halves of the image (along the Y-axis) with those found on the X-axis. The $Merge$ function integrates these individual findings into a single image, highlighting the cusps in the digital dental model.

This strategy enables the localization of essential markers on the dental arch, facilitating a representation of the cusps to be analyzed in subsequent phases.

2.3.3 Third phase: curve fitting

In the third phase, the focus is on selecting mathematical models to accurately describe the shape of the dental arch. This phase utilizes 21 reference points, which are based on morphometric templates and arranged following the Fibonacci series, for mathematical approximation. Various mathematical models were evaluated, but a sixth-order polynomial model was chosen due to its balance between adaptability and the right mix of computational efficiency and precision in modeling the dental arch form (23–25).

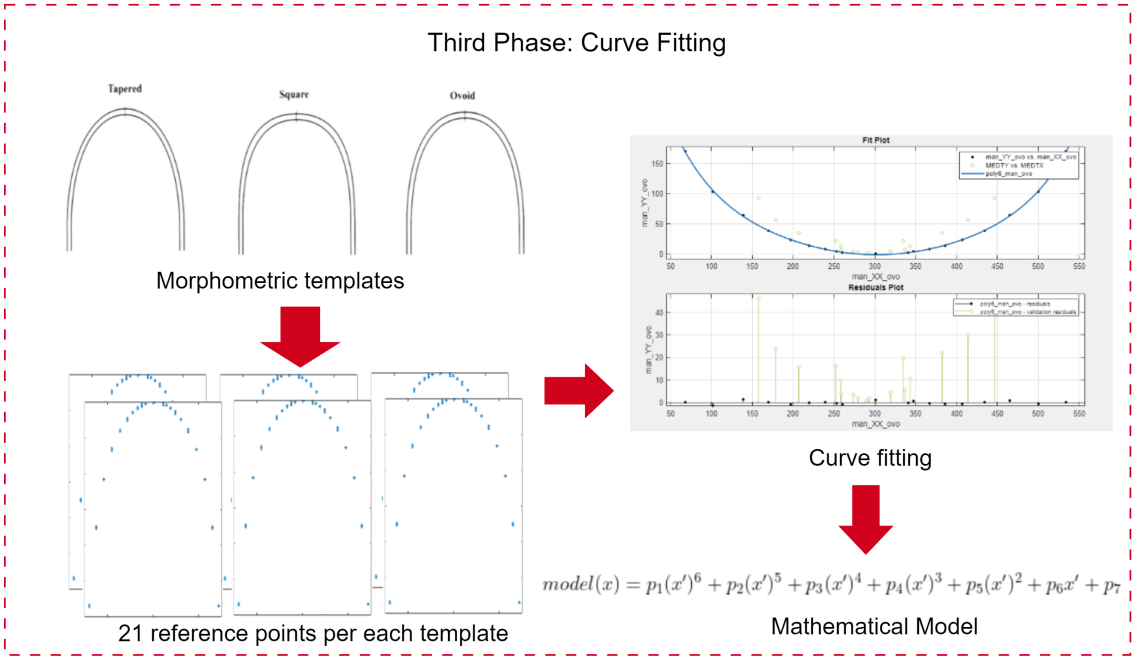


Figure 4. Third phase: curve fitting. This diagram demonstrates how the models for the dental arch shapes are created using templates (tapered, square, ovoid) with 21 reference points. These points enable precise curve fitting with a sixth-order polynomial, representing each arch.

Annotation with reference points. Each template is annotated with 21 reference points to represent the geometry of the dental arch. This includes a central point, complemented by ten points on each side, placed according to the Fibonacci sequence or golden ratio, thus reflecting the natural proportions observed in biological forms.

Exploration of mathematical models. A variety of mathematical models are evaluated, including polynomial functions, Fourier series, exponential models, and Gaussian distributions, in order to identify the most suitable representation of the dental arch shape.

Application to main arch forms. Our curve-fitting process is applied to morphometric templates representing the three main arch forms in orthodontics (ovoid, square, and tapered). This categorization is applied to both the lower and upper jaw, thereby recognizing the unique characteristics and clinical significance of each arch form in the dental model. This approach ensures that the developed mathematical models are broadly applicable and reflect the diverse arch geometries observed in clinical practice.

Selection of the sixth-order polynomial model. The sixth-order polynomial model was chosen for its ability to capture the curvatures of dental arch shapes, offering a balance between model complexity and computational tractability. This decision was supported by extensive literature documenting the

effectiveness of polynomial models in modeling dental forms. The model is defined as follows:

$$f(x) = p_1x^6 + p_2x^5 + p_3x^4 + p_4x^3 + p_5x^2 + p_6x + p_7 \tag{11}$$

where $f(x)$ represents the curve fitting function, with the coefficients $p_1, \ldots, p_7$ optimized to align with the arch shapes.

2.3.4 Fourth phase: optimization of arch shape selection through model comparison

This phase is characterized by a detailed comparative analysis between the cusp images and the sixth-order polynomial models fitted based on the 21 reference points annotated on the morphometric templates.

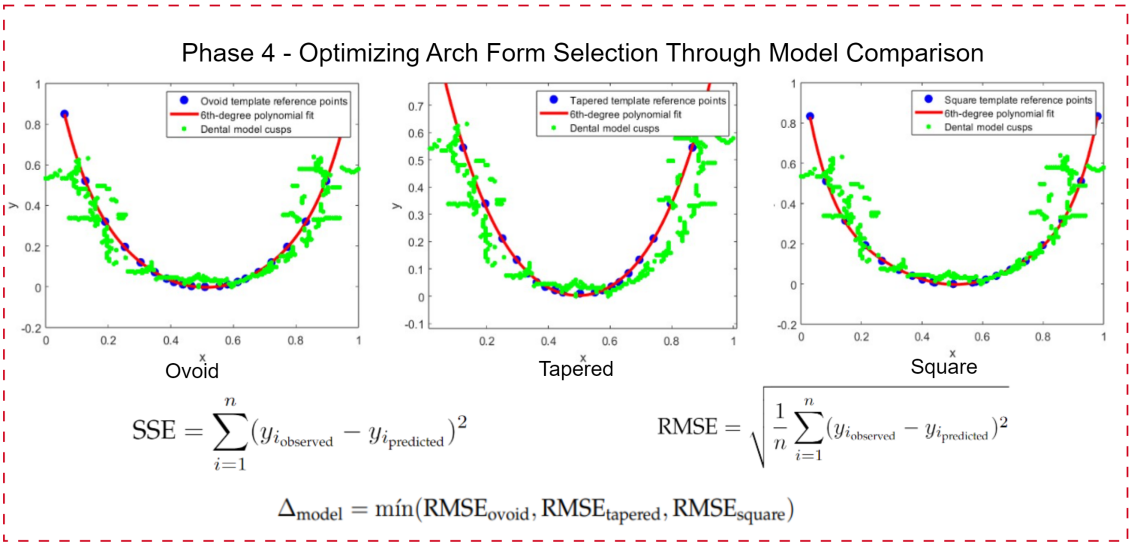


Figure 5. Comparative analysis of dental arch shapes. This diagram illustrates the optimization process of selecting the most accurate mathematical model for each dental arch shape, using error metrics (SSE and RMSE) against reference points from the cusp image.

The comparative analysis was conducted using advanced curve-fitting techniques, where the similarity between theoretical models and empirical data was assessed using error metrics such as the sum of squared errors (SSE) and the root mean square error (RMSE). These metrics provide a quantitative measure of the models' fidelity in replicating the actual shapes of the arch as outlined by the cusp images.

Below is a description of the aforementioned error metrics used to quantify the alignment between the mathematical models and the cusp images:

- **Sum of squared errors (SSE).** This metric aggregates the squared differences between the observed cusp locations in the images and the corresponding predictions made by the mathematical model.

$$\text{SSE} = \sum_{i=1}^{n} (y_{i_{\text{observed}}} - y_{i_{\text{predicted}}})^2 \tag{12}$$

where $y_{i,\text{actual}}$ are the actual cusp positions and $y_{i,\text{model}}$ are the positions predicted by the model.

- **Root mean square error (RMSE).** This metric provides a measure of the average magnitude of the error, offering a more interpretable assessment of model performance.

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_{i_{\text{observed}}} - y_{i_{\text{predicted}}})^2} \tag{13}$$

Interpretation of SSE and RMSE in arch shape comparison. The interpretation of the SSE and the RMSE in the context of dental arch shape comparison extends beyond mere numerical values. While high SSE and RMSE values might initially indicate discrepancies between the mathematical models and the actual cusp images, these metrics play a broader and combined role in our analysis:

- Rather than outright disqualifying models due to high error values, the SSE and RMSE provide insights into which arch shape (ovoid, tapered, or square) best fits the three-dimensional structure of the dental arch under study.
- Through comparative analysis, these metrics enable the evaluation of relative accuracy among models, identifying the one that most accurately reflects the specific geometries of the empirical data.

Therefore, the essence of SSE and RMSE in our research is to enable a comparative evaluation, highlighting the arch shape model that most faithfully represents the arch form:

$$\Delta_{\text{model}} = \min(\text{RMSE}_{\text{ovoid}}, \text{RMSE}_{\text{tapered}}, \text{RMSE}_{\text{square}}) \tag{14}$$

Here, Δ_{model} signifies the model with the lowest RMSE, hence the closest approximation to the actual dental arch shape, emphasizing comparative rather than absolute error analysis to identify the anatomically most accurate model.

For a better understanding of accuracy compared to the other models, we calculated the normalized error percentage for each model while considering the total SSE (the sum of the SSE across all models) and adjusting each model's error in relation to this total:

$$\text{SSE}_{\text{total}} = \text{SSE}_{\text{ovoid}} + \text{SSE}_{\text{tapered}} + \text{SSE}_{\text{square}} \tag{15}$$

$$\text{SSE}_{\text{model\_percentage}} = \left(1 - \frac{\text{SSE}_{\text{model}}}{\text{SSE}_{\text{total}}}\right) \times 100 \tag{16}$$

This percentage reflects the model's error relative to the collective error, where values closer to 100 indicate higher model precision. This numerical representation offers an understandable metric for selecting the most accurate model, which would be the one closest to 100.

2.4 Validation process

This section outlines the methodologies adopted for validating the predictive accuracy of the dental arch form classification model, which integrated attribute agreement analysis, statistical validation metrics, and a Gage R&R study to furnish a thorough evaluation framework.

2.4.1 Attribute agreement analysis

Attribute agreement analysis examines the consistency of assessments by Evaluators 1, 2, 3, and our tool, both within and against a known standard. Fleiss's *kappa* quantifies this agreement:

$$\kappa = \frac{P_o - P_e}{1 - P_e} \tag{17}$$

where P_o is the observed agreement, and P_e is the expected agreement by chance.

2.4.2 Statistical validation metrics

As previously mentioned, in order to quantify prediction accuracy, we used the RMSE and SSE, *i.e.*, as expressed in Eqs. (13) and (12), respectively.

2.4.3 Gage R&R study - XBar/R method

The Gage R&R study assesses the measurement system's variability, which is crucial for understanding the sources of variability in the measurements:

$$\text{Total Variation} = \text{Repeatability} + \text{Reproducibility} + \text{Part-to-Part Variation} \tag{18}$$

Combining these methods ensures a comprehensive and multi-faceted validation of the model's performance, affirming its reliability for clinical application.

3 Results

This research sought to establish an automated process for the selection of dental arch shapes. By employing a comparative analysis of mathematical models, the aim was to autonomously determine a model that exhibits a high degree of alignment with the arch forms commonly observed in clinical orthodontics. The methodology was anchored in the use of morphometric templates and the expertise of evaluators, whose detailed annotations laid the groundwork for model validation. The analysis was divided into four distinct phases, each crafted to achieve the goal of selecting the optimal arch shape.

3.1 Expert labeling

Expert evaluators played an indispensable role in the validation process, providing labeled data that served as the benchmark for comparing models. The precision of their annotations established the gold standard against which the performance of the tool called *Automated 3D Dental Arch Shape Prediction* was measured.

For the labeling of the refined set of models, they were categorized as belonging to the lower or upper jaw. To facilitate an exploratory examination of the results obtained by the evaluators, run charts are presented for both the lower (Fig. 6) and the upper jaw (Fig. 7). These charts aim to visually represent the consistency and variance of expert evaluations, providing insights into the diagnostic agreement across different parts of the dental arch.

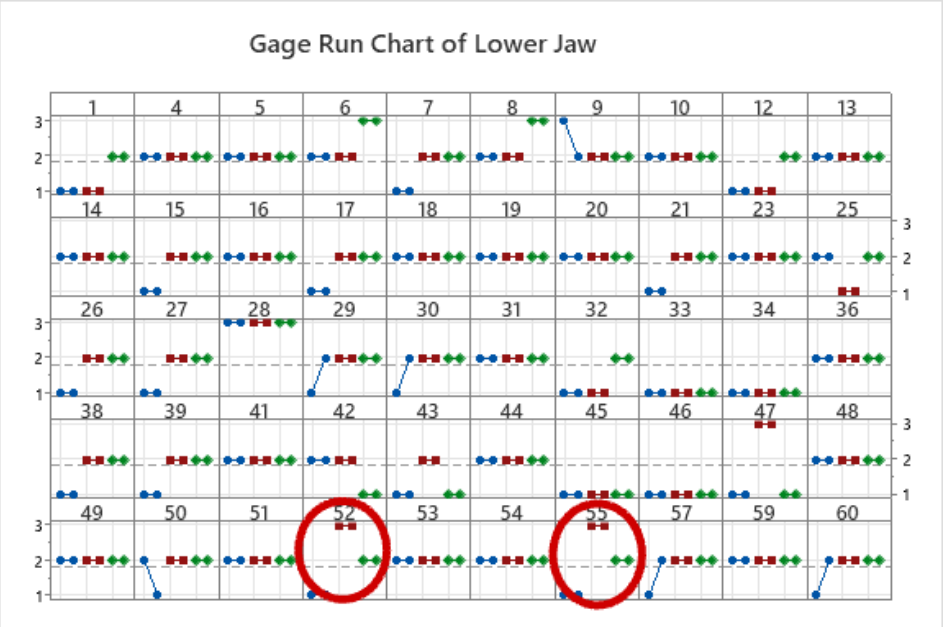


Figure 6. A run chart comparing lower jaw arch form evaluations by three experts at two time points. Arch shapes are coded as 1 (square), 2 (ovoid), and 3 (tapered), with experts labeled as *Evaluator 1* (blue), *Evaluator 2* (red), and *Evaluator 3* (green). Models 52 and 55 resulted in a tie, resolved by a majority weighted vote, where Evaluator 2, the most experienced one, made the final selection.

Comparative Gage R&R study for lower and upper jaw evaluations

The Gage repeatability and reproducibility (R&R) study, utilizing the XBar/R method, was conducted to assess the measurement systems for both lower and upper jaw analyses, showcasing the dynamics between repeatability, reproducibility, and part-to-part variability.

Table II. Comparative variance components from the Gage R&R study for the lower and upper jaws

	Lower Jaw		Upper Jaw	
Source	**VarComp**	**% Contribution**	**VarComp**	**% Contribution**
Total Gage R&R	0.029268	12.90%	0.037706	21.55%
Repeatability	0.001257	0.55%	0.004224	2.41%
Reproducibility	0.028011	12.34%	0.033482	19.13%
Part-to-part	0.197693	87.10%	0.137287	78.45%
Total variation	0.226961	100.00%	0.174993	100.00%

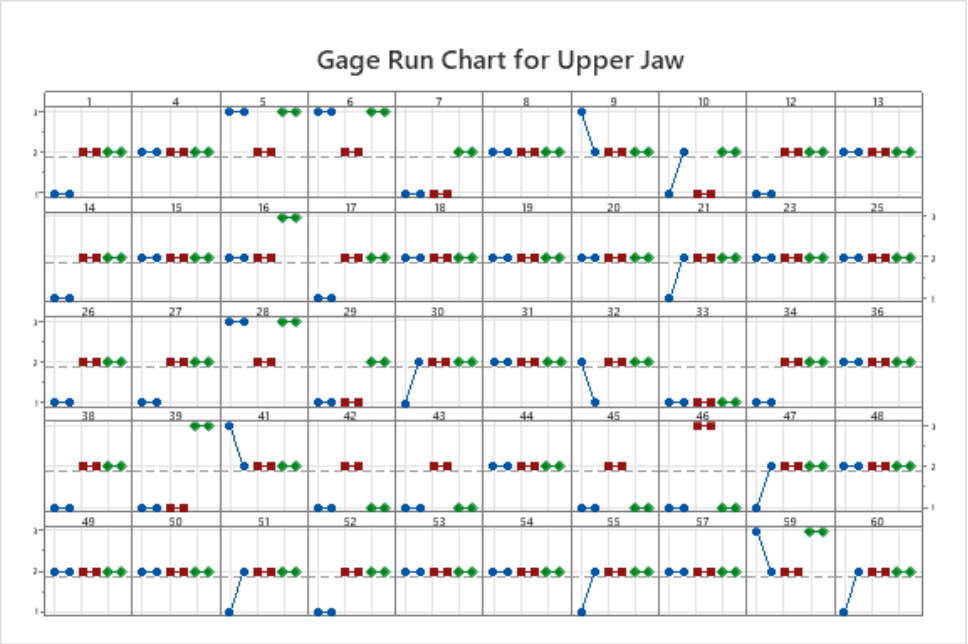


Figure 7. Run chart comparing expert evaluations for upper jaw arch form at two different time points. Arch forms are coded as 1 (square), 2 (ovoid), and 3 (tapered), with evaluators represented by color: *Evaluator 1* (blue), *Evaluator 2* (red), and *Evaluator 3* (green).

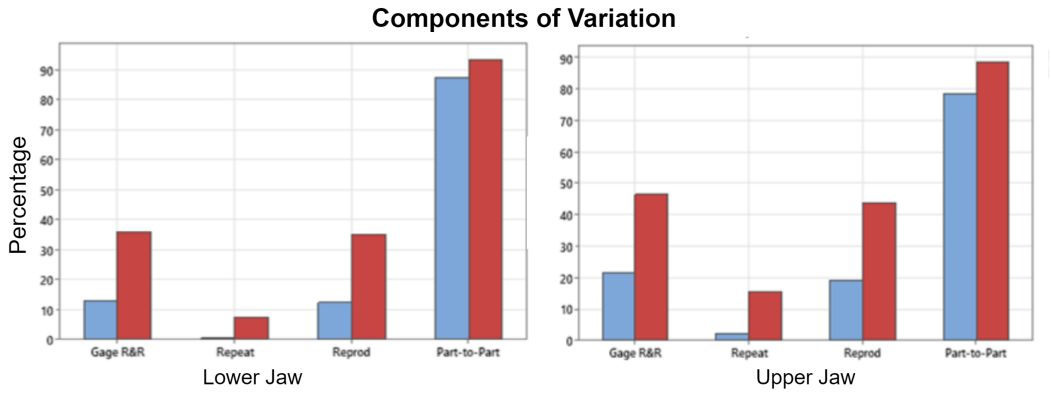


Figure 8. Bar chart illustrating the components of variation in Gage R&R studies for the lower and upper jaws, where blue bars indicate the % contribution and red bars represent the % study variation

3.2 Results of the software tool

Our study evaluated the software comprehensively rather than in isolated segments. From automatic dental alignment to cusp detection and the application of curve fitting techniques, each phase contributed to a unified system. The results provide a comprehensive perspective on the tool's effectiveness in predicting dental arch shapes and highlight its potential impact in the field of dentistry.

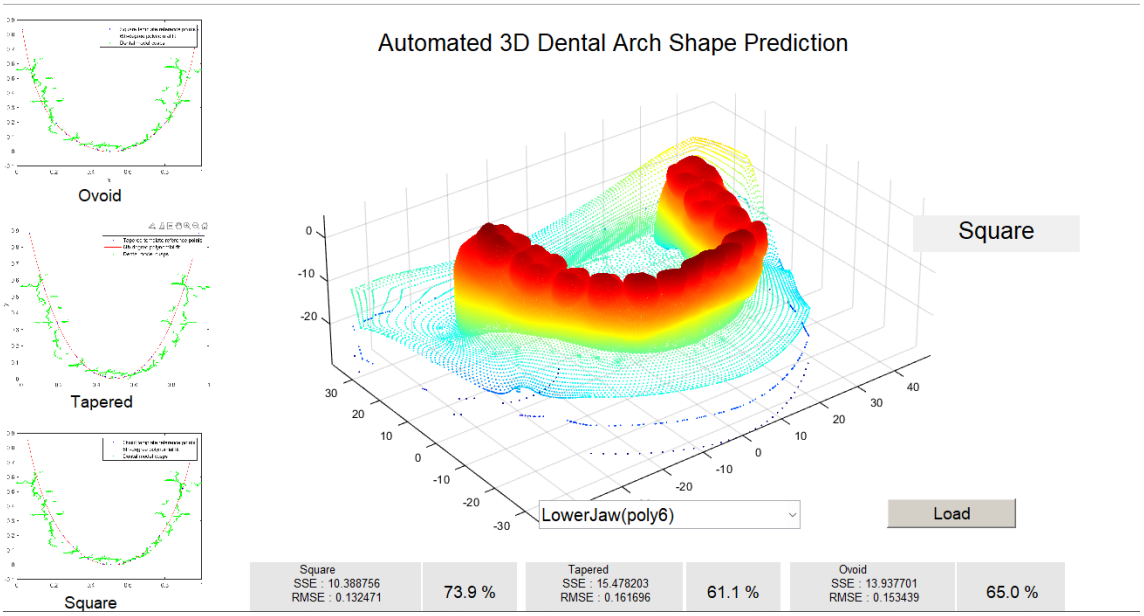


Figure 9. Interface of the tool demonstrating the fit accuracy for different dental arch shapes

3.2.1 Attribute agreement analysis

In the above-presented attribute agreement analysis, the consistency of three evaluators and our method was assessed for both the lower and upper jaws. The analysis aimed to measure the reliability of the evaluators' measurements against each other and against a standard measurement.

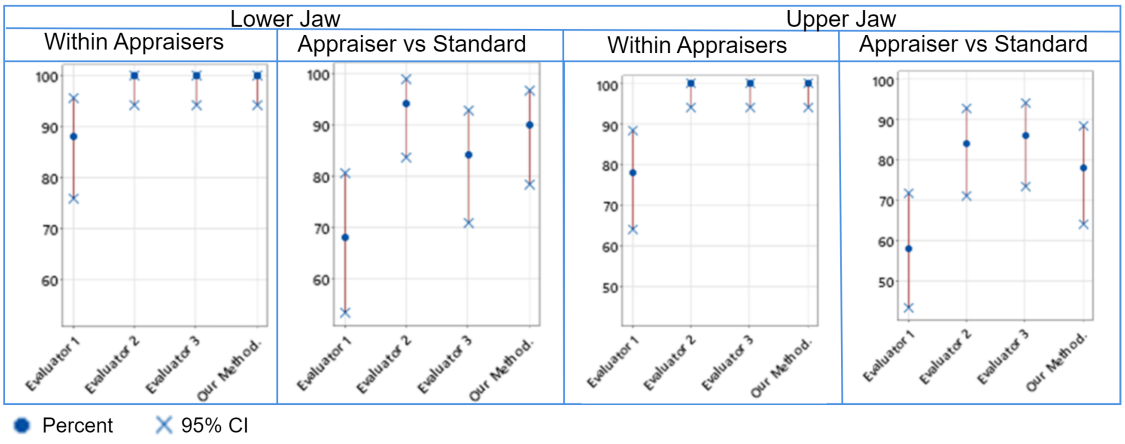


Figure 10. Comparison of intra- and inter-rater reliability in dental arch assessments for lower and upper jaws. The left panels display the within-evaluator agreement (*i.e.*, Eva 1, Eva 2, Eva 3, Our Method) with the corresponding percentages and 95% CIs. The right panels show evaluator *vs.* standard agreement, including the percentage and the 95% CIs for both jaws.

Each evaluator *vs.* the standard. This portion assessed the agreement level for each evaluator, providing *kappa* statistics, standard errors, Z-scores, and p-values for each response and overall (Table III).

All evaluators *vs.* the standard. This portion summarized the collective agreement, offering *kappa* statistics for each response and overall, along with standard errors, Z-scores, and p-values, demonstrating the group's consistency with the standard (Table III).

3.2.2 Mathematical model fit and comparison summary

The mathematical models for predicting dental arch shapes were evaluated with a focus on polynomial and non-linear models. The comparison was based on the SSE and the RMSE, aiming to identify a model that accurately captured the complexity of dental arches while maintaining a balance to prevent overfitting (Table IV).

4 Discussion

This study's introduction of a four-phase methodology for analyzing dental arch shapes combines imaging techniques with mathematical modeling to enhance the precision of dental arch shape evaluations.

The initial phase automates the 3D to 2D conversion of dental models, simplifying the analysis by reducing spatial complexity. This step is crucial for standardizing the orientation of all dental models, ensuring consistent evaluation throughout the research. Our study revealed that manual adjustments were necessary for seven dental models (five from the lower jaw and two from the upper jaw) due to their initial horizontal orientation, which rendered the automatic alignment process ineffective. This limitation arose because the alignment algorithm, designed to process models within a vertical orientation range of -90 to 90º, as shown in Eq. (6), solely relies on the Y-axis for adjustments. Consequently, models that are not vertically aligned fall outside the algorithm's capability, necessitating manual correction.

The second phase, focusing on cusp detection, demonstrates the methodology's ability to extract detailed morphological features from the dental arches.

In the third phase of our analysis, the use of a sixth-degree polynomial (poly6) model with 21 reference points for dental arch shape modeling is thoroughly justified by an examination of model fitness data, as detailed in Table IV. This strategy achieves a harmonious balance between accuracy and model complexity. The high R^2 values, nearing unity for all arch shapes (ovoid, square, and tapered), underscore the poly6 model's efficacy in capturing a significant portion of the variance in the data, thereby indicating an outstanding fit. Additionally, the model's relatively low SSE and RMSE underscore its precision in predictions with minimal average errors. The poly6 model's moderate com-

Table III. Attribute agreement analysis for the jaw

Each evaluator *vs.* the standard											
	Lower jaw						**Upper jaw**				
Evaluator	**Response**	***Kappa***	**SE *kappa***	**Z**	**P(vs > 0)**	**evaluator**	**Response**	***Kappa***	**SE *kappa***	**Z**	**P(vs > 0)**
Evaluator 1	1	0.411189	0.100000	4.1119	0.0000	Evaluator 1	1	0.345875	0.100000	3.45875	0.0003
	2	0.517622	0.100000	5.1762	0.0000		2	0.345023	0.100000	3.45023	0.0003
	3	0.423794	0.100000	4.2379	0.0000		3	0.816850	0.100000	8.16850	0.0000
	Overall	0.462560	0.087779	5.2696	0.0000		Overall	0.406666	0.082762	4.91368	0.0000
Evaluator 2	1	0.787385	0.100000	7.8738	0.0000	Evaluator 2	1	0.607843	0.100000	6.07843	0.0000
	2	0.890351	0.100000	8.9035	0.0000		2	0.545159	0.100000	5.45159	0.0000
	3	0.846390	0.100000	8.4639	0.0000		3	-0.041667	0.100000	-0.41667	0.6615
	Overall	0.845600	0.080041	10.5646	0.0000		Overall	0.499687	0.084550	5.90993	0.0000
Evaluator 3	1	0.702381	0.100000	7.0238	0.0000	Evaluator 3	1	0.607843	0.100000	6.07843	0.0000
	2	0.533800	0.100000	5.3380	0.0000		2	0.671053	0.100000	6.71053	0.0000
	3	0.290780	0.100000	2.9078	0.0018		3	0.633700	0.100000	6.33700	0.0000
	Overall	0.558499	0.081112	6.8856	0.0000		Overall	0.642675	0.077394	8.30398	0.0000
Our method	1	0.645641	0.100000	6.4564	0.0000	Our method	1	0.362155	0.100000	3.62155	0.0001
	2	0.780702	0.100000	7.8070	0.0000		2	0.378882	0.100000	3.78882	0.0001
	3	0.846390	0.100000	8.4639	0.0000		3	0.645390	0.100000	6.45390	0.0000
	Overall	0.742666	0.080041	9.2785	0.0000		Overall	0.412707	0.081497	5.06406	0.0000
All evaluators *vs.* the standard											
Response		***Kappa***	**SE *kappa***	**Z**	**P(vs > 0)**	Response		***Kappa***	**SE *kappa***	**Z**	**P(vs > 0)**
1		0.636649	0.0500000	12.7330	0.0000	1		0.480929	0.0500000	9.6186	0.0000
2		0.680619	0.0500000	13.6124	0.0000	2		0.485029	0.0500000	9.7006	0.0000
3		0.601839	0.0500000	12.0368	0.0000	3		0.513568	0.0500000	10.2714	0.0000
Overall		0.652331	0.0411533	15.8513	0.0000	Overall		0.490434	0.0407967	12.0214	0.0000

Table IV. Model fitness by shape

Shape	Model	R-square	SSE	DFE	Adj-R-sq	RMSE	#Coeff.
Ovoid	fourier4	0.99995	2.9067	9	0.99989	0.5683	10
Ovoid	poly6	0.99982	9.6444	12	0.99973	0.89649	7
Ovoid	poly4	0.99914	45.944	14	0.9989	1.8116	5
Ovoid	fourier2	0.99915	45.924	13	0.99882	1.8795	6
Square	fourier4	0.99998	1.1142	9	0.99996	0.35185	10
Square	poly6	0.99986	8.2757	12	0.99978	0.83045	7
Square	poly4	0.99832	98.527	14	0.99784	2.6529	5
Square	fourier2	0.9983	97.393	13	0.99764	2.7371	6
Tapered	fourier4	0.99998	0.99108	9	0.99997	0.33184	10
Tapered	fourier2	0.99978	12.823	13	0.9997	0.99315	6
Tapered	poly6	0.99976	14.168	12	0.99964	1.0866	7
Tapered	poly4	0.9995	29.509	14	0.99935	1.4518	5

plexity, indicated by a number of coefficients (#Coeff.= 7), is pivotal in preventing overfitting, thus ensuring that the model faithfully represents the dental arch shapes without undue influence from data noise. The high adjusted R^2 values across various arch shapes further highlight the model's generalization ability, making the poly6 model the preferred choice for dental arch shape analysis within our study.

The fourth phase of our work concentrated on improving the selection of dental arch shapes through an in-depth comparison of cusp images against sixth-order polynomial models. By employing error metrics such as the SSE and RMSE, this approach enables a comprehensive evaluation aimed at determining which model most faithfully reflects the true architecture of the dental arch.

Expert evaluators played a crucial role in the validation process by providing labeled data that served as a benchmark for comparing models. Their precise annotations established the gold standard against which the performance of the automated 3D dental arch shape prediction tool was measured.

The Gage R&R run charts (Figs. **6** and **7**) for lower and upper jaw evaluations by the three experts illustrate the usefulness of majority voting in standardizing dental arch shape classification. Variations in the lower jaw consensus suggest a learning curve or interpretation issues among the evaluators, whereas the upper jaw shows more consistency, hinting at clearer classification or increased evaluator confidence. Weighted voting by experience lends authority to the consensus in cases of disagreement, enhancing the validity of the decision-making process. However, the observed variability, especially in the lower jaw, underscores the need for more definitive guidelines or training to ensure uniformity across evaluations, thus improving the reliability and effectiveness of orthodontic assessments.

The comparative analysis (Table **II**) between the lower and upper jaw evaluations highlighted a discernible difference in measurement consistency, with the upper jaw exhibiting a higher total

variability (21.55%) compared to the lower jaw (12.90%). Notably, the reproducibility component accounted for the majority of this variability, indicating a need for improved evaluator agreement in upper jaw assessments. These results underscore the critical importance of considering anatomical differences between dental arches when developing and validating predictive models.

The attribute agreement analysis conducted in this study underscores a robust intra-rater and inter-rater reliability in the classification of dental arch shapes. The consistent findings within individual evaluators and between each of them and the majority vote standard suggest a high level of agreement (Fig. **10**). This concordance points to the reliability of the evaluators and the assessment method, reinforcing the majority vote as a credible standard for dental arch shape classification. It also validates our method as a consistent and reliable approach when compared to expert opinions, which is instrumental for standardizing evaluations in orthodontic practice. In Table **III**, our method aligns well with expert evaluations in classifying jaw shapes, as evidenced by generally high *kappa* values. While Evaluator 2 and Evaluator 3 show the highest agreement with the majority vote for the lower and upper jaws, respectively, negative *kappa* values in certain cases highlight the need for standardized methods to minimize assessment variability and ensure accuracy.

5 Conclusion

Attribute agreement analysis. A robust intra- and inter-rater reliability in the classification of dental arch shapes was demonstrated. The consistency observed both within individual evaluators and between each evaluator and the majority vote standard suggests a high level of agreement, thus validating the proposed tool as a consistent and reliable approach for dental arch shape classification.

Comparative Gage R&R study. Discernible differences in measurement consistency between lower and upper jaw evaluations were highlighted, with the upper jaw exhibiting higher total variability. This underscores the need to improve the agreement between evaluators in upper jaw assessments, as well as the importance of considering anatomical differences between dental arches when developing predictive models.

Mathematical model selection and fit. Mathematical models for predicting dental arch shapes were evaluated with a focus on polynomial and nonlinear models.The sixth-degree polynomial model achieved a harmonious balance between accuracy and model complexity, highlighting its efficacy in capturing a significant portion of the variation in the data.

Clinical applicability. The tool developed in this study can have a significant impact on the field of dentistry by providing an automated tool for selecting dental arch shapes, which could improve treatment outcomes and our understanding of oral health.

6 Author contributions

J.D. Tamayo-Quintero conceived the idea and conducted the background research. J.D. Tamayo-Quintero, J.B. Gómez-Mendoza, and S.V. Guevara-Pérez were responsible for data collection, methodological development, and formal analysis. All three authors contributed to the writing and revision of this manuscript, ensuring a robust and comprehensive presentation of the findings of this research.

References

[1] S.-J. Lee, S. Lee, J. Lim, H.-J. Park, and T.T. Wheeler, "Method to classify dental arch forms," *Amer. J. Orthodon. Dentofac. Orthoped.*, vol. 140, no. 1, pp. 87-96, 2011. https://doi.org/10.1016/j.ajodo.2011.03.016 ↑ 3

[2] H.-X. Yang, F.-L. Li, and L.-M. Li, "Comparison of maxillary anterior mathematical proportions among 3 dental arch forms," *J. Prosthetic Dent.*, vol. 130, no. 4, pp. 614-619, 2023. https://doi.org/10.1016/j.prosdent.2021.11.025 ↑ 3

[3] J. Tomášik *et al.*, "AI and face-driven orthodontics: A scoping review of digital advances in diagnosis and treatment planning," *AI*, vol. 5, no. 1, pp. 158–176, 2024. https://www.mdpi.com/2673-2688/5/1/9 ↑ 3

[4] K. Arai and L. A. Will, "Subjective classification and objective analysis of the mandibular dental-arch form of orthodontic patients," *Amer. J. Orthodon. Dentofac. Orthoped.*, vol. 139, no. 4, pp. e315-e321, 2011. https://www.sciencedirect.com/science/article/pii/S0889540610010243 ↑ 3

[5] R. S. Bharadwaj, S. H. Kalgeri, A. T. Shivakumar, V. Doddawad, P. M. Shankar, and S. K. B. B., "Revolutionizing dental health care: An in-depth exploration of technological advancements," *Euro. J. Gen. Dent.*, vol. 13. bo. 1, pp. 001-014, 2024. http://www.thieme-connect.com/products/ejournals/abstract/10.1055/s-0043-1776321 ↑ 3

[6] X. Yang, K. Huang, D. Yang, W. Zhao, and X. Zhou, "Biomedical big data technologies, applications, and challenges for precision medicine: A review," *Global Chall.*, vol. 8, no. 1, art. 2300163, 2024. https://doi.org/10.1002/gch2.202300163 ↑ 3

[7] A. Sikri, J. Sikri, V. Piplani, and Y. Thakur, "Applications of artificial intelligence in dentistry: A narrative review," *South Asian Res. J. Oral. Dent. Sci.*, vol. 6, no. 1, pp. 1–10, 2024. https://sarpublication.com/media/articles/SARJODS_61_1-10.pdf ↑ 3

[8] Y. Cen *et al.*, "Application of three-dimensional reconstruction technology in dentistry: a narrative review," *BMC Oral Health*, vol. 23, no. 1, art. 630, 2023. https://doi.org/10.1186/s12903-023-03142-4 ↑ 3

[9] Z. H. Ahmed *et al.*, "Artificial intelligence and its application in endodontics: A review," *J. Contemp. Dental Prac.*, vol. 24, no. 11, pp. 912-917, 2024. https://thejcdp.com/doi/JCDP/pdf/10.5005/jp-journals-10024-3593 ↑ 3

[10] J. Liu, Y. Chen, S. Li, Z. Zhao, and Z. Wu, "Machine learning in orthodontics: Challenges and perspectives," *Adv. Clin. Exp. Med.*, vol. 30, no. 10, pp. 1065–1074, 2021. https://advances.umw.edu.pl/en/article/2021/30/10/1065/ ↑ 3

[11] C. Elkhill, J. Liu, M. G. Linguraru, S. LeBeau, D. Khechoyan, B. French, and A. R. Porras, "Geometric learning and statistical modeling for surgical outcomes evaluation in craniosynostosis using 3D photogrammetry," *Comp. Meth. Prog. Biomedicine*, vol. 240, art. 107689, 2023. https://doi.org/10.1016/j.cmpb.2023.107689 ↑ 3

[12] M. Vodanović, M. Subašić, D. P. Milošević, I. Galić and H. Brkić, "Artificial intelligence in forensic medicine and forensic dentistry," *J. Forensic Odonto-Stomatology*, vol. 41, no. 2, pp. 30-41, 2023. https://pmc.ncbi.nlm.nih.gov/articles/PMC10473456/ ↑ 3

[13] Pretty, I. A., "Forensic dentistry: 1. Identification of human remains," *Dental Update*, vol. 34, no. 10, pp. 621-630, 2007. https://doi.org/10.12968/denu.2007.34.10.621 ↑ 3

[14] S. R. Dutta, P. Singh, D. Passi, D. Varghese, and S. Sharma, "The role of dentistry in disaster management and victim identification: An overview of challenges in Indo-Nepal scenario," *J. Maxillofac. Oral Surg.*, vol. 15, no. 4, pp. 442-448, Dec. 2016. https://doi.org/10.1007/s12663-016-0896-4 ↑ 3

[15] R. Ren, H. Luo, C. Su, Y. Yao, and W. Liao, "Machine learning in dental, oral and craniofacial imaging: A review of recent progress," *PeerJ*, vol. 9, e11451, 2021 https://peerj.com/articles/11451/ ↑ 4

[16] M. Cicciù *et al.*, "3D digital impression systems compared with traditional techniques in dentistry: A recent data systematic review," *Materials*, vol. 13, no. 8, art. 1982, 2020. https://www.mdpi.com/1996-1944/13/8/1982 ↑ 4

[17] S. Kapila, S. R. Vora, S. Rengasamy Venugopalan, M. H. Elnagar, and S. Akyalcin, "Connecting the dots towards precision orthodontics," *Orthodont. Craniofac. Res.*, vol. 26, pp. 8–19, 2023. https://onlinelibrary.wiley.com/doi/full/10.1111/ocr.12725 ↑ 4

[18] F. Schwendicke, W. Samek, and J. Krois, "Artificial intelligence in dentistry: Chances and challenges," *J. Dental Res.*, vol. 99, no. 7, pp. 769–774, 2020. https://journals.sagepub.com/doi/full/10.1177/0022034520915714 ↑ 4

[19] L. T. Reyes, J. K. Knorst, F. R. Ortiz, and T. M. Ardenghi, "Scope and challenges of machine learning-based diagnosis and prognosis in clinical dentistry: A literature review," *J. Clin. Translat. Res.*, vol. 7, no. 4, pp. 523, 2021. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8445629/ ↑

[20] I. Shafi, A. Fatima, H. Afzal, I. de la Torre Díez, V. Lipari, J. Breñosa, and I. Ashraf, "A comprehensive review of recent advances in artificial intelligence for dentistry e-health," *Diagnostics*, vol. 13, no. 13, art. 2196, 2023. https://www.mdpi.com/2075-4418/13/13/2196 ↑ 4

[21] G. Tullock, "Problems of majority voting," *J. Pol. Econ.*, vol. 67, no. 6, pp. 571-579, 1959. https://www.journals.uchicago.edu/doi/abs/10.1086/258244 ↑ 5

[22] A. Tiwari, A. Garg, B. Virang, S. Sahu, N. Shah, and N. Verma, "Arch form in orthodontics: A review," *J. App. Dent. Med. Sci.*, vol. 4, no. 1, p. 118, 2018. ↑ 5

[23] M. Memarpour, M. Oshagh, and M. R. Hematiyan, "Determination of the dental arch form in the primary dentition using a polynomial equation model," *J. Dent. Child.*, vol. 79, no. 3, pp. 136-142, 2012. https://www.ingentaconnect.com/content/aapd/jodc/2012/00000079/00000003/art00003 ↑ 10

[24] S. AlHarbi, E. A. Alkofide, and A. AlMadi, "Mathematical analyses of dental arch curvature in normal occlusion," *The Angle Orthodontist*, vol. 78, no. 2, pp. 281-287, 2008. https://doi.org/10.2319/121806-516.1 ↑ 10

[25] X. Wang *et al.*, "Accuracy of dental arch form in customized fixed labial orthodontic appliances," *Amer. J. Orthodon. Dentofac. Orthoped.*, vol. 162, no. 2, pp. 173-181, 2022. https://doi.org/10.1016/j.ajodo.2021.02.026 ↑ 10

Juan David Tamayo Quintero

Juan David Tamayo Quintero is an electronics engineer with a Master's degree in Industrial Automation. He holds a PhD in Automatic Engineering from Universidad Nacional de Colombia (2024). As a full-time faculty member at Tecnológico de Antioquia, he serves in the Department of Engineering's Software Engineering program. He has also lectured in Systems Engineering at EAFIT University. His research and professional interests focus on artificial intelligence and the development of computer programs for medical applications and related fields.

Email: jdtamayoq@unal.edu.co

Juan Bernardo Gómez Mendoza

Juan Bernardo Gómez Mendoza is an electronics engineer who graduated from Universidad Nacional de Colombia in 2002. He earned his Master's degree in Industrial Automation from the same university in 2004. Furthering his education, he obtained a PhD in Engineering with a focus on Automation from Universidad Nacional de Colombia in 2012, as well as a PhD in Applied Sciences from INSA de Lyon in the same year. He is an expert in artificial vision and image processing, with advanced knowledge in artificial intelligence and robotics.

Email: jbgomezm@unal.edu.co

Sonia Victoria Guevara Pérez

Sonia Victoria Guevara Pérez holds a Master of Dental Surgery (MDS) and a Master of Science in Biological Anthropology. She completed her PhD in Biomechanics at the Laboratoire de Biomécanique Appliquée (LBA). Her skills and expertise lie in 3D imaging, maxillofacial development, dentofacial orthopedics, and clinical orthodontics. Sonia is multilingual, fluent in both Spanish and French. She is currently affiliated with Universidad Nacional de Colombia in Bogotá, Colombia.

Email: svguevarap@unal.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.22185

Research

Optimal Selection of Intrinsic Mode Functions Applied to Seizure Detection

Selección óptima de funciones de modo intrínseco aplicada a la detección de convulsiones

Luis Daladier Guerrero-Otoya[1]*, Maximiliano Bueno-López[2], Eduardo Giraldo[2], Marta Molinas[3]

[1]Master's student in Automation at Universidad del Cauca (Popayán, Colombia)
[2]Professor at Universidad Tecnológica de Pereira (Pereira, Colombia)
[4]Professor at the Department of Engineering Cybernetics, NTNU (Trondheim, Norway)

Abstract

Context: Epilepsy is a severe chronic neurological disorder with considerable incidence due to recurrent seizures. These seizures can be detected and diagnosed noninvasively using an electroencephalogram. Empirical mode decomposition has shown excellent results in identifying epileptic crises.
Method: This study addressed a significant gap by proposing a novel approach for the automated selection of the most relevant intrinsic mode functions (IMFs) using empirical mode decomposition and discrimination metrics such as the Minkowski distance, the mean square error, cross-correlation, and the entropy function. The main objective was to address the challenge of determining the optimal number of IMFs required to accurately reconstruct brain activity signals.
Results: The results were promising, as they facilitated the identification of IMFs that contained the most relevant information, marking a significant advancement in the field. To validate these findings, standard methods including the correlation coefficient, the p-value, and the Wasserstein distance were employed. Additionally, an EEGLAB-based brain mapping was conducted, adding robustness and credibility to the results obtained.
Conclusions: Our method is a fundamental tool that enhances epileptic seizure identification from EEG signals, with significant clinical implications in the diagnosis and treatment of epilepsy.

Keywords: seizure identification, empirical mode decomposition, optimal selection of IMFs, intrinsic mode functions, discrimination metrics

Article history

Received: 17th/May/2024

Modified: 5rd/March/2025

Accepted: 8th/April/2025

Ing, vol. 30, no. 1, 2025. e22185

* **Correspondence:** luisdaladier@unicauca.edu.co

Resumen

Contexto: La epilepsia es un trastorno neurológico crónico grave con una incidencia considerable debido a convulsiones recurrentes. Estas convulsiones pueden ser detectadas de manera no invasiva y diagnosticadas mediante un electroencefalograma. La descomposición modal empírica ha mostrado excelentes resultados en la identificación de crisis epilépticas.
Métodos: Este estudio abordó una brecha significativa al proponer un enfoque novedoso para la selección automatizada de las funciones de modo intrínseco (IMF) más relevantes utilizando descomposición empírica de modo y métricas de discriminación tales como la distancia de Minkowski, el error cuadrático medio, la correlación cruzada y la función de entropía. El objetivo primario fue abordar el desafío de determinar el número óptimo de IMF requeridas para reconstruir con precisión las señales de actividad cerebral.
Resultados: Los resultados fueron prometedores, pues facilitaron la identificación de IMF que contenían la información más relevante, marcando un avance significativo en el campo. Para validar estos hallazgos, se emplearon métodos estándar, incluyendo el coeficiente de correlación, el valor p y la métrica de Wasserstein. Además, se realizó un mapeo cerebral con EEGLAB, lo que agregó robustez y credibilidad a los resultados obtenidos.
Conclusiones: Nuestro método es una herramienta fundamental que permite mejorar la identificación de convulsiones epilépticas a partir de señales de EEG, con importantes implicaciones clínicas en el diagnóstico y tratamiento de la epilepsia.

Palabras clave: Identificación de convulsiones, descomposición modal empírica, selección óptima de IMFs, funciones modales intrínsecas, métricas de discriminación

Table of contents

1 Introduction

Epilepsy is a neurological disorder that affects approximately 50 million people around the world (1,2). To address this issue, electroencephalography (EEG) is used to measure the electrical activity of the brain, which has become the standard of care in epilepsy treatment (3). Nevertheless, due to its non-stationary nature and its nonlinearity, this is still an open problem in the field of neuroscience (4–7), which has been explored using different pattern recognition techniques derived from classical approaches, *e.g.*, Fourier transforms (8–10), Wavelet transforms (11–13), Stockwell transforms (14), and empirical mode decomposition (EMD) (15–18).

Many research studies have been carried out to enhance the detection and classification of epileptic seizures in real time. For instance, (19) presented a new method for the classification of epileptic seizures which applies EMD to select the two most relevant intrinsic mode functions (IMFs) to extract two features of energy and two features of spectral dimension, while (20) proposed a multi-objective optimization approach to solving two non-restricted objectives, using NSGA-II and NSGA-III algorithms to improve classification accuracy and reduce the number of channels. On the other hand, (21) presented a novel method for the classification of epileptic seizures that exclusively uses Tsallis entropy at a low computational cost, and (22) presented the thresholded offset Gaussian (TOG) statistical algorithm, which acts as a noise filter and extracts seven features.

This work presents a method for the optimal selection of IMFs from real EEG signals. This approach is data-driven and applies EMD in the pre-processing stage. The selection of the most relevant IMFs is based on a threshold and a set of discrimination metrics such as the Minkowski distance, cross-correlation, the mean square error (MSE), and entropy. The IMFs with the most relevant information are selected for brain mapping and to identify active electrodes.

The remainder of this paper is organized as follows: Section 2 outlines the dataset used, EMD, and the proposed method; Section 3 presents and discusses the results; and Section 4 states the conclusions of this work.

2 Methodology

This section outlines the methodology used in this study. Fig. **1** illustrates our method in the form of a block diagram. The methodology comprises two main components: a description of the dataset used and an explanation of the EMD method and our proposal, which employs restriction metrics. The overarching objective of this research endeavor was to minimize the number of IMFs required for the reconstruction of EEG signals while retaining relevant information. This systematic approach facilitates a comprehensive understanding of the experimental setup and the analytical framework employed in this work.

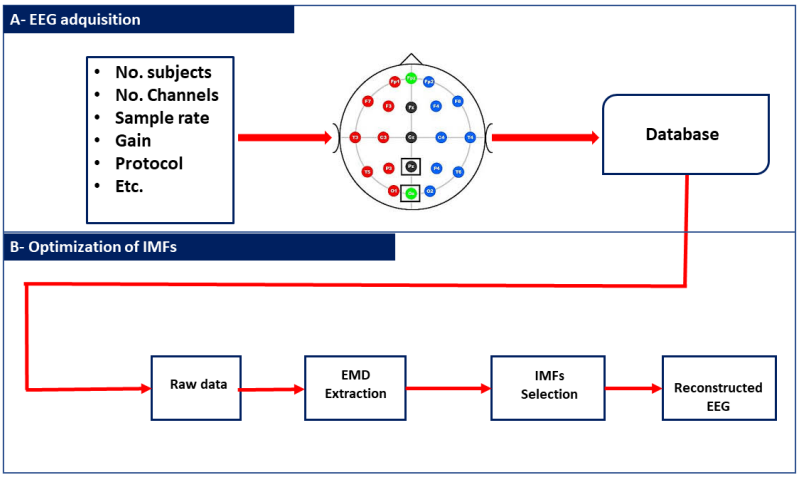


Figure 1. Implemented methodology

2.0.1 Data description

The dataset used in this work was the CHB-MIT Scalp EEG database (23,24), which contains 939 hours of recordings that featured 173 seizures (clonic, atonic, tonic). This data set was developed at Boston Children's Hospital, involving patients with intractable seizures who were monitored for several days after the withdrawal of antiseizure medication, in order to characterize their seizures and assess their candidacy for surgical intervention. The dataset consists of bipolar EEG signals from 24 patients that were recorded using 23 channels (FP1-F7, F7-T7, T7-P7, P7-O1, FP1-F3, F3-C3, C3-P3, P3-O1, FP2-F4, F4-C4, C4-P4, P4-O2, FP2-F8, F8-T8, P8-O2, FZ-CZ, CZ-PZ, P7-T7, T7-FT9, FT9-FT10, FT10-T8, and T8-P8). All EEG data were sampled at 256 Hz and 16-bit resolution, using the 10-20 international system and an electrical power line of 120 V and 60 Hz. Fig. 2, presents the raw ictal EEG signals from all 23 channels during an epileptic seizure period of 40 seconds, belonging to the third instance of patient 1.

2.1 Empirical mode decomposition

EMD is an adaptive data analysis method presented by Norden Huang in 1998 as a solution to a problem with the Hilbert transform (25). It essentially aims to empirically identify intrinsic oscillatory modes by their characteristic time scales and then decompose the data accordingly (26). The pseudocode of EMD is shown in Algorithm 1. According to Huang, each IMF should satisfy two conditions (27,28):

- In the whole dataset, the number of extrema and zero-crossings should be the same or differ at most by one.
- At any instance, the mean value of the envelope of the local maxima and minima should be null.

The signal $x(t)$ can be represented and reconstructed using Eq. (1), where n is the number of IMF.

$$x_{(t)} = \sum_{i=1}^{n} IMF_{i(t)} + r_{n(t)} \tag{1}$$

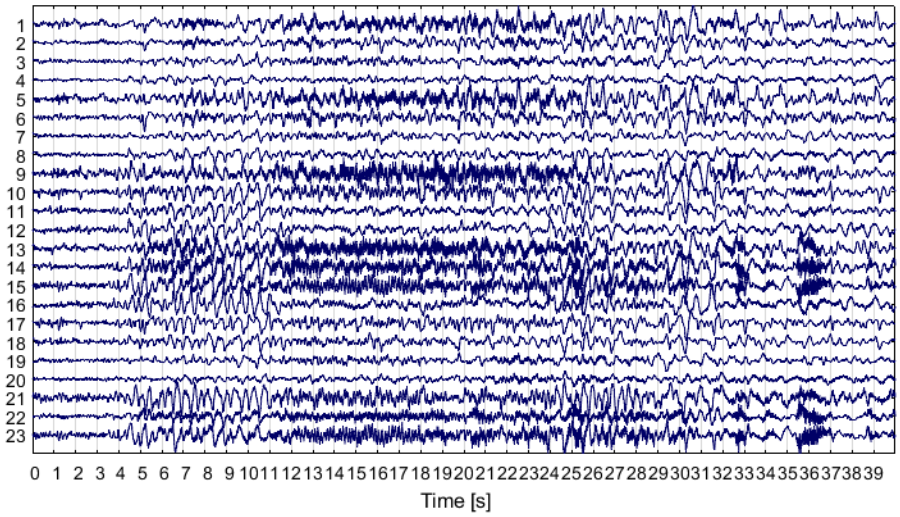


Figure 2. Raw ictal EEG signals from all channels in the third instance of patient 1 of the CHB-MIT dataset

The computational complexity of EMD for a signal of length N is $O(NlogN)$, which makes it a computationally demanding method, especially when dealing with large datasets (29).

2.2 EMD optimized for the selection of relevant IMFs

This approach aims use EMD to select the most relevant IMFs for the detection of epileptic seizure segments from a raw EEG signal. Since IMFs are assumed to be nearly orthogonal components of the original signal, each relevant IMF should be relatively well correlated with the original signal. This presupposes that irrelevant IMFs are poorly correlated with the original signal and can thus be discriminated.

$$\tau_e = \frac{max(metric) - min(metric)}{2} + min(metric) \tag{2}$$

Fig. 3 depicts the proposed methodology for selecting the most relevant IMFs in a raw EEG signal. First, the EMD method decomposes the raw EEG signal into a different number of IMFs. Then, four metrics of discrimination are computed from each IMF, *i.e.*, the Minkowski distance, the MSE, cross-correlation, and the entropy function. To obtain the most relevant IMFs, a threshold τ_e defined by Eq.(2) is applied to each metric. To select the IMFs, their values for the first three metrics should not surpass the threshold τ_e, while the cross-correlation metric should exceed it. Finally, two relevant sets of IMFs are formed (union and intersection) for rebuilding the EEG signal and evaluating the best option.

Algorithm 1 Sifting process for EMD

Input: The raw signal $x_{(t)}$, for $m = 1..., M$.

Output: $IMFs\ array$

Sifting: $True$;

while sifting = True **do**
  $Extract\ extrema\ (maxima\ and\ minima)\ from$ x(t).
  $Interpolate\ to\ obtain\ the\ envelope\ of\ extrema$
  $Compute\ the\ local\ mean$ m(t) = (emin + emax)/2.
  $Compute\ the\ modal\ function$ d(t) = x(t) - m(t).
  if $d(t)$ satisfies the stopping criterion **then**
    $IMFi \leftarrow d(t)$
    $sifting = False$
  else
    $x(t) \leftarrow d(t)$
    $sifting = True$
  end if
  if $x(t)$ is not monotonic **then**
    Continue;
  else
    Break;
  end if
end while=0

return $IMFs\ array$

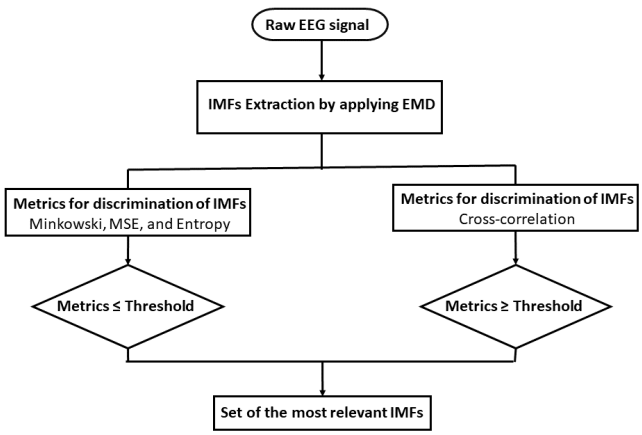


Figure 3. Proposed method for selecting relevant IMFs from raw EEG signals

We selected Minkowski distance, MSE, cross-correlation, and entropy as the primary metrics because of their ability to effectively quantify signal similarity and information content. Minkowski distance generalizes multiple distance measures, which makes it adaptable to different signal characteristics. The MSE evaluates the error between reconstructed and original EEG signals, ensuring minimal loss of relevant information. Cross-correlation measures temporal similarity, which is essential for identifying seizure patterns, while entropy quantifies signal complexity, enabling the discrimination of structured and noisy IMFs. Additionally, we compared these metrics against alternative methods, such as spectral entropy and statistical moment-based measures, concluding that our selection provides the best trade-off between sensitivity and computational efficiency.

2.2.1 Minkowski distance

This metric is a measure of similarity and distance (30–32). Also called *Euclidean distance*, it helps to determine the distance between two vectors X and Y. It can be calculated as follows (33):

$$d_{mink(\tau_e)} = \sqrt{\sum_{i=1}^{n} |x_i - y_i|^2} <= \tau_e, \tag{3}$$

where x_i and y_i are the $i-th$ samples of the original EEG signal and the extracted IMF. According to (34), redundant IMFs differ from the original signal in shape and frequency content, which means that irrelevant ones exhibit a maximum value of d_{mink}. The computational complexity of Minkowski distance is $O(N)$, making it efficient for large-scale EEG analysis.

2.2.2 Cross-correlation

This is a measure of similarity between two signals x and y. The cross-correlation R estimated using Eq. (4):

$$R_{j(\tau_e)} = \sum_{j=1}^{n} x_j^* y_{(i+j)} >= \tau_e, \tag{4}$$

where x_j and $y_{(i+j)}$ are the $j-th$ samples of the original EEG signal and the extracted IMF, respectively. In this case, the computational complexity is $O(NlogN)$.

2.2.3 Mean square error

The MSE measures the average of the squares of the errors, *i.e.*, the difference of the average squares for the estimated and actual values of the signal. This is given by Eq. (5).

$$MSE_{(\tau_e)} = \frac{1}{n} \sum_{i=1}^{N} |x_i - y_i| <= \tau_e, \tag{5}$$

where x_i and y_i are the $i-th$ samples of the original raw EEG signal and the extracted IMF. The computational complexity is $O(N)$.

2.2.4 Entropy function

An entropy-based cost function is applied to each IMF (4,35), as defined by Eq. (6).

$$e_{i(\tau_e)} = -\sum_{k}^{n} ||IMF_i(t_k)||_2^2 log(||IMF_i(t_k)||_2^2) <= \tau_e. \tag{6}$$

To reconstruct the estimated EEG signal $\hat{x}_{(t_k)}$, an automatic selection of the set of IMFs with the lowest entropies $O = e_i$, delimited by the threshold τ_e, is performed. The computational complexity is $O(N)$.

$$\hat{x}_{(t_k)} = \sum_{i \in O} IMF_i(t_k) \tag{7}$$

2.3 Accuracy assessment

The Wasserstein metric (W_m) (also known as either the *Kantorovich-Rubinstein metric* or *earth-movers distance*) (15,36), was used as a quality criterion of signal reconstruction accuracy. This metric provides a comparison between the truth $x_{(t)}$ and the estimated EEG signal $\hat{x}_{(t_k)}$. The index W_m is defined by (8). A lower W_m value represents a better signal reconstruction accuracy.

$$W_m = \sqrt{\frac{1}{n}\sum_{i=1}^{n} |x_{i(t)} - \hat{x}_{i((t_k)}|^2} \tag{8}$$

The correlation coefficient r is a measure of similarity between two signals. It is defined by the ratio of covariance to variance, as shown in (9).

$$r_{x\hat{x}_{(t_k)}} = \frac{S_{x\hat{x}_{(t_k)}}}{S_x S_{\hat{x}_{(t_k)}}} \tag{9}$$

Unlike the wavelet and Fourier transforms, which rely on predefined basis functions (37), the EMD-optimized method adaptively decomposes signals based on their intrinsic oscillations. This adaptive nature makes it particularly effective for analyzing nonstationary EEG signals. While deep learning models have yielded promising results in seizure detection, their lack of interpretability remains a significant challenge. In contrast, our approach enhances explainability by explicitly selecting IMFs based on signal characteristics, thus ensuring a more transparent and interpretable feature extraction process. Given that these selection metrics are applied to K IMFs, the computational complexity of the selection process is $O(KNlogN)$.

3 Results

This section analyzes and discusses the outcomes derived from the implementation of our methodology, as well as the validation tests conducted in accordance with the procedures outlined above. The implementation was carried out using the MATLAB R2020b software, and the validation was conducted using EEGLAB. This study focused on channel CZPZ of patient 1, for which two signals were analyzed: a normal EEG signal from the first instance and an ictal EEG signal from the third instance.

Fig. 4 presents a comparison between raw normal and ictal EEG signals, revealing a significant increase in the power of the ictal EEG signal within the low-frequency range, specifically between 0 and 6 Hz, in contrast to the normal signal. This increase in power can be attributed to the presence of synchronized neuronal activity during ictal episodes. In the context of epileptic seizures, abnormal neuronal activity often manifests as an excessive synchronization of neurons in specific brain regions. This synchronization leads to the generation of low-frequency electrical signals, contributing to the power increase observed in the frequency spectrum of the ictal EEG signal when compared to the normal one. These findings characterize the distinctive hyperexcitatory neuronal activity of epileptic seizures, providing valuable insights for the detection and understanding of these neurological events.

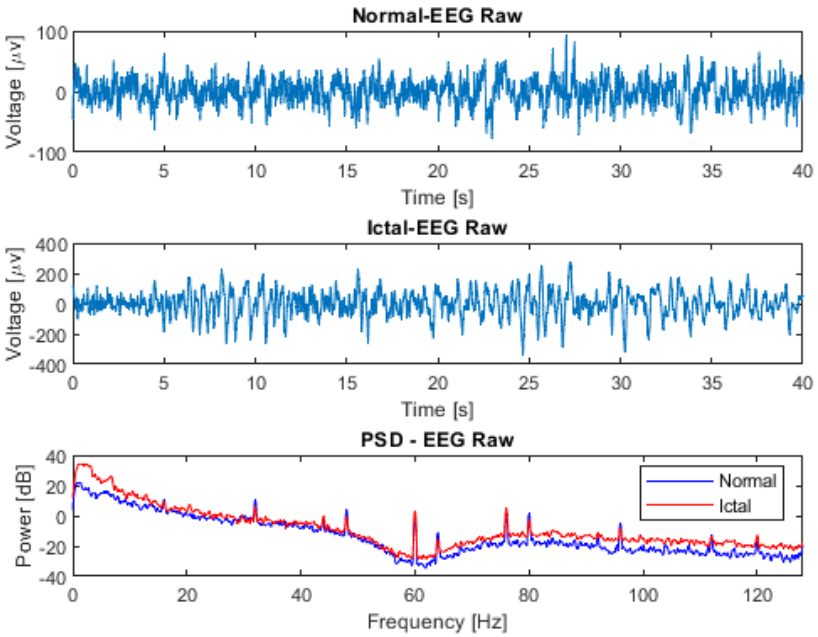


Figure 4. Comparison of raw EEG data from normal and ictal EEG signals and their power spectral density (PSD), as obtained from channel CZPZ of the first and third instances of patient 1 of the CHB-MIT dataset

3.1 Analysis of EMD

The initial phase of the analysis involved the application of EMD to the raw EEG signals, hereafter referred to as the *EEG-normal* and *EEG-ictal signals*. Each raw EEG signal underwent decomposition into a series of IMFs, arranged according to its frequency components from highest to lowest (designated as IMF1 to IMF8 in this instance). Upon conducting the database analysis via EMD, eight IMFs were derived per signal, as illustrated in Figs. 5 and 6. This systematic decomposition process facilitated the extraction of signal components across varying frequency ranges, enabling a comprehensive examination of the EEG dataset.

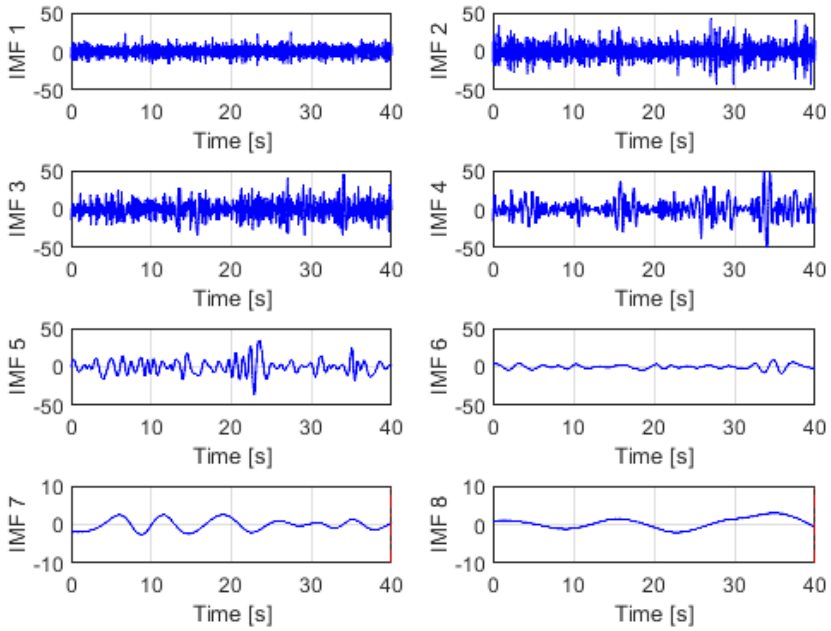


Figure 5. EMD of an EEG-normal signal using eight IMFs of the CZPZ channel corresponding to the first instance of patient 1 in the CHB-MIT dataset

The second step of the analysis was reconstruction, which utilized the initial six IMFs extracted from each signal. This reconstruction process aimed to synthesize the signals based on the dominant frequency components captured by the selected IMFs. The reconstructed signals were then subjected to PSD analysis in order to discern the signal power distribution across different frequency bands. This analytical approach facilitated the examination of the frequency characteristics inherent in the EEG signals, providing insights into the underlying neuronal activity patterns associated with normal and ictal states. Fig. 7 illustrates the PSD diagrams obtained from the reconstructed signals, which describe the frequency composition of the EEG data under study.

The reconstructed EEG-normal and EEG-ictal signals, as observed in Fig. 7, exhibit a remarkably similar behavior to that shown in Fig. 4, despite the processing conducted through EMD. This finding suggests the presence of high-frequency noise that was not eliminated during the reconstruction process, as no pre-processing stage was applied. Consequently, there is an evident need to employ a reliable alternative to supplement this pre-processing stage in EEG signals. Such an approach would enhance reconstruction quality and facilitate a more accurate interpretation of the obtained data.

3.2 Analysis of the optimal selection of IMFs (optimized EMD)

The primary objective of the proposed methodology, termed *optimized EMD*, was to optimize the selection of IMFs containing relevant information, aligning the raw EEG signals within frequency

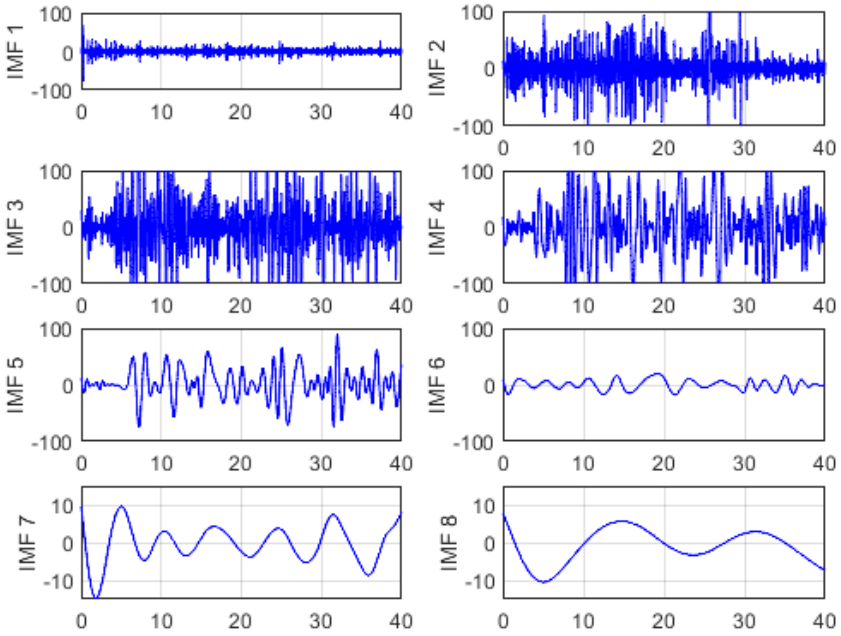


Figure 6. EMD of an EEG-ictal signal using eight IMFs of the CZPZ channel corresponding to the third instance of patient 1 in the CHB-MIT dataset

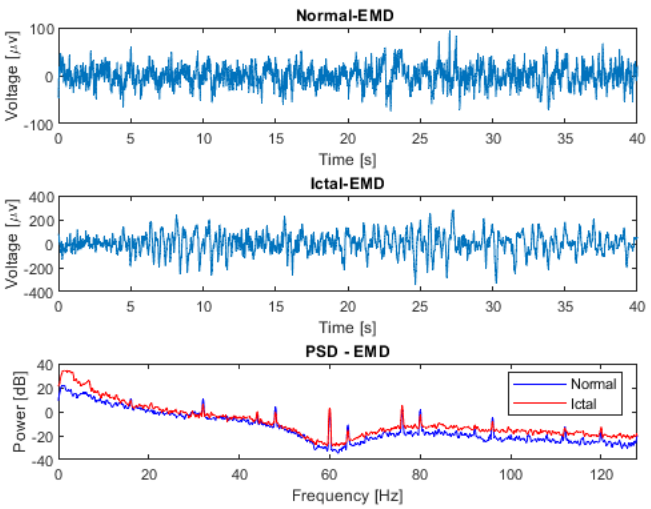


Figure 7. Comparative analysis of EMD applied to the first six IMFs for the reconstruction of normal and ictal EEG signals corresponding to the CZPZ channel of the first and third instances of patient 1

bands for the study of epileptic seizure detection. To this effect, four discrimination metrics based on a predefined threshold were employed. The outcomes of these discrimination techniques, encompassing Minkowski distance, cross-correlation, MSE, and entropy, alongside the evaluation metrics used for accuracy assessment (the correlation coefficient, p-value, and Wm), are systematically presented in Tables I and II. In addition, they are depicted in Figs. 8 and 9, where the threshold τ_e is denoted by a dashed red horizontal line. The results evince a superior performance for the EEG-ictal signal. This meticulous analysis facilitates a comprehensive evaluation of the efficacy and reliability of the proposed method in discerning relevant IMFs for epileptic seizure detection.

Table I. Metrics for the selection of relevant IMFs in the normal EEG signal from channel CZPZ of patient 1's first instance

Metric	IMFs retained after invoking threshold	Correlation coefficient with original signal	p-Value	Wm
Minkowski	2,3,4,5	0.9680	0	5.1979
Cross-corr	2,3,4,5	-0.0153	0.1207	86.9605
MSE	2,3,4,5	-0.0153	0.1207	86.9605
Entropy	3,4	-0.0068	0.4928	77.7654

Table II. Metrics for the selection of relevant IMFs in the ictal EEG signal from channel CZPZ from patient 1's third instance

Metric	IMFs retained after invoking threshold	Correlation coefficient with original signal	p-value	Wm
Minkowski	3,4	0.8852	0	39.1013
Cross-corr	3,4	0.8852	0	39.1013
MSE	3,4	0.8852	0	39.1013
Entropy	3,4	0.8852	0	39.1013

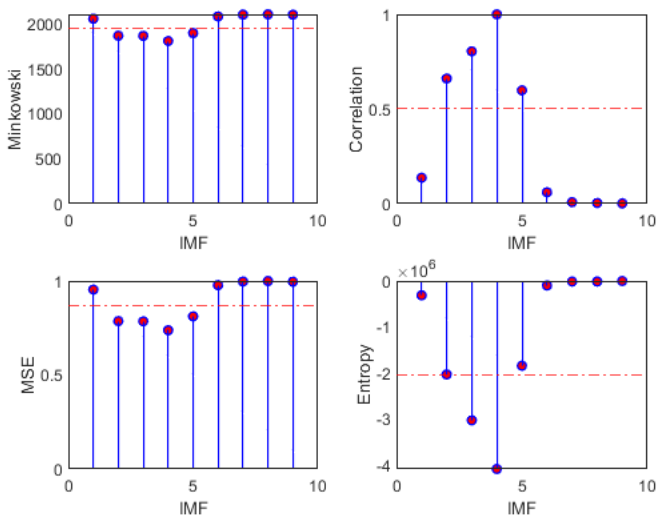


Figure 8. Metrics for IMF selection from channel CZPZ's normal EEG signal

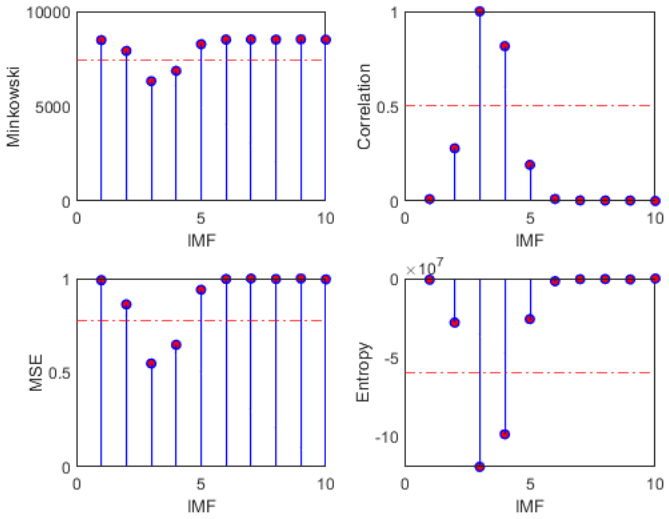


Figure 9. Metrics for IMF selection from channel CZPZ's ictal EEG signal

Figs. **10** and **11** provide a visual comparison between the original raw EEG-normal and EEG-ictal signals, alongside the reconstructed (rEEG) signals corresponding to channel CZPZ. The reconstruction outcomes are succinctly summarized in Table **III**. Remarkably, the reconstructed signals exhibit a notable absence of noise. These findings underscore the effectiveness of the proposed methodology in attenuating noise and enhancing signal fidelity, thereby contributing to the accurate representation of EEG activity in both normal and ictal states.

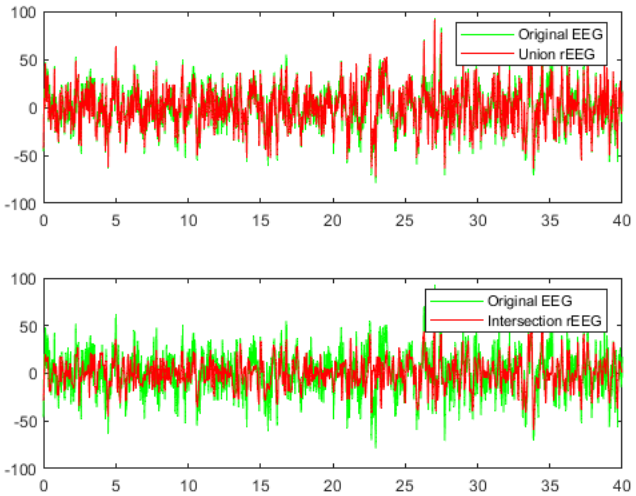


Figure 10. Comparison between the original raw and reconstructed EEG-normal signals using union and intersection on channel CZPZ

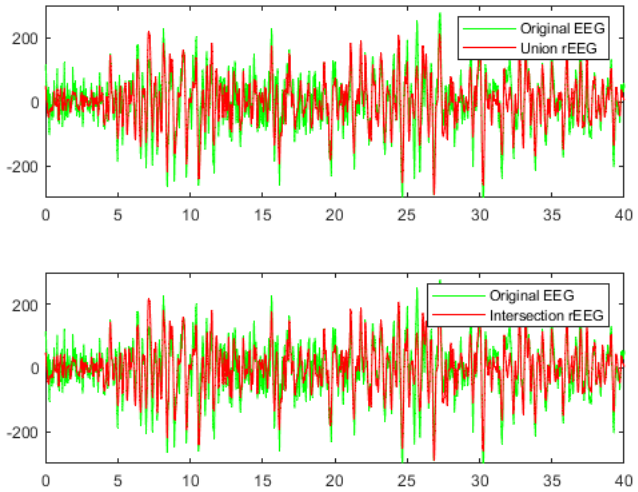


Figure 11. Comparison between the original raw and reconstructed EEG-ictal signals using union and intersection on channel CZPZ

Table III. Accuracy assessment of optimized EMD on channel CZPZ

Operation	Wm		Correlation coefficient		Optimal IMFs	
	Normal EEG	Ictal EEG	Normal EEG	Ictal EEG	Normal EEG	Ictal EEG
Union	5.1979	39.1013	0.9680	0.8852	2,3,4,5	3,4
Intersection	13.7360	39.1013	0.7484	0.8852	3,4	3,4

An assessment of signal reconstruction employing optimized EMD is presented in Table III and illustrated in Fig. 12. The findings reveal a superior signal reconstruction for the EEG-ictal signal utilizing only two IMFs, resulting in enhanced performance index values, including a correlation coefficient of 0.88 and a Wm of 39.10 for both operation modes (union and intersection). The PSD diagram depicted in Fig. 12 underscores two key facts. Firstly, the proposed methodology effectively confines the signal within the operational bandwidth, serving as a low-pass filter to mitigate noise originating from high frequencies, such as the 120 V and 60 Hz power line interference. Secondly, it elucidates the augmented power distribution in low-frequency bands within the EEG-ictal signal compared to the EEG-normal one. This comprehensive evaluation highlights the efficacy of the proposed method in enhancing signal fidelity and discerning features relevant to epileptic seizure activity.

To assess the efficacy of optimized EMD, performance indices such as the correlation coefficient and the Wm were employed in a comprehensive evaluation. This evaluation was conducted on ictal EEG recordings obtained from all channels corresponding to the third instance of patient 1. The results of this evaluation, conducted with a statistical significance threshold of $\rho < 0.05$, are presented in Table IV. Such rigorous testing provides valuable insights into the performance and reliability of optimized EMD in the context of epileptic seizure detection.

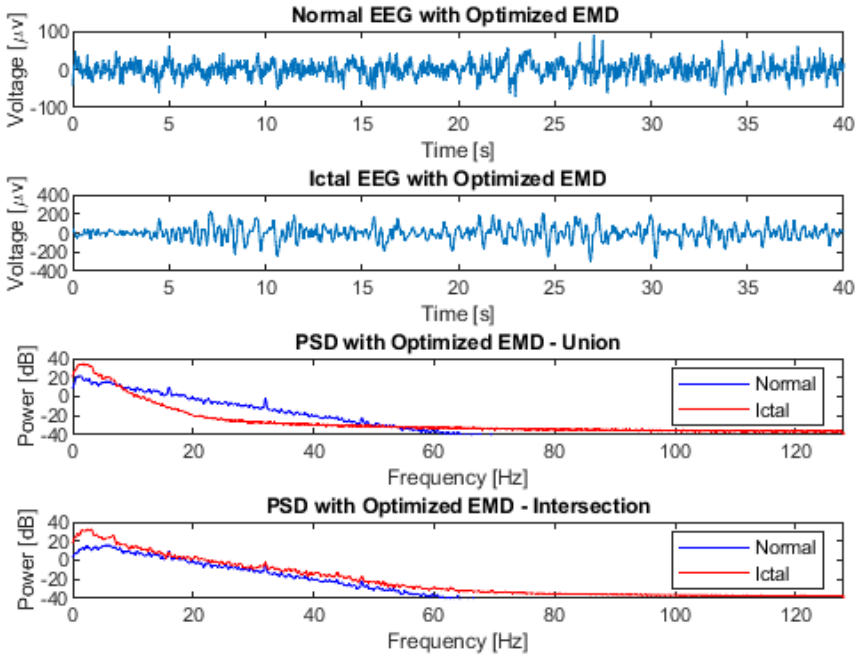


Figure 12. Comparative analysis of the use of optimized EMD on EEG-normal and EEG-ictal signals from channel CZPZ of the third instance of patient 1 in the CHB-MIT dataset

The use of optimized EMD to pre-process the raw EEG signal constitutes a novel approach to epileptic seizure detection. When applied to raw EEG signals without a pre-processing stage, a comparison of power spectra between raw and EMD-processed EEG instances does not reveal significant disparities, as illustrated in Figs. 4 and 7. In both scenarios, analogous high-frequency power peaks are discerned, encompassing the 60 Hz frequency associated with the power line. When applying the proposed method, a significant improvement was observed in identifying the range of frequencies of interest (Fig. 12). In this context, the method acts as a low-pass adaptive filter, eliminating the need for a reference signal, which is common in conventional adaptive filters. This technique enables the effective suppression of high frequencies, including the frequency of the electrical line, which does not contribute to the acquisition of relevant information on the studied pathology.

By decomposing signals into IMFs, the proposed method facilitates the extraction of essential information for precise seizure identification. The process of selecting the most relevant IMFs, guided by thresholds and discrimination metrics such as the Minkowski distance, cross-correlation, MSE, and entropy function, is of paramount importance. This meticulous process plays a pivotal role in enhancing the accuracy and reliability of seizure detection. This ensures that only IMFs that contain significant information relevant to epileptic activity are retained, thus reducing computational complexity and

Table IV. Accuracy assessment of optimized EMD for all channels from patient 1's third instance

Channel	Wm		Correlation coefficient		Optimal IMFs	
	Union	Intersection	Union	Intersection	Union	Intersection
FP1-F7	33.93	33.93	0.83	0.83	3,4	3,4
F7-T7	23.69	23.69	0.87	0.87	3,4,5	3,4
T7-P7	57.05	57.05	0.85	0.85	3,4	3,4
P7-O1	88.94	104.08	0.75	0.63	4,5	4,5
FP1-F3	41.22	41.22	0.91	0.91	3,4	3,4
F3-C3	69.53	69.53	0.71	0.71	3,4	3,4
C3-P3	33.93	33.93	0.83	0.83	3,4	3,4
P3-O1	23.69	23.69	0.87	0.87	3,4,5	3,4
FP2-F4	57.05	57.05	0.85	0.85	1,3,4	3,4
F4-C4	88.94	104.08	0.75	0.63	3,4	3,4
C4-P4	41.22	41.22	0.91	0.91	3,4	3,4
P4-O2	69.53	69.53	0.71	0.71	3,4	3,4
FP2-F8	33.93	33.93	0.83	0.83	4	4
F8-T8	23.69	23.69	0.87	0.87	4	4
T8-P8	57.05	57.05	0.85	0.85	3,4	3,4
P8-O2	88.94	104.08	0.75	0.63	3,4	3,4
FZ-CZ	41.22	41.22	0.91	0.91	3	3
CZPZ	39.10	39.10	0.88	0.88	3,4	3,4
P7-T7	69.53	69.53	0.71	0.71	3,4	3,4
T7-FT9	33.93	33.93	0.83	0.83	4,5	4,5
FT9-FT10	23.69	23.69	0.87	0.87	3,4	3,4
FT10-T8	57.05	57.05	0.85	0.85	1,2,3,4,5	4,5
T8-P8	88.94	104.08	0.75	0.63	3,4	3,4

improving detection performance. The results presented in Table IV and depicted in Fig. 12 evidence notable performance indices and demonstrate the capability of discriminating irrelevant information in both union and intersection modes.

In the union mode, even when employing a single channel such as Fp2-F8, F8-T8, and FZ-CZ, correlation coefficients ranging between 0.83 and 0.91 and Wm values of 23.69-41.22 were observed. These values either surpassed or were at least equivalent to those obtained from channels utilizing more than three IMFs, such as F7-T7, P3-O1, FP2-F4, and FT10-T8. In intersection mode, the majority of channels required only two IMFs for effective operation, particularly emphasizing IMF3 and IMF4, with the exception of channels Fp2-F8, F8-T8, and FZ-CZ, which demonstrated efficacy with a single IMF, particularly IMF4. Furthermore, the outcomes derived from brain mapping facilitated the validation of our proposal by enabling the identification of the characteristic frequency range of epileptic seizures. Moreover, the localization of active electrodes represents a noteworthy advancement. Through the identification of IMFs harboring the most relevant information, this method enhances the identification

of frequency bands in seizure detection, offering valuable insights into the underlying neural dynamics of seizure events.

3.3 Analysis of brain activity mapping

To assess the effectiveness of the proposed method, a graphical evaluation was conducted using the EEGLAB software, focusing on brain mapping for ictal EEG signal recordings from all channels belonging to the third instance of patient 1. Two scenarios were examined: firstly, ictal EEG signals without EMD processing, and, secondly, ictal EEG signals reconstructed via optimized EMD. Intersection mode data from Table **IV** were utilized for signal reconstruction. For each channel, a digital band-pass filter (0.1-55 Hz) was implemented, and the segments that were visibly affected by technical artifacts were removed, such as eye movements, muscle movements, and sweating, *etc.* Subsequently, for each EEG, the channel exhibiting the highest power was identified through spectral response analysis.

The results depicted in Fig. **13** indicate the occurrence of maximum power magnitudes at high frequencies (1, 16, 32, and 48 Hz), primarily attributable to additive noise rather than genuine neuronal activity. Consequently, the identification of electrodes with higher brain activity was hindered. Conversely, Fig. **14** demonstrates that the application of optimized EMD acts as a low-pass filter, effectively attenuating the high frequencies associated with noise interference and other artifacts. This facilitated the identification of heightened brain activity in electrodes F7, F10, and Pz.

From these observations, it can be inferred that the application of optimized EMD to EEG signals satisfactorily serves as a pre-processing stage, substantially enhancing signal reconstruction by confining it within the requisite bandwidth for epileptic seizure detection.

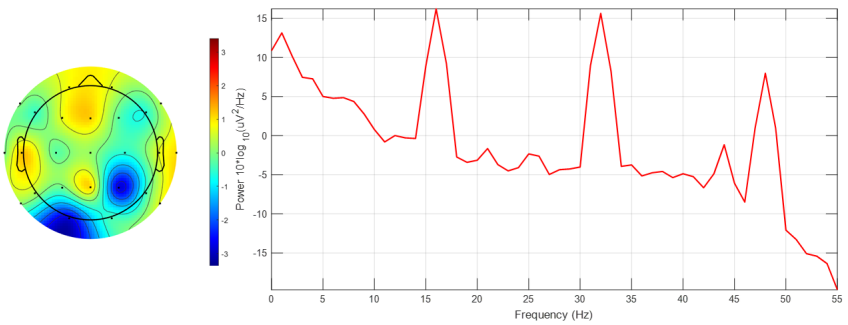


Figure 13. Brain mapping and activity power spectrum of the ictal EEG signal obtained from the third instance of patient 1 without EMD

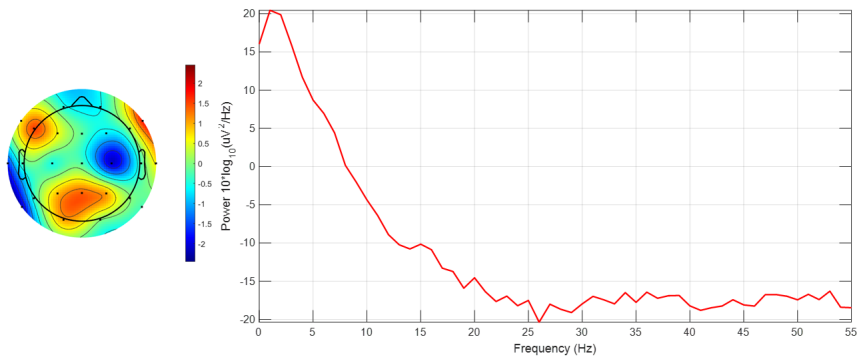


Figure 14. Brain mapping and activity power spectrum of the ictal EEG signal obtained from the third instance of patient 1 after applying optimized EMD

In forthcoming endeavors, our attention will be directed towards validating the proposed methodology through feature extraction techniques seeking to predict and classify epileptic seizures using larger and more diverse datasets. In addition, we plan to conduct clinical trials in order to assess the performance of our method in real-world scenarios. Furthermore, we will explore the potential integration of machine learning algorithms to enhance the accuracy of seizure detection. Such endeavors hold promise for the advancement of epilepsy diagnosis and treatment.

3.4 Robustness analysis: optimal IMF selection in diverse EEGs of epilepsy patients

This study evaluated the robustness of an optimal IMF selection method for epilepsy detection, analyzing EEG data from four patients in the CHB-MIT database, including cases of focal and generalized seizures. Three of these patients had focal seizures (patients 1, 2, and 3), while the fourth experienced generalized seizures. The analysis was performed using the 23 EEG channels available in each patient to compare the optimal selection of IMFs.

The results reveal a marked inter- and intra-patient variability in the channel-specific optimal IMFs, underscoring the necessity of personalized approaches in EEG analysis (Table **V**). Specifically, the first three patients, with focal seizures, exhibit a predominance of IMFs 3 and 4 across most channels. In contrast, the fourth patient, experiencing generalized seizures, displays a more diverse selection of optimal IMFs, reflecting the increased complexity of the EEG signals associated with this seizure type. The identification of patient- and channel-specific optimal IMFs, predominantly within the range of IMFs 1 to 6, demonstrates their relevance for epilepsy detection and suggests that personalized IMF selection can significantly improve the accuracy of detection algorithms, as well as the differentiation between focal and generalized seizures.

Table V. Channel-specific optimal IMFs for epilepsy patients

Channel	Optimal IMFs			
	Patient 1	**Patient 2**	**Patient 3**	**Patient 4**
FP1-F7	3,4	3,4	4,5	4
F7-T7	3,4	4	3,4	1,2,4,5
T7-P7	3,4	2,3,4	1,2,3,4	1,2,4,5
P7-O1	4,5	3,4	3,4	1,4
FP1-F3	3,4	3,4	3,4	3,4
F3-C3	3,4	3,4	3,4	4
C3-P3	3,4	3	3,4	1,2,3,4
P3-O1	3,4	1,2,3,4	4,5	1,2
FP2-F4	3,4	3,4	3,4	4,5
F4-C4	3,4	3,4	4,5	1,2,4
C4-P4	3,4	2,3	3,4	1,2,4
P4-O2	3,4	2,3	3,4	1,2,3,4,5
FP2-F8	4	3,4	3,4	1,2
F8-T8	4	3,4	4,5	1,2,4,5
T8-P8	3,4	3	1,2	1,2
P8-O2	3,4	2,3	1,4	1,2,3,4
FZ-CZ	3	3,4	4,5	4
CZPZ	3,4	2,3	3	3,4
P7-T7	3,4	2,3,4	1,2,3,4	1,2,4,5
T7-FT9	4,5	4	4,5	1,2,6
FT9-FT10	3,4	3,4	4,5	3,4
FT10-T8	4,5	3,4,5	1,5	1,2,4
T8-P8	3,4	3	1,2	1,2

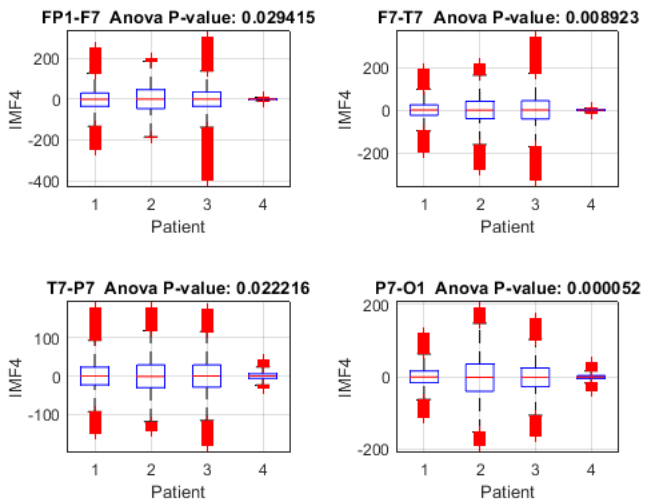


Figure 15. Variance analysis of IMF 4 across channels Fp1-F7, F7-T7, T7-P7, and P7-O1 between focal seizure patients and a seizure-free subject

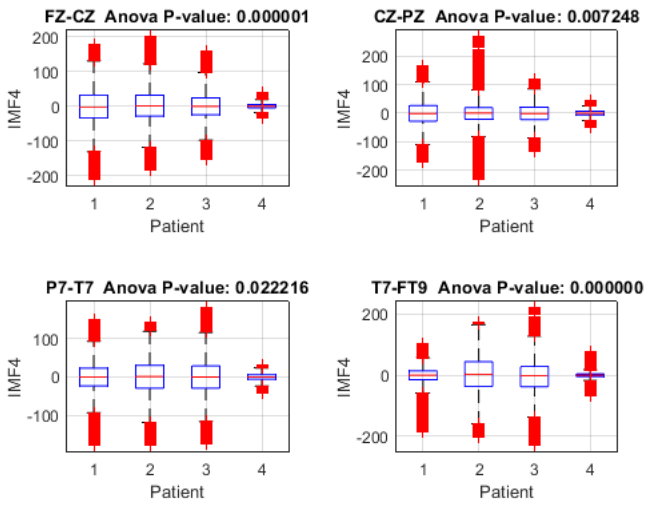


Figure 16. Variance analysis of IMF 4 across channels Fz-Cz, Cz-Pz, P7-T7, and T7-FT9 between focal seizure patients and a seizure-free subject

To assess the proposed method's capacity for discriminating between focal seizure (patients 1, 2, and 3) and seizure-free (patient 4) signals, an analysis of variance (ANOVA) was performed on IMF 4 across the Fp1-F7, F7-T7, T7-P7, and P7-O1 channels (Fig. **15**), revealing significant differences between the three patients with focal seizures and the seizure-free subject ($p < 0.05$ in all channels), with the highest significance observed in P7-O1 ($p = 0.000052$). These findings suggest that IMF 4 may capture relevant information for identifying focal epilepsy, particularly in the occipital region. Moreover, the ANOVA of IMF 4 in the FZ-CZ, CZ-PZ, P7-T7, and T7-FT9 channels (Fig. **16**) revealed significant differences between patients with focal seizures and the seizure-free subject ($p < 0.05$ in all channels), with the most pronounced differences in FZ-CZ ($p = 0.000001$) and T7-FT9 ($p < 0.000001$). These findings highlight the relevance of IMF 4 in distinguishing focal epilepsy, particularly in the central and temporal regions, suggesting the need to complement the analysis with power estimations in order to quantify the magnitude of these differences.

To further investigate this differentiation, the analysis was complemented by estimating the power of IMF 4 in each channel and patient. The power distribution in channels FP1-F7, F7-T7, T7-P7, and P7-O1 revealed notable differences between patients with focal seizures and the seizure-free subject (Fig. **17**). Patients 1, 2, and 3 exhibited higher power across all evaluated channels, whereas patient 4 showed considerably lower values. Notably, the frontotemporal channels (FP1-F7 and F7-T7) reported the highest values, suggesting that IMF 4 captures relevant information in these regions, likely related to the propagation of epileptic activity. Similarly, the power distribution of IMF 4 in channels FZ-CZ, CZ-PZ, P7-T7, and T7-FT9 exhibits comparable trends, further supporting the relevance of this component in distinguishing focal epilepsy. Notably, patients 1, 2, and 3 showed higher power values compared to patient 4, reinforcing the association between IMF 4 and epileptiform activity.

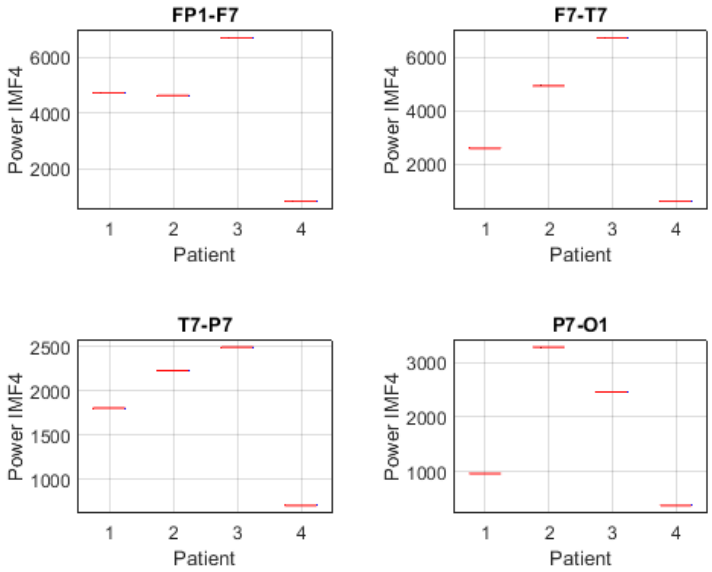


Figure 17. Power distribution of IMF 4 across channels Fp1-F7, F7-T7, T7-P7, and P7-O1 between focal seizure patients and a seizure-free subject

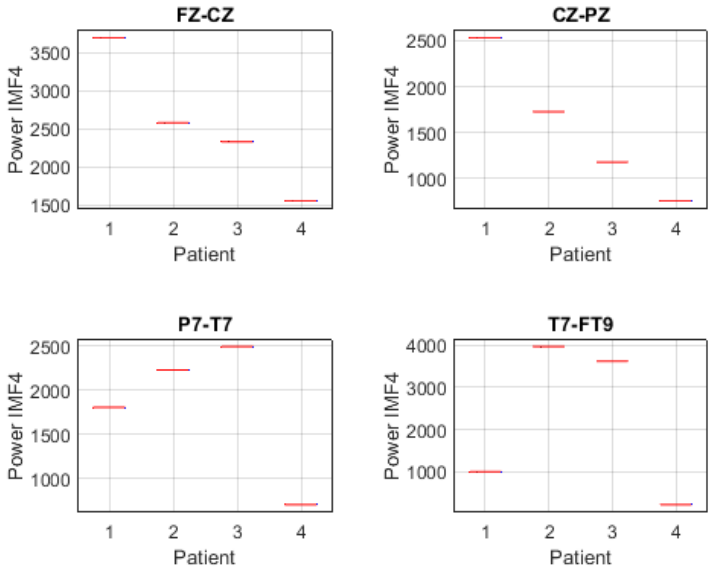


Figure 18. Power distribution of IMF 4 across channels Fz-Cz, Cz-Pz, P7-T7, and T7-FT9 between focal seizure patients and a seizure-free subject

Fig. 19 illustrates the distribution of IMF 4 in patients with focal (1-3) and generalized (4) seizures, revealing significant inter-subject differences according to the ANOVA ($p < 0.05$). Channels T7-P7 ($p = 0.000000$) and P7-O1 ($p = 0.000267$) exhibited the highest variability, suggesting that IMF 4 captures distinctive brain activity patterns. The presence of outliers indicates possible transient events or epileptiform discharges, particularly in T7-P7. These findings highlight the potential of IMF 4 for distinguishing between seizure types, although further validation with a larger patient cohort is recommended to ensure the robustness of the method.

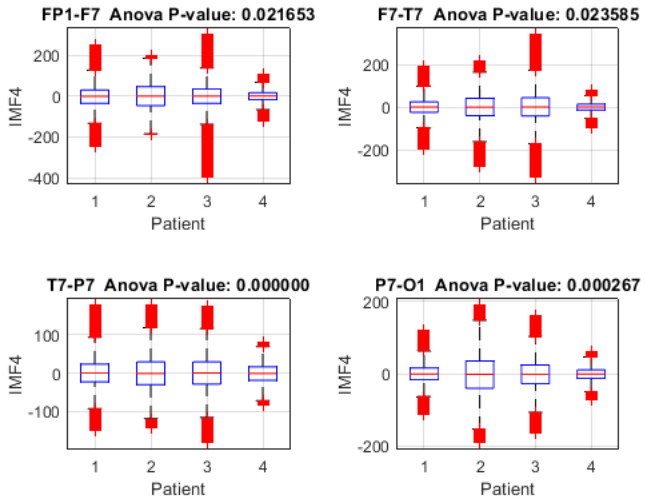


Figure 19. Variance analysis of IMF 4 across channels Fp1-F7, F7-T7, T7-P7, and P7-O1 between focal seizure patients and a generalized seizure patient

4 Conclusions

This study proposed an optimized method for the selection of intrinsic mode functions using empirical mode decomposition for epileptic seizure detection in raw electroencephalogram signals. The results demonstrated that IMF 4 plays a crucial role in distinguishing focal epilepsy, as indicated by the power distribution across different EEG channels. Patients with focal seizures exhibited significantly higher power values in specific frontotemporal and central-parietal channels in comparison with seizure-free subjects, reinforcing the relevance of IMF 4 in the characterization of epileptiform activity.

The proposed method provides an alternative to traditional EEG pre-processing techniques, potentially reducing computational cost while preserving critical signal information for classification. Moreover, the selection process, based on discrimination metrics such as Minkowski distance, cross-correlation, and entropy, showcased an effective way to identify relevant IMFs with a computational complexity of $O(KNlogN)$, making our proposal suitable for large-scale EEG datasets. However, optimized EMD necessitates multiple iterations (K), which increases the required processing

times. For real-time execution, optimization with digital signal processors (DSPs) or graphics processing units (GPUs) is essential for parallel computations. Future research should explore the feasibility of this method in real-time applications, particularly for intracranial EEG (iEEG) signals, which exhibit different noise characteristics and resolution. Additionally, expanding the evaluation to larger and more diverse datasets will further validate the robustness and generalization of this approach.

Ultimately, this methodology contributes to the advancement of adaptive signal processing techniques for epileptic seizure detection, offering a more interpretable and computationally efficient alternative to deep learning-based approaches.

5 Author contributions

Luis Daladier Guerrero and Maximiliano Bueno López conceived the idea and conducted the background research. Maximiliano Bueno and Eduardo Giraldo supported the experiments. The collaborative efforts of Maximiliano Bueno López, Eduardo Giraldo, and Marta Molinas involved problem identification and data curation. Maximiliano Bueno López reviewed and supported the editing of the paper. Maximiliano BUeno López, Eduardo Giraldo, and Marta Molinas supervised the research and provided critical feedback. It is important to note that all authors have rigorously reviewed the content and have provided their collective consent and approval for the final manuscript.

References

[1] S. M. Zuberi *et al.*, "Ilae classification and definition of epilepsy syndromes with onset in neonates and infants: Position statement by the ilae task force on nosology and definitions," *Epilepsia*, vol. 63, no. 6, pp. 1349–1397, 2022. https://doi.org/10.1111/epi.17239 ↑ 3

[2] E. C. Wirrell *et al.*, "Methodology for classification and definition of epilepsy syndromes with list of syndromes: Report of the ilae task force on nosology and definitions," *Epilepsia*, vol. 63, no. 6, pp. 1333–1348, 2022. https://doi.org/10.1111/epi.17237 ↑ 3

[3] S. Wong *et al.*, "Eeg datasets for seizure detection and prediction—a review," *Epilepsia Open*, vol. 8, no. 2, pp. 252–267, 2023. https://doi.org/10.1002/epi4.12704 ↑ 3

[4] B.-L. Maximiliano, M.-G. P. Andrés, G. Eduardo, and M. M. M. Cabrera, "Electroencephalographic source localization based on enhanced empirical mode decomposition," *IAENG Int. J. Comp. SCi.*, vol. 46, p. 11, 2019. https://www.iaeng.org/IJCS/issues_v46/issue_2/IJCS_46_2_11.pdf ↑ 3, 8

[5] D. Wu, J. Wei, P.-P. Vidal, D. Wang, Y. Yuan, J. Cao, and T. Jiang, "A novel seizure detection method based on the feature fusion of multimodal physiological signals," *IEEE Internet Things J.*, vol. 11, no. 16, pp. 27 545–27 556, 2024. https://doi.org/10.1109/JIOT.2024.3398418 ↑ 3

[6] Z. Wang *et al.*, "Research progress of epileptic seizure prediction methods based on eeg," *Cogn. Neurodyn.*, vol. 18, no. 5, pp. 2731–2750, 2024. https://doi.org/10.1007/s11571-024-10109-w ↑ 3

[7] V. R., M. S., S. C., and K. S., "Quadrature response spectra deep neural based behavioral pattern analytics for epileptic seizure identification," *Meas. Sci. Rev.*, vol. 24, no. 2, pp. 67–71, Apr. 2024. https://journals.savba.sk/index.php/msr/article/view/2065 ↑ 3

[8] B. Zhang *et al.*, "Cross-subject seizure detection in eegs using deep transfer learning," *Comput. Math. Methods Med.*, vol. 2020, p. 7902072, 2020. https://doi.org/10.1155/2020/7902072 ↑ 3

[9] Y. Du, J. Jin, Y. Liu, and Q. Wang, "Classification of seizure eegs based on short-time fourier transform and hidden markov model," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Ann. Summit Conf. (APSIPA ASC)*, 2020, pp. 875–880. https://ieeexplore.ieee.org/document/9306434 ↑ 3

[10] M. Amiri, H. Aghaeinia, and H. R. Amindavar, "Automatic epileptic seizure detection in eeg signals using sparse common spatial pattern and adaptive short-time fourier transform-based synchrosqueezing transform," *Biomed. Signal Process. Control*, vol. 79, p. 104022, 2023. https://doi.org/10.1016/j.bspc.2022.104022 ↑ 3

[11] S. A. El-Gindy *et al.*, "Efficient communication and eeg signal classification in wavelet domain for epilepsy patients," *J. Ambient Intell. Humaniz. Comput.*, vol. 12, no. 10, pp. 9193–9208, 2021. https://doi.org/10.1007/s12652-020-02624-5 ↑ 3

[12] M. Shen, P. Wen, B. Song, and Y. Li, "Real-time epilepsy seizure detection based on eeg using tunable-q wavelet transform and convolutional neural network," *Biomed. Signal Process. Control*, vol. 82, p. 104566, 2023. https://doi.org/10.1016/j.bspc.2022.104566 ↑ 3

[13] D. Sunaryono, J. Siswantoro, R. Sarno, R. I. Susilo, and S. I. Sabilla, "Epilepsy detection using combination dwt and convolutional neural networks based on electroencephalogram," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, 2023, pp. 411–416. https://doi.org/10.1109/ISITIA59021.2023.10221031 ↑ 3

[14] F. A. Jibon *et al.*, "Epileptic seizure detection from electroencephalogram (eeg) signals using linear graph convolutional network and densenet based hybrid framework," *J. Radiat. Res. Appl. Sci.*, vol. 16, no. 3, p. 100607, 2023. https://doi.org/10.1016/j.jrras.2023.100607 ↑ 3

[15] A. Soler, P. A. Muñoz-Gutiérrez, M. Bueno-López, E. Giraldo, and M. Molinas, "Low-density eeg for neural activity reconstruction using multivariate empirical mode decomposition," *Front. Neurosci.*, vol. 14, p. 175, 2020. https://www.frontiersin.org/articles/10.3389/fnins.2020.00175/full ↑ 3, 8

[16] V. R. Carvalho, M. F. D. Moraes, A. P. Braga, and E. M. A. M. Mendes, "Evaluating five different adaptive decomposition methods for eeg signal seizure detection and classification," *Biomed. Signal Process. Control*, vol. 62, p. 102073, 2020. https://doi.org/10.1016/j.bspc.2020.102073 ↑ 3

[17] L.-D. Guerrero, L. D. Romero, and M. Bueno-López, "A review of epileptic seizure detection using eeg signals analysis in the time and frequency domain," in *Proc. IEEE 21st Int. Conf. Commun. Technol. (ICCT)*, 2021, pp. 1363–1367. https://doi.org/10.1109/ICCT52962.2021.9657835 ↑ 3

[18] M.-G. Murariu, F.-R. Dorobanțu, and D. Tărniceriu, "A novel automated empirical mode decomposition (emd) based method and spectral feature extraction for epilepsy eeg signals classification," *Electronics*, vol. 12, no. 9, p. 1958, 2023. https://www.mdpi.com/2079-9292/12/9/1958 ↑ 3

[19] L. A. Moctezuma and M. Molinas, "Classification of low-density eeg for epileptic seizures by energy and fractal features based on emd," *J. Biomed. Res.*, vol. 34, no. 3, pp. 180–190, 2019. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7324275/ ↑ 3

[20] L. A. Moctezuma and M. Molinas, "Eeg channel-selection method for epileptic-seizure classification based on multi-objective optimization," *Front. Neurosci.*, vol. 14, p. 593, 2020. https://www.frontiersin.org/articles/10.3389/fnins.2020.00593/full ↑ 3

[21] M. Thilagaraj, M. P. Rajasekaran, and N. A. Kumar, "Tsallis entropy: As a new single feature with the least computation time for classification of epileptic seizures," *Cluster Comput.*, vol. 22, no. 6, pp. 15 213–15 221, 2019. https://doi.org/10.1007/s10586-018-2549-5 ↑ 3

[22] V. S. Jebakumari, D. S. Saravanan, and D. Devaraj, "Seizure detection in eeg signal with novel optimization algorithm for selecting optimal thresholded offset gaussian feature," *Biomed. Signal Process. Control*, vol. 56, p. 101708, 2020. https://doi.org/10.1016/j.bspc.2019.101708 ↑ 3

[23] A. L. Goldberger *et al.*, "Physiobank, physiotoolkit, and physionet: Components of a new research resource for complex physiologic signals," *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000. https://www.ahajournals.org/doi/abs/10.1161/01.CIR.101.23.e215 ↑ 4

[24] A. H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Massachusetts Institute of Technology, Cambridge, MA, 2009, ph.D. dissertation. https://dspace.mit.edu/handle/1721.1/54669 ↑ 4

[25] M. Chen, D. P. Mandic, P. Kidmose, and M. Ungstrup, "Qualitative assessment of intrinsic mode functions of empirical mode decomposition," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, 2008, pp. 1905–1908. https://doi.org/10.1109/ICASSP.2008.4518007 ↑ 4

[26] E. Mäteling and W. Schröder, "Analysis of spatiotemporal inner–outer large-scale interactions in turbulent channel flow by multivariate empirical mode decomposition," *Phys. Rev. Fluids*, vol. 7, no. 3, p. 034603, Mar 2022. https://doi.org/10.1103/PhysRevFluids.7.034603 ↑ 4

[27] N. E. Huang *et al.*, "The empirical mode decomposition and the hilbert spectrum for nonlinear and non-stationary time series analysis," *Proc. Roy. Soc. London A.*, vol. 454, pp. 903 – 995, 1998. ↑ 4

[28] D. P. Mandic, N. U. Rehman, Z. Wu, and N. E. Huang, "Empirical mode decomposition–based time–frequency analysis of multivariate signals: The power of adaptive data analysis," *IEEE Signal Process. Mag.*, vol. 30, no. 6, pp. 74–86, 2013. https://doi.org/10.1109/MSP.2013.2267931 ↑ 4

[29] Y.-H. Wang, C.-H. Yeh, H.-W. V. Young, K. Hu, and M.-T. Lo, "On the computational complexity of the empirical mode decomposition algorithm," *Physica A*, vol. 400, no. 0, pp. 159–167, 2014. https://doi.org/10.1016/j.physa.2014.01.020 ↑ 5

[30] H. K. Fatlawi and A. Kiss, "Similarity-based adaptive window for improving classification of epileptic seizures with imbalance eeg data stream," *Entropy*, vol. 24, no. 11, p. 1641, 2022. https://doi.org/10.3390/e24111641 ↑ 7

[31] F. Zhao, Y. Gao, X. Li, Z. An, S. Ge, and C. Zhang, "A similarity measurement for time series and its application to the stock market," *Expert Syst. Appl.*, vol. 182, p. 115217, 2021. https://doi.org/10.1016/j.eswa.2021.115217 ↑ 7

[32] A. Degirmenci and O. Karal, "Efficient density and cluster based incremental outlier detection in data streams," *Information Sciences*, vol. 607, pp. 901–920, 2022. https://doi.org/10.1016/j.ins.2022.06.013 ↑ 7

[33] J. Han, J. Pei, and H. Tong, *Data Mining: Concepts and Techniques*, 4th ed. San Francisco, CA: Morgan Kaufmann, 2022. https://www.elsevier.com/books/data-mining/han/978-0-12-8117606 ↑ 7

[34] D. Boutana, M. Benidir, and B. Barkat, "On the selection of intrinsic mode functions in emd method: Application on heart sound signal," in *Proc. 3rd Int. Symp. Appl. Sci. Biomed. Commun. Technol. (ISABEL)*, 2010, pp. 1–5. https://doi.org/10.1109/ISABEL.2010.5702895 ↑ 7

[35] P. A. Muñoz, E. Giraldo, M. B. López, and M. Molinas, "Automatic selection of frequency bands for electroencephalographic source localization," in *Proc. 9th IEEE Int. Conf. Neural Eng. (NER)*, pp. 1179–1182. https://doi.org/10.1109/NER.2019.8716979 ↑ 8

[36] Y. Rubner, C. Tomasi, and L. J. Guibas, "The earth mover's distance as a metric for image retrieval," *Int. J. Comput. Vis.*, vol. 40, no. 2, pp. 99–121, 2000. https://doi.org/10.1023/A:1026543900054 ↑ 8

[37] M. Bekbalanova, A. Zhunis, and Z. Duisebekov, "Epileptic seizure prediction in eeg signals using emd and dwt," in *Proc. 15th Int. Conf. Electron., Comput. Computation (ICECCO)*, pp. 1–4. https://doi.org/10.1109/ICECCO48375.2019.9043270 ↑ 8

Luis Daladier Guerrero Otoya

He received a Bachelor's degree in Mechanical Engineering from Universidad Autónoma de Occidente, as well as a specialization in Industrial Automation from Universidad del Valle. He is a Master's degree candidate in Automatics from Universidad del Cauca and served for three years as a professor involved in research processes related to Industrial Engineering at Corporación Universitaria Comfacauca, Popayán, Colombia.
Email: luisdaladier@unicauca.edu.co

Maximiliano Bueno López

He received a Bachelor's and a Master's degree in Electrical Engineering from Universidad Tecnológica de Pereira, Colombia. PhD in Engineering from Universidad Nacional Autónoma de México. He served as a postdoctoral researcher at the Norwegian University of Science and Technology (NTNU), in the Department of Engineering Cybernetics. He currently serves as an associate professor at Universidad Tecnológica de Pereira, Pereira, Colombia.
Email: m.bueno3@utp.edu.co.edu.co

Eduardo Giraldo Suarez

He received his BSc and MSc degrees in Electrical Engineering from Universidad Tecnológica de Pereira, Colombia, in 2004 and 2006, respectively. He holds a PhD in Engineering from Universidad Nacional de Colombia, Manizales, Colombia. He currently serves as a professor at the Department of Instrumentation

and Control of Universidad Tecnológica de Pereira. His research interests include systems identification, adaptive control, optimal control, dynamic inverse problems, and state estimation and its applications to a large number of physical systems, from biological systems to power systems.
Email: egiraldos@utp.edu.co

Marta Molinas

She received a BSc in Electromechanical Engineering from Universidad Nacional de Asunción, Paraguay, in 1992. She earned a Master's degree in Engineering from the University of the Ryukyus, Nishihara, Japan, in 1997. She completed a PhD in Electrical Engineering at the Tokyo Institute of Technology, Tokyo, Japan, in 2000. Since 2014, she serves as a professor of industrial electrotechnics and electrical signal analysis at the Department of Engineering Cybernetics (NTNU), Trondheim, Norway. Her research interests include the area of nonlinear and nonstationary signal analysis, from the domain of harmonics in power systems to that of biological systems, specifically dealing with EEG signals.
Email: marta.molinas@ntnu.no

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.22420

Research

Structural Assessment of Flexible Pavements Based on the Level of Detail of Management Functions

Evaluación estructural de pavimentos flexibles basada en el nivel de detalle de las funciones de gestión

Edwin-Antonio Guzmán-Suárez[1], Diego-Fernando Gualdrón-Alfonso[2], and Jorge-Andrés Sarmiento-Rojas[2]*

[1]Fundación Universitaria Juan de Castellanos (Tunja, Colombia).

[2]Universidad Pedagógica y Tecnológica de Colombia, Tunja campus (Colombia).

Abstract

Context: Pavement condition data are a fundamental component of pavement management systems (PMS) and play a critical role in structural evaluation. The quality of these data directly influences decision-making processes at the network, project, or research level particularly regarding the pavement project life cycle.
Method: This study aimed to assess 18 techniques for evaluating the structure of flexible pavements, utilizing both non-destructive (NDT) and destructive (DT) testing. Following a comprehensive review of the consulted techniques, proprietary models were developed and implemented across multiple projects to structurally evaluate in-service pavements. Statistical analysis was employed to determine the relationships between parameters, distinguishing between those based on empirical and mechanistic approaches.
Results: The application of evaluation techniques revealed that parameters such as radial strain (εrca), vertical strain (εzsr), and the structural number exhibit a strong correlation when categorized within the same approach. Conversely, their correlation is moderately strong when differing approaches are used. Additionally, models relying solely on deflection basin data demonstrated high correlation with rigorous methods that incorporate thickness data.
Conclusions: These findings underscore the practical value of the developed models in pavement management at the network level, offering cost-effective solutions that enhance the detection of structural deficiencies and inform maintenance and rehabilitation strategies.

Keywords: pavement, testing, structural models

Article history

Received: June 25th, 2024

Modified: March 18th, 2025

Accepted: March 30th, 2025

Ing, vol. 30, no. 1, 2025, e22420

* **Correspondence:** jorge.sarmiento02@uptc.edu.co

Resumen

Contexto: Los datos sobre el estado del pavimento son un componente fundamental de los sistemas de gestión (SGP) y desempeñan un papel importante en la evaluación estructural. Dependiendo de la calidad de la información, estos datos son valiosos para la toma de decisiones a nivel de red, de proyecto o de investigación en relación con la gestión del ciclo de vida del pavimento.
Método: El objetivo de este estudio fue evaluar 18 técnicas de evaluación estructural de pavimentos flexibles basadas en ensayos no destructivos (END) y destructivos (DT). Después de una revisión integral de las técnicas consultadas, se desarrollaron modelos propios que fueron implementados en múltiples proyectos para la evaluación estructural de pavimentos en servicio. Se empleó el análisis estadístico para determinar las relaciones entre parámetros, distinguiendo entre aquellos que se basan en enfoques empíricos y mecanísticos.
Resultados: La aplicación de las técnicas de evaluación reveló que parámetros como la deformación radial (εrca), la deformación vertical (εzsr) y el número estructural presentan una fuerte correlación cuando pertenecen al mismo enfoque. Por otro lado, su correlación es moderadamente fuerte si el enfoque difiere. Además, los modelos que se basan únicamente en datos de cuencas de deflexión se correlacionan muy bien con métodos rigurosos que también requieren datos de espesores.
Conclusiones: Los resultados resaltan el valor práctico de los modelos desarrollados para la gestión de pavimentos a nivel de red, proporcionando una solución eficiente en costos que facilita la detección de deficiencias estructurales y mejora las estrategias de mantenimiento y rehabilitación.

Palabras clave: pavimento, pruebas, modelos estructurales

Table of contents

1. Introduction

Structural evaluations are conducted to assess the characteristics of pavement, focusing on the bearing capacity of the subgrade and the structural resistance of the layers composing the pavement structure. These evaluations are essential for estimating the remaining service life of the pavement (1). Various methodologies are used for structural assessment, categorized as *non-destructive* (NDT) and *destructive* (DT) tests, which allow for both qualitative and quantitative evaluations of pavement performance.

One of the most widely used NDT techniques for structural evaluation is the falling weight deflectometer (FWD), which provides the surface deflection basin – a critical dataset for identifying structural behavior parameters. These parameters facilitate qualitative and quantitative assessments of pavement conditions (2).

According to (3) qualitative evaluations are conducted using deflection bowl parameters such as the surface curvature index (SCI), the base damage index (BDI), the base curvature index (BCI), the radius of curvature (Rc), the AREA, the normalized area ratio (A'_r), the area under the pavement profile (Aupp), area indices (AI1, AI2, AI3, AI4), shape factors (F1, F2, F3), the structural index (Ie), the extensibility factor (S), the curvature factor (CF), the deflection slope (SD), the deflection ratio (DR), and the load extensibility index (LSI).

For quantitative pavement assessment, rigorous techniques utilize both deflection (NDT) and thickness (DT) data. These include the AASHTO NDT I method and the mechanistic approaches, which employ computational models to back-calculate the modulus of the layers and estimate the stress states within the structure. Additionally, some methods use closed-form direct calculation formulas based on NDT and DT data, as outlined in (4–6).

These correspond to the AASHTO 93, AASHTO NDT II, Rohde, Wimsatt, and Howard methods. Other direct calculation formulas rely exclusively on NDT deflection data as the input (5,7–12).

The application of techniques such as such as the AASHTO 93, Rohde, and mechanistic methods has been confined to project-level evaluation events for determining maintenance or rehabilitation strategies in highway concessions. Thus, their application in pavement management Systems (PMS) for planning, decision-making, or deterioration modeling remains limited.

While rigorous methods offer higher precision, their implementation demands time and financial resources, restricting their applicability to project-level management. Conversely, direct calculation methods provide advantages such as simplified formulations, rapid applicability, and reduced reliance on specialized software. In addition, by eliminating the need for thickness data, they offer lower cost and avoid the challenges associated with geotechnical exploration, which makes them valuable for network-level evaluations.

The main objective of this research is to test and validate structural evaluation techniques, which are explored in conjunction with original models and applied to in-service pavement structures. This work also introduces assessment methodologies at both the network and project levels. These methodologies facilitate performance monitoring and help to implement actions related to maintenance, rehabilitation, programming, and planning. Consequently, this document presents the results obtained through empirical and mechanistic approaches utilizing NDT and/or DT data for the assessment of structural conditions.

2. Methodology

This study employs a multi-step approach to the structural evaluation of pavements. First, a comprehensive review of various structural evaluation techniques is conducted, followed by the development of original models based on the conceptual frameworks of the reviewed methods. These techniques are classified as *empirical or mechanistic* depending on key parameters such as the structural number (SN) or the stress states (stresses and deformations), respectively.

The reviewed and developed models are then applied to pavements with an asphalt concrete wearing course supported by untreated granular materials. It should be noted that these pavements correspond to various national roads currently in service. Finally, this study assesses the feasibility of implementing these models in network- and project-level management functions within PMS.

To compare and validate the selected models, multiple structural evaluation methods are explored, incorporating both empirical and mechanistic approaches through NDT and DT testing, including the Noureldin, Yonapave, Rohde, and Howard methods, among others referenced in the introduction. While various techniques were tested – such as AASHTO NDT I and mechanistic approaches for estimating the remaining service life – the final selection of AASHTO 93 and the mechanistic (rational) approach using BISAR 3.0 was guided by three key criteria:

- *Broad international recognition*. This provides a solid and reliable basis for evaluating accuracy and applicability.
- *Extensive prior validation in the local context.*
- *Established performanc*e of other models considered and developed in this research, including those based solely on NDT data.

The studied models were constructed using an initial reference, defining ranges of variation in pavement layer thickness and modulus values. While various software tools based on linear elastic theory are available for estimating pavement stress states (*e.g.*, DEPAV, KENLAYER, WESLEA), BISAR 3.0 was selected for this study since it assesses the behavior of the studied parameters. BISAR is a well-established program suitable for the multi-layer elastic analyses aimed at effectively determining the stress states of modeled structures.

The evaluated parameters include the radial tensile strain in the asphalt pavement base (εrca), the vertical compressive strain in the subgrade (εzsr), and indices such as the SCI, the BDI, the BCI, the

Table I. Parameters of the pavement structures modeled

Layer	Thickness (cm)	Modulus (MPa)	Poisson coefficient
Asphalt concrete	5 - 10 - 15 - 20 - 30 - 45 - 60	750 - 1500 - 2250 - 3000 - 3750 - 4500	0.35
Granular	15 - 30 - 45 - 60	75 - 150 - 225 - 300 - 375 - 450	0.40
Subgrade	-	25 - 50 - 75 - 100 - 125 - 150	0.45

Aupp, and the overall AREA. This combination of parameters, as presented in Table I, resulted in a database comprising over 6000 theoretically modeled structures.

The studied roads span different geographical regions and vary in road hierarchy, including both dual and single carriageways serving primary or secondary functions, with an average length of 75 km per carriageway. These in-service structures consist of typical flexible pavements, featuring an asphalt concrete wearing course with thicknesses ranging from 5 to 40 cm, supported by granular materials of varying widths. The structural layers rest upon fine or fine-grained subgrade soils.

To perform the structural evaluation, each carriageway was segmented using the cumulative differences method, with deflection as the response parameter. This method, as outlined in the AASHTO 1993 guidelines and INVÍAS standards (*e.g.*, INV 821-13), categorizes sections with a similar structural behavior by analyzing changes in the slope of the cumulative sum parameter (Zx), derived from deflection measurements along the roadway.

Within each homogeneous unit, assuming that the deflection data approximated a normal distribution, statistical analysis was performed using metrics such as the mean, the standard deviation, and the coefficient of variation. This allowed selecting a characteristic deflection basin that represented the predominant structural behavior, resulting in a total of 160 homogeneous sections defined for this work.

The consulted structural evaluation techniques provide insights into pavement performance through both qualitative and quantitative parameters. Qualitative parameters offer an initial assessment of structural condition, serving as a foundation for quantitative modeling. These quantitative models, which incorporate condition indices, determine the pavement's structural integrity and aid in identifying whether reinforcement measures are required. The following sections will provide a detailed analysis of these models.

2.1. Empirical approach

These techniques rely on parameters thar are empirically obtained. One of these parameters is the SN, which was originally developed in the 1950s through the AASHO road test (13).

2.1.1. Structural capacity of existing or in-service pavement

It is essential to distinguish between the standard structural number (SN) and the effective structural number (SNeff). The standard SN is typically calculated during the pavement design stage, based on specified material properties including layer and drainage coefficients and layer thickness. This value serves as a preliminary estimate for dimensioning the pavement structure. In contrast, SNeff is obtained from field measurements from the existing pavement, reflecting its actual condition and structural capacity after construction, usage, and deterioration – this is traditionally determined using methods like FWD testing. Therefore, SNeff is used to evaluate the current structural adequacy and determine potential rehabilitation needs.

Acknowledging this distinction between design and evaluation parameters, the standard SN remains a fundamental concept in pavement. SN has been widely recognized as a parameter for assessing structural capacity and quantifying pavement strength. It has been adopted as a key criterion in design and evaluation processes to determine the need for reinforcement, and it has been involved in most of the deterioration models employed in PMS (14).

The effectiveness of using the SN and its variants, such as the modified structural number (SNP) and the SNeff, lies in its ability to integrate the mechanical properties of pavement layers into a quantifiable measure of resistance against traffic loads and environmental conditions. These parameters allow for structural capacity evaluation and deterioration prediction. Globally, the SNP is widely used in management models such as HDM-4, a critical tool for road investment and maintenance planning, as it can forecast performance indicators, including crack propagation and the international roughness index (IRI), support optimal resource allocation, and extend the service life of pavement.

In Colombia, consulting firms responsible for structural evaluations in road concessions routinely apply empirical approaches for maintenance management. These firms often incorporate the SNeff (typically derived using the AASHTO 93 method) into their analysis algorithms in order to assess the structural capacity and remaining service life of pavement, which aids in maintenance and rehabilitation decision-making.

Table **II** outlines the main concepts of the methods used for estimating the structural capacity of pavements in service, which are often associated with determining the SNeff.

2.1.2. Required pavement structural capacity

The structural capacity required to accommodate projected traffic demands, which considers subgrade strength, weather conditions, and performance criteria, is quantified using the structural number required (SNreq). This parameter serves as a critical measure in pavement design and evaluation, ensuring that the pavement structure meets the expected load-bearing capacity over its service life. In this study, direct calculation expressions are utilized to estimate it. These include one model developed within this research and another identified in the literature review (Table **III**).

Table II. Structural capacity of pavements in service

AASHTO 1993	
Subgrade strength	**Structural capacity of pavement**
$M_{rr} = \frac{0.24 \cdot P}{D_r \cdot r}$ (1) $r \geq 0.7 \cdot a_e$ $\geq 0.7 \cdot \left[\sqrt{a^2 + \left(H_p^3 \sqrt{\frac{E_p}{M_{rr}}}\right)}\right]$ (2)	$D_0 = 1.5 \cdot p \cdot a \cdot \left\{ \frac{1}{M_{rr} \cdot \sqrt{1 + \left(\frac{H_p}{a} \cdot \sqrt[3]{\frac{E_p}{M_{rr}}}\right)^2}} + \left\lvert \frac{1 - \frac{1}{\sqrt{1 + \left(\frac{H_p}{a}\right)^2}}}{E_p} \right\rvert \right\}$ (3) $SN_{eff} = 0.0045 \cdot H_p^3 \cdot \sqrt[3]{E_p}$ (4)
According to the AASHTO (American Association of State Highway and Transportation Officials), the back-calculated resilient modulus (Mrr) of the subgrade can be determined from the deflections obtained using FWD, via Eq. (1) (15). An adjustment factor (C) is applied to this expression to convert Mrr into an equivalent value measured in the laboratory (Mr). This laboratory value is intended for use in the design procedures described by the AASHTO guidelines. Further details on this can be found in (3). The deflection used for back-calculating Mrr is determined using Eq. (2).	The equivalent modulus, determined through Eq. (3), reflects the structural capacity of the existing pavement, i.e., the equivalent stiffness provided by the asphalt layer and the materials beneath it (15). The methodology proposed by AASHTO employs the SNeff to assess a pavement's structural capacity. This SNeff can be back-calculated from deflection measurements using FWD, as described in Eq. (4).
South African method	
Structural capacity of the pavement-subgrade system	
$SNC = e^{5.12} \cdot BL^{0.31} \cdot A_{UPP}^{-0.78}$ (5) $SNC = \sum_{i=1}^{n} a_i \cdot h_i + SN_{sg} = SN_{eff} + SN_{sg}$ (6) $SN_{sg} = 3.51(\log CBR) - 0.85(\log CBR)^2 - 1.43$ (7) For CBR ≥ 3 for CBR < 3 equals 0 (zero)	$ICS = D_0 - D_{30}$ (8) $A_{UPP} = \frac{5D_0 - 2D_{30} - 2D_{60} - D_{90}}{2}$ (9)
Through a multiple regression procedure, Schnoor and Horak developed a method for assessing the structural capacity of flexible pavements using the modified structural number (SNC) and deflection basin parameters measured via FWD. This resulted in the Regression Eq. (5), which is supported by Eqs. (8) and (9) (16). It should be noted that the SNC accounts for the subgrade's structural contribution. Therefore, to express pavement capacity in terms of the SNeff, the subgrade structural number (SNsg) must be subtracted, as shown in Eqs. (6) and (7) (4).	
Studied model	
Pavement structural capacity	

$SN_{eff} = 1.2 + 7.707H_p - 0.0013D_0 - 0.00061A_{UPP}$	(10)
Eq. (10) correlates specific parameters from the deflection basin with others that provide information about the structural response. The correlation between parameters was statistically determined based on the results of modeling various theoretical structures using the BISAR 3.0 software (3).	

Table III. Structural capacity required

Expression found in the literature	Expression developed in this study
$SN_{req} = 0.05716(\log N - 2.32 \log M_r + 9.07605)^{2.36777}$ (11)	$SN_{req} = 12.547 \cdot M_r^{-0.6567} N^{0.0238 \cdot \ln M_r + 0.0685}$ (12)
Eq. (11) corresponds to a study on structural capacity conducted by the Virginia Transportation Research Council (VTRC) (17).	Using curve fitting methods, an expression was developed which relates SNreq to the design parameters of the method. This is presented in Eq. (12).

2.2. Mechanistic approach

This approach utilizes pavement mechanics to determine the relationship between the cause of loading (traffic loads) and the structural response, which is characterized by stress states (stresses, deformations, and deflections). The radial tensile strain in the base of the asphalt pavement (εrca) and the vertical compressive strain in the subgrade (εzsr) are the parameters considered in estimating structural and functional failure repetitions, respectively.

2.2.1. In-service pavement loadings

Pavement service deformations are calculated using computer programs that simulate stress states through multilayer elastic systems. In this study, the BISAR 3.0 software was employed to calculate the control parameters involved in the mechanistic method (εrca, εzsr), which are associated with fatigue and rutting deterioration (18). To calculate these parameters, regression expressions were also tested in order to predict the structural response based on certain deflection basin parameters obtained via FWD (Table **IV**).

2.2.2. Allowable pavement loads

In-service loads are estimated employing the transfer functions and correlation models presented in Table **V**.

Table IV. Loadings calculated using regression expressions

Literature model	Study model
$\log(\varepsilon_{rca}) = 0.821 \cdot \log(A_{UPP}) + 1.210$ (13)	$\varepsilon_{rca} = 10.443 + 0.4556 \cdot A_{UPP} - 0.4936 \cdot ICS$ $R^2 = 81.64\,\%$ (14)
$\log(\varepsilon_{zsr}) = 1.017 \cdot \log(ICB) - 0.042$ $\cdot \log(H_{ca}) - 0.494 \cdot \log(H_{gr}) + 2.624$ (15)	$\varepsilon_{zsr} = 3.027 \cdot IDB^{1.037}$ $R^2 = 84.54\,\%$ (16)
$ICB = D_{60} - D_{90}$ (17)	$IDB = D_{30} - D_{60}$ (18)
To estimate the pavement's remaining service life, Eqs. (13) and (15) are employed, which yield reasonable results, unlike others that were assessed (19).	Through simple or multiple regression analysis, the prediction models in Eqs. (14) and (16) were developed for the in-service loads, based on parameters of the deflection basin (3).

Table V. Loads calculated by regression expressions

Transfer functions	Correlation equation study
$\varepsilon_{radm} = (0.856 \cdot Vb + 1.08) \cdot E_1^{-0.36} \cdot \left(\frac{N}{RF}\right)^{-0.2}$ (19)	$N_{8.2T\,adm} = 6.2064 \times 10^{13} \cdot ICS^{-3.3575}$ (20)
$\varepsilon_{zadm} = 0.021 \cdot N^{-0.25}$ (21)	$N_{8.2T\,adm} = 3.0310 \times 10^{15} \cdot IDB^{-4.3265}$ (22)
There are several transfer functions to estimate the admissible load values. In this study, 20 of them were tested, ten for fatigue control and ten for rutting. All of them yield different results: some are more permissible, and others are more restrictive. Finally, we decided to work with the Shell Transfer Functions (19) and (21) (20).	To indirectly estimate the remaining pavement life based on deflection basin parameters, this study generated Correlation Eqs. (20) and (22) between the ICS and IDB layer indices and the consulted transfer functions for fatigue and rutting control, respectively.

3. Indices for the structural condition and remaining service life of pavement

Pavement condition is estimated with the indices presented in Table **VI**.

4. Results

Fig. **1** presents the correlation results for the different models studied within the mechanistic and empirical approaches. In addition, it shows how the selected structural indices correlate (IVRF/IRVA *vs.* SCI) depending on the method taken as a reference.

Table VI. Indices for the condition and remaining service life of pavement

Pavement condition and remaining service life indices	**Remaining life ratio**
$SCI = \frac{SN_{eff}}{SN_{req}}$ (23)	Index for fatigue control $\text{IVRF} = \frac{\varepsilon_{radm}}{\varepsilon_{rca}} \circ \frac{N_{8.2T\,adm}}{N}$ (24) Index for rutting control $\text{IVRA} = \frac{\varepsilon_{zadm}}{\varepsilon_{zsr}} \circ \frac{N_{8.2T\,adm}}{N}$ (25)
The pavement's structural condition is evaluated based on the SNeff, determined using NDT from deflections obtained via FWD (South Africa method) and, in some methods, from pavement thickness (AASHTO 1993 and this work). This can be estimated using the structural condition index (SCI) proposed by the Texas Department of Transportation (TxDOT), as shown in Eq. (23) (21).	These indices, obtained using Eqs. (24) and (25), represent the remaining service life of the pavement, in a manner similar to the SCI. They are derived from the ratio between the permissible loads for the stress states or load repetitions, obtained from the transfer functions, vs. the in-service values, calculated from the pavement mechanics or the traffic study. They are the inverse of the deformation (εrca, εzsr) or loading stress consumption. The IVRF and IVRA indices, corresponding to Eqs. (24) and (25), represent ratios comparing allowable and actual in-service conditions, accounting for traffic variability through the estimated number of in-service load repetitions or through the representative loads used to calculate in-service strains. These indices integrate environmental factors, particularly their influence on layer moduli, which is crucial for these calculations. In standard applications, modeling long-term material property changes (aging) is often simplified, which may limit accuracy in capturing progressive pavement deterioration. While the index concept remains adaptable, its application across different pavement types requires specific structural models and appropriately calibrated transfer functions (ε_{adm}, N_{adm}) for those conditions, limiting direct extrapolation without such validation. Reliability considerations are generally embedded within the selected transfer functions or failure criteria.

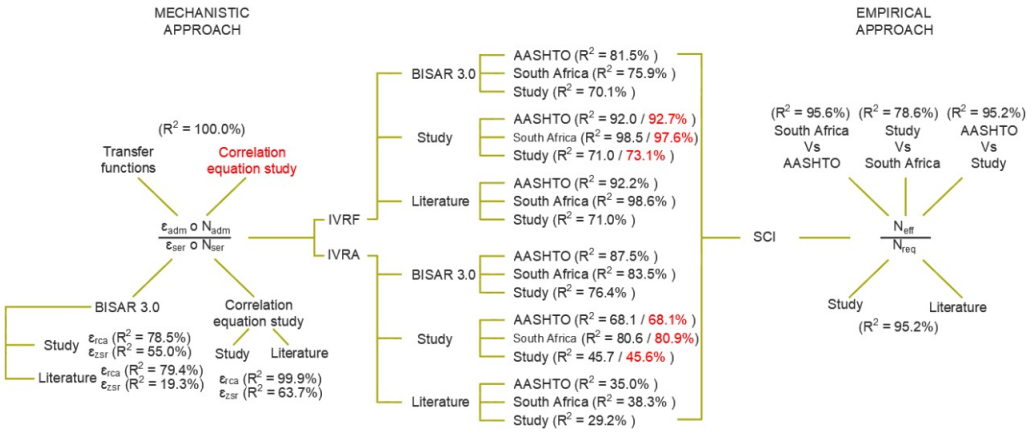


Figure 1. Coefficient of determination for the different models under study

Fig. 2 demonstrates how our model strongly predicts the radial tensile deformation at the base of the asphalt layer (R2=78.5 %). However, the fit is less accurate with the vertical compressive deformation in the subgrade, and the relationship between parameters is moderate (R2=55.0 %). A similar trend is observed in the literature-reported model and the control parameters for fatigue and rutting, with R2 values of 79.4 and 19.3 %. Both models, the one consulted and the one developed, predict the radial

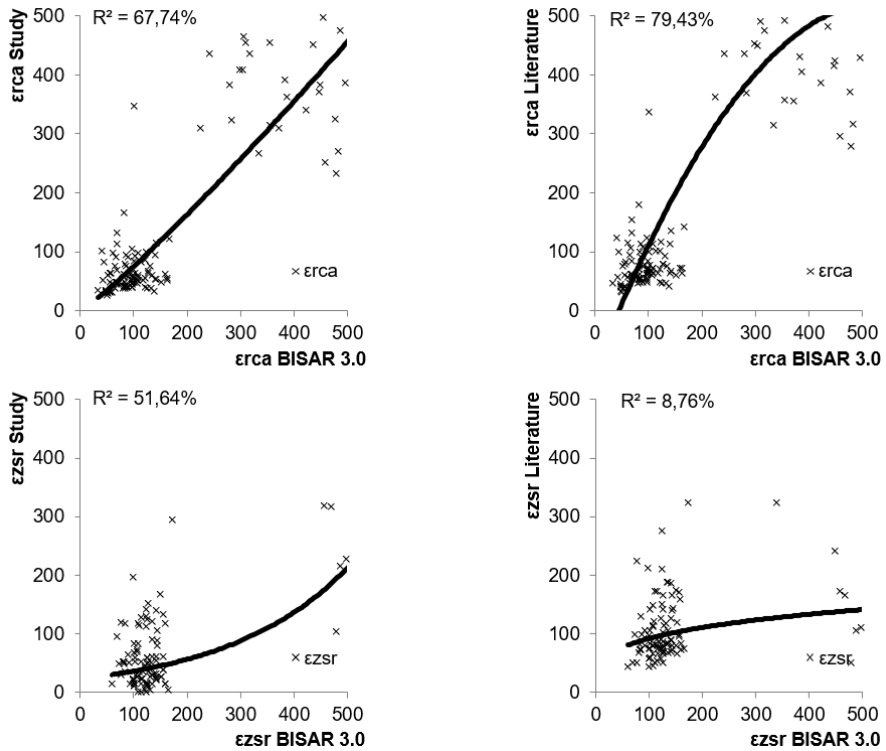


Figure 2. Relationship between εrca and εzsr in the studied models

deformation of the asphalt layer in the same way (R2=99.9 %). However, for the vertical deformation of the subgrade, the relationship is not as strong, and each model estimates the behavior of this parameter differently (R2=63.7 %) (Fig. 3).

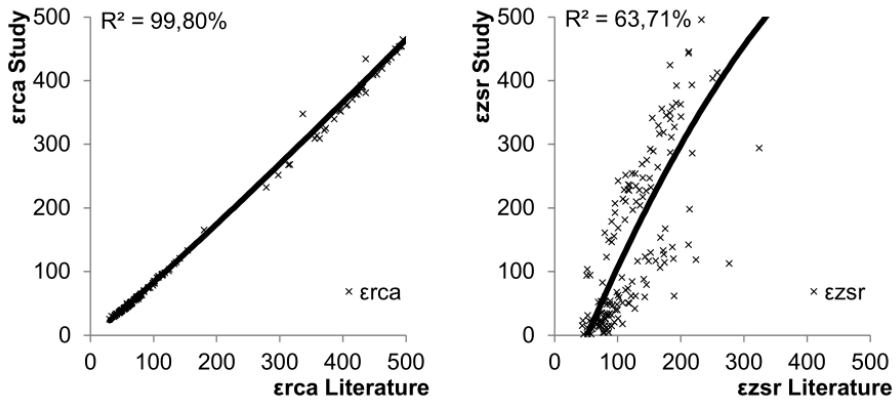


Figure 3. Relationship between εrca and εzsr, as determined in this study and the literature-reported model

To assess the structural capacity of the pavement, the AASHTO 93 method was used as reference due to its recognized international acceptance, extensive application, and proven reliability in the local context. As illustrated in Fig. 4, the AASHTO SNeff exhibits a strong relationship with the predictions obtained with the study's proposed methods and the South Africa approach, with R2 = 95,2 and 93,8 %, respectively. Similarly, a strong relationship is observed when determining the SNeff using our model and the South Africa method (R2=69.7 %).

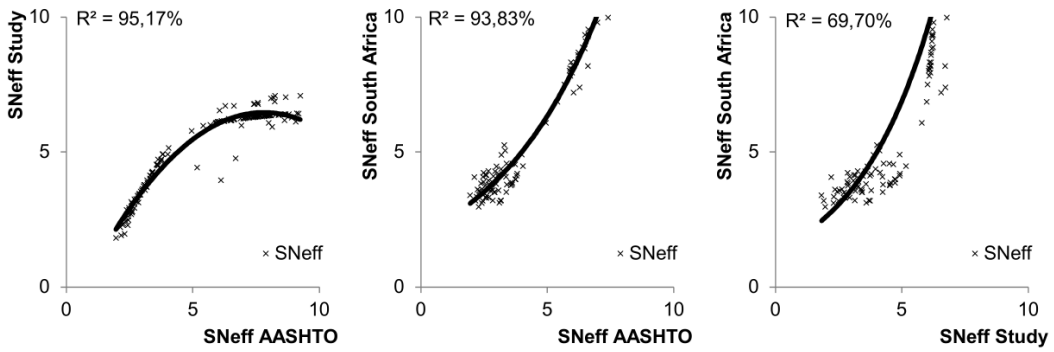


Figure 4. Relationship between the SNeff values determined through the different methods under study

To calculate the SCI, an expression was developed for determining the SNreq. The derived expression exhibits a strong correlation with an existing model in the literature (R2=95.2 %). Fig. 5 illustrates these relationships. Given the strong fit between our model, the literature-based model, and the general AASHTO equation, the SNreq parameter can be accurately predicted.

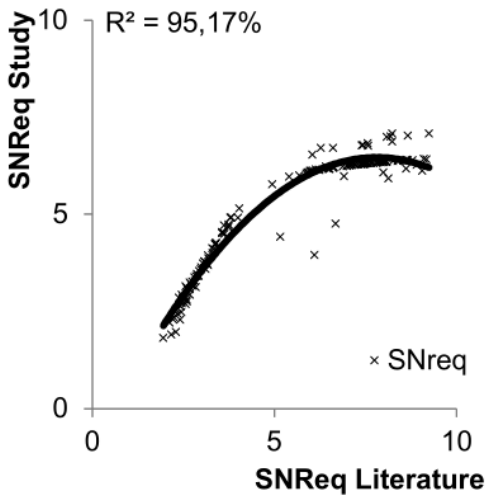


Figure 5. Relationship between the SNreq values determined using our model and the literature-reported approach

The methods selected as the foundation for this study are among the most rigorous in their calculations, particularly in estimating the remaining service life of pavement based on the relationship between the admissible loads (computed using transfer functions) and the service life (determined via BISAR 3,0). Additionally, as discussed above, the AASHTO 93 method was chosen as a reference framework due to its wide recognition, high acceptance, and extensive validation in the local context. As illustrated in Figs. **6** and **7** a strong correlation is observed between the structural conditions estimated for both the mechanistic and empirical approaches by means of the defined indices (IVRF/IVRA and SCI). These approaches consistently indicate whether pavement reinforcement is required, helping to identify structural deficiencies and mitigate pavement failure risks caused by fatigue and rutting due to repeated traffic loads.

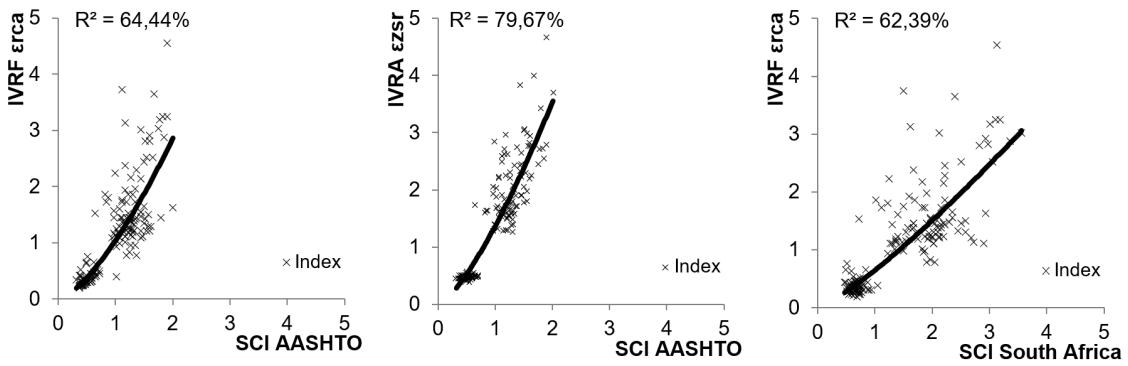


Figure 6. Relationship between IVRF (BISAR 3.0) and SCI, as determined through the methods under study

Similarly, Figs. 6 and 7 illustrate the SCI obtained with both models, showing a strong relationship with the IVRF/IVRA indices. This relationship is slightly stronger with the SCI obtained from the South Africa SNeff model. This behavior indicates that the structural condition of the pavement can be adequately predicted by applying any of the models.

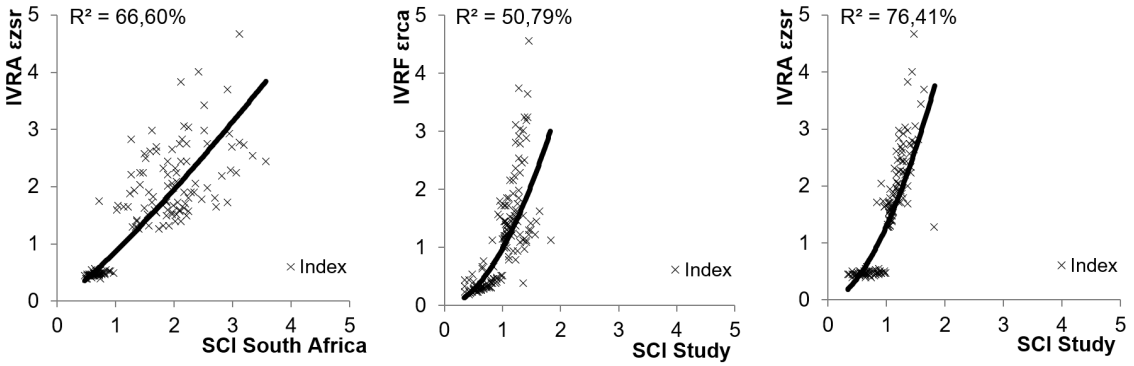


Figure 7. Relationship between IVRA (BISAR 3.0) and SCI, as determined through the methods under study

During this research, alternative methods were tested in which the structural condition is estimated exclusively based on deflection basin parameters, without relying on thickness values obtained from geotechnical or georadar exploration. These techniques include correlation equations to predict εrca and εzsr (mechanistic approach) and the South Africa method (empirical approach). As illustrated in Fig. 8, the relationship between SCI and IVRF in both methods is strong, confirming that a preliminary diagnosis of the pavement can be conducted by monitoring its structural condition and remaining service life. This eliminates the need for invasive testing, making these methods particularly useful for network-level pavement studies. Alternatively, this study proposes estimating the remaining service life by relating the permissible traffic loads (calculated based on the ICS and IDB layer indices) to the traffic study values. The correlation between the IVRF/IVRA and SCI determined through the AASHTO and South Africa methods indicates remarkable consistency. Fig. 9 illustrates the correlation results obtained with the South Africa method. The application of NDT methods can effectively aid in assessing pavement condition in terms of both structural capacity and remaining service life.

5. Discussion

In this study, the application of newly developed models for the structural assessment of pavement was tested while following the same approach and resolution principles as the techniques established in the literature. The results of the proposed expressions for evaluating structural parameters demonstrated strong alignment with findings from previously consulted methods, including correlation equations for predicting the radial tensile strain and the vertical compressive strain (mechanistic approach) as well as the structural number (empirical approach).

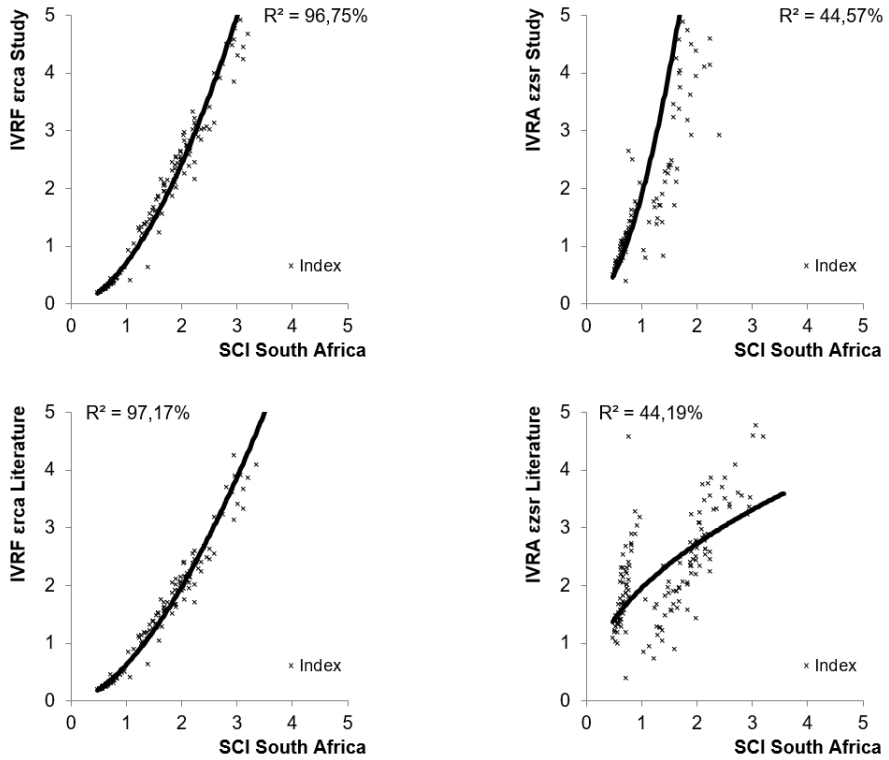


Figure 8. Relationship between IVRF/IVRA (correlation equations) and SCI, as determined with the South Africa method

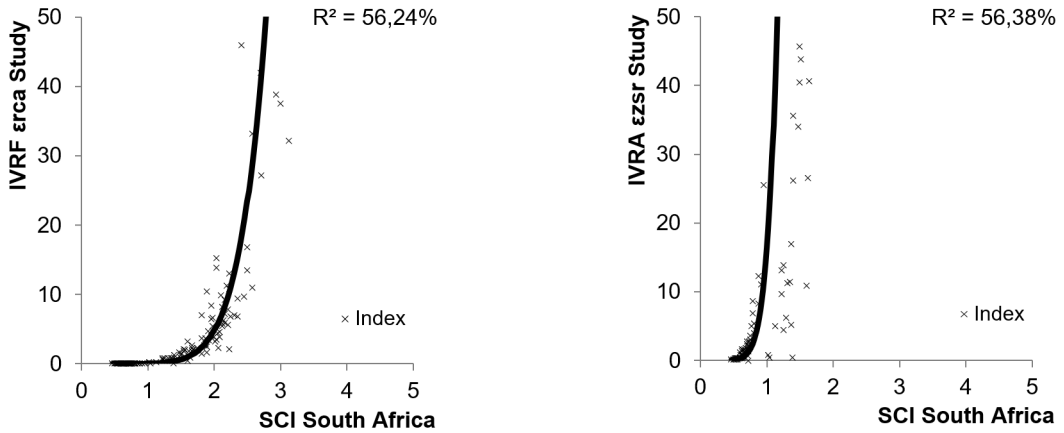


Figure 9. Relationship between IVRF/IVRA (traffic) and SCI, as determined using the South Africa method

In the specialized literature, (21) conducted a study on the structural condition index, concluding that it correlates strongly with the pavement's remaining service life as a function of fatigue and rutting.

This study successfully validated this relationship for the evaluated in-service pavement sections, which consisted of asphalt concrete wearing courses supported by untreated granular materials.

It is critical to note that these findings apply specifically to the flexible pavement structures evaluated in this study, *i.e.*, asphalt concrete layers over untreated granular materials, supported by various subgrade types. The formulations presented are based on traditional design methodologies developed under specific loading and environmental conditions. Therefore, direct application in significantly different contexts – such as varied climatic conditions, traffic loads, or pavement structures incorporating treated bases – requires careful consideration and likely adjustments through local calibration. This calibration should address structural layer coefficients, drainage coefficients, and reliability levels in order to ensure accurate predictions and facilitate adaptation to similar scenarios.

For network-level pavement assessments, the South Africa method is strongly recommended, as it demonstrates high accuracy in predicting the AASHTO SN. A key advantage of this method is its exclusive reliance on parameters derived from FWD deflection bowl measurements, thereby eliminating the need for thickness data. Unlike the current study, which developed models using theoretical structural frameworks, the South Africa method was formulated using field thickness data and FWD deflection bowl analyses. Consequently, a significant limitation of this study is the lack of model validation using real-world in-service pavement data.

There are several alternative, simple-to-use methods for estimating structural capacity and the remaining service life of pavement based solely on FWD data, without requiring thickness measurements. These approaches eliminate the costs and challenges associated with geotechnical exploration or georadar surveys while maintaining strong correlations with traditional rigorous methods typically applied at the project level. Furthermore, the indices employed in this study exhibit a good mutual correlation, thus supporting structural deficiency detection and collectively indicating whether reinforcement is required to meet structural and remaining life expectations.

6. Future research recommendations

For future studies, we recommend developing a model that exclusively relies on FWD-derived parameters, built upon detailed geotechnical and deflectometric data rather than on theoretical structural models, as was the case with this study. This new model should be designed to evaluate asphalt pavement structures with both treated and untreated granular layers.

7. Author contributions

Edwin Antonio Guzmán Suárez: conceptualization, methodology, software, formal analysis, data curation, writing (original draft)
Diego Fernando Gualdrón Alfonso: validation, formal analysis, writing (original draft)
Jorge Andrés Sarmiento Rojas: writing (review & editing), visualization, supervision

8. Conflict of interest

The authors declare that there are no conflicts of interest that could have influenced the research process or the interpretation of the results. This declaration is made in compliance with the standards of transparency and integrity in academic endeavors, ensuring that the study and its conclusions are based solely on objective data and analysis.

References

[1] Z. Liu, Q. Yang, and X. Gu, "Assessment of pavement structural conditions and remaining life combining accelerated pavement testing and ground-penetrating radar," *Remote Sens.*, vol. 15, no. 18, 2023. https://doi.org/10.3390/rs15184620 ↑3

[2] L. Fuentes, K. Taborda, X. Hu, E. Horak, T. Bai, and L. F. Walubita, "A probabilistic approach to detect structural problems in flexible pavement sections at network level assessment," *Int. J. Pavement Eng.*, vol. 23, no. 6, pp. 1867–1880, 2022. https://doi.org/10.1080/10298436.2020.1828586 ↑3

[3] E. A. Guzmán Suárez, "Evaluación estructural de pavimentos flexibles," Fundación Universitaria Juan de Castellanos, 2023. https://doi.org/10.38017/9789588966557 ↑3, 7, 8, 9

[4] G. T. Rohde, "Determining pavement structural number from FWD testing," *Transp. Res. Rec.*, no. 1448, pp. 61–68, 1994. http://onlinepubs.trb.org/Onlinepubs/trr/1994/1448/1448-008.pdf ↑3, 7

[5] Z. Zhang, L. Manuel, I. Damnjanovic, and Z. Li, "Development of a new methodology for characterizing pavement structural condition for network level applications," Center for Transportation Research, Bureau of Engineering Research, University of Texas at Austin, Austin, TX, USA, Rep. no. FHWA/TX-04/0-4322-1, 2003. [Online]. Available: https://ctr.utexas.edu/wp-content/uploads/pubs/0_4322_1.pdf ↑3

[6] G. T. Rohde and A. Hartman, "Comparison of procedures to determine structural number from FWD deflection," in *Combined 18th ARRB Transp. Res. Conf, Transit New Zealand Land. Transp. Symp.*, 1996, pp. 99–115. https://trid.trb.org/view/470309 ↑3

[7] A. L. Crook, S. R. Montgomery, and W. S. Guthrie, "Using falling-weight deflectometer data for network-level flexible pavement management," *Transp. Res. Rec.*, vol. 2304, no. 1, pp. 75–83, 2012. https://doi.org/10.3141/2304-09 ↑3

[8] COST Action 336, "Use of falling weight deflectometers in pavement evaluation: final report," Off. Publ. Eur. Communities, Luxembourg, Rep. 392, 2005. ↑3

[9] HTC Infrastructure Management Ltd., "Implementation of predictive modelling for road management – establishing pavement strength for use with dTIMS," HTC Infrastruct. Manage. Ltd., Ontario, Canada, Tech. Rep., 2000. ↑3

[10] J. Crovetti, "Development of rational overlay design procedures for flexible pavements," Wis. Highway Res. Program, Madison, WI, Rep. WHRP 05-12, 2005. ↑3

[11] A. Kavussi, M. Abbasghorbani, F. Moghadas Nejad, and A. Bamdad Ziksari, "A new method to determine maintenance and repair activities at network-level pavement management using falling weight deflectometer," *J. Civ. Eng. Manag.*, vol. 23, no. 3, pp. 338–346, 2017. https://doi.org/10.3846/13923730.2015.1073173 ↑3

[12] G. Salt and D. Stevens, "Pavement performance prediction: determination and calibration of structural capacity (SNP)," in *20th ARRB Transp. Res. Conf.*, 2001, art. 15. ↑3

[13] N. Vitalis and M. Ephather, "Improvement of the algorithm for computing adjusted structure number for determination of backlog maintenance," *World J. Eng. Technol.*, vol. 10, no. 02, pp. 158–178, 2022. https://doi.org/10.4236/wjet.2022.102009 ↑5

[14] H. S. Abd El-Raof, R. T. Abd El-Hakim, S. M. El-Badawy, and H. A. Afify, "Structural number prediction for flexible pavements using the long term pavement performance data," *Int. J. Pavement Eng.*, vol. 21, no. 7, art. 16, 2018. https://doi.org/10.1080/10298436.2018.1511786 ↑6

[15] AASHTO, "Guide for design of pavement structure," 1993. [Online]. Available: https://habib00ugm.files.wordpress.com/2010/05/aashto1993.pdf ↑7

[16] E. A. Guzmán Suárez, *Evaluación estructural de pavimentos flexibles*. Tunja, Colombia: Fundación Universitaria Juan de Castellanos, 2023. https://doi.org/10.38017/9789588966557 ↑7

[17] E. Horak, J. Maina, and A. Hefer, "Structural number determined with the falling weight deflectometer and used as benchmark methodology," in *Int. Conf. Civil Eng. Energy and Environ. Struct.*, 2014, art. 6165. https://doi.org/10.13140/2.1.1459.6165 ↑8

[18] J. M. Bryce, G. W. Flintsch, S. W. Katicha, B. K. Diefenderfer, and A. Sarant, "Development of pavement structural capacity requirements for innovative pavement decision-making and contracting: phase II," Va. Transp. Res. Council, Charlottesville, VA, Rep. FHWA/VTRC 16-R20, 2016. ↑8

[19] O. Mora *et al.*, "Comparative analysis on strains in asphalt pavement design using linear elastic and viscoelastic theories," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 1126, no. 1, art. 012028, 2021. https://doi.org/10.1088/1757-899x/1126/1/012028 ↑9

[20] Y. R. Kim and H. Park, "Use of falling weight deflectometer multi-load data for pavement strength estimation," 2002. [Online]. Available: https://rosap.ntl.bts.gov/view/dot/5253 ↑9

[21] INVÍAS, "Manual de diseño de pavimentos asfálticos en vías con medios y altos volúmenes de tránsito," 2015. ↑10, 15

[22] S. S. S. Peddibhotla, M. Murphy, and Z. Zhang, "Validation and implementation of the structural condition index (SCI) for network-level pavement evaluation," 2011. [Online]. Available: http://ctr.utexas.edu/wp-content/uploads/pubs/5_4322_01_1.pdf ↑

Edwin Antonio Guzmán Suárez

Engineer in Transportation and Roads, UPTC; specialist in Road Infrastructure, UPTC; Specialist in Geographic Information Systems, FUJC; MSc in Engineering with an emphasis on Road Infrastructure, UPTC; MSc in Traffic, Mobility, and Road Safety, UCJC. Consultant with more than 15 years of national and international experience in road infrastructure design and pavement rehabilitation projects in countries such as Honduras, Peru, and Equatorial Guinea. Undergraduate and graduate university professor.

Email: eaguzman@jdc.edu.co

Diego Fernando Gualdron

Civil engineer with specializations in Environmental Engineering, Project Management, and Geographic Information Systems. He has a Master's degree in Construction Project Management and has authored several books and articles in indexed journals, including *Aplicaciones SIG para la Ingeniería Civil*. He has more than ten years of teaching experience and an extensive professional career in both the public and private sector.
Email: diego.gualdron@uptc.edu.co

Jorge Andrés Sarmiento Rojas

Civil engineer, PMP; MBA with an emphasis on Finance; MSc in Integrated Project Management; PhD in Project Management. He has experience in project management in the Colombian public and private sector. Dedicated researcher and undergraduate and postgraduate teacher. He has led a variety of research projects, integrating techniques and tools that have led to their success, and he has authored publications such as research and training books as well as high-impact technical and scientific documents. Due to his leadership, commitment, ethics and communication skills, he has innovated multiple processes and procedures in the field of business management.
Email: jorge.sarmiento02@uptc.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.22528

Research

Analysis of the Physical and Mechanical Behavior of Soil Reinforced with Banana Fibers

Análisis del comportamiento físico y mecánico del suelo reforzado con fibras de plátano

Favio Osmar Schreiber Robles[1], Sócrates Pedro Muñoz Pérez[2]*, Juan Martín García Chumacero[1], Elver Sánchez Díaz[1], Carlos Arturo Damiani Lazo[3], Juan De Dios Malpartida Iturregui[1], Ángel Antonio Ruiz Pico[2], Edwin Adolfo Díaz Ortiz[2], Ernesto Dante Rodríguez Lafitte[1], and Ana Paula Bernal Izquierdo[1]

[1]Universidad Señor de Sipán (Chiclayo, Peru)

[2]Universidad César Vallejo (Trujillo, Perú)

[3]Universidad Nacional de San Agustín de Arequipa (Arequipa, Peru)

Abstract

Context: Soils reinforced with natural fibers such as banana fibers (BF) constitute a promising alternative for improving the geotechnical properties of the soil, especially in rapidly growing urban contexts like Peru.
Methods: This study was structured into four stages: the extraction and preparation of soil samples; the evaluation of the physical characteristics of the fibers; mixing with proportions of 0.5, 1, 1.5, and 2% BF relative to the soil dry weight; and physical and mechanical tests to assess the effects on geotechnical properties.
Results: The addition of 1% BF optimized the properties of the modified soil: the maximum dry density remained stable, the California bearing ratio increased by 5.95%, and the unconfined compressive strength increased by 23.81% compared to natural soil.
Conclusions: The use of BF-treated soil meets the local standards for application in infrastructure such as roads and pavements, thus promoting sustainable construction practices and contributing to the development of resilient and environmentally responsible infrastructure.

Keywords: soils, natural fibers, banana fibers, geotechnical properties, physical and mechanical behavior

Article history

Received: July 26th, 2024

Modified: November 25th, 2024

Accepted: March 1st, 2025

Ing, vol. 30, no. 1, 2025, e22528

*Correspondence: mperezsp@ucvvirtual.edu.pe

Resumen

Contexto: Los suelos reforzados con fibras naturales como las fibras de plátano (FP) representan una alternativa prometedora para mejorar las propiedades geotécnicas de los suelos arcillosos de baja plasticidad, especialmente en contextos urbanos de rápido crecimiento como el Perú.
Métodos: Este estudio se estructuró en cuatro etapas: la extracción y preparación de muestras del suelo; la evaluación de las características de la fibra; la mezcla con proporciones de 0.5, 1, 1.5 y 2 % de FP respecto al peso seco del suelo; y la realización de pruebas físicas y mecánicas del suelo estabilizado.
Resultados: La adición de 1 % de FP optimizó las propiedades del suelo natural, pues la máxima densidad seca se mantuvo estable, el índice de soporte de California aumentó en un 5.95 % y la resistencia a la compresión no confinada incrementó en un 23.81 % en comparación con el suelo natural.
Conclusiones: El uso del suelo tratado con FP cumple con los estándares locales para aplicaciones en infraestructuras como carreteras y pavimentaciones, lo que promueve prácticas constructivas sostenibles y contribuye al desarrollo de infraestructuras resilientes y ambientalmente responsables.

Palabras clave: suelos, fibras naturales, fibras de plátano, propiedades geotécnicas, comportamiento físico y mecánico

Table of contents

1. Introduction

Soil stabilization remains a critical focus in construction and environmental engineering, with natural fibers (NF) emerging as a sustainable and cost-effective reinforcement technique. (1) highlight their advantages: low cost, environmental benefits, and high availability, which makes them a viable alternative to conventional materials (2, 3) emphasize their importance in reducing the environmental impacts of construction projects. Low soil tensile strength often leads to structural failures (4), but fibers improve stability by leveraging their tensile properties, reducing cracks and degradation (5, 6). This reinforcement occurs through interlocking and friction, thereby enhancing resistance (7). Natural fibers are subject to degradation processes such as biodegradation, moisture exposure, and temperature fluctuations, which can affect their structural integrity and, consequently, the stability of the reinforced soil. It is essential to investigate how these factors influence the longevity of the fibers to ensure that the observed geotechnical improvements remain effective over time.

Global banana production is estimated at nearly 30.5 million tons per year, making it the third most important export product after fuel and wood (8). Furthermore, it is expected to increase at an annual rate of 1.5 %, reaching 135 million tons by 2028 (9). India leads global production, accounting for approximately 31.5 % of the total, followed by mainland China, Indonesia, Brazil, and Ecuador (10). In 2023, Peru's commercial banana (plantain) production reached 2 414 382 tons (11). The country has 160 000 hectares dedicated to banana and plantain cultivation, with more than 70 % concentrated in the Amazon region (12). Banana fibers contain approximately 40 % cellulose, 21.5 % hemicellulose, 24 % lignin, 0.34 % pectin, 7.2 % extractable lipophilic sugars, and 7.36 % water-soluble sugars (8).

The use of randomly distributed natural fibers to reinforce soil is an effective and simple method, similar to cement or lime stabilization (13). For example, banana fibers, with densities of 1.2-1.6 g/cm^3, an elongation of 5.3 %, and a tensile strength between 355 and 754 MPa, are biodegradable (14, 15). While synthetic fibers offer superior performance, natural fibers are more cost-effective, accessible, and environmentally friendly, making them a valuable option for sustainable projects (16, 17). Currently, chemical pretreatments such as sodium hydroxide, calcium hydroxide, hydrogen peroxide, sulfuric acid, silane, enzymes, epoxy resins, and calcium sulfate dihydrate are used to enhance these organic fibers, increasing their durability over time.

Banana fibers (BFs) are preferred for their high tensile strength (355-754 MPa), biodegradability, availability in tropical regions like Peru, and low cost. They are more affordable than jute or coir since they are agricultural byproducts. Additionally, their use helps to reduce agricultural waste and mitigate environmental impacts, contributing to sustainable and eco-friendly projects.

The use of fibers in soils offers multiple benefits and is suitable for various geotechnical applications, including subbase layers, foundations, slope protection, and earthquake- resistant structures (18). Many studies have demonstrated the influence of incorporating BF into the soil. For example, (19) observed that a proportion of 0.4 % BF increased the liquid limit by 25 %, and the plastic limit substantially increased by 125.45 % compared to natural soil. (20) found that 2 % BF increased both the liquid and

plastic limits by 21 %. However, further research is needed on the physical behavior of soil with different proportions of BF by dry weight of soil (2).

The addition of natural fibers (NFs) significantly influences soil properties, particularly the maximum dry density (MDD), the California bearing ratio (CBR), and the unconfined compressive strength (UCS). Studies have shown mixed effects on MDD: while 0.5 % fiber increased MDD by 1.58 and 6.67 % (21, 22), higher proportions such as 2 % led to reductions of 7.78 and 2.01 % (19, 23). This decline was attributed to excess voids acting as physical barriers at higher fiber contents (24). Meanwhile, the addition of 0.4 % BF increased the optimum moisture content (OMC) by up to 20 % compared to natural soil (19). Similarly, (21) demonstrated that adding 0.5 % BF stabilized the soil, achieving an OMC of 16 % relative to natural soil. Additionally, (22) found that, at the same fiber proportions, the OMC increased by 41.03 % compared to natural soil. Small additions of fiber, *i.e.*, 0.3 and 0.4 %, increased the CBR by 23.71 and 21.55 %, respectively (19, 24). Larger proportions like 1 and 1.25 % achieved more substantial improvements, enhancing the CBR by up to 226.92 % (25, 26). This enhancement is due to the improved soil cohesion and stress redistribution provided by fibers. Similarly, the UCS results showed significant strength gains, with 0.3 and 1.5 % fiber increasing the reported values by 26.41 and 100 %, respectively (24, 27). These improvements result from enhanced internal cohesion and shear strength, as fibers reinforce the soil structure, boosting its load-bearing capacity under compression (28).

This study addresses a significant knowledge gap, as no prior research in Peru has explored the incorporation of BF to improve soil behavior, despite the existence of international studies. The objective is to analyze the physical and mechanical behavior of soil reinforced with BF while focusing on four percentages, *i.e.*, 0.50, 1, 1.5, and 2 % of soil dry weight, in order to evaluate their effect. A key novelty of this research lies in using natural waste materials like BF, promoting sustainability by reducing the environmental impact of improper waste disposal. This aligns with Sustainable Development Goal (SDG) 12 (responsible consumption and production) by encouraging resource-efficient practices and with SDG 9 (industry, innovation, and infrastructure) by fostering green industries and innovation.

However, the limitations of this work include the focus on laboratory-scale tests without considering field conditions, as well as the results being specific to the chemical composition and properties of the BF used. Further research is needed to validate the findings regarding diverse soil types and environmental conditions. This research provides a foundation for integrating BF as a sustainable material in geotechnical applications, offering insights into its potential to enhance soil properties and contribute to more environmentally responsible construction practices.

2. Materials and methods

2.1. Materials Soil

The soil samples used in this study were collected from Pueblo Joven Las Dunas, in the city of Lambayeque, Peru. Once the location had been identified, soil samples were excavated from a depth of 1.50 m and placed in airtight bags. Approximately 250 kg of soil were extracted from these pits.

The selected road is strategically important for the movement of agricultural products and connections between surrounding communities. The main geotechnical characteristics of its soil are presented in Table I.

Table I. Geotechnical properties of the natural soil studied

Test	Characteristics	Standard	Results
Depth of sample taken (m)	0.5-1.5	-	
Grain size distribution	Gravel >4.75 mm (%)	ASTM D422 (29)	-
	Sand 0.075 – 4.75 mm (%)		20.82
	Fines <0.075 mm (%)		79.18
Soil type	Classification (USCS)	ASTM D2487 (30)	CL
Atterberg limits	Liquid limit (%)	ASTM D4318 (31)	20.3
	Plastic limit (%)		19.2
Modified Proctor tests	Maximum dry density (g/cm³)	ASTM D1557 (32)	1.8
	Optimum moisture content (%)		12.7
Strength test	California bearing ratio at 95 % MDD (%)	ASTM D1883 (33)	33.6
	California bearing ratio at 100 % MDD (%)		34.4
	Unconfined compressive strength after 28 days (MPa)	ASTM D2166 (34)	0.21

2.1.1. Banana fiber

Locally available BF (Fig. 1) was used at various stages, and proper treatment was conducted before incorporating it into the soil mix. The fibers were treated with a 5 % sodium hydroxide (NaOH) solution for 24 hours and then air-dried (35). This sodium hydroxide treatment increased the tensile strength of the natural fibers (36–44). In the application process, the soil was prepared, and BF was mixed in proportions of 0.50, 1, 1.5, and 2 % by dry weight of soil, using suitable mixing equipment to ensure a homogeneous distribution.

Various physical properties of the BF are described in Table II. Additionally, based on an analysis using X-ray fluorescence testing as per ASTM E1621-16 (45), the chemical components identified in the fibers are shown in Table III and Fig. 2.

2.2. Methods

2.2.1. Liquid limit, plastic limit, and modified Proctor test

The liquid and plastic limits (LL and PL) were determined according to ASTM D4318 (31). The modified Proctor test for both natural and modified soil was conducted in accordance with ASTM D1557 (32). Fig. 3 depicts the physical tests conducted on the soil.

Figure 1. Banana fiber

Table II. Physical properties of the banana fibers used

Test description	Results	Standard
Length (mm)	40 - 50 mm	-
Diameter (mm)	0.40 - 2.0 mm	-
Density (g/cm^3)	1.25	
Loose unit weight (kg/m^3)	0.26	ASTM C 188 (46)
Tensile strength	250 MPa	ASTM D 3822 (47)

Table III. Chemical properties of the banana fibers used

Element	% mass
Al_2O_3	0.338
SiO_2	1.644
P_2O_5	14.595
SO_2	7.806
ClO_2	4.973
K_2O	66.567
CaO	4.764
TiO_2	5.028
MnO	11.019

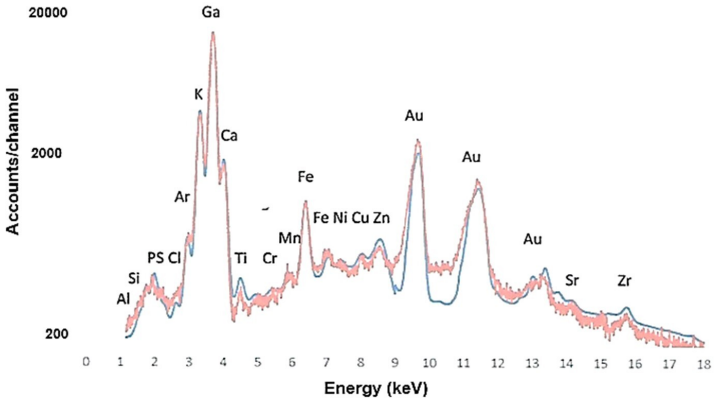


Figure 2. X-ray fluorescence analysis of the banana fibers used

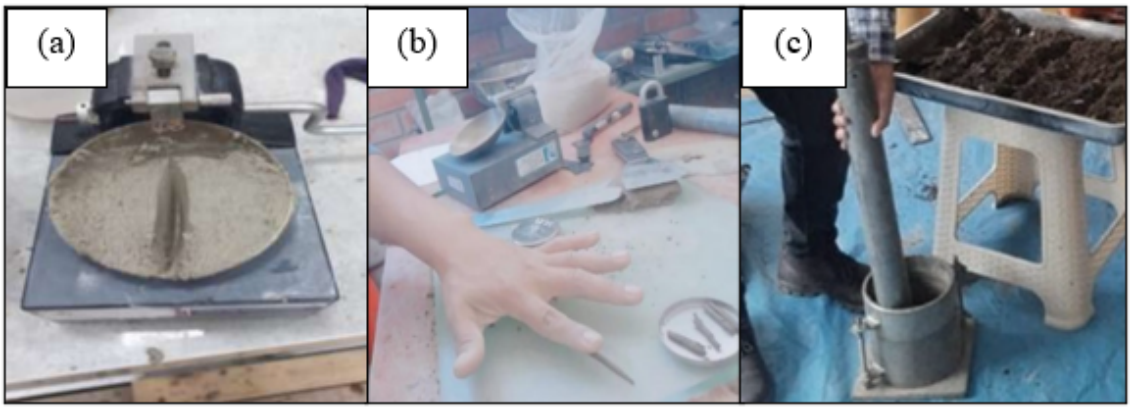


Figure 3. a) Liquid limit, b) plastic limit, and c) modified Proctor tests

2.2.2. California bearing ratio and unconfined compressive strength

Fig. 4 shows the mechanical tests performed on the soil containing BF. The CBR was evaluated according to ASTM D1883 (33), and the UCS was evaluated according to ASTM D2166 (34).

2.2.3. Experimental program

The methodology developed for this research was divided into four stages. Fig. 5 illustrates its process flow.

Phase I involved soil sampling and processing, with samples extracted from a depth of 1.50 m in the Lambayeque region of Peru, totaling approximately 250 kg. This material was divided for physical and mechanical tests in its natural state, which included grain size analysis, Atterberg limits, specific gravity,

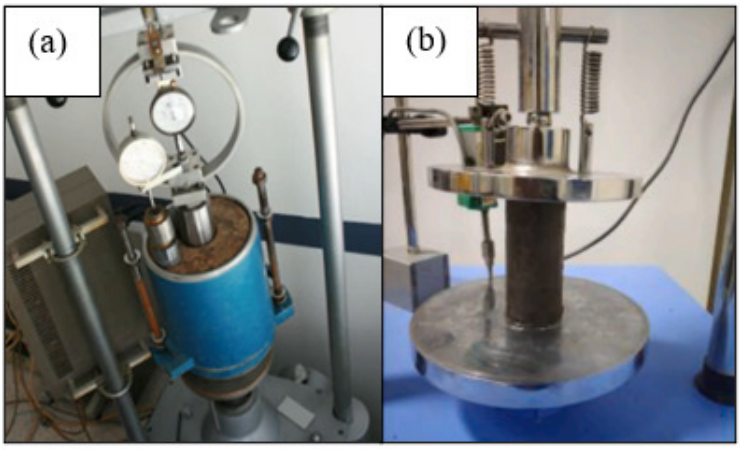


Figure 4. a) California bearing ratio and b) unconfined compressive strength testing

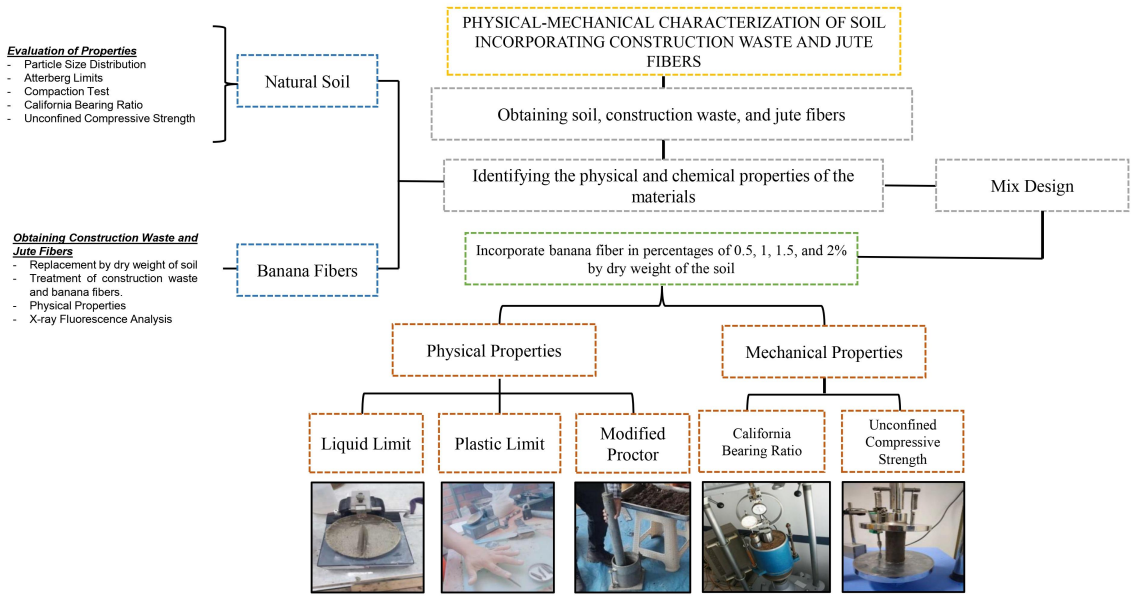


Figure 5. Research process

compaction, CBR, and UCS. Soil stabilization in this study was both mechanical and chemical. This combination was chosen due to its synergistic benefits, as mechanical stabilization with fibers enhances soil cohesion and internal strength, while chemical treatment optimizes the interaction between the fibers and the soil, improving the structure and load-bearing capacity of the material.

Phase II evaluated the physical characteristics of the BF, which were first cleaned and then treated with a 5% NaOH solution for 24 hours to enhance their mechanical and bonding properties. After treatment, the fibers were thoroughly rinsed with water, air-dried, and cut into lengths of approximately 40-50 mm for uniformity.

Phase III involved the preparation of soil-BF mixtures, where the fibers were randomly distributed into the dry sample in proportions of 0.50, 1.0, 1.5, and 2.0 % by weight of dry soil. Care was taken to ensure even mixing for homogeneity. The treated organic fiber was randomly placed in the natural soil due to its ease of application during the stabilization process.

Finally, phase IV focused on conducting physical and mechanical tests on the experimental soil-fiber mixtures to assess their suitability for light traffic subbase applications. Table **IV** presents the composition of these mixtures.

Table IV. Description of each mixture

Nomenclature	Description
D-0	Natural soil
D-05BF	Natural soil + 0.5 % BF
D-10BF	Natural soil + 1.0 % BF
D-15BF	Natural soil + 1.5 % BF
D-20BF	Natural soil + 2.0 % BF

3. Results and discussion

3.1. Physical properties

3.1.1. Liquid and plastic limits

According to Fig. **6**, and based on ASTM D4318 (31), the natural soil exhibited a LL of 20.3 % and a PL of 19.2 %. An analysis of the sample with 0.5 % BF revealed an increase of 12.32 % in the LL and 10.42 % in the PL compared to natural soil. With 1 % BF, the LL increased by 5.91 %, and the PL showed a slight increase (1.56 %) compared to the 0.5 % BF sample. The samples with 1.5 and 2 % BF showed LL reductions of 4.93 and 6.77 %, respectively, in comparison with natural soil, while the PL decreased by 6.40 and 9.38 %. These results suggest a reduction in plasticity, especially with the addition of 1.5 and 2 % BF.

Fig. **6** also suggests that the addition of low BF concentrations increases the LL and PL in soil samples, which can be attributed to the presence of K_2O and P_2O_5, which allow enhancing soil cohesion and water retention capacity. However, at higher concentrations (1.5 and 2 %), the reduction in the LL and PL indicate that an excess of these components may negatively affect plasticity.

These findings are similar to those of (19), who demonstrated that 0.4 % BF increased the LL by 25 %. Likewise, the PL substantially increased by 125.45 % compared to natural soil. However, (20) observed that 2 % fiber increased both the LL and PL by 21 %, which contrasts with our research findings. These discrepancies may be due to differences in the specific properties of the fibers used, the nature of the base soil, and the methodologies applied in the experiments.

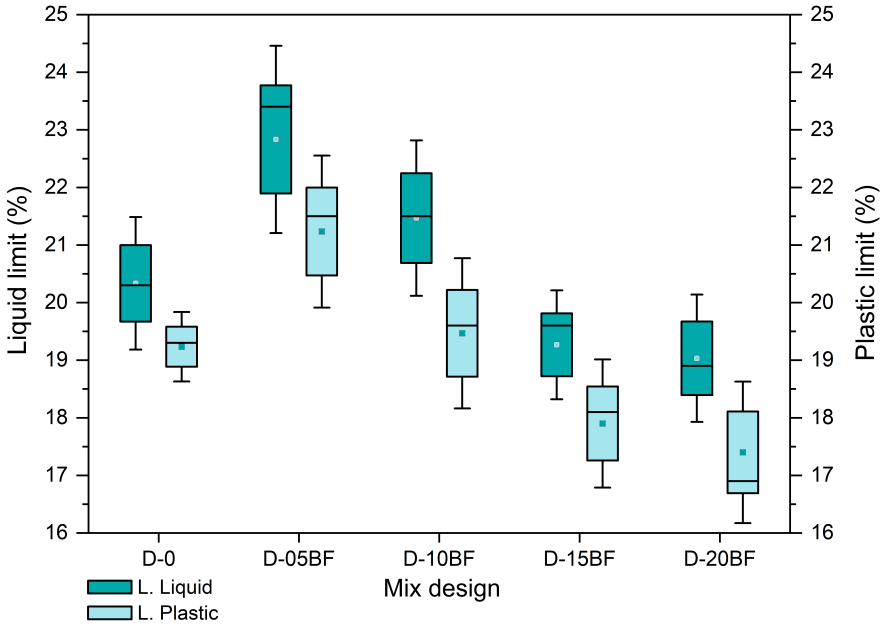


Figure 6. Variation in consistency limits for different banana fiber proportions

3.1.2. Modified Proctor test

According to the compaction test (Fig. 7), which was conducted according to ASTM D1557 (32), the studied natural soil has a MDD of 1.83 % and an OMC of 12.7 %. When adding 0.5 % BF, the MDD increases by 0.55 % and the OMC by 3.94 % compared to natural soil. Similarly, with 1 % BF, the MDD remains stable, but the OMC increases slightly (0.79). With 1.5 % BF, the MDD decreases by 0.55 % and the OMC by 7.87 %. Finally, with 2 % BF, the MDD decreases by 1.64 % and the OMC by 9.45 % compared to natural soil. These results indicate that the addition of BF has a variable effect on MDD and OMC, with increases at low BF doses and decreases at higher doses (Fig. 7). This could be related to the presence of Al_2O_3 (0.338 %) and SiO_2 (1.644 %) in the fibers, which may interfere with the uniform compaction of the soil by forming agglomerations, thereby reducing its density and water absorption capacity.

(19) found that the soil's MDD increased by 7.78 % with 2 % BF. In addition, the OMC decreased by 20 % with the addition of 0.4 % BF compared to natural soil. Similarly, (21) showed that the incorporation of 0.5 % NF improved the MDD by 1.58 %, while the OMC remained stable at 16 % in comparison with untreated soil. However, other studies reported that 2 and 3 % NF reduced the MDD and OMC by 2.01 and 55.84 %, respectively (23, 48). Higher concentrations of NF can introduce voids and destabilize the soil structure, decreasing its density and moisture retention capacity. The observed variations may be due to differences in the properties of the fibers, the nature of the soil, and the experimental methodologies used.

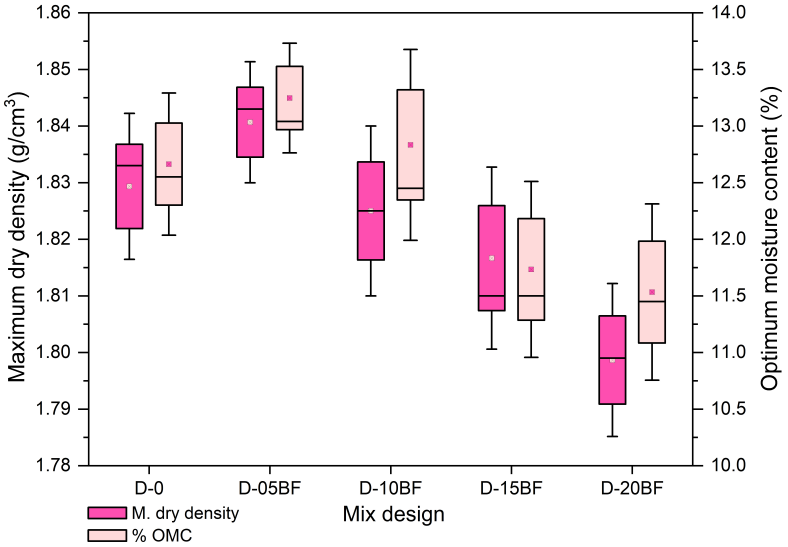


Figure 7. Variation in MDD and OMC when using different banana fiber proportions

3.2. Mechanical properties

3.2.1. California bearing ratio

As shown in Fig. 8, the CBR value at 100 % for natural soil is 33.6 %. Adding 0.5 and 1 % BF increases this value by 3.87 and 5.95 %, respectively. However, increasing the BF content to 1.5 % slightly decreases the CBR by 1.49 %. Similarly, with 2 % BF, the CBR decreases significantly (13.10 %) when compared to untreated soil. These results indicate that the addition of BF enhances the bearing capacity of the soil, albeit up to a certain point, after which its effectiveness decreases.

The CBR value presented in Fig. 8 improves with up to 1 % BF, suggesting that MnO and CaO (4.764 %) contribute to an increased resistance to applied loads. However, the decrease in CBR at higher concentrations may indicate that an excessive proportion of these compounds reduces soil cohesion, thereby affecting stability.

These findings are similar to those of (24), who found that 0.3 % BF increased the CBR by 23.71 % compared to the natural sample. On the other hand, with 0.4 % BF, the CBR of the soil increased by 21.55 % (19). Additionally, (28) demonstrated that the incorporation of 1 % BF improved the soil support by 32.6 %. Other studies have shown increases of 68 and 171 % with the same fiber proportions compared to untreated soil (29, 30). These results suggest that the addition of BF significantly improves the soil's bearing capacity by increasing its cohesion and internal strength, with more pronounced increases as the fiber proportion approaches optimal levels.

3.2.2. Unconfined compressive strength

As shown in Fig. 9, during UCS testing, optimal results were observed after 21 days with different doses of BF. The natural soil reported a UCS of 0.20 MPa, while adding 0.5 % BF increased this value by

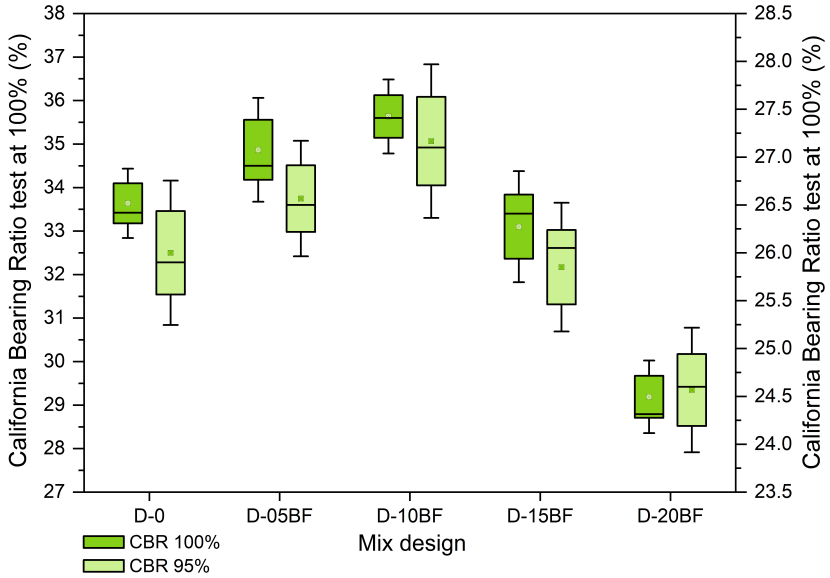


Figure 8. Variation in CBR for different banana fiber proportions

14.29 %. With an addition of 1 % BF, the strength further increased by 23.81 %. However, increasing BF content to 1.5 % returned the strength to 0.20 MPa, equal to that of natural soil. Finally, with 2 % BF, the strength decreased significantly (14.29 %). These results indicate that BF addition enhances soil strength up to a certain point, beyond which its effectiveness diminishes. Fig. **9** shows that the UCS increases significantly with up to 1 % BF due to the improved load transfer facilitated by the presence of SO_2 (7.806 %) and TiO_2 (5.028 %), which enhance rigidity and durability. However, at higher concentrations, the UCS decreases, likely due to the heterogeneous distribution of the fibers, which affects the structural integrity of the soil.

The incorporation of BF into the soil improves its physical and chemical adhesion, optimizing stress transfer and increasing its load-bearing capacity. The NaOH treatment enhances adhesion and strength, while the fibers distribute stresses uniformly, reducing localized failures and improving ductility. An optimal proportion of 1 % BF significantly improves both the CBR and the UCS, although higher proportions may negatively impact soil compaction and strength.

By comparing these results against previous research, it is observed that low to moderate doses of BF (*e.g.*, 0.5 and 1 %) significantly improve soil strength, which is consistent with the findings of (24), who found a 26.41 % increase in UCS with 0.3 % BF compared to natural soil. However, our results differ from those reported by (27), who observed a 100 % increase in strength with 1.5 % BF, contrasting with the strength stabilization observed in our study. Various studies have shown that organic materials significantly enhance soil properties (49–51), which makes them valuable for the civil construction industry (52–58). These discrepancies could be due to differences in the nature of the soil, the properties of the fibers used, and the experimental methodologies applied.

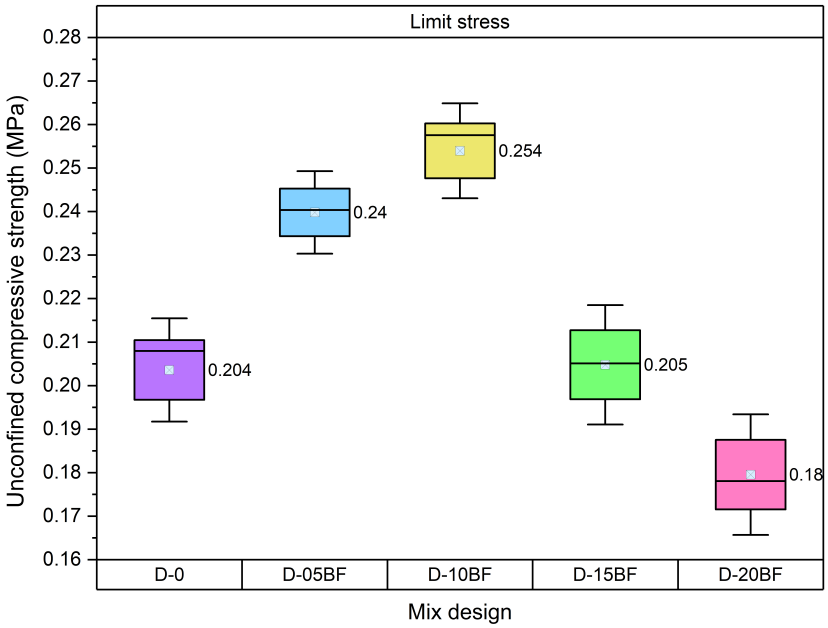


Figure 9. Variation in consistency limits for different banana fiber proportions

3.3. Statistical analysis

3.3.1. Consistency limit

An ANOVA analysis indicated significant LL differences between the mix designs (p=0.0089), unlike the PL (p=0.2075). The Tukey test showed that D-05BF and D-10BF have significantly higher values than D-15BF and D-20BF. This suggests that a moderate amount of BF improves the material's cohesion, but excessive amounts reduce its plasticity.

3.3.2. Maximum dry density and optimum moisture content

Both of these parameters are significant according to the ANOVA (p=0.0207 and p=0.0248, respectively). The Tukey test indicated that the MDD increases with small amounts of fiber (D-05BF and D-10BF) and decreases with higher additions (D-15BF and D-20BF). Similarly, the moisture content is higher in D-05BF and D-10BF but decreases in D-15BF and D-20BF, suggesting that fiber initially enhances water retention before losing its effect.

3.3.3. CBR at 100 and 95 %

The ANOVA showed significant differences for the soil tested with CBR at 100 and 95 % (p=0.0011 and p=0.0006). The Tukey test confirmed that D-05BF and D-10BF have higher CBR values than D-0, while D-15BF and D-20BF show a notable decrease. This indicates that fiber enhances structural strength up to a certain point, but excessive amounts negatively affect material compaction.

3.3.4. Unconfined compressive strength

This parameter is highly significant (p=0.0001). The Tukey test indicated that D-10BF has the highest strength, while D-20BF has the lowest. This suggests that a moderate fiber addition reinforces soil cohesion, but an excess reduces the material's load-bearing capacity.

In conclusion, adding BF improves mechanical properties by up to 10 %, but higher amounts result in losses regarding cohesion and compaction.

4. Conclusions

The incorporation of BF has a significant influence on the physical properties of the soil. The conclusions drawn from the findings are presented below.
The chemical characteristics of the natural material (BF) can vary due to the diversity of its sources and the treatment it undergoes before being incorporated as reinforcement into the soil. These variations can affect the quality of the soil, with a direct impact on the geotechnical characteristics of the experimental sample.

This analysis highlights that the chemical properties of BF, particularly the high levels of K_2O, P_2O_5, and MnO, play a crucial role in enhancing the geotechnical properties of the soil.

The optimal proportion of BF was 1.0 %, as it exhibited the most significant values in relation to the geotechnical properties of the soil. The MDD remained stable, the CBR increased by 5.95 %, and the UCS increased by 23.81 % compared to natural soil. These results emphasize that the inclusion of BF can strengthen the soil, making it more resistant and effective for construction applications, especially in rapidly growing urban areas like Peru. Additionally, the use of BF-treated soil meets the local standards for application in projects such as road construction and paving, thus ensuring proper regulatory compliance and contributing to the environmental sustainability of infrastructures. However, it is crucial to consider that higher doses of BF could have adverse effects on the mechanical behavior of the soil.

We suggest conducting additional microstructural tests to delve deeper into the specific interaction between BF and the soil, which will allow for a more thorough evaluation of the stability and durability of the modified material. These tests could reveal the way in which fibers influence the soil's microstructure, including the formation of particle bridges and fiber orientation under load. Evaluating how environmental conditions (*e.g.*, moisture and temperature) affect fiber adhesion and long-term behavior is crucial to ensuring the material's sustainability in real-world applications.

5. Acknowledgements

We are grateful to the Professional School of Civil Engineering of Universidad Señor de Sipán for providing us with the facilities to carry out this research.

6. Access to research data

Should additional access be required, data supporting the findings of this study can be obtained from the corresponding author.

7. Conflict of interest

The authors declare that they have no competing interests that could influence the outcomes and/or discussion of this work.

8. Author contributions

Favio Osmar Schreiber Robles: conceptualization, investigation, methodology, writing (original draft). **Sócrates Pedro Muñoz Pérez:** conceptualization, investigation, methodology, writing (original draft). **Juan Martín García Chumacero:** conceptualization, investigation, methodology. **Elver Sánchez Diaz:** investigation, writing (review and editing). **Carlos Arturo Damiani Lazo:** conceptualization, methodology, validation, formal analysis. **Juan De Dios Malpartida Iturregui** and **Ana Paula Bernal Izquierdo:** investigation, writing (review and editing). **Ángel Antonio Ruiz Pico:** validation, visualization, writing. **Edwin Adolfo Diaz Ortiz:** conceptualization, methodology, visualization. **Ernesto Dante Rodríguez Lafitte:** visualization, supervision, project administration.

References

[1] S. Hasnah *et al.*, "A review on natural fiber reinforced polymer composites (NFRPC) for sustainable industrial applications," *Polymers*, vol. 14, no. 17, art. 3698, 2022. https://doi.org/10.3390/polym14173698 ↑3

[2] N. Faizah, M. AlHamidi, F. Mansor, and S. Anuar, "Influence of banana fiber on shear strength of clay soil," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 864, no. 1, art. 012099, 2020. https://doi.org/10.1088/1757-899X/864/1/012099 ↑3, 4

[3] R. Kiani *et al.*, "Sustainability in construction projects: A systematic literature review," *Sustainability*, vol. 13, no. 4, art. 1932, 2021. https://doi.org/10.3390/su13041932 ↑3

[4] G. Alhakim, L. Jaber, O. Baalbaki, and F. Barraj, "Utilization of fan palm, date palm, and phragmites australis fibers for improving the mechanical behavior of sandy soil," Geomech. *Energy Environ.*, vol. 33, art. 100427, 2023. https://doi.org/10.1016/j.gete.2022.100427 ↑3

[5] A. Temitayo, O. Olugbemiga, O. Olufemi, and E. Temidayo, "Physico-chemical, thermal and micro-structural characterization of four common banana pseudo-stem fiber cultivars in Nigeria," *J. Nat. Fibers*, vol. 20, no. 1, art. 2167031, 2023a. https://doi.org/10.1080/15440478.2023.2167031 ↑3

[6] W. Qamar, A. Hassan, Z. Rehman, and Z. Masoud, "Sustainable application of wool- banana bio-composite waste material in geotechnical engineering for enhancement of elastoplastic strain

and resilience of subgrade expansive clays," *Sustain.*, vol. 14, no. 20, art. 13215, 2022. https://doi.org/10.3390/su142013215 ↑3

[7] A. Samir, F. Ashour, A. Abdel, and M. Bassyouni, "Recent advances in biodegradable polymers for sustainable applications," *npj Mater. Degrad.*, vol. 6, art. 68, 2022. https://doi.org/10.1038/s41529-022-00277-7 ↑3

[8] S. Fogue *et al.*, "Investigation of chemical, physical and morpho-mechanical properties of banana-plantain stalk fibers for ropes and woven fabrics used in composite and limited-lifespan geotextile," *Heliyon*, vol. 3, no. 8, art. e29656, 2024. https://doi.org/10.1016/j.heliyon.2024.e29656 ↑3

[9] V. Kumar *et al.*, "Potential of banana based cellulose materials for advanced applications: A review on properties and technical challenges," *Carbohydr. Polym. Technol. Appl.*, vol. 6, art. 100366, 2023. https://doi.org/10.1016/j.carpta.2023.100366 ↑3

[10] P. Badanayak, J. Seiko, and G. Bose, "Banana pseudostem fiber: A critical review on fiber extraction, characterization, and surface modification," *J. Nat. Fibers.*, vol. 20, no. 1, art. 2168821, 2023. https://doi.org/10.1080/15440478.2023.2168821 ↑3

[11] J. C. León Carrasco, "Producción de banano de Perú alcanzó las 2.414.382 toneladas en 2023," Agraria.pe, 2024. [Online]. Available: https://agraria.pe/noticias/produccion-de-banano-de-peru-alcanzo-las-2-414-382-toneladas-37112#:~:text=Producci%C3%B3n%20de%20banano%20de%20Per%C3%BA%20alcanz%C3%B3%20las%202.414.382%20toneladas%20en%202023,-Enviar%20Imprimir&text=La%20principal%20regi%C3%B3n%20productora%20de,y%20Hu%C3%A1nuco%20con%20233.775%20toneladas ↑3

[12] Minagri, "MINAGRI desarrollará tecnología que proteja al banano orgánico de letal hongo Fusarium Oxysporium," inia, 2020. [Online]. Available: https://www.inia.gob.pe/2020-nota-040/ ↑3

[13] D. Dutta and N. Sit, "Preparation and characterization of potato starch-based composite films reinforced by modified banana fibers and its application in packaging of grapes," *Int. J. Biol. Macromol.*, vol. 254, part 2, art. 127791, 2024. https://doi.org/10.1016/j.ijbiomac.2023.127791 ↑3

[14] A. Temitayo, O. Olugbemiga, O. Olufemi, E. Temidayo, and M. Hussain, "Banana pseudo stem fiber, hybrid composites and applications: A review," *Hybrid Adv.*, vol. 4, art. 100101, 2023b. https://doi.org/10.1016/j.hybadv.2023.100101 ↑3

[15] A. Karimah *et al.*, "A review on natural fibers for development of eco-friendly bio- composite: Characteristics, and utilizations.," *J. Mater. Res. Technol.*, vol. 13, pp. 2442- 2458, 2021. https://doi.org/10.1016/j.jmrt.2021.06.014 ↑3

[16] A. Vinod, S. Suchart, and P. Jyotishkumar, "Renewable and sustainable biobased materials: an assessment on biofibers, biofilms, biopolymers and biocomposites," *J. Clean. Prod.*, vol. 258, art. 120978, 2020. https://doi.org/10.1016/j.jclepro.2020.120978 ↑3

[17] S. Rangappa, S. Siengchin, and H. Dhakal, "Green-composites: Ecofriendly and Sustainability," *Appl. Sci. Eng. Prog.*, vol. 13, no. 3, art. 183-184, 2020. https://doi.org/10.14416/j.asep.2020.06.001 ↑3

[18] T. Nguyen and B. Indraratna, "Natural fibre for geotechnical applications: Concepts, achievements and challenges," *Sustain.*, vol. 15, no. 11, art. 8603, 2023. https://doi.org/10.3390/su15118603 ↑3

[19] P. Ellappan, V. Arumugam, and N. Muthukumaran, "Influence of natural fibres in strengthening of black cotton soil," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 955, art. 012066, 2020. https://doi.org/10.1088/1757-899X/955/1/012066 ↑3, 4, 9, 10, 11

[20] D. Lal, R. Pravalika, S. Kumar, and G. Rao, "Stabilization of expansive soil by using Jute fiber," *3rd Int. Congr. Adv. Mech. Sci.*, vol. 998, art. 012045, 2020. https://doi.org/10.1088/1757-899X/998/1/012045 ↑3, 9

[21] S. Prasanna and N. Macedon, "Application of jute fiber in soil stabilization," *Sustain. Agri. Food Environ. Res.*, vol. 11, no. X, pp. 1-7, 2023. https://doi.org/10.20944/preprints202008.0534.v1 ↑4, 10

[22] N. Bawadi, N. Ahmad, A. Mansor, S. Anuar, and M. Rahim, "Effect of natural fibers on the soil compaction characteristics," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 476, art. 012043, 2020. https://doi.org/10.1088/1755-1315/476/1/012043 ↑4

[23] L. Patil and S. Pusadkar, "MDD & OMC of black cotton soil reinforced with randomly distributed banana fibers," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 970, art. 012029, 2020. https://doi.org/10.1088/1757-899X/970/1/012029 ↑4, 10

[24] P. Somnath and S. Dipankar, "Performance evaluation of natural fiber reinforced Laterite soil for road pavement construction," *Mater. Today Proc.*, vol. 62, part 2, art. 1246- 1251, 2022. https://doi.org/10.1016/j.matpr.2022.04.534 ↑4, 11, 12

[25] S. Sridhar, H. Guruprasad, D. Naveenkumar, and D. Sangeetha, "Influence of treated coir fiber on durability properties of black cotton soil," *Mater. Today Proc.*, vol. 80, part. 2, pp. 1611-1616, 2023. https://doi.org/10.1016/j.matpr.2023.02.127 ↑4

[26] S. Kumar, A. Kumar, and S. Naval, "Influence of Jute fibre on CBR value of expansive soil," *Civil Eng. J.*, vol. 6, no. 6, pp. 1180-1194, 2020. https://doi.org/10.28991/cej-2020-03091539 ↑4

[27] M. Faisal, M. Khan, and K. Ali, "Effect of artificial and natural fibers on behavior of soil," *Mater. Today Proc.*, vol. 64, part 1, pp. 481-487, 2022. https://doi.org/10.1016/j.matpr.2022.04.954 ↑4, 12

[28] M. Wasim, J. Israr, K. Farooq, and H. Mujtaba, "Strength mechanism of a swelling soil improved with Jute fibers: A laboratory treatment," *Geotech. Geol. Eng.*, vol. 41, no. 7, pp. 1-14, 2023. https://doi.org/10.1007/s10706-023-02517-2 ↑4, 11

[29] *Standard test method for particle-size analysis of soils*, ASTM D422-63, ASTM International, 2022. ↑5, 11

[30] *Standard practice for classification of soils for engineering purposes (unified soil classification system)*, ASTM D2487-17, ASTM International, 2020. ↑5, 11

[31] *Standard test methods for liquid limit, plastic limit, and plasticity index of soils*, ASTM D4318, ASTM International, 2010. ↑5, 9

[32] *Standard test methods for laboratory compaction characteristics of soil using modified effort (56,000 ft-lbf/ft3 (2,700 kN-m/m3))*, ASTM D1557-12, ASTM International, 2021. ↑5, 10

[33] Método de prueba estándar para la relación de carga de California (CBR) de suelos compactados en laboratorio, ASTM D1883-14, ASTM International, 2022. ↑5, 7

[34] Standard test method for unconfined compressive strength of cohesive soil, ASTM D2166-06, ASTM International, 2010. ↑5, 7

[35] C. Contreras, J. Albuja-Sánchez, O. Proaño, C. Ávila, and A. Damián-Chalán, "The influence of abaca fiber treated with sodium hydroxide on the deformation coefficients Cc, Cs, and Cv of Organic Soils," *Fibers*, vol. 12, no. 10, art. 89, 2024. https://doi.org/10.3390/fib12100089 ↑5

[36] S. Ramírez, "Influencia de la concentraión de NaOH en la resistencia mecánica a la tensión de un material compuesto reforzado con fibra de bambú," Doctoral Thesis, Universidad ECCI, Bogotá, Colombia, 2017. [Online]. Available: https://repositorio.ecci.edu.co/entities/publication/fbcd1dba-7d84-44ce-89c3-fee611d744a7 ↑5

[37] Vivek, R. K . Dutta, and R. Parti, "Effect of chemical treatment on the tensile strength behavior of coir geotextiles," *J. Nat. Fibers.*, vol. 17, no. 4, pp. 542-556, 2020. https://doi.org/10.1080/15440478.2018.1503132 ↑5

[38] J. Nisar, M. Mir, and Vivek, "Study on optimal preparation and rheological characteristics of waste low density polyethylene (LDPE)/styrene butadiene styrene (SBS) composite modified asphalt binder," *Constr. Build. Mater.*, vol. 407, art. 133459, 2023. https://doi.org/10.1016/j.conbuildmat.2023.133459 ↑5

[39] Vivek, R. K. Dutta, and R. Parti, "Application potential of treated coir geotextiles in unpaved roads," *J. Nat. Fibers.*, vol. 17, no. 10, pp. 1454-1467, 2020. https://doi.org/10.1080/15440478.2019.1578718 ↑5

[40] Vivek, R. Dutta, and R. Parti, "Effect of chemical treatment of the coir geotextiles on the interface properties of sand–/clay–coir geotextile interface," *J J. Inst. Eng. (India) Series A*, vol. 100, pp. 357-365, 2019. https://doi.org/10.1007/s40030-018-0348-x ↑5

[41] Vivek and R. Dutta, "Bearing ratio behavior of sand overlying clay with treated coir geotextiles at the interface," *J. Nat. Fibers.*, vol. 19, no. 14, pp. 7534-7541, 2021. https://doi.org/10.1080/15440478.2021.1952135 ↑5

[42] P. Jaswall, Vivek and S. Sinha, "Investigation on tensile strength characterisation of untreated and surface treated coir geotextiles," *J. Ind. Text.*, vol. 52, 2022. https://doi.org/10.1177/15280837221118847 ↑5

[43] P. Jaswal and Vivek, "Experimental study on monotonic behaviour of two-layered unpaved road model reinforced with treated coir geotextiles," *Int. J. Pavement Res. Technol.*, vol. 17, pp. 1377-1388, 2023. https://doi.org/10.1007/s42947-023-00293-z ↑5

[44] P. Jaswal, Vivek, and S. Sinha, "Improvement in the performance of two layered model pavement with treated coir geotextile at the interface," *J. Ind. Text.*, vol. 52, 2022. https://doi.org/10.1177/15280837221114161 ↑5

[45] *Standard guide for x-ray emission spectrometric analysis*, ASTM E1621-94(1999), ASTM International, 2017. ↑5

[46] *Standard test method for density of hydraulic*, ASTM C188, ASTM International, 2017. ↑6

[47] *Standard test method for tensile properties of single textile fibers*, ASTM D3822, ASTM International, 2020. ↑6

[48] V. Sagar and T. Kaki, "Effects of coir fibre on the strength of black cotton soil," *Int. J. Eng. Res. Mech. Civil Eng. (IJERMCE)*, vol. 10, no. 6, pp. 7-10, 2023. https://doi.org/10.1016/j.matpr.2023.02.127 ↑10

[49] J. García Chumacero, P. Acevedo Torres, C. Corcuera La Portilla, and S. Muñoz Pérez, "Influence of artificial intelligence on the optimization of the dosage of natural hydraulic lime, plastic and metallic fibers on the geological characteristics of a treated soil," *Rev. Ing. Construcc.*, vol. 38, no. 3, pp. 473-484, 2023. https://doi.org/10.7764/RIC.00080.21 ↑12

[50] C. J. M. García, P. L. Acevedo Torres, C. C. Corcuera la Portilla, S. P. Muñoz Pérez, and L. I. Villena Zapata, "Effect of the reuse of plastic and metallic fibers on the characteristics of a gravelly soil with clays stabilized with natural hydraulic lime," *Innov. Infrastruct. Solut.*, vol. 8, art. 185, 2023. https://doi.org/10.1007/s41062-023-01155-0 ↑12

[51] J. M. García Chumacero, J. L. Gonzales Macedo, and D. J. Sánchez Castillo, "Contribution of agricultural ashes and HDPE as a waste material for a sustainable environment applied to the stabilization of a low plasticity clay soil," *Innov. Infrastruct. Solut.*, vol. 9, art. 67, 2024. https://doi.org/10.1007/s41062-024-01372-1 ↑12

[52] J. García, G. Arriola, L. Villena, and S. Muñoz, "Strength of concrete using partial addition of residual wood ash with respect to cement," *Rev. Politéc.*, vol. 52, no. 1, pp. 45- 54, 2023. https://doi.org/10.33333/rp.vol52n1.05 ↑12

[53] S. Muñoz, L. Villena, F. Tesen, Y. Coronel, J. García, and C. Brast, "Influence of coconut fiber on mortar properties in masonry walls," *Electron. J. Struct. Eng.*, vol. 23, no. 4, pp. 52-58, 2023. https://doi.org/10.56748/ejse.23391 ↑12

[54] J. Bereche and J. García, "Replacement of fine aggregate with refractory brick residue in concrete exposed to elevated temperatures," *Rev. Politéc.*, vol. 53, no. 2, pp. 79- 88, 2024. https://doi.org/10.33333/rp.vol53n2.08 ↑12

[55] C. J. Burgos, B. E. Cubas, and C. J. García, "Analysis of the combination of glass and polypropylene fibers on the mechanical properties of mortar," *J. Build. Pathol. Rehabil.*, vol. 10, art. 22, 2024. https://doi.org/10.1007/s41024-024-00532-1 ↑12

[56] J. Chaname, J. García, and G. Arriola, "Mejoramiento de las propiedades mecánicas del concreto estructural utilizando microporoso etileno acetato de vinilo," *Rev. Politéc.*, vol. 53, no. 2, pp. 17-26, 2024. https://doi.org/10.33333/rp.vol53n2.02 ↑12

[57] G. Cabanillas Hernández, J. García Chumacero, L. Villegas Granados, G. Arriola Carrasco, and N. Marín Bardales, "Sustainable use of wood sawdust as a replacement for fine aggregate to improve the properties of concrete: A Peruvian case study," *Innov. Infrastruct. Solut.*, vol. 9, no. 7, art. 233, 2024. https://doi.org/10.1007/s41062-024-01567-6 ↑12

[58] A. L. Y. Cadenas, H. R. M. Jacinto, C. J. M. García, V. N. R. Salinas, and C. C. O. Chávez, "Experimental study of hybrid concrete blocks based on sawdust and calcined clay using cement as a binder," *Innov. Infrastruct. Solut.*, vol. 9, art. 445, 2024. https://doi.org/10.1007/s41062-024-01768-z ↑12

Favio Osmar Schreiber Robles
Bachelor of Civil Engineering from Universidad Señor de Sipán. He has significant experience in state projects in the Lambayeque region. He currently resides in Chiclayo, Peru.
Email: sroblesf@uss.edu.pe

Sócrates Pedro Muñoz Pérez
Civil Engineer graduated from Universidad Nacional Pedro Ruiz Gallo. He holds a Master's degree in Earth Sciences with an emphasis in Geotechnics from Universidad Nacional de San Agustín, a PhD in Civil Engineering from Universidad Nacional del Santa, and a PhD in Public Management and Governance from Universidad César Vallejo, as well as a Diploma in Applied Geology in Mining from the Chamber of Mines of Peru.
Email: socrates.munoz@untrm.edu.pe

Juan Martín García Chumacero
Civil Engineer graduated from Universidad Señor de Sipán. He works as a consultant and researcher in various areas of the civil engineering sector. He has significantly contributed to the drafting of research reports, scientific articles, and systematic reviews, and has also served as a reviewer for indexed journals. His experience also includes supervising educational infrastructure projects and real estate projects at the national level.
Email: gchumacerojuanm@uss.edu.pe

Elver Sánchez Díaz
Professional School of Civil Engineering, Faculty of Engineering, Architecture and Urbanism, Universidad Señor de Sipán. He holds a Master's degree in Business Administration (MBA) from Universidad Privada César Vallejo, Peru. Additionally, he is a Bachelor of Environmental Civil Engineering from Universidad Católica Santo Toribio de Mogrovejo, Peru. Currently, he is dedicated to teaching in Chiclayo, Peru. His research interests focus on civil and environmental engineering.
Email: sanchezdelv@uss.edu.pe

Carlos Arturo Damiani Lazo
Civil Engineer with an MSc in Project Formulation, Management, and Evaluation, as well as a PhD in Education. He served as the Director of the Professional School of Civil Engineering of Universidad Alas Peruanas and has extensive teaching and professional experience in the engineering sector, excelling as a consultant and supervisor of works in various regions of Peru.
Email: cdamiani@unsa.edu.pe

Juan de Dios Malpartida Iturregui
Civil Engineer graduated from Universidad Señor de Sipán. His experience includes supervising educational infrastructure projects and real estate projects at the national level. He resides in Chiclayo, Peru.
Email: miturreguijuan@uss.edu.pe

Ángel Antonio Ruíz Pico
A PhD in Civil Engineering, he is a distinguished professional with experience in groundwater hydrology and water management. He is the founder of the INIAI Research Group and the president of the Ángel Antonio Ruiz Foundation. He directed the Science and Technology Park of Universidad Señor de Sipán in Peru and has been a professor and researcher at various universities. He is the author of numerous scientific articles and holds patents

for water source devices and construction materials.
Email: aaruizpico@gmail.com

Edwin Adolfo Díaz Ortiz

Civil Engineer specializing in environmental management. He holds a PhD in Science and Engineering. He has experience in developing environmental management instruments and has served in various engineering projects as a consultant and supervisor of works. He is a university professor in the fields of soil mechanics, roadworks, works supervision, construction quality, civil project management and administration, and scientific research projects
Email: edwin.diaz@untrm.edu.pe

Ernesto Dante Rodríguez Lafitte

Agronomist graduated from Universidad Nacional de Cajamarca, Bachelor of Administration, with a Master's degree in University Teaching and an Master's degree in Business Administration and International Relations from Universidad César Vallejo, where he also obtained a PhD in Public Management and Governance. He has held academic positions such as Academic Vice-Principal, Dean, and School Director at Universidad Señor de Sipán.
Email: rlafitte@uss.edu.pe

Ana Paula Bernal Izquierdo

Bachelor of Civil Engineering from Universidad Señor de Sipán. She has significant experience in state projects in the Lambayeque region. She currently resides in Chiclayo, Peru.
Email: bizquierdoanapa@uss.edu.pe

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.21930

Research

Deep Learning and Time Series for the Prediction of Monthly Precipitation. A Case Study in the Department of Boyacá, Colombia

Aprendizaje profundo y series temporales para la Predicción de la Precipitación Mensual. Estudio de caso: Departamento de Boyacá-Colombia

Yesid Esteban Duarte[1], Marco Javier Suárez Barón[1]*, Oscar Javier García Melagrejo[2], César Augusto Jaramillo Acevedo [2], and Carlos Augusto Meneses Escobar[2]

[1]Department of Systems and Computing, Universidad Pedagógica y Tecnológica de Colombia (Sogamoso, Colombia)
[2]Department of Geology, Universidad Pedagógica y Tecnológica de Colombia (Sogamoso, Colombia)
[3]Department of Systems and Computing, Universidad Tecnológica de Pereira (Pereira, Colombia)

Abstract

Context: This paper proposes a method for the prediction of monthly precipitation in the department of Boyacá using models based on deep neural networks (DNNs). These approaches have achieved significant improvements in prediction accuracy when compared to traditional methods.
Method: Data with a spatial resolution of 0.5° were extracted from CHIRPS 2.0 and subsequently preprocessed for the implementation of two approaches based on recurrent neural networks (RNNs) with long short-term memory (LSTM) and ConvLSTM architectures, aiming to provide accurate predictions of monthly precipitation in the studied region.
Objectives: The goal of this time series analysis is to predict monthly precipitation and develop accurate models that can forecast future rainfall patterns based on historical data. This aids in water resource management and agricultural planning, as well as in mitigating the impacts of droughts or floods.
Results: According to the results obtained, the LSTM model stands out for its robustness in terms of performance metrics, such as a lower mean squared error, a lower root mean squared error, and a coefficient of determination closer to 1. This demonstrates its higher accuracy compared to the ConvLSTM model.
Conclusions: Deep learning models, especially RNNs with LSTM, are effective tools for predicting crucial climate data.
Keywords: deep learning, neural networks, LSTM, ConvLSTM, time series

Acknowledgments: The authors would like to thank the Research Vice-Principalship of Universidad Pedagógica y Tecnológica de Colombia for funding project SGI 3535.

Article history

Received: July 11th, 2023
Modified: October 16th, 2024
Accepted: November 5th, 2024

Ing, vol. 30, no. 1, 2025, e21930

* **Correspondence:** marco.suarez@uptc.edu.co

Resumen

Contexto: Este trabajo propone un método para la predicción de la precipitación mensual en el departamento de Boyacá utilizando modelos basados en redes neuronales profundas (DNNs). Estos enfoques han logrado mejoras significativas en la precisión de la predicción en comparación con los métodos tradicionales.
Método: Datos con una resolución espacial de 0.5° fueron extraídos de CHIRPS 2.0 y posteriormente preprocesados para la implementación de aproximaciones basadas en redes neuronales recurrentes (RNNs) con memoria de largo plazo (LSTM) y ConvLSTM para proporcionar predicciones precisas de precipitación. mensual en la región.
Objetivos: El objetivo de este análisis de series temporales es predecir la precipitación mensual y desarrollar modelos precisos que pronostiquen los patrones de precipitación futuros basándose en datos históricos. Esto contribuye a la gestión de los recursos hídricos, la planificación agrícola y la mitigación de los efectos de sequías o inundaciones.
Resultados: Los resultados obtenidos muestran que el modelo LSTM se destaca por su robustez en términos de métricas de rendimiento, tales como un menor error cuadrático medio, una menor raíz del error cuadrático medio y un coeficiente de determinación más cercano a 1. Esto demuestra su mayor precisión en comparación con el modelo ConvLSTM.
Conclusiones: Los modelos de aprendizaje profundo, especialmente los modelos RNN con LSTM, son herramientas eficaces para predecir datos climáticos cruciales.
Palabras clave: aprendizaje profundo, redes neuronales, LSTM, ConvLSTM, series temporales

Agradecimientos: Los autores agradecen al Vicerrectorado de Investigación de la Universidad Tecnológica de Pereira por la financiación del proyecto SGI 3535.

Table of contents

1. Introduction

Precipitation is an essential factor that influences various aspects of life, from water resource management to decision-making in multiple sectors such as agriculture (1). In this context, understanding and analyzing precipitation patterns over time can answer crucial questions: What are the wettest and driest months in a given region? How is precipitation distributed in different geographical areas? What are the climatic trends of a given region? (2).

The department of Boyacá, Colombia, is characterized by its climatic variability and importance in agriculture and water resource management (3). The lack of accurate precipitation forecasts can have a significant impact on decision-making in sectors such as agriculture, aqueduct management, and water supply planning (4). These sectors could benefit from decision-making based on accurate data and reliable precipitation predictions, which in turn can have a positive impact on economic and environmental sectors.

The central problem addressed in this document is not only limited to the need to accurately predict monthly precipitation in the department of Boyacá; it also involves every aspect of the process, from collecting data to selecting the methods, algorithms, and metrics used to make predictions.

Regarding data collection, historical precipitation data were extracted from the CHIRPS 2.0 dataset, which has a spatial resolution of 0.5°. The selection of this data source was based on the quality and availability of the information, which is essential to ensuring accurate predictions.

Regarding the methods and algorithms used, we decided to implement LSTM and ConvLSTM neural networks. This choice was based on the nature of the collected monthly precipitation data, which exhibit temporal and spatial patterns (5). The use of these specific methods allowed capturing these patterns effectively, which was essential in generating accurate forecasts.

Finally, the selection of suitable evaluation metrics (MSE, RMSE, R^2) played an essential role in the project. These metrics were chosen for the purpose of objectively measuring the accuracy and predictive capacity of the models, providing a solid basis for evaluating their performance.

The contribution of this work lies in the application of models based on LSTM and ConvLSTM neural networks in the prediction of monthly precipitation in the department of Boyacá. To this effect, the performance of said models was evaluated in terms of precision metrics and predictive capacity, providing an interactive visual representation of historical data and predicted values through

a visualization scheme, thereby contributing to decision-making oriented to sectors such as agriculture and water resources management in the department.

This article is structured as follows. It begins with the *Background* section, which contextualizes the issue of precipitation forecasts in Boyacá. Then, the *Model development* section details the models used. This is followed by a description of the *Evaluation metrics* that were used to evaluate their performance. Next, in the *Results and discussion* section, the results obtained by the models are presented. Then, the Data visualization section introduces an interactive Power BI dashboard to visualize historical data and predictions. In addition, two specific test cases are examined. Finally, the *Conclusions and future work* section summarizes the main conclusions and proposes future lines of research in the field of climate forecasting in Boyacá.

2. Background

Monthly precipitation predictions are crucial in various applications such as water management, agricultural planning, and natural disaster forecasting (6). In recent years, deep learning has proven its worth in predicting climatic events (7), including precipitation. This section analyzes the most relevant literature on the use of deep learning in this field.

Convolutional neural networks (CNNs) have been used to analyze climatic maps and extract relevant spatial characteristics to predict monthly precipitation (8). Furthermore, CNNs can capture complex patterns in climate data by applying convolution operations (9). Using temperature, atmospheric pressure, and humidity data, CNNs can identify significant spatial correlations that influence precipitation. A study by (10) applied a CNN architecture to predict precipitation in China, exhibiting high precision. The results showed a significant improvement in predictive ability compared to traditional methods.

Recurrent neural networks (RNNs) have also been shown to be effective in time series prediction, including the specific case of monthly precipitation. According to (11), RNNs can model temporal dependencies through recurring connections, which allows them to capture patterns over time in climate data.

The use of deep learning through approaches such as CNNs, RNNs, and GAN has proven its effectiveness in predicting monthly precipitation (12). These approaches have achieved significant improvements in prediction accuracy when compared to traditional methods. However, there are still challenges that need to be addressed, such as the interpretation of deep learning models and the uncertainty associated with the predictions.

In summary, the use of deep learning in the prediction of monthly precipitation has shown promise. Approaches based on CNNs, RNNs, and GANs have achieved significant improvements in prediction accuracy and in the generation of realistic climate data. However, more research is needed

to improve the interpretation of the models and address the challenges associated with uncertainty and generalizability. Deep learning has the potential to revolutionize monthly precipitation prediction and provide valuable information for decision-making in various areas of application.

3. Materials and methods

This section describes the methodology for monthly precipitation analysis implemented in our research. In addition, the materials, methods, and tools applied are presented.

3.1. Case study

This project was carried out in the department of Boyacá, which is located in the center of Colombia, specifically in the Andean region. Boyacá is known for its agricultural activity, which highlights the importance of precipitation forecasting in decision-making concerning this sector as well as water resources management. Fig. 1 shows the geographical location of Boyacá in Colombia.

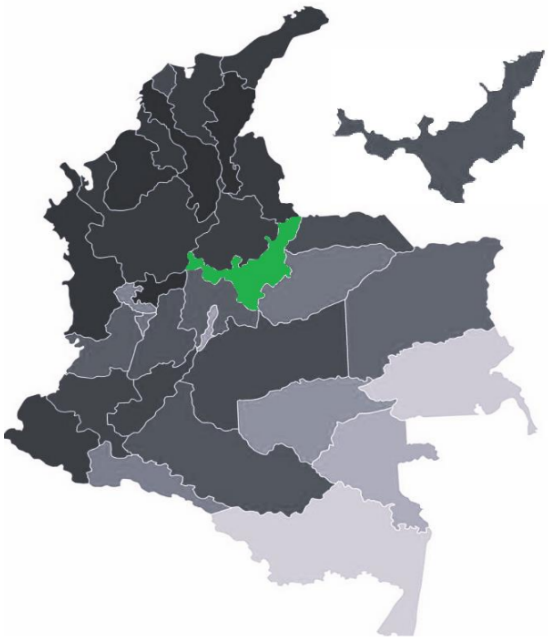


Figure 1. Geographic location of the department of Boyacá

3.2. Information flow

Our work began with the extraction of climate files from CHIRPS 2.0, covering precipitation records from 1981 to May 2023, with a spatial resolution of 0.05°. Data preparation included monthly consolidation and outlier retention. Fig. 2 shows the workflow applied in creating the prediction models.

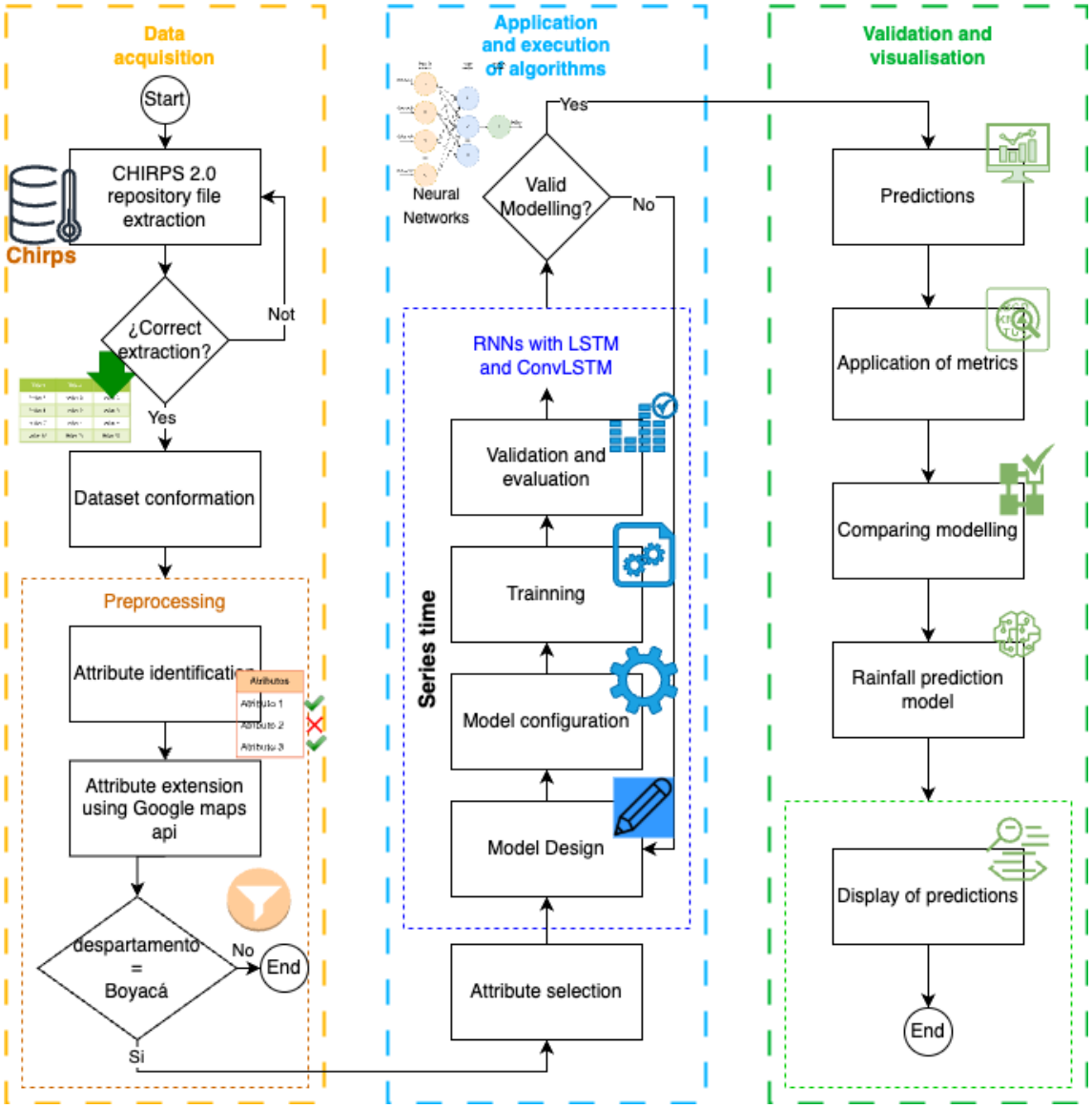


Figure 2. Workflow of the model

Afterwards, the databases were merged, and geographic filtering was performed to obtain the records for Boyacá. It is worth highlighting that, in attribute selection, the latitude, longitude, and date (year and month) are fundamental to considering local and temporal variations in precipitation forecasts.

Finally, regression models based on RNNs with LSTM and ConvLSTM layers were applied to predict monthly precipitation. These models were trained and validated using metrics such as the mean squared error (MSE) and the coefficient of determination (R^2).

As previously mentioned, Fig. 2 shows the information flow used, which is divided into three main phases:

- Data acquisition
- Application and execution of algorithms
- Validation and visualization

3.2.1. Data acquisition

For this task, files were downloaded from a web repository that contains links to climate files in netCDF format (.nc). Libraries such as *requests* and *BeautifulSoup* were used to request the web page and analyze its content. Then, each file identified in the links was downloaded and stored.

This work used precipitation data from CHIRPS 2.0 (*i.e.*, Climate Hazards Group InfraRed Precipitation with Station data, version 2.0) (13), a recognized source in global climate research. The data, with a spatial resolution of 0.05°, span from 1981 to May 2023. These data were preprocessed by consolidating precipitation values on a monthly basis through the sum of daily data, which allowed for a more relevant analysis of climatic trends over the course of the months. To this effect, daily data from 1981 were downloaded and combined into netCDF files (.nc) in order to facilitate processing.

3.2.2. Dataset conformation

Daily precipitation data from various CHIRPS 2.0 sources were consolidated and organized to establish a consistent data structure. This integration enabled a complete view of the precipitation records over time in the department of Boyacá, improving the understanding of its climatic patterns. The .nc files were loaded with precipitation data from 1981 to May 2023, the necessary fields were selected, and they were combined into a single dataset.

3.2.3. Preprocessing

The geospatial data contained in the netCDF files were processed. The workflow began by uploading the files downloaded from CHIRPS 2.0. Then, through the Google Maps API, the municipality and the department were obtained from the geographical coordinates, which provided additional information on the geographical points of interest for later filtering, leaving only the data for the department of Boyacá and finally storing them in a new netCDF file.

It is important to note that, at this stage of preprocessing, no data cleansing had been performed which involved the elimination of outliers or other irregularities. This decision was based on the consideration that outliers may be inherently significant to the climatic and geographic context of Boyacá. For example, they may indicate particular conditions in *páramo* areas or other specific weather events. Maintaining these values in the dataset contributes to a more accurate and authentic representation of the actual climatic conditions in the region.

3.2.3.1 Geographic filtering

To exclusively limit the analysis to the department of Boyacá, a rigorous geographic filtering process was implemented. This process was carried out using the Google Maps API and the *Geopandas* library, which made it possible to obtain the geographical coordinates of specific points in the department. Through this method, it was possible to identify and retain only the records related to the region of interest, excluding data from other geographic locations.

3.2.4. Feature selection

In analyzing information and making predictions, the characterization and description of the different characteristics that make up the data play a fundamental role. Each attribute not only represents an inherent feature of the information but also contributes significantly to the conclusions and predictions derived from the analysis. In the context of monthly precipitation prediction, the accurate identification and characterization of properties is essential to understanding the underlying causes that influence precipitation patterns, ultimately improving the accuracy of predictive models.

Table **I** describes the variables used in constructing the models for precipitation prediction. Each variable is detailed in terms of its data and variable type, and a brief description of its relevance in the climatic context is provided. These variables play a key role in the forecasting process and provide a better understanding of local and seasonal weather patterns.

Table I. Variables used in the construction of the models

Variables used in the construction of the models		
Variable	**Data type**	**Description**
Latitude	Numerical	North/South coordinate
Length	Numerical	East/West coordinate
Time	Date	Represents the date of the measurement record.
Precipitation	Numerical	Amount of precipitation in millimeters
Municipality	Categorical	Name of the geographic location where the data were recorded

Each of the identified attributes has a unique role in the construction and validation of predictive models. Their selection and characterization allow understanding the complex relationships between different spatiotemporal factors and how they influence precipitation patterns. These attributes are fully considered in the creation of models that can capture variability and trends in precipitation data, ultimately leading to more accurate and informed predictions.

3.2.5. Feature scaling

The feature scaling technique was applied to ensure that all features were on the same scale before entering them into the machine learning models. The *StandardScaler* method was used, which transforms each feature in such a way that it has a mean of 0 and a standard deviation of 1, thus achieving a uniform

scale and an adequate distribution. Table II presents a detailed description of this technique, its purpose, its benefits, and its importance in the analysis of climate data.

Table II. Feature scaling technique

Feature	Description
Scaling technique	StandardScaler
Description of the technique	It transforms each feature so that it has a mean of 0 and a standard deviation of 1.
Purpose	To make sure that all the features are on the same scale before entering them into the models.
Benefits	It prevents a given feature from being disproportionately weighted due to its magnitude, ensuring that all features are regarded as equally important in the analysis.
Importance	It ensures that machine learning models are able to capture the relationships between variables effectively and that differences in scales do not negatively affect the quality and accuracy of predictions.

3.3. Regression model applied to time series

In this section, the regression methodology used for climate analysis is described, which uses RNNs with LSTM and ConvLSTM layers. These architectures were specifically adapted to predict precipitation in the department of Boyacá and are widely used in sequential and spatial data processing.

3.3.1. Model architecture

Fig. 3 provides an overview of the architecture used in the models based on RNNs with LSTM (long short-term memory) and ConvLSTM layers, which were designed to address a regression problem related to precipitation prediction. This architecture was configured to model and anticipate monthly precipitation data in a geospatial context.

This graphical representation clearly illustrates the main layers and components of the neural network, including the LSTM input layer, the dense output layer, and the way in which they are connected to each other. Through this graphical representation, the flow of data is shown as it is processed and propagated throughout the network.

3.3.2. RNNs with LSTM

In implementing the model based on RNNs with LSTM layers, the main objective was to capture temporal patterns in monthly precipitation data for the department of Boyacá. Here is a detailed description of the architecture and design decisions of the model:

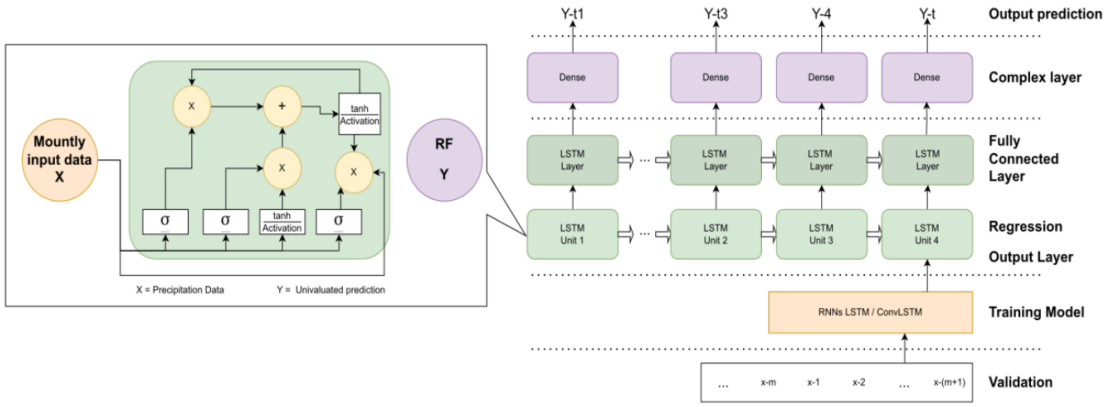


Figure 3. Architecture of the proposed model

⟹ *LSTM layer with 64 units:* The selection of an LSTM layer with 64 units was based on the complexity of the studied problem. In forecasting precipitation, it is crucial to capture the complex temporal relationships present in the data. By using 64 units in this layer, the model is allowed to effectively learn these relationships. However, the risk of overfitting should be taken into account, so this number of units is considered to be an appropriate balance.

⟹ *ReLU (rectified linear unit) activation function:* A ReLU activation function was applied to this LSTM layer since it introduces nonlinearities into the model and allows the network to capture temporal and nonlinear relationships efficiently. This is essential for understanding and predicting complex weather patterns.

⟹ *Dense output layer:* A dense layer was used with a single unit in the output, since the main task is a regression that seeks to predict a continuous value, *i.e.*, the precipitation level. A single unit in the output layer is sufficient to generate this prediction.

⟹ *Adam optimizer and MSE loss function:* To adjust the model weights during training, the Adam optimizer was selected. Adam is an efficient and effective choice that is well-suited for a variety of machine learning problems. As for the loss function, the MSE was chosen.

Instead of a time window, each data point was considered independently, and predictions were made for the next 12 months starting in June 2023. The dataset was divided into 80 % for training and 20 % for testing in order to avoid overfitting.

3.3.3. ConvLSTM neural network

A ConvLSTM neural network model was designed with the objective of capturing spatial-temporal patterns in precipitation data partitioned into 80 % for training and 20 % for testing. The specific design decisions are detailed below:

⟹ *ConvLSTM layer with 64 filters and a kernel of size (1, 3):* This choice was based on the need to incorporate both spatial and temporal information in the modeling process. The decision to use 64 filters allowed the model to detect a variety of spatial characteristics and patterns in the data,

while the kernel of size (1, 3) was used to apply convolutions on the time axis (months), thus capturing temporal patterns in three-month windows.

$\Longrightarrow$ *Flatten layer:* After the ConvLSTM layer, a flatten layer was applied to transform the data into a format suitable for the dense layer. This was necessary because the ConvLSTM layer produces outputs in 2D format, while the dense layer requires inputs in 1D format to make predictions.

$\Longrightarrow$ *Dense layer and optimization:* Similar to the approach used in the LSTM model, a dense layer with a single unit was incorporated to make the final prediction. The Adam optimizer and the MSE loss function were also selected since they have proven to be effective in model compilation and parameter adjustment.

3.4. Evaluation metrics

Once the preprocessing of the data and the construction of the forecasting models had been completed, the models were validated and evaluated to ensure that they were capable of providing valid monthly precipitation predictions for the department of Boyacá. To this effect, metrics and visualization techniques were used, aiming to evaluate the performance and quality of the models.

3.4.1. MSE

This metric has become a fundamental tool in evaluating the precision of precipitation prediction models. Eq. (1) allows measuring how close the model predictions are to the actual precipitation values in a dataset.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \tag{1}$$

where

$\rightarrow$ MSE represents the mean squared error, a numerical value that indicates how much the predictions of the model vary with respect to the actual precipitation data.

$\rightarrow$ n represents the total number of observations in the dataset. These observations correspond to precipitation measurements from different locations and times.

$\rightarrow$ i represents the real precipitation value of observation i.

$\rightarrow$ $\hat{y}_i$ denotes the predicted precipitation level that the model has generated for observation i.

The MSE provides a quantitative measure of how accurate the precipitation prediction model is. The lower the MSE value, the closer the model predictions to the actual precipitation data. This means that minimizing the MSE would improve the model's ability to accurately predict the precipitation level at different locations and times.

3.4.2. Root mean squared error (RMSE)

This is a metric that is derived directly from the MSE and is used to evaluate the quality of model predictions. It provides a more interpretable measure of the magnitude of the average error between the predictions and the actual values.

The RMSE formula is as follows:

$$RMSE = \sqrt{MSE} \tag{2}$$

→ The RMSE represents the root mean squared error, a numerical value that indicates how much the model predictions deviate from the actual precipitation measurements on average.

Unlike the MSE, which measures the error on a square scale, the RMSE provides a more intuitive and practical assessment of prediction quality. In summary, the RMSE is valuable for understanding how close or far the model predictions are from the actual precipitation values in terms of the amount of rainfall measured.

3.4.3. R^2

The coefficient of determination (R^2) indicates how well the model's predictions fit the actual variability regarding the amount of rainfall. An R2 value close to 1 suggests that the model is effective in capturing and explaining most variations in the precipitation data.

In Eq. (3), the MSE of the model is compared against the variance of the actual precipitation values, *i.e.*, $Var(y)$. The difference between the two provides a measure of the quality of the model in terms of its ability to explain data variability.

$$R^2 = 1 - \frac{MSE}{Var(y)} \tag{3}$$

3.4.4. Residuals

The residuals indicator denotes the difference between the predictions and the actual values (residuals) based on observations. It is essential to identify any systematic patterns in the errors of the model. If the residuals are randomly scattered around 0 and show no discernible pattern, this suggests that the model is adequate for the data. However, if patterns are observed in the residuals, it may be necessary to review and adjust the model.

4. Results and discussion

4.1. Applying time series in the development of the models

In this project, time series play an essential role regarding the prediction of monthly precipitation data for the department of Boyacá. This section outlines how time series were applied in the development of the forecasting model.

⟹ *Temporal aggregation:* The daily data collected (from 1981 to May 2023) were aggregated in monthly time series, highlighting seasonal and annual patterns.

⟹ *Data preprocessing:* The time series were normalized to improve training and divided into training and test sets.

⟹ *Modeling with RNNs (LSTM) and ConvLSTM:* LSTM models were used to capture temporal patterns, and ConvLSTM models were used to incorporate spatial and temporal characteristics in the precipitation data.

⟹ *Model evaluation with time series metrics:* Specific metrics such as the MSE, the RMSE, and the R^2 were used to evaluate the precision of the models.

⟹ *Visualizing the results in the form of time series:* The results were displayed in prediction vs. reality graphs to assess the fit of the models over time.

In summary, time series played an essential role throughout the process, from the preparation of the data to the evaluation and visualization of the results, enabling the understanding and prediction of the evolution of monthly precipitation in Boyacá over time.

4.2. Obtained dataset

Daily precipitation data for 1981 to May 2023 were obtained from a total of 753 precipitation measurement stations within the department of Boyacá, with a spatial resolution of 0.05°. Fig. 4 presents the measurement points and a visual representation of their location.

The final dataset (13) contained the latitude, longitude, precipitation, date, year, month, and municipality data necessary for the creation of the recurrent neural network models with LSTM and ConvLSTM layers.

4.3. Exploratory analysis

Fig. 5 shows the department of Boyacá and all its coordinates, with the aim of facilitating the visualization of the study area.

The percentage of precipitation measurement stations was extracted, which is directly proportional to the area of each municipality. Therefore, the larger the area, the more measurement points. In this case, the municipality of Puerto Boyacá is the largest and includes 6.8 % of the total distribution, followed by Cubara and Aquitania, as shown in Fig. 6.

The dataset taken from CHIRPS 2.0 indicates that there is a distance of 5.53150863126391 km between each measuring station. To exemplify the above, Fig. 7a shows the monthly precipitation by measurement points for May 2023, with a scale of 0 to 500. The highest precipitation for a given point is denoted by a darker blue color.

Fig. 7b shows the total monthly precipitation for each point within a municipality, in this case for Sogamoso in 2023, providing more detailed information about this specific point and facilitating more accurate conclusions.

An analysis of the monthly precipitation in the department of Boyacá for 2022, as represented in the heatmaps of Fig. 8, reveals some key points:

- The spatial distribution of precipitation can be observed, revealing patterns of rainfall concentration in different areas across months. The wettest months, *i.e.*, May, June, and July, with rainfall of up to 500 mm, can be clearly identified, as well as the driest months, *i.e.*, December, January, and February.
- The topography of the department of Boyacá seems to influence the distribution, since mountainous areas, valleys, and plains exhibit different precipitation patterns.
- The precipitation distribution across different months can be compared, which allows understanding how weather conditions change from month to month.
- In addition, some potential climatic anomalies can be identified, such as unusually dry or humid months in comparison with the general pattern. In the case of 2022, according to Fig. 8, no significant anomalies are observed.

In general, the analysis and visualization of monthly precipitation by coordinates allow obtaining detailed information on local climatic conditions and how they vary over time, which can be useful

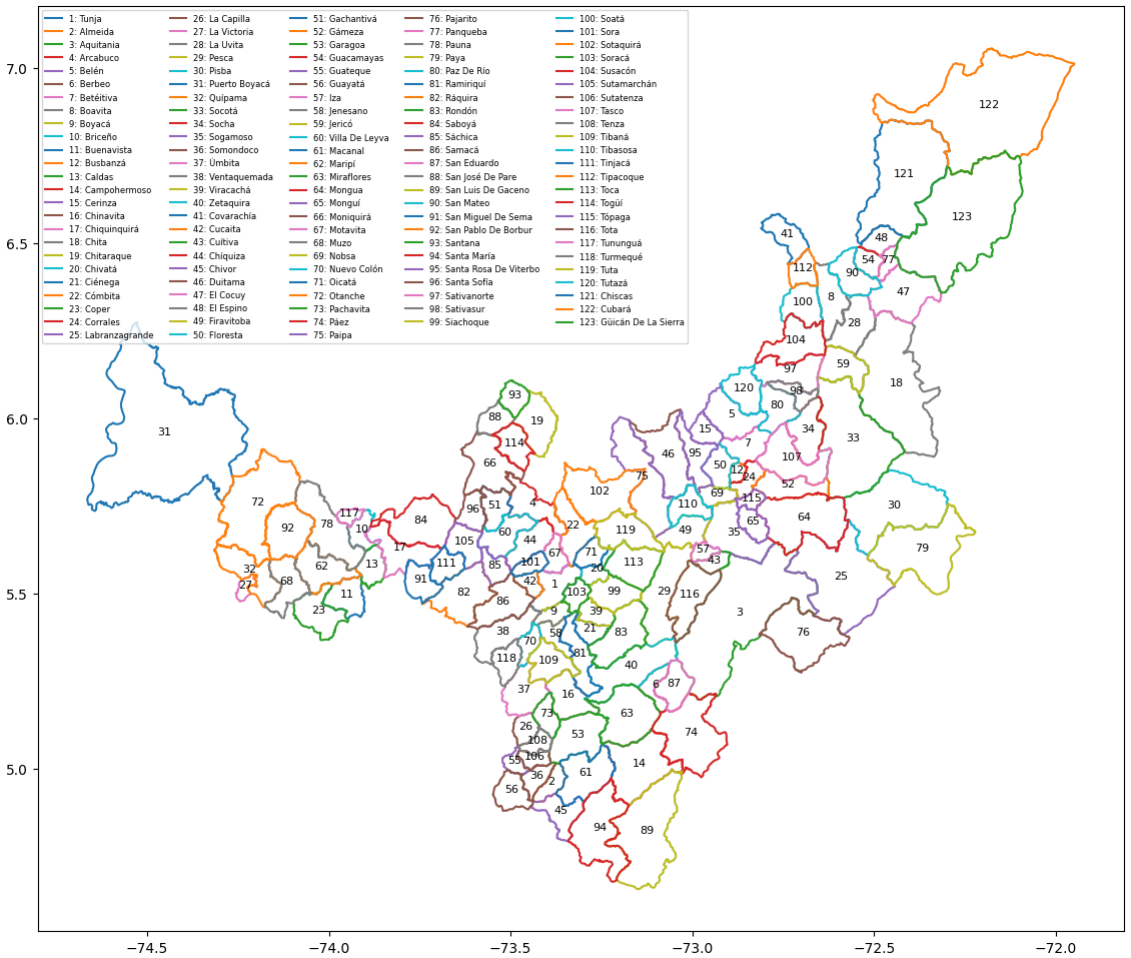


Figure 4. Precipitation measurement points

for making informed decisions in relation to sectors such as agriculture, water management, and urban planning.

4.4. Evaluating the models

Evaluating monthly precipitation forecasting models is essential to understanding their ability to provide accurate estimates of climatic conditions. In this work, various metrics were applied to quantify the prediction quality of two models based on RNNs with LSTM and ConvLSTM architectures. The metrics used included the MSE, RMSE, and R^2. These metrics provided critical information on the models' accuracy and explanatory power regarding prediction variability. Table **III** presents the numerical calculations of these metrics and the corresponding results, allowing for a detailed comparison and analysis of the evaluated models.

4.4.1. Comparison of predictions against real values (prediction *vs*. reality graphs)

This section presents a detailed assessment of how the model predictions align with the actual precipitation values in the test set. This evaluation was conducted by visually comparing the results with the 45° reference line, which represents the ideal scenario, *i.e.*, when the predictions perfectly coincide with the actual values.

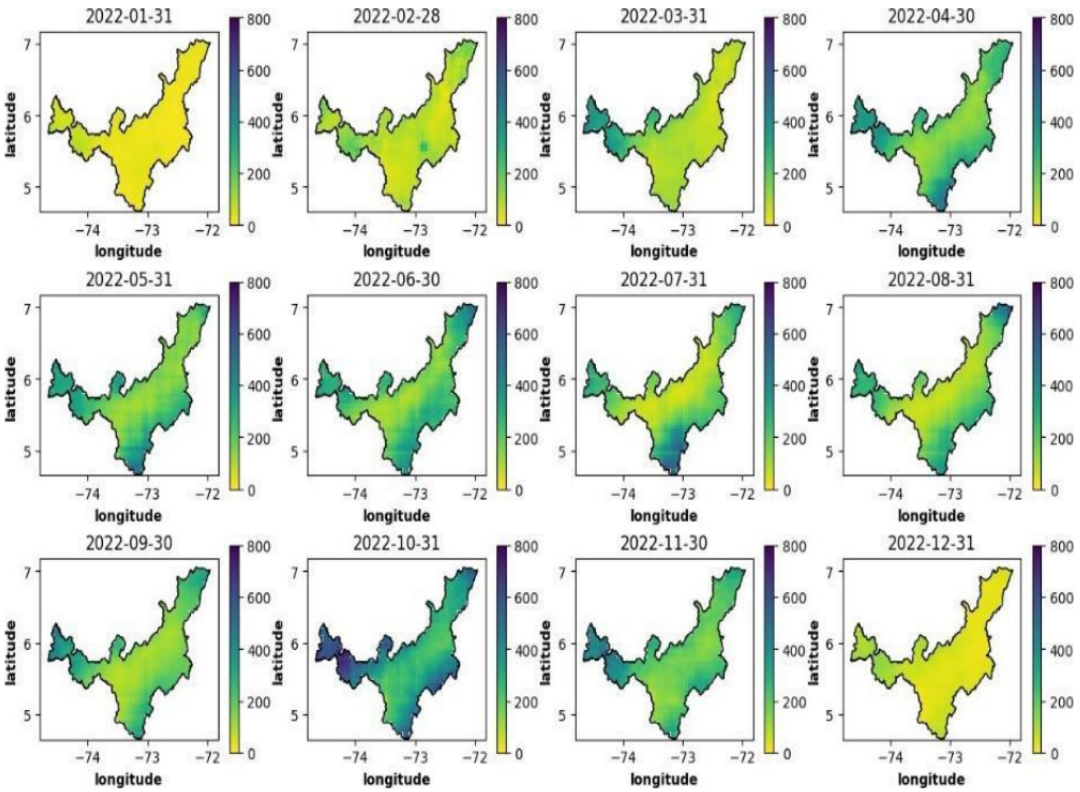


Figure 5. Coordinates of the department of Boyacá

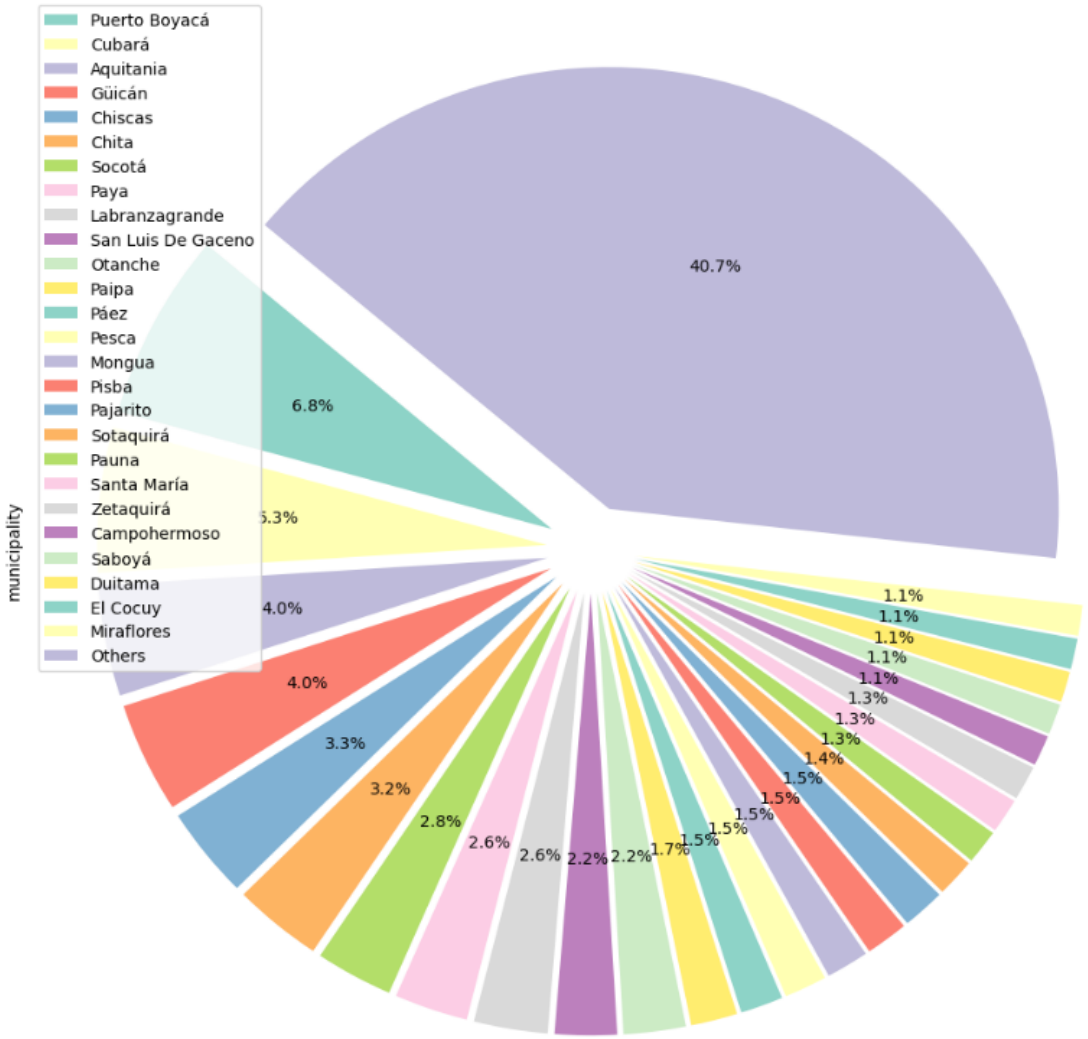


Figure 6. Coordinate distribution by municipality

In the evaluation process, predictions were generated by the RNNs with LSTM and ConvLSTM for the monthly precipitation of the test set and compared with the actual values. Then, another visual comparison was performed using separate graphs for each model, with the 45° line as a reference to assess the accuracy of the predictions.

4.4.1.1. Prediction vs. reality graphs

Fig. 9a shows that the predictions of the RNN with LSTM are remarkably aligned with the 45° line. As the actual precipitation value increases, the model's predictions tend to stay close to this baseline. This behavior indicates that the RNN with LSTM is capable of accurately capturing trends and variations in real values, even in situations involving higher precipitation values. The consistency

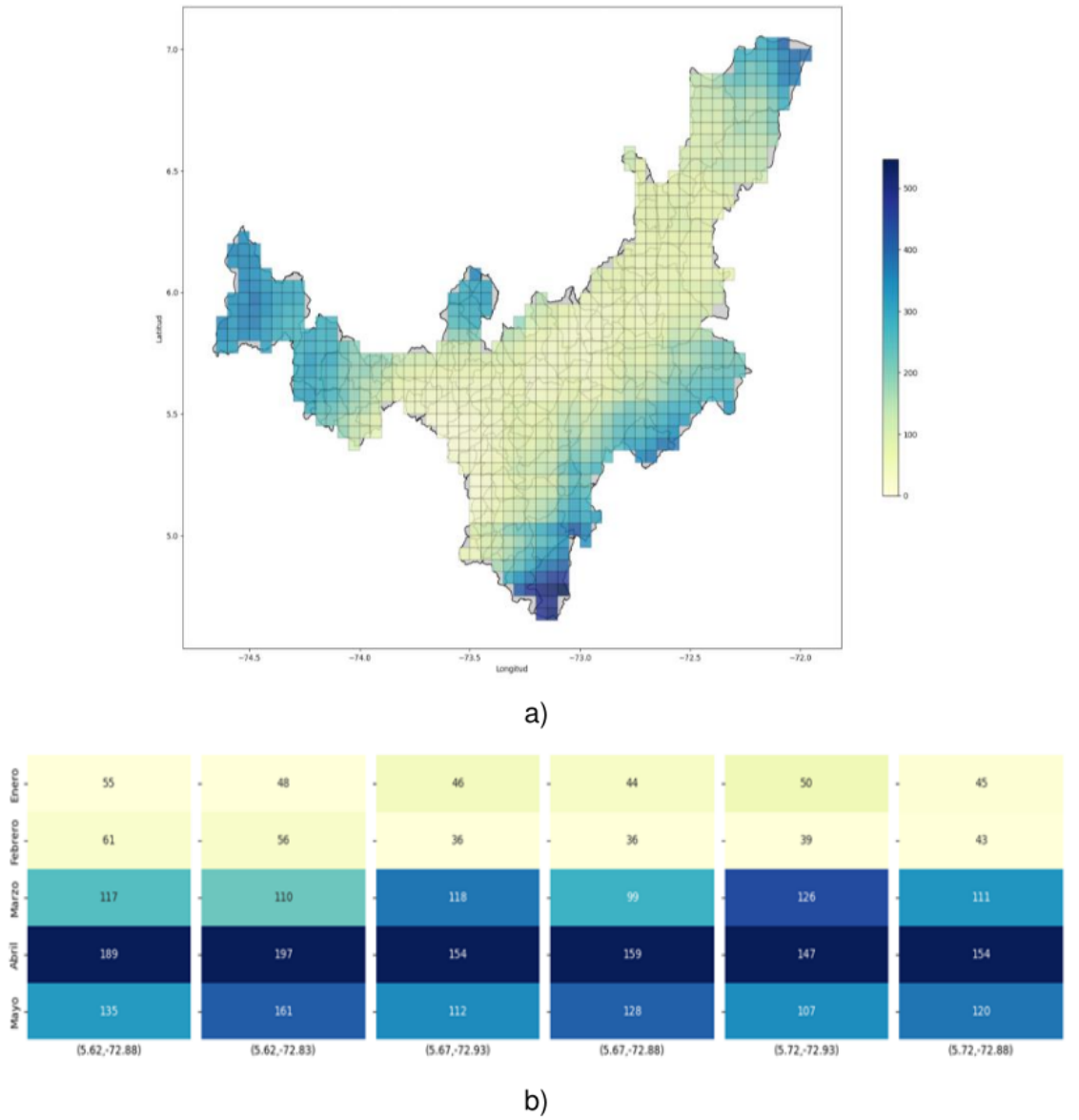


Figure 7. a) Monthly precipitation by measurement points for May 2023. b) Total monthly precipitation by coordinates in Sogamoso, 2023

in the alignment with the 45° line suggests a good fit and agreement between the predictions and the actual values.

Fig. 9b presents the alignment of the predictions generated by the ConvLSTM model with respect to the 45° line. Although this model also shows a trend close to the reference line, note that, as the real precipitation value increases, the predictions seem to gradually separate from the red-dotted line. This observation suggests that the ConvLSTM model may have a more difficult time capturing variations in actual values as they increase. The gradual separation of the predictions indicates that the

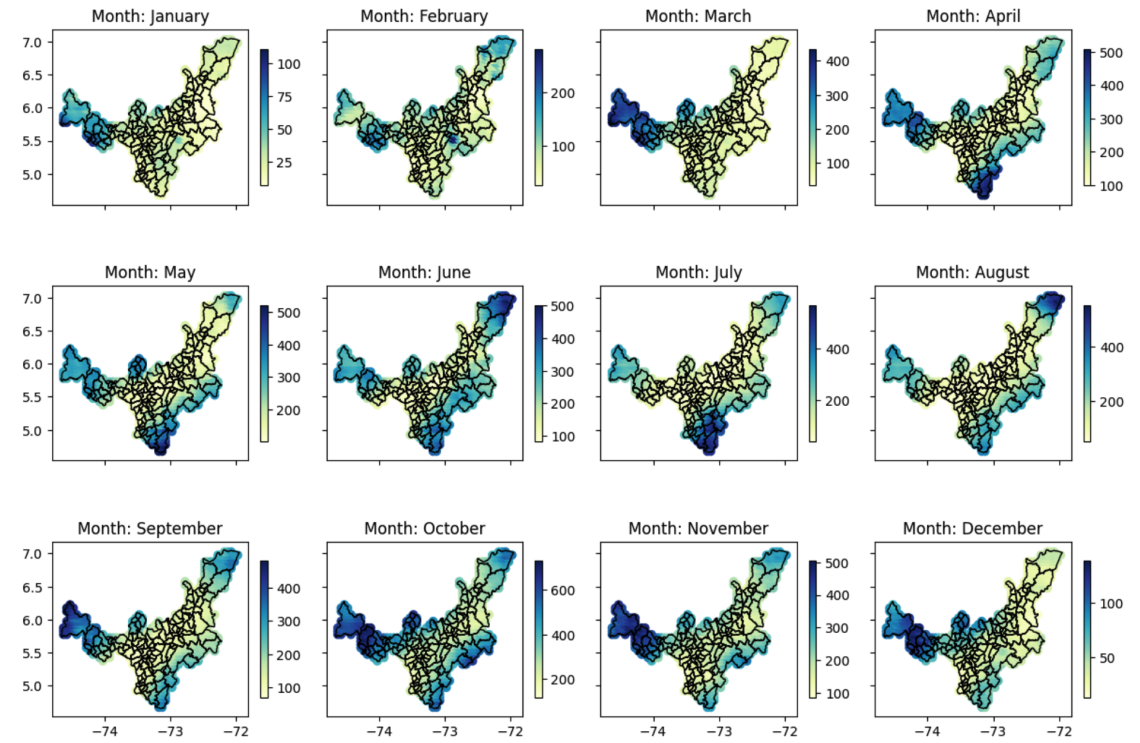


Figure 8. Total monthly precipitation by coordinates in 2022

ConvLSTM model could benefit from additional adjustments to improve its ability in predicting higher precipitation values.

Although both figures show points close to the 45° line, note that the RNN with LSTM performs better regarding its alignment with higher values. In contrast, the Convolutional model ConvLSTM l shows a gradual divergence.

4.4.2. Residuals analysis

Residual graphs are an essential tool in understanding models' quality and coherence regarding their ability to predict real precipitation values. This subsection presents the step-by-step process of how these graphs were elaborated and their specific interpretation.

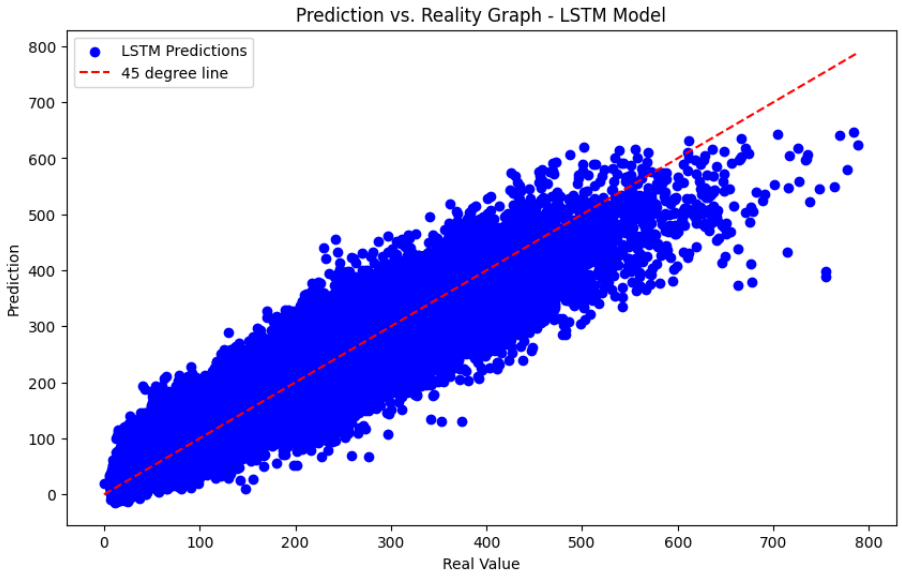
The evaluation process consisted of three fundamental steps. First, precipitation predictions were generated using the RNN with LSTM and the ConvLSTM model. These predictions were compared against the corresponding real values. Then, the residuals were calculated, obtaining the discrepancies between the predictions and the actual values. Finally, residual plots were constructed for each model, allowing for an effective visualization of the distribution and error patterns in relation to the actual values.

Table III. Scaling technique

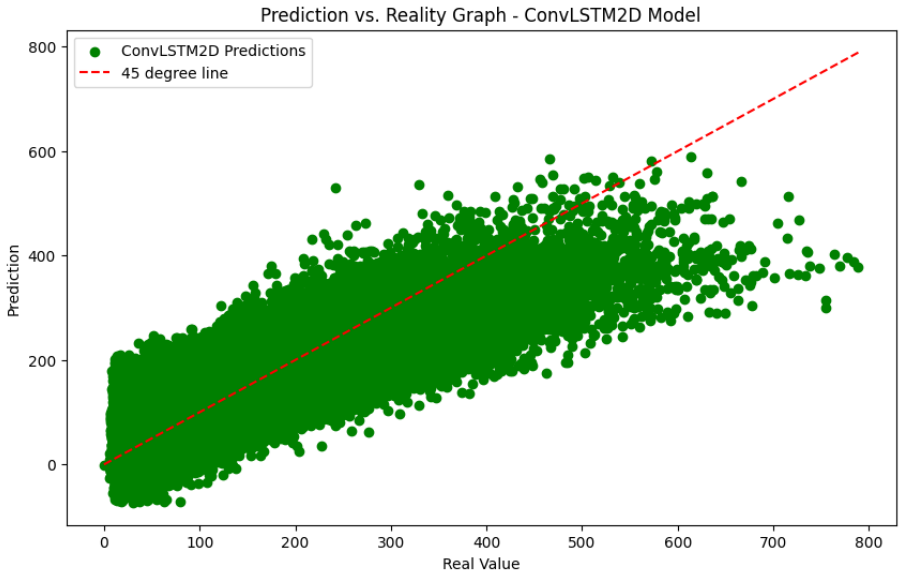
Metric	Numerical calculation	Result (analysis and discussion)
MSE	RNN with LSTM: 1401.37	The MSE of 1401.37 indicates that, on average, the predictions have a mean square error of approximately 1401.37 units of precipitation when compared to the actual values. This suggests a moderate dispersion in the predictions.
	ConvLSTM model: 3152.08	The higher MSE of the ConvLSTM model (3152.08) indicates that the predictions tend to have a higher MSE compared to the RNN with LSTM, which suggests a greater dispersion in the predictions of this model.
RMSE	RNN with LSTM: 37.43	The RMSE of 37.43 mm of precipitation for the RNN model with LSTM implies that, on average, the predictions have an error of approximately 37.43 mm when compared to the actual values. This indicates a relatively low dispersion with respect to the true values.
	ConvLSTM model: 56.14	The higher RMSE of the ConvLSTM model (56.14) suggests that the predictions have a greater dispersion compared to the RNN with LSTM, which indicates a margin of error of 56.14 mm in the predictions.
R^2	RNN with LSTM: 0.876	The R^2 value (0.876) for the RNN with LSTM indicates that this model explains 87.6 % of the variability in the precipitation data, which suggests a good performance in the explanation and prediction of said variability.
	ConvLSTM model: 0.722	The R^2 value of the ConvLSTM model (0.722) indicates that it explains approximately 72.2 % of the variability in the precipitation data, which indicates a good performance, although slightly lower compared to the RNN with LSTM.

4.4.2.1. Evaluating the prediction errors of the studied models

In the residual graph of the RNN with LSTM (Fig. **10**a), an interesting pattern can be observed. Most of the residuals are concentrated within the range of -60 to 60, which suggests that the model's predictions mostly fall within this margin of error. However, as the actual precipitation values increase, the points begin to disperse and move away from the reference line. This could indicate that the RNN with LSTM has difficulties in accurately predicting extreme precipitation values, which translates into more notable errors as the real values increase.



(a)



(b)

Figure 9. Prediction *vs.* reality graphs: a) RNN with LSTM, b) ConvLSTM model

Regarding the ConvLSTM2D model, the residual graph (Fig. 10b) also exhibits a pattern, albeit with distinct characteristics. Here, most of the residuals are in a wider range, between -100 and 100. This suggests that the predictions of the ConvLSTM2D model may have greater variability in terms of errors when compared to the RNN with LSTM. As with the previous model, as the actual precipitation values increase, the points tend to disperse and move away from the reference line.

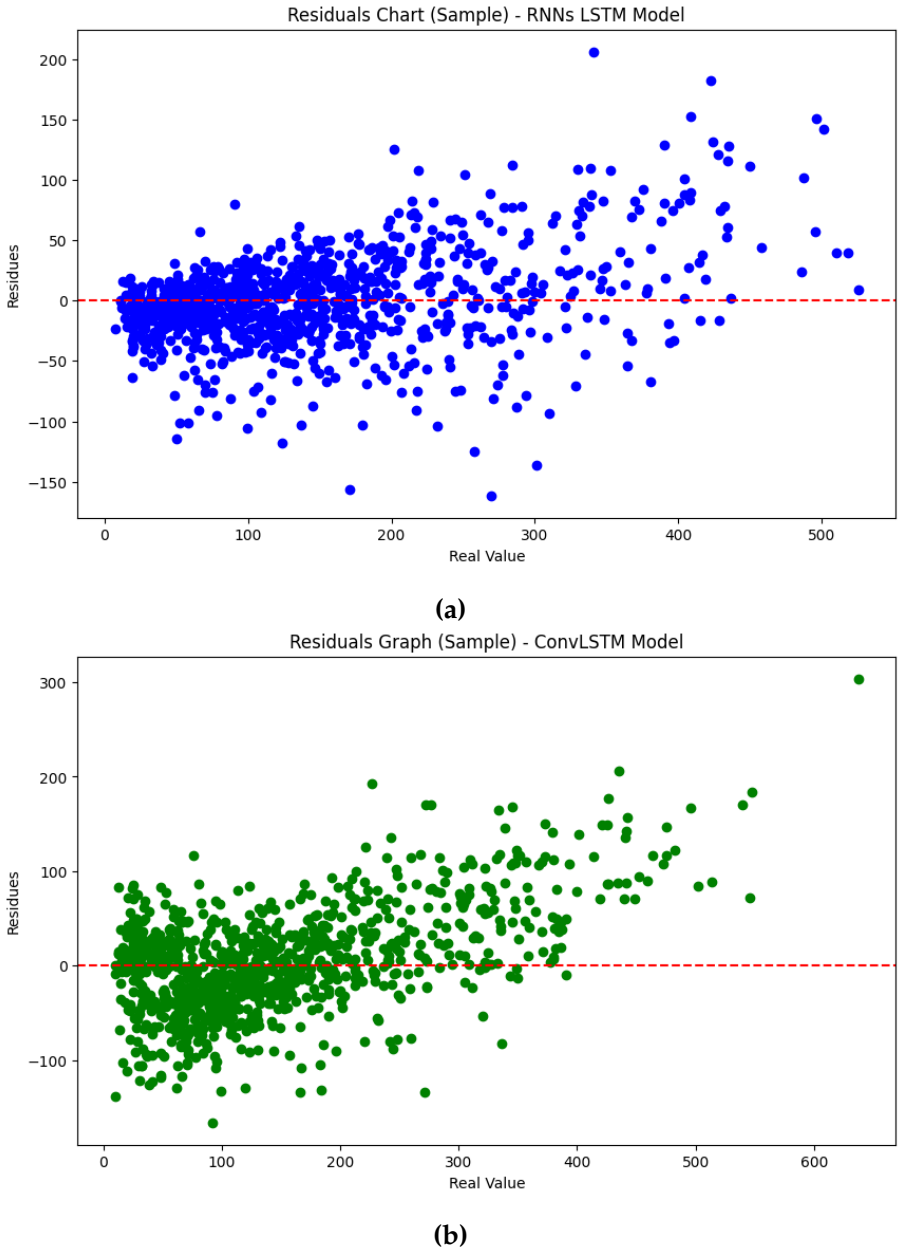


Figure 10. Residual plots: a) RNN with LSTM, b) ConvLSTM model

Both models exhibit a tendency towards the dispersion of the residuals as the real values increase. This dispersion suggests that both the RNN with LSTM and the ConvLSTM model may have difficulties in accurately capturing and predicting extreme precipitation values. However, it is important to note that most of the residuals remain within reasonable ranges in both models, indicating that the predictions are generally consistent with the actual values.

5. Data visualization

In the context of our monthly precipitation forecasting project for the department of Boyacá, an interactive dashboard was developed using Power BI to visualize the results.

The dashboard's development process began by defining the objectives and requirements, establishing what information to display, how to visualize historical data and predictions, and what type of user interaction to allow. Then, the historical monthly precipitation data and the predictions of the RNN models with LSTM and ConvLSTM model were collected and organized for later visualization in Power BI. The dashboard design was structured in three columns, each focused on a key category: historical data, the predictions of the RNN with LSTM, and the predictions of the ConvLSTM model. Specific visualizations were created for each column, including line graphs, treemaps, and geographic maps to represent trends, data by municipality and month, and spatial distributions, respectively. In addition, interactive filters were configured to allow users to dynamically explore the data and observe temporal and spatial patterns. The dashboard is illustrated in Fig. 11.

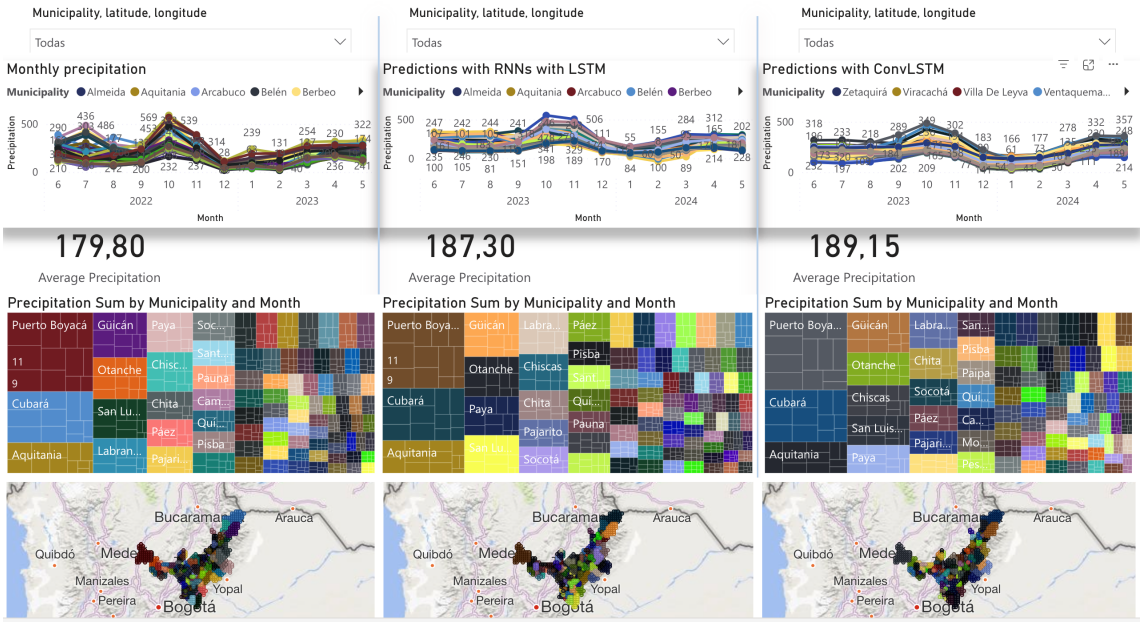


Figure 11. Dashboard structure

The description of the dashboard structure is described to continue:

⟹ *Real historical monthly precipitation:* This component shows a line graph with historical precipitation data, allowing for comparisons against the predicted values.

⟹ *Predictions generated by the RNN with LSTM:* This section explains the predictions made by this model, which are superimposed with historical data to evaluate its predictive capacity.

⟹ *Predictions generated by the ConvLSTM model:* This section presents the predictions generated by the ConvLSTM model, comparing them against the real data in an attempt to determine which

model offers the most accurate predictions.

⟹ *Precipitation treemap by municipality and month:* This section shows the precipitation level for each municipality across different months, facilitating the identification of patterns and distributions.

⟹ *Point distribution map:* This section allows visualizing the location of 753 measure points in Boyacá along with their precipitation level. This provides a spatial perspective of variations in precipitation.

6. Practical contributions

This project's first case study focused on the municipality of Sogamoso (latitude 5.72 and longitude -72.93). A detailed analysis of the time series was carried out, and the results of the studied models were compared against actual historical precipitation data. Both models closely followed the precipitation pattern, identifying three months with higher precipitation (April, October, and November) and three with lower levels (January, February, and December). In addition, a significant increase in the estimates with respect to historical values was noted for the 12 subsequent months, with the RNN with LSTM projecting an average of 124.09 mm (*vs.* the historical value of 111.65 mm) and the ConvLSTM model anticipating an even higher average (145.73 mm), suggesting a substantial increase in the expected rainfall.

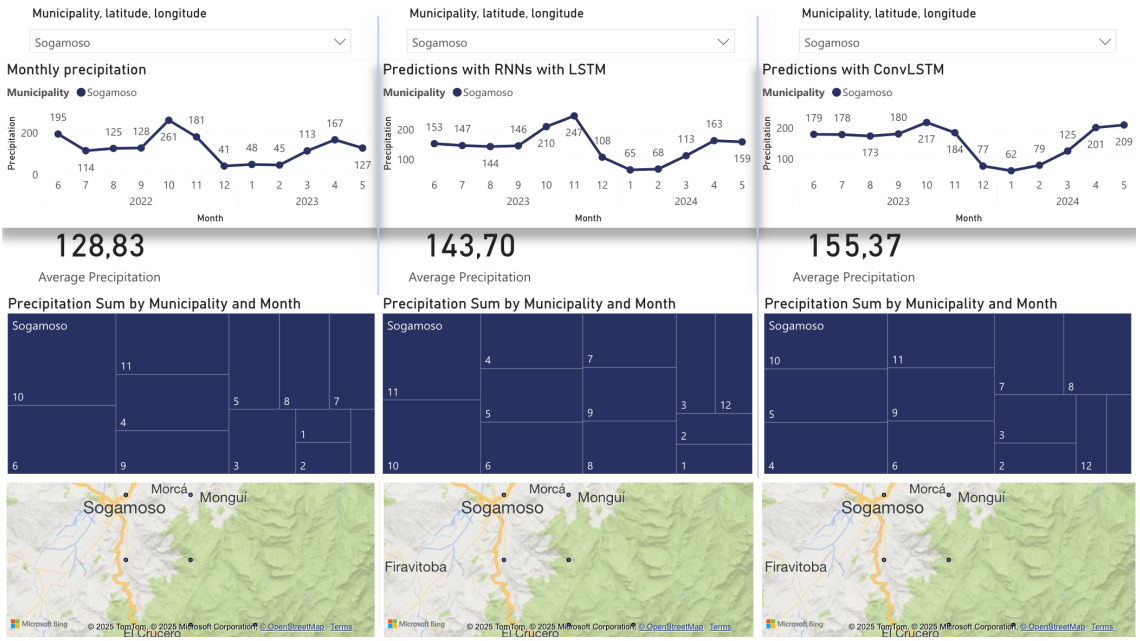


Figure 12. Data visualization, municipality of Sogamoso

In the second case study, involving the municipality of Tibasosa (latitude 5.77 and longitude -73.03), a similar precipitation pattern was observed, with April, October, and November recording the highest rainfall. This coincidence could be attributed to the geographical proximity of the two

municipalities. The actual data show a monthly average precipitation of 108.45 mm in this municipality, but the RNN with LSTM predicts an average of 117.39 mm, indicating an increase in comparison with historical values. Furthermore, the ConvLSTM model predicts an even higher average (141.87 mm). This significant increase has important implications for water resource management and planning in Tibasosa.

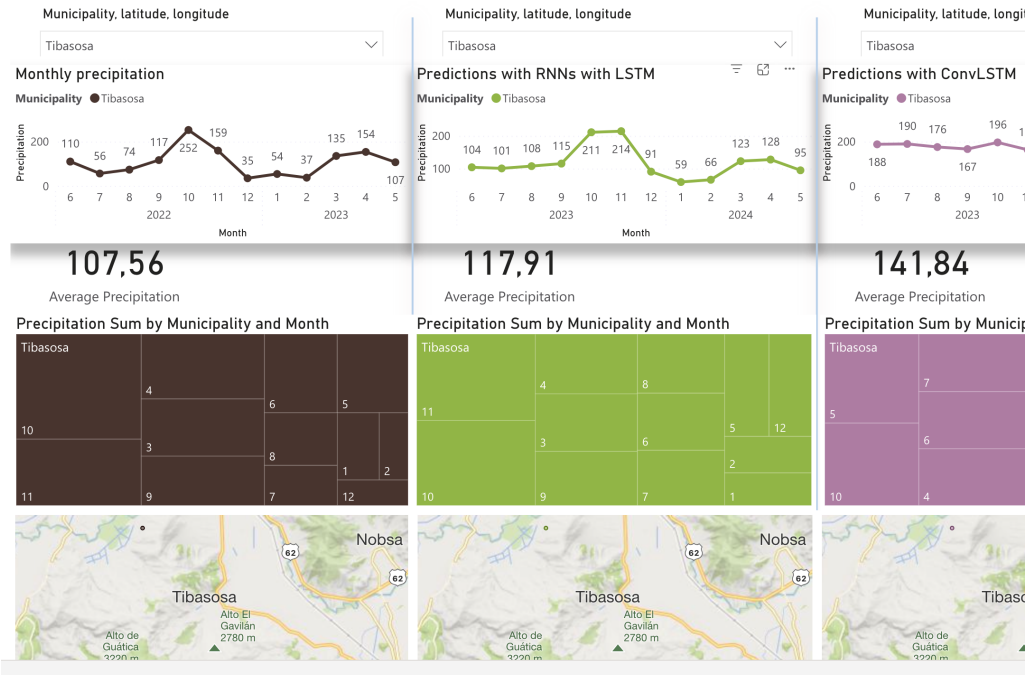


Figure 13. Data visualization, municipality of Tibasosa

It is important to note that, according to the metrics used to evaluate the performance of the models, the RNN with LSTM exhibits a superior performance in this particular scenario. This observation highlights this model's ability to make more accurate predictions compared to the ConvLSTM model in the specific context of the municipality of Sogamoso.

Geographical filtering allowed limiting the dataset to the specific area selected for the study, making it possible to analyze the information of interest and determine the maximum and minimum precipitation values in the region of Boyacá from 1981 to 2023. This included a detailed observation of the precipitation levels generated by El Niño and La Niña in 2022 and early 2023.

7. Conclusions and future work

This project undertook the task of predicting monthly precipitation data using models based on neural networks. Two approaches were employed: one using a long short-term memory architecture and the other implementing convolutional long short-term memory. The results obtained demonstrate the predictive capacity of both models, generating scenarios based on the precipitation patterns of the region.

Both models were subjected to an evaluation that allowed quantitatively understanding their performance. The metrics used provided valuable information on the accuracy of the predictions and the tunability of each model. The comparison showed that the LSTM model achieves a lower MSE and RMSE, as well as an R^2 closer to 1, indicating a better fit than the ConvLSTM model.

The prediction *vs.* reality graphs confirmed the solidity of the models. Both approaches showed points near the 45° line, suggesting that the predictions were in line with the actual values. However, it was noted that the LSTM model exhibits a greater proximity to the points of this line, which is indicative of its greater overall precision.

An examination of the residual plots revealed important patterns in the error distribution. For the LSTM model, most of the residuals oscillated between -60 and 60, while the ConvLSTM model showed a somewhat greater dispersion (-100 to 100). These results indicate that the LSTM model can better capture the data variability.

It is worth stating that this work successfully evaluated and developed monthly precipitation forecasting models for the department of Boyacá. The predictions provided by both models offer valuable insights for decision-making in sectors such as agriculture, water management, and urban planning. In addition, a dataset was generated, with predictions for 12 months after June 2023, allowing for continuous analysis and a better understanding of the weather patterns in the region.

In addition, the dashboard, developed using Power BI, constitutes a robust data visualization strategy to understand historical data, model predictions, and spatial distributions concerning precipitation in the department of Boyacá. This interactive tool is essential for communicating results effectively and supporting informed climate and meteorological decision-making.

The RNN with LSTM proved to be more effective than the ConvLSTM model in predicting monthly precipitation in the study area, providing reliable and accurate information for decision-making in critical sectors. Its ability to capture weather patterns and its performance regarding assessment metrics position it as a promising tool in research and practical application. However, it is important to note that, like any model, it faces challenges, such as the need for a larger amount of historical data to further improve its accuracy, as well as the exploration of regularization techniques to avoid possible overfitting issues.

Nonetheless, the ConvLSTM model, despite its slightly lower performance in this specific context, is still relevant and could find specific applications in other climate scenarios or geographic regions.

In the future, the combination of multiple models could be considered in order to take advantage of their individual strengths and further improve predictions. In addition, expanding the network of meteorological stations in the region and incorporating additional data could enrich and improve climate predictions, thus contributing to more effective decision-making and mitigation of climate-related risks in Boyacá.

Finally, this project demonstrates that deep learning models, especially RNNs with LSTM, are effective tools for predicting crucial climate data. The results obtained provide confidence in the studied models' ability to contribute significantly to research and decision-making in the context of monthly precipitation in Boyacá.

8. Author contributions

Yessid Esteban Duarte Prieto: conceptualization, formal analysis, research, programming.
Marco-Javier Suárez-Barón: conceptualization, formal analysis, research, methodology, software, validation, writing (original draft).
Oscar Javier García Cabrejo: conceptualization, formal analysis, research, methodology, software, validation, writing (original draft).
César Augusto Jaramillo: software validation, writing, methodology.
Carlos Augusto Meneses Escobar: software validation, writing, methodology.

References

[1] C. E. González Orozco, M. Porcel, D. F. Alzate Velásquez, and J. O. Orduz Rodríguez, "Extreme climate variability weakens a major tropical agricultural hub," *Ecol. Ind.*, vol. 111, art. 106015, Apr. 2020. https://doi.org/10.1016/j.ecolind.2019.106015 ↑3

[2] A. K. Smith, B. L. Johnson, and C. D. Williams, "A deep learning approach for weather forecasting prediction using convolutional neural networks," *J. Clim. Atm. Sci.*, vol. 15, no. 3, pp. 456-472, 2022. ↑3

[3] M. Rocío Sánchez, H. Romero Zepeda, and M. C. Sánchez, "We protect our moors or peasants," *Estancias*, vol. 2, no. 4, pp. 265-279, 2022. ↑3

[4] J. C. Valderrama Balaguera, H. F. Castro Silva, and C. A. Dávila Carrillo, "Pronósticos de variables climatológicas mediante los modelos de punto de cambio y Holt-Winters," *Mundo FESC*, vol. 11(S2), pp. 337-352, 2021. https://doi.org/10.61799/2216-0388.986 ↑3

[5] A. K. Smith, B. L. Johnson, and C. D. Williams, "Short-term rainfall forecasting using machine learning-based approaches of PSO-SVR, LSTM and CNN," *J. Hydromet.*, vol. 25, no. 4, pp. 789-805, 2023. ↑3

[6] E. Morales Rojas, E. A. Díaz Ortiz, L. García, and M. Milla Pino, "Monthly rainfall forecast: A case study in the native communities of Peru," *Pakamuros*, vol. 9, no. 3, pp. 71-85, 2021. https://doi.org/10.37787/pakamuros-unj.v9i3.219 ↑4

[7] D. M. Herrera Posada and E. Aristizábal, "Artificial intelligence and machine learning model for the spatial and temporal prediction of drought events in the department of Magdalena, Colombia," *INGE CUC*, vol. 18, no. 2, pp. 249-265, 2022. https://doi.org/10.17981/ingecuc.18.2.2022.20 ↑4

[8] B. Zhao, H. Lu, S. Chen, J. Liu, and D. Wu, "Convolutional neural networks for time series classification," *Syst. Eng. Elec.*, vol. 23, pp. 162-169, 2017. https://doi.org/10.1109/IAEAC.2017.8053998 ↑4

[9] J. Zhang, Y. Li, X. Wang, and H. Chen, "Infrared precipitation estimation using convolutional neural network for FengYun satellites," *IEEE Trans. Geosci. Rem. Sens.*, vol. 60, no. 5, pp. 1-12, 2022. ↑4

[10] J. Xia, C. Wang, C. Wang, Q. Sun, and W. Fang, "Hybrid LSTM-GRU with attention mechanism and residual convolutional neural network for raindrop nowcasting," *Int. J. Rem. Sens.*, pp. 1362-1383, 2018. ↑4

[11] T. Nan, W. Cao, Z. Wang, Y. Gao, L. Zhao, X. Sun, and J. Na, "Evaluation of shallow groundwater dynamics after water supplement in North China Plain based on attention-GRU model," *J. Hydrol.*, vol. 615, pp. 128-145, 2023. https://doi.org/10.1016/j.jhydrol.2023.130085 ↑4

[12] CHIRPS, "data.chc.ucsb.edu," [Online]. Available: https://data.chc.ucsb.edu/products/CHIRPS-2.0/global_daily/netcdf/p05/ ↑4

[13] Y. E. D. Prieto, "Rainfall data for the department of Boyaca," Kaggle, 2023. [Online]: Available: https://www.kaggle.com/datasets/estebanduarte/rainfall-data-for-the-department-of-boyaca ↑7, 13

Yessid Esteban Duarte Prieto

Systems and computing engineer, UPTC, Sogamoso, Colombia. Research assistant for the GALASH-UPTC group.
Email: yessid.duarte@uptc.edu.co

Marco Javier Suárez Barón

Systems engineer, PhD in Strategic Planning and Technology Management, UPAEP, Mexico; Master of Information Management, Escuela Colombiana de Ingeniería. Associate professor at UPTC (Colombia). Head of the GALASH research group.
Email: marco.suarez@uptc.edu.co

Oscar Javier García Cabrejo

Geologist from Universidad Nacional de Colombia; PhD in Civil Engineering from the University of Illinois at Urbana-Champaign. He works in hydrologic modeling while applying machine learning. Researcher of the INGEOLOG group. Associate professor at UPTC (Colombia).
Email: oscar.garcia04@uptc.edu.co

César Augusto Jaramillo Acevedo

Systems and computing engineer; MSc in Systems and Computing Engineering from Universidad Tecnológica de Pereira (Colombia). He has been a professor-researcher of UTP for more than 12 years. He has been an active member of research groups, and his areas of interest and teaching are related to software engineering, compilers, AI, IoT systems, the cloud, distributed systems, and the Industry 4.0.
Email: swokosky@utp.edu.co

Carlos Augusto Meneses Escobar

Systems and Computing engineer from Universidad de los Andes; MSc in Physical Instrumentation from UTP. He has been a professor-researcher of Universidad Tecnológica de Pereira for more than 25 years. He has been an

active member of research groups, and his areas of interest and teaching are related to software engineering, AI, the Industry 4.0, and digital image processing.
Email: cmeneses@utp.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.21898

Research

Design and Implementation of a Virtual Laboratory for the Analysis and Synthesis of Mechanisms

Diseño e implementación de un laboratorio virtual para el análisis y síntesis de mecanismos

Luz Adriana Mejía Calderón[1]* and Carlos Alberto Romero Piedrahita[2]

[1]Faculty of Applied Mechanics, Universidad Tecnológica de Pereira (Pereira, Colombia)

[2]Faculty of Technology, Universidad Tecnológica de Pereira (Pereira, Colombia)

Abstract

Context: This paper presents the design and implementation of a virtual laboratory developed as a pedagogical support tool for courses in mechanisms analysis and synthesis within Mechanical Engineering programs.
Method: Our virtual laboratory was developed using GeoGebra, Moodle, and Matlab. The design methodology was structured into four key stages: requirements analysis, conceptual design, content development, and implementation and evaluation.
Results: The outcome is a comprehensive virtual laboratory framework comprising 12 interactive exercises focused on the analysis and synthesis of planar mechanisms. This paper outlines the functional features of the virtual laboratory, including its graphical user interface, user guide, instructional practices, and assessment procedures – demonstrating its value as an effective tool to support and enhance both teaching and student learning.
Conclusions: The virtual laboratory allowed students to conceptualize through the visualization of kinematic and dynamic variables, offering a complementary and interactive alternative to traditional theoretical instructions. This aligns with active learning strategies, as the teacher reported greater student engagement. Students advanced at their own pace through the exercises and were able to observe how geometric and functional modifications in mechanisms influence their kinematic and dynamic behavior.

Keywords: virtual laboratory, kinematics, dynamics, synthesis, mechanism

Article history

Received:
February 19th, 2024

Modified:
March 16th, 2025

Accepted:
April 16th, 2025

Ing, vol. 30, no. 1, 2025, e21898

* **Correspondence:** adriamec@utp.edu.co

Resumen

Contexto: Este trabajo presenta el diseño y la implementación de un laboratorio virtual desarrollado como soporte pedagógico en la enseñanza de la teoría de máquinas y mecanismos en ingeniería mecánica.
Método: Nuestro laboratorio virtual fue desarrollado utilizando Geogebra, Moodle y Matlab. La metodología de diseño se estructuró en cuatro etapas clave: análisis de requerimientos, diseño conceptual, desarrollo de contenido e implementación y evaluación.
Resultados: El resultado es un marco integral de laboratorio virtual que contiene 12 ejercicios centrados en el análisis y la síntesis de mecanismos planos. Este artículo presenta las características funcionales del laboratorio virtual, incluyendo su interfaz gráfica de usuario, guía de usuario, prácticas institucionales y procedimientos de evaluación —demostrando su valor como herramienta efectiva para apoyar y mejorar la enseñanza y el aprendizaje de los estudiantes.
Conclusiones: El laboratorio virtual permitió a los estudiantes conceptualizar mediante la visualización de variables cinemáticas y dinámicas, ofreciendo una alternativa complementaria e interactiva a la instrucción teórica tradicional. Esto se alinea con las estrategias de aprendizaje activo, pues el docente reportó un mayor compromiso por parte de los estudiantes. Los estudiantes avanzaron a su propio ritmo a través de los ejercicios y pudieron observar cómo las modificaciones geométricas y funcionales en los mecanismos influyen en su comportamiento cinemático y dinámico.

Palabras clave: laboratorio virtual, cinemática, dinámica, síntesis, mecanismos

Table of contents

1. Introduction

In the teaching-learning of the analysis and synthesis of mechanisms, graphic and analytical methodologies are regularly used. These approaches are complementary: the graphical methodology provides a visual representation of velocity, acceleration, and force vectors, allowing students to better understand the physics of motion; while the analytical methodology relies on mathematical models, laying the foundation for simulation and optimization processes. Additionally, it enables the parameterization, systematization, and solution of multiple configurations of the mechanisms under study.

The creation and implementation of virtual laboratories are widespread trends in engineering education. These trends aim to establish flexible university curricula and promote diverse student-centered teaching-learning methodologies, thereby fostering autonomous learning. The application of new teaching methodologies has reshaped classrooms, workspaces, scheduling, and student engagement in their educational journey. This transformation has been particularly evident in recent years, accelerating the shift from traditionl face-to-face instruction to virtual formats. Consequently, technology has emerged as a fundamental mediator in the evolving teaching landscape (1).

Information and communication technologies (ICTs) and learning and knowledge technologies (LKTs) have expanded the scope of learning and participation, enabling the exposition, manipulation, analysis, synthesis, application, and evaluation of ideas. These activities are enhanced through direct observation and reflective engagement in virtual environments. In this context, virtual laboratories play a pivotal role, providing students with immersive, interactive, and adaptable learning experiences that bridge the gap between theoretical concepts and practical applications.

Virtual laboratories (VLs) serve as an electronic workspace designed for remote collaboration and experimentation in research and other creative pursuits (2). They facilitate the production and dissemination of academic results by leveraging state-of-the-art information and communication techniques. Particularly, during the COVID-19 pandemic of 2020, VLs proved to be powerful educational technological tools (3).

Some of their main advantages include:

- Interactivity, enabling learners to manipulate and transform instructional material, data, and information
- Accessibility
- Methodological variability
- Flexibility in terms of time and space
- Active student participation
- Encouragement of self-learning

VLs are computational systems that can be accessed via the Internet. They are based on theoretical models, simulate processes, and provide dynamic object visualization in a way similar to conventional laboratories (4). These systems can be implemented as standalone units containing multiple virtual experiments on the same platform, or as integrations of experiments from different platforms. Tools such as GeoGebra, Matlab Live, MatlabApps, Java, Flash, CGIs, PHP, and Wolfram facilitate the development of online applications, thus contributing to the creation of VLs.

Various open-access interactive simulation sources, including Walter-Fendt (https://www.walter-fendt.de/html5/phes), Educaplus (http://educaplus.org), and PhET (https://phet.colorado.edu), offer numerous science and mathematics applets. Some of these

applets can be easily embedded in platforms like Moodle, which is widely used in university settings. Prior to the pandemic, the use of simulations in VLs for physics teaching in engineering courses had already demonstrated benefits in fostering autonomous learning (5). This trend continued during and after the pandemic, leading to increased pedagogical research on the perception of VLs in traditional education (6–9). The growing interest in VLs suggests their potential adoption by higher education institutions and even their incorporation into state policies, as exemplified by India's implementation under the theory of diffusion of innovation (10). Research findings indicate that the acceptance, intention to use, and perceived relevance of this technology significantly influence user perception.

Over the past three years, numerous implementations of VLs across various fields of engineering have been reported. In (11–13), VLs have been applied to circuit-related experiments in electronics engineering. In civil engineering, (14) explores the effectiveness of a virtual laboratory designed to prepare students for concrete laboratory sessions, allowing them to become familiar with experimental procedures beforehand. The study reported a 16 % reduction in time spent in the physical lab.

In mechanical engineering, (15) addresses the development and implementation of a virtual simulation platform for teaching wind turbine assembly practices. Given the large size and high cost of turbine components, this platform offers a valuable alternative for practical training, which is often limited in traditional lab settings.

In the field of biotechnology, (16) evaluates the impact of an immersive virtual reality (IVR) intervention on student learning and academic engagement, particularly in the instruction of infrared (IR) spectroscopy. This study found that students who participated in an IVR experience scored higher on IR-related exam questions compared to control groups and reported greater academic engagement as well as a positive attitude towards the use of IVR in learning.

In chemical engineering, (17) explores the integration of VLs into chemistry education to enhance students' understanding and practical skills. The results indicated that VLs were effective in reducing common misconceptions among students.

Finally, in safety engineering, (18) describes the development and implementation of a virtual simulation platform for teaching emergency ventilation systems in subway environments – a type of experiment that cannot be conducted in real-world conditions due to its hazardous nature.

In the field of mechanisms analysis and synthesis, applications have predominantly been developed using GeoGebra, a software that parameterizes geometric figures based on relationships between and the properties of entities (*e.g.*, the links of a mechanism and velocity and acceleration vectors). This enables the creation of interactive applets. Collections of individual applets, such as the one presented by (19), provide substantial support for teaching courses. However, their use has been historically limited to kinematic analysis, without incorporating dynamic analysis, which is crucial in the design process.

Since machine design and construction involve topological design, kinematic and dynamic analysis, simulation, and mechanisms optimization, this work primarily aims to implement numerical methodologies for computational, graphical, and analytical solutions. This objective is fulfilled through the development of a VL dedicated to teaching the dimensional synthesis of mechanisms, as well as their kinematic and dynamic analysis.

The following sections describe the design methodology of the VL for mechanisms analysis and synthesis, detailing the requirements analysis, conceptual design, content development, implementation, and evaluation stages. Subsequently, initial evaluation comments and feedback gathered from user surveys regarding performance, content, learning value, presentation, challenges, and recommendations are discussed. The VL was implemented within the Mechanical Engineering program at the Universidad Tecnológica de Pereira as part of its educational innovation efforts. Finally, this paper concludes by presenting key findings relevant to the authors.

2. Methodology

The design of VLs, conceived as interactive digital objects within the academic environment, requires a structure that considers not only general and specific objectives and content, but also aesthetic and functional attributes. According to (20), designing a VL requires a clear understanding of its purposes and context. The appropriate technology must be selected to ensure efficient media design and use. The level of realism should align with the intended purpose and target users, and learning guides should be provided both before and after the VL session.

(21) further emphasizes that, in order for a digital object to be effective, its design must account for user type, task analysis, content structure, content presentation, and navigation modes – both within the object itself and across the network.

Building upon these foundational principles, a structured methodology for the design and implementation of a VL was established. This process generally includes the following stages: requirements analysis, conceptual design, practice content development, implementation, and evaluation.

2.1. Requirements analysis

This stage defines the general characteristics that the VL should possess, including the number and scope of practices, the format of the graphical interface, interactivity adapted to the subject, applications, and potential users. Additionally, it encompasses the documentation and information required for the access, use, and further development of the laboratory.

The designed laboratory is intended for students enrolled in the courses titled *Theory of Machines and Mechanisms*, and *Synthesis of Mechanisms*. It consists of practices developed as individual learning objects within the same environment. These objects are later integrated into a unified website hosted on

the institution's Moodle platform.

Given the subject matter and the learning objectives of the practices, the requirements for each learning object are as follows:

- *Graphic interface:* Includes dialog boxes and control bars for adjusting input parameters, enabling the real time visualization of changes in the polygons that represent the kinematic and dynamic variables of the mechanism. This interface also allows for simulations of the analyzed mechanisms.
- *Interactivity:* Allows users to navigate between the different types of analysis (kinematic, dynamic, and simulation) through tabs. Users can modify the lengths and positions of links and frames. For dimensional synthesis, interactivity facilitates the exploration of design parameters (precision points), synthesis type (guided, trajectory or function), and visualization of rotation criteria, transmission angles, and the simulation of the synthesized mechanism.
- *Documentation and help resources:* Provides guidance for each practice.

2.2. Conceptual design of the laboratory

During the conceptual design stage, the platform, structure, and layout of the VL are defined. Considerations regarding the selection of suitable software for practice development and implementation are also addressed.

The laboratory's operating platform (the LMS Moodle V3.11 package) was selected based on the institution's existing virtual infrastructure. This learning management system (LMS) enables the planning and implementation of modules for academic courses, as well as diverse evaluation strategies that facilitate student progress tracking (22).

The structure of the virtual laboratory is specifically designed for subjects related to the theory of mechanisms. It was conceptualized as a combination of codes using both open- source and licensed software – some for the graphical development of mechanisms analysis and synthesis (*e.g.*, the creation of velocity, acceleration, and force polygons; graphical dimensional synthesis), and others for analytical processing using mathematical expressions with numerical resolution.

The conceptual design includes the creation and hosting of interactive applications executable via a web browser, eliminating the need for additional software installation. The diagramming process of the practices considers the learning levels, sequence, interrelation, and complexity of the exercises. This structured approach ensures that the course begins with simple exercises and progressively introduces more complex mechanical systems along with advanced programming, visualization, and simulation tools.

2.3. Design and content

During the design and content development stage, solution algorithms are programmed in line with the laboratory's objectives, intended applications, and user base. These algorithms are tailored to the

specific requirements of each practice, while help and information materials are seamlessly integrated. Additionally, interactivity and navigability are carefully structured to ensure a comprehensive user experience.

Table I presents the available practice selections, categorized by study type: mechanism analysis, and mechanism synthesis. The choice of specific mechanisms was based on their relevance to the design of machines. Many industrial machines are derived from modifications of the four-bar, crank-rocker, and crank-slider mechanisms or their combinations. These mechanisms are fundamental in various applications, including positive displacement pumps, compressors, internal combustion engines, presses, die-cutting machines, lifting platforms, and rectilinear motion systems. Practices focused on mechanism synthesis are closely aligned with the course content and were defined according to this criterion.

Table I. Practices in the proposed virtual laboratory

Type of study	Practice
Kinematic and dynamic mechanism analysis	Four-bar linkage mechanism, slider crank mechanism, scotch yoke mechanism (collisa mechanisms), Whitworth quick return mechanism, cam mechanism, gear mechanism
Dimensional synthesis	Three-precision-point kinematic synthesis (guidance, established supports, and trajectory), four-precision-point kinematic synthesis, four-bar function generator, slider-crank function generator, cognate mechanisms, synthesis of cam mechanisms

The practices within the graphical environment were developed using Geogebra and integrated as applets into their corresponding section. One such example is the practice involving a four-bar linkage mechanism, which is depicted in Fig. 1. In this section, students can validate the mechanism's rotation criteria through an interactive application, where they can manipulate the link lengths and verify compliance with Grashof's law. Additionally, students are prompted to simulate the motion of the mechanism by rotating the input link according to the given configuration. The documentation for this section covers the concept of rotatability and its associated criteria. Moreover, students are encouraged to modify the link dimensions with predefined values, allowing them to observe of various kinematic inversions achieved.

For all kinematic analyses, interactive applets allow users to modify the dimensions of the mechanisms. Users can adjust the magnitudes and directions of the velocities and angular accelerations of the input links using sliders (Fig. 2), which allows them to observe the resulting kinematic changes.

In the previous stage, the accompanying documentation explains the construction of velocity and acceleration polygons. Through guided practice exercises, students are encouraged to introduce kinematic changes and observe how the shapes and dimensions of these polygons are transformed. Directions, suggestions, and questions – such as those exemplified in Fig. 3 – are incorporated to enhance students' understanding and internalization of the material.

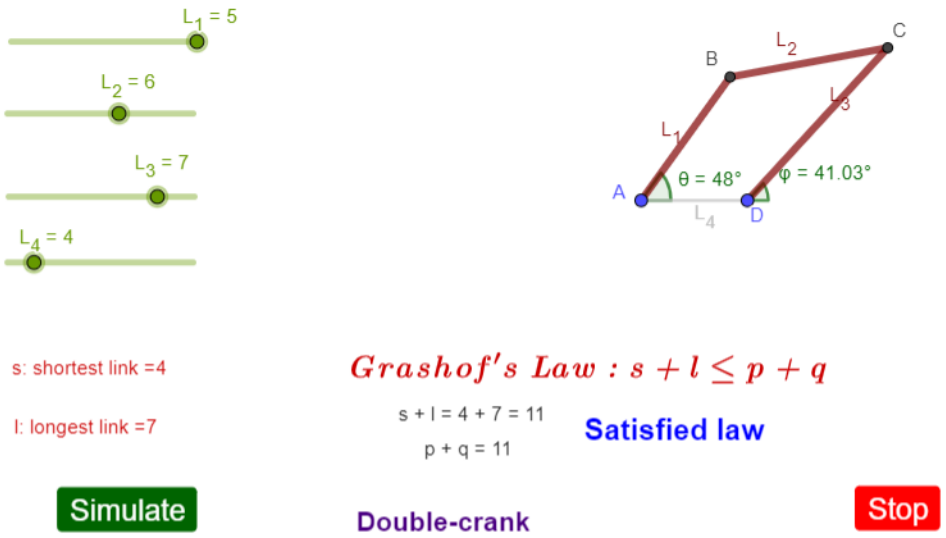


Figure 1. Graphical user interface of the Grashof module for the four-bar linkage mechanism

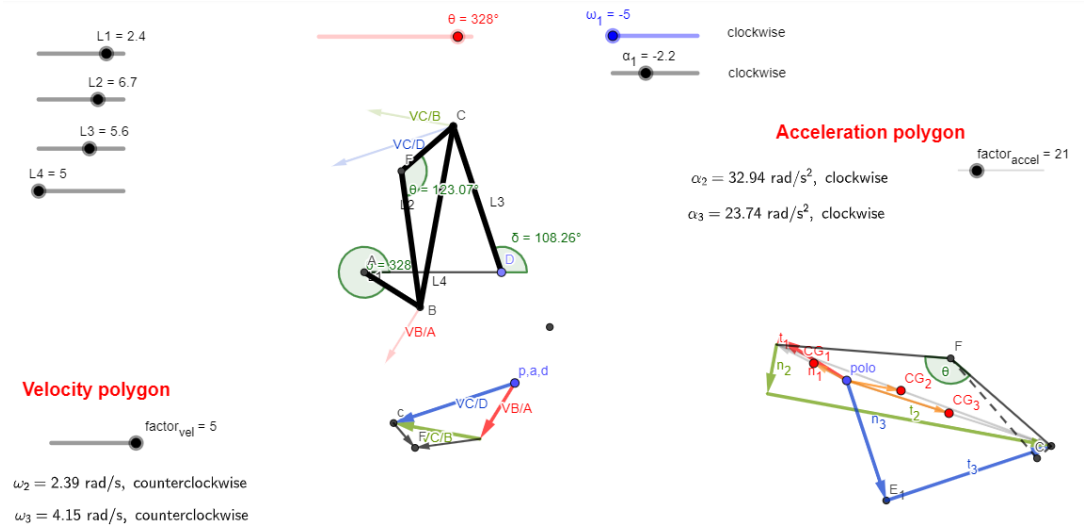


Figure 2. Graphical user interface for the kinematic analysis of a four-bar linkage mechanism

In the sections related to graphical kinetostatics, the process begins with the calculation of inertial loads added to the mechanism, allowing students to track how these loads change as the geometric parameters are varied. Subsequently, the forces within the kinematic pairs are calculated for each Assur group. Finally, the compensatory moment in the primary or input link of the mechanism is determined. Fig. 4 displays the kinetostatic calculation module for the slider-crank mechanism. As previously mentioned, adjustments can be made to both the mechanisms' dimensions and the kinematic variables of the input element (*e.g.*, angular velocity and acceleration).

Practice exercise

1. For the Four-Bar Linkage Mechanism presented in the Geogebra app, set the links (L1 to L4) to their minimum values.

2. Set angular velocity ω_1=1 rad/s and angular acceleration $\alpha_1 = 0.2$ of the crank, and vary its angle of rotation θ.
   - Check the change of the directions and magnitudes of the velocity of points B and C of the mechanisms. You can adjust the velocity scale factor
   - Verify the perpendicularity of the relative velocities $V_{C/B}$ and $V_{C/D}$ with the lines BC y CD respectively.

   For velocity direction $v_{C/B}$. What is the direction of the angular velocity of link 2, clockwise or counterclockwise?

   Fotr velocity direction $v_{C/D}$, What is the direction of the angular velocity of link 3, clockwise or counterclockwise?

   - Check the directions of normal and tangential accelerations of point C relative to points B and D. You can adjust the accelerations scaling factor.

   What are the directions of the normal accelerations?

   What directions do the tangential accelerations have?

   What direction of rotation does the angular acceleration of link 2 carry?

   What direction of rotation does the angular acceleration of link 3 carry?

3. Set the crank angle $\theta = 0°$.

Figure 3. Follow-up practice guidelines related to kinematic analysis of a four-bar linkage

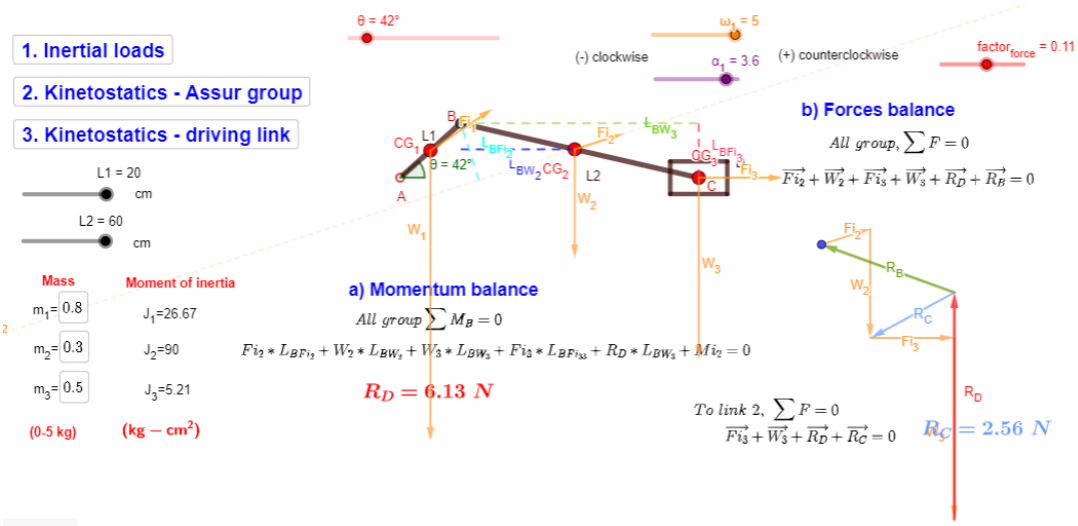


Figure 4. Graphical user interface for the kinetostatics of the slider crank mechanism

Parameterizing constraint forces using these graphical methods is complex and requires numerous auxiliary variables to determine the magnitudes, directions, and signs of the forces and moments analyzed in each section. The parametric graphical method implemented in this VL allows users to instantly visualize the points where the maximum forces occur. This real-time visualization facilitates the analysis of variations in kinematic or dynamic parameters.

The developed VL also addresses the kinematics of gear trains. Within the gear practice modules, students determine transmission ratios and analyze speed variations across different types of gear trains: simple, compound, and epicyclic. Fig. 5 illustrates two of the implemented modules.

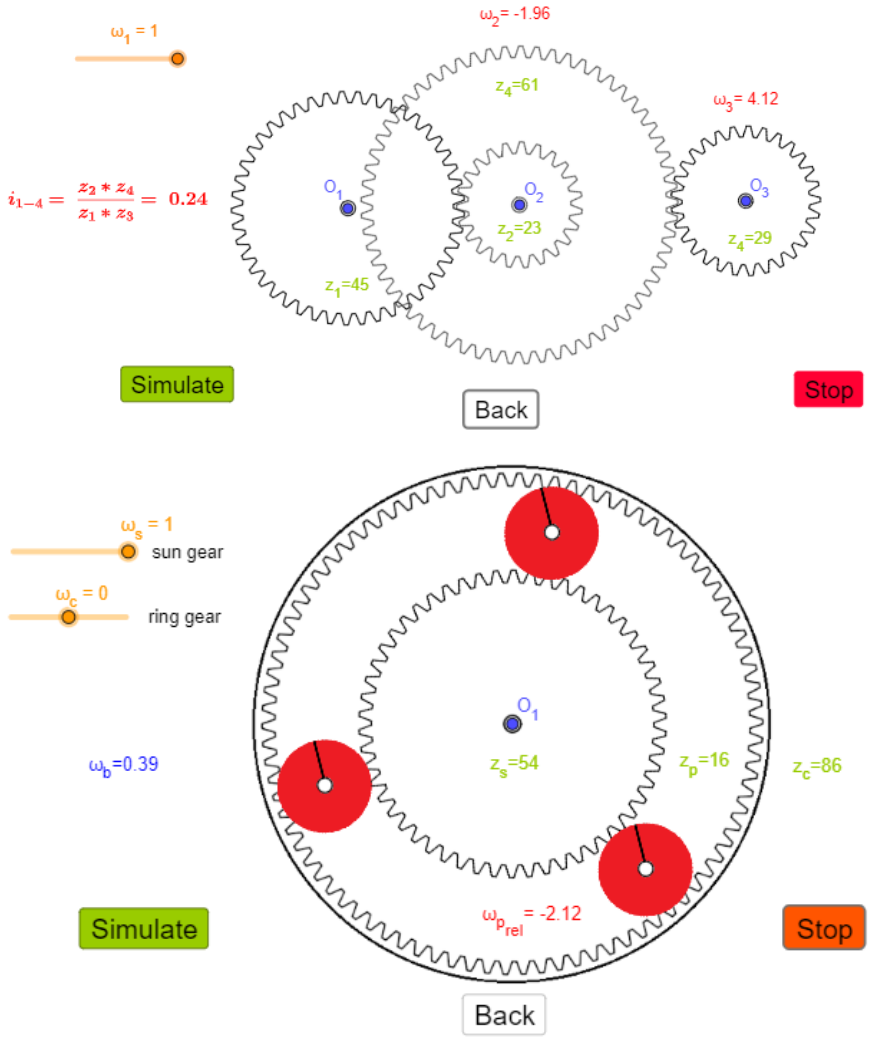


Figure 5. Graphical user interface of the module for compound and epicyclic gear trains

Regarding graphical dimensional synthesis, the exercises are designed to let students select the type of synthesis they wish to perform and modify pre-defined design points. Construction lines are displayed to guide them in following the explanations provided in the documentation, facilitating a step-by-step approach to the requested synthesis, as presented in Fig. 6.

A crucial parameter in dimensional synthesis is the minimum achievable transmission angle of the synthesized mechanism, which serves as an indicator of mechanical advantage. In graphical synthesis, this value can be determined upon completion of the design process. This application enables students to quickly generate alternative mechanisms by adjusting points of interest, modifying mechanism frames, or altering assumed link lengths. Such adjustments help to optimize motion transmission while considering factors like Grashof's law compliance and transmission angle minimization.

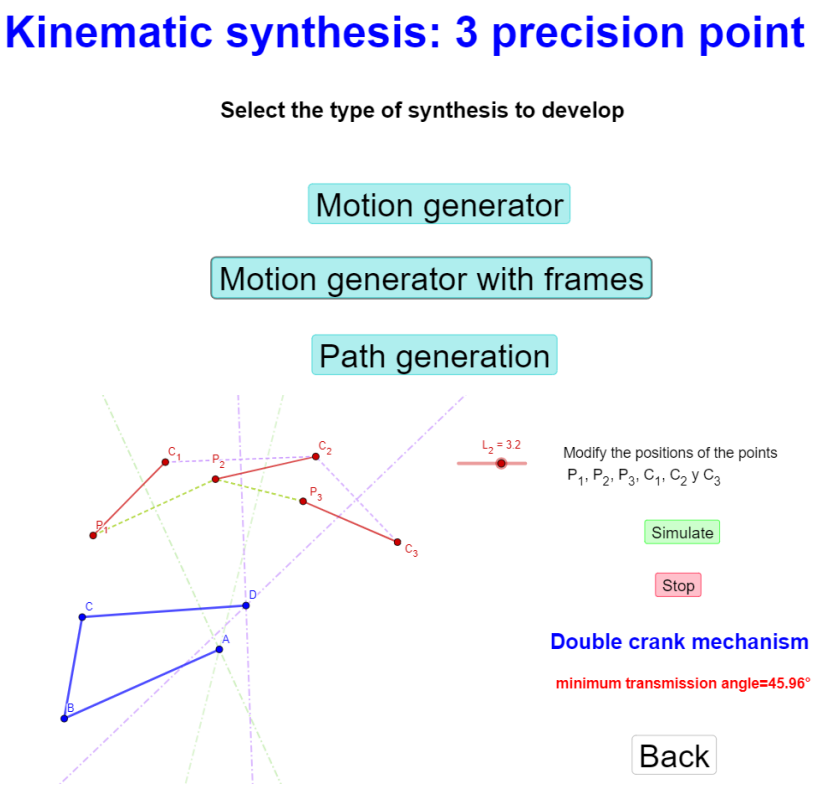


Figure 6. Graphical user interface of the dimensional synthesis module (using 3 precision points)

In the context of analytical methodologies, software selection should consider the students' level of training within the courses in *Theory of Machines and Mechanisms, and Mechanisms Synthesis.* Since students have not yet completed their coursework in programming and numerical methods, Matlab – which is available online on our institution's platform – is preferred. Matlab offers built-in functions for solving nonlinear equation systems (*solve*, *fsolve*), which are essential in position analysis. Additionally, the availability of Matlab Online ensures that students can access the software from any location at any time, significantly enhancing the VL's role as a self-training and self-learning tool.

For analytical methodology practices, interactive executable documents (Matlab Live) have been developed. These documents integrate program code, images, and explanatory texts to simplify the understanding of the implemented algorithms. Fig. 7 presents one of the executable documents developed within the VL.

To facilitate access to Matlab-based practices, dedicated sections with access instructions have been provided on the Moodle platform. These sections allow students to download the necessary files and store them in their personal Matlab Drive folder. Students can work on these files remotely using Matlab Online. Fig. 8 presents the access instructions for one of the practice exercises.

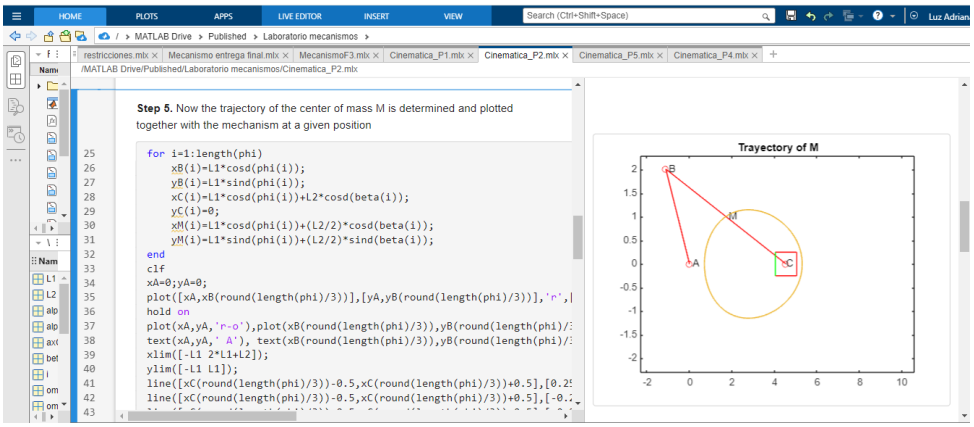


Figure 7. Executable document for the kinematic analysis of a slider-crank mechanism

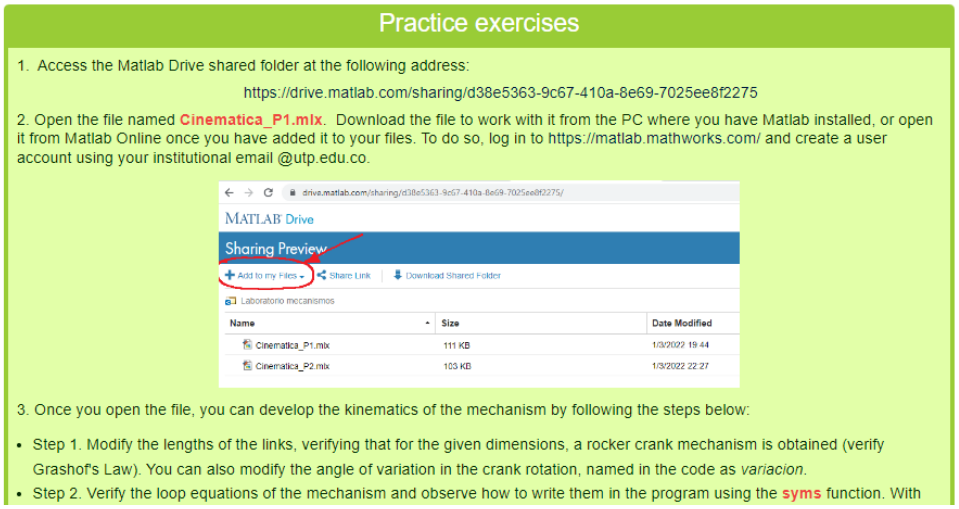


Figure 8. Follow-up practice guidelines for accessing analytical practices

Each practice includes documentation detailing the formulation of analytical expressions required for kinematic and kinetostatic analyses. For instance, Fig. 9 presents selected sections of the development of a four-bar linkage mechanism.

2.4. Implementation

The virtual laboratory has been implemented on the Moodle platform, utilizing tabs in order to allow students to seamlessly navigate through various practice exercises. Each tab contains sections with detailed descriptions of the mechanism and the corresponding analyses – both kinematic and kinetostatic – as illustrated in Fig. 10. Within each section, students can access supporting documentation, interactive applets (created using GeoGebra), and executable documents (created in Matlab) to facilitate practice.

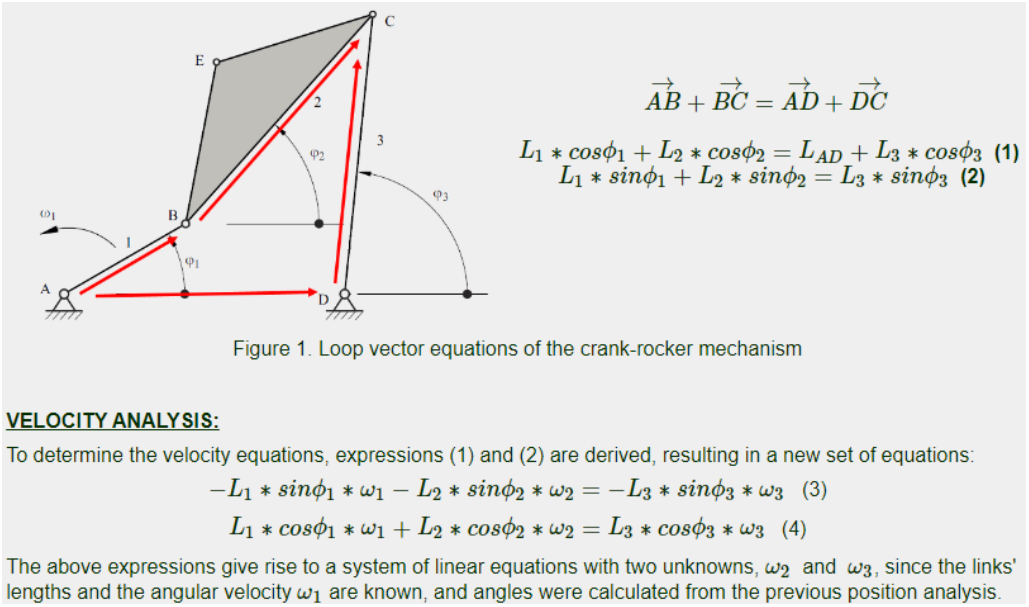


Figure 9. Documentation for the kinematic analysis of a practice within the VL

Figure 10. Implementation of the VL on the Moodle platform

2.5. Evaluation

The assessment of a digital educational resource's value to users is essential to determine its impact on the learning process, along with other factors that define its quality. Evaluation becomes particularly relevant once the digital resource is in use. Various educational software evaluation models can be found in the literature, with notable contributions made by (23), (24), and (25). In general, these evaluations aim to measure both the technical and educational quality of the material. Table II presents a summary of the criteria incorporated into the evaluation instrument.

Table II. Features for evaluating educational software

Category	Dimension	Features
General information	Documentation	The purpose and objectives are clearly stated. Instructions are clear and easily understandable.
	Technical and pedagogic aspects	*System requirements.* Adequate technical specifications are provided. *Ease of use.* The software is user-friendly and intuitive. Information accuracy. The content is accurate and reliable. *Appropriate difficulty level*. The software aligns with the intended level of difficulty.
Specific information	Software for practice	Provides flexible timing options for practice. Generates interest and engagement among users. Offers clear and comprehensible instructions. Includes realistic and relevant examples for practice.

The main page of the VL features a section for user feedback and general recommendations, as well as a digital form that students and other users can complete. Table **III** presents a selection of the questions included in this form.

Table III. Sample questions from the VL digital evaluation form

Question	Rating
Are the objectives of the VL clearly documented?	1-5
Is the planning process evident in the lab's development, including its sequence and use of computational resources?	1-5
Is the content of the VL aligned with the course objectives?	1-5
Does the sequence of practice exercises facilitate an appropriate study of the topics?	1-5
Is the practical material useful for the course?	1-5
How would you rate the presentation and organization?	1-5
Is there a clear connection between practical and theoretical content?	1-5
Was the allocated time sufficient to complete the exercises in their entirety?	1-5
Did you face difficulties in completing the exercises due to a lack of previous knowledge in the studied subjects?	1-5
How easy was it to access the exercises?	1-5
Were the computer programs used in the practice exercises appropriate?	1-5
Please mention some positive aspects of the VL:	Open answer
Please mention some negative aspects of the VL:	Open answer
Do you have any recommendations for improving the VL?	Open answer

3. Results and discussion

This evaluation facilitated the identification of the key challenges and benefits associated with the VL. Fig. 11 presents some of the categorized observations. Notably, users assign high ratings to the documentation provided in the laboratory. Additionally, the ease of access and the perceived value of the contents in student's academic training are evident.

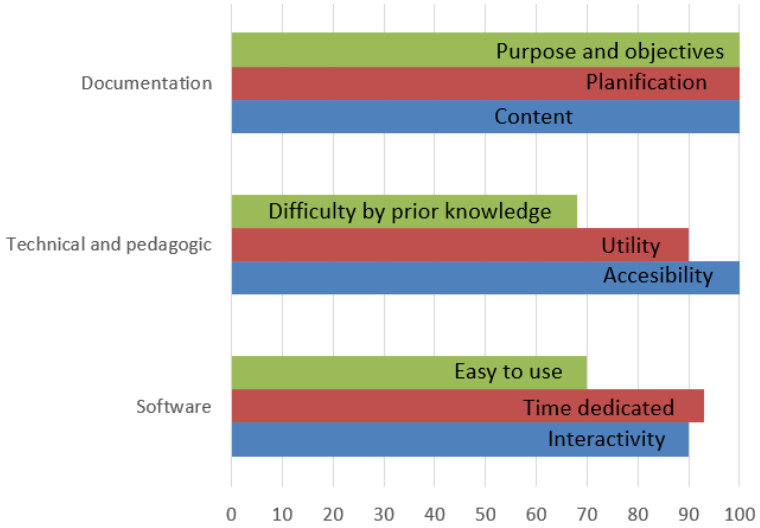


Figure 11. Some evaluative aspects of the VL

The findings are consistent regarding several aspects with those reported in a comprehensive study on the effectiveness of VLs in engineering education (26). This study aggregated data from 46 individual investigations extracted from 22 publications. Various educational outcomes were evaluated using the Hedges g, a widely accepted effect size metric employed in meta-analysis and statistical research to quantify the standardized difference between two groups, such as experimental and control groups.

According to conventional interpretation, a Hedges g value of approximately 0.2 indicates a small effect, 0.5 a moderate effect, and values greater than 0.8 a large effect.

One of the most notable results of the study is the strong positive impact of VLs on student motivation and engagement (Hedges g=3.571), a perception that was also reflected in the evaluation of the VL developed in our work. The reported usefulness and the favorable feedback provided by students support the conclusion that VLs foster greater student involvement in the learning process.

In contrast, a more modest effect on academic performance and skill acquisition was reported, with effect sizes ranging between 0.6 and 0.7, as noted in (26). This trend was also observed in the feedback from students participating in the VL on mechanisms: a considerable number of individuals indicated difficulties stemming from insufficient prior knowledge, particularly regarding computational tools for solving equations and the analytical modeling of mechanical systems.

Additionally, (26) identified a negative cognitive load effect in certain scenarios (Hedges g=-2.107), suggesting that students may struggle with information overload. A comparable issue was observed in our implementation, which was primarily related to the extensive time investment required and the lack of prior training in the software tools used in the practical exercises.

Nevertheless, in terms of content presentation and instructional design, students expressed positive attitudes towards the interactivity and flexibility offered by the virtual environment.

4. Conclusions

This paper presented a virtual laboratory designed to enhance learning in courses on the analysis and synthesis of mechanisms within undergraduate mechanical engineering programs. The development methodology integrated Geogebra, Moodle, and MatLab platforms, covering the entire process from conceptual design to implementation. The design adhered to the learning objectives of the *Theory of Mechanisms and Machines and Mechanisms Synthesis* courses, emphasizing content organization, user interactivity, autonomous learning, and learning outcomes assessment.

As a new initiative at our institution, the VL currently includes 12 practical exercises focused on the analysis and synthesis of planar mechanisms. The digital platform allows for future expansion by users and provides a framework for developers to apply user feedback for iterative improvement.

The effectiveness of the VL as a pedagogical tool to support and enhance both teaching and learning has been evaluated during its initial deployment in courses on mechanisms. Our VL empowers students to better understand concepts by visualizing kinematic and dynamic variables in various types of mechanisms. It serves as an alternative and complementary tool to traditional lectures, supporting active learning methodologies that foster student engagement. Students can progress at their own pace, observing how geometric and functional changes affect the kinematic and dynamic behavior of mechanisms. User feedback highlights the benefits of repeating exercises as needed, as well as the convenience of accessing the platform at any time, from any computer.

In response to the evolving demands of virtual and blended education, the virtual laboratory promotes the integration of e-learning strategies. Moreover, acknowledging the necessity of customized pedagogical tools, the platform allows instructors to create their own theoretical materials and manage the simulation environment. Accompanying didactic guides with a constructivist approach further supports the achievement of targeted learning outcomes in courses on mechanism analysis and synthesis.

5. Author contributions

Luz Adriana Mejía: conceptualization, methodology, software, formal analysis, writing (original draft, review, and editing).

Carlos Alberto Romero: conceptualization, supervision, writing (original draft, review, and editing)

References

[1] W. Ali, "Online and remote learning in higher education institutes: A necessity in light of COVID-19 Pandemic," *Higher Ed. Stud.*, vol. 10, no. 3, art. 16, 2020. https://doi.org/10.5539/hes.v10n3p16 ↑3

[2] United Nations Educational, Scientific, and Cultural Organization, *Report of the Expert Meeting on Virtual Laboratories*. Paris, France: UNESCO, 2000. ↑3

[3] C. Infante, "Propuesta pedagógica para el uso de laboratorios virtuales como actividad complementaria en las asignaturas teórico-prácticas," *Rev. Mex. Inv. Ed.*, vol. 19, no. 62, pp. 917-937, 2014. http://www.redalyc.org/articulo.oa?id=14031461013 ↑3

[4] A. Lorandi, G. Hermida, J. Hernández, and E. Ladrón de Guevara, "Los laboratorios virtuales y laboratorios remotos en la enseñanza de la ingeniería," *Rev. Int. Ed. Ing.*, vol. 4, pp. 24-30, 2011. ↑3

[5] J. Cabrera Medina and I. Sánchez Medina, "Laboratorios virtuales de física mediante el uso de herramientas disponibles en la Web," *Mem. Cong. UTP*, vol. 1, no. 1, pp. 49-55, 2016. https://revistas.utp.ac.pa/index.php/memoutp/article/view/1296 ↑4

[6] R. Radhamani *et al.*, "What virtual laboratory usage tells us about laboratory skill education pre- and post-COVID-19: Focus on usage, behavior, intention and adoption," *Educ. Inf. Technol.*, vol. 26, pp. 7477-7495, 2021. https://doi.org/10.1007/s10639-021-10583-3 ↑4

[7] L. Mishra, T. Gupta, and A. Shree, "Online teaching-learning in higher education during lockdown period of COVID-19 pandemic," *Int. J. Ed. Res. Open*, vol. 1, art. 100012, 2020. https://doi.org/10.1016/j.ijedro.2020.100012 ↑4

[8] F.D. Syahfitri *et al.*, "The development of problem based virtual laboratory media to improve science process skills of students in biology," *Int. J. Res. Rev.*, vol. 6, no. 6, pp. 64-74, 2020. https://doi.org/10.20961/ijpte.v2i0.24952 ↑4

[9] N. Kapilan, P. Vidhya, and X.Z. Gao, "Virtual laboratory: A boon to the mechanical engineering education during COVID-19 pandemic," *High. Ed. Fut.*, vol. 8, no. 1, pp. 31-46, 2021. https://doi.org/10.1177/2347631120970757 ↑4

[10] K. Achuthan, P. Nedungadi, V. Kolil, S. Diwakar, and R. Raman, "Innovation adoption and diffusion of virtual laboratories," *Int. J. Online Biomed. Eng.*, vol. 16, no. 9, pp. 4-25, 2020. https://doi.org/10.3991/ijoe.v16i09.11685 ↑4

[11] G. Singh, A. Mantri, O. Sharma, and R.,Kaur, "Virtual reality learning environment for enhancing electronics engineering laboratory experience," *Comp. App. Eng. Ed.*, vol. 29, no. 1, pp. 229-243 2021. https://doi.org/10.1002/cae.22333 ↑4

[12] N. Tuli, G. Singh, A. Mantri, and S. Sharma, "Augmented reality learning environment to aid engineering students in performing practical laboratory experiments in electronics engineering", *Smart Learn. Environ.*, vol. 9, art. 26, 2022. https://doi.org/10.1186/s40561-022-00207-9 ↑4

[13] C.Hao, A. Zheng, Y.Wang, and B.Jiang, "Experiment information system based on an online virtual laboratory," *Future Internet*, vol. 13, no. 2, art. 27, 2021. https://doi.org/10.3390/fi13020027 ↑4

[14] F. Vahdatikhaki, I. Friso-van Den Bos, S. Mowlaei, and B. Kolloffel, "Application of gamified virtual laboratories as a preparation tool for civil engineering students," *Eur. J. Eng. Ed.*, vol. 49, no. 1, pp. 164-191, 2024. https://doi.org/10.1080/03043797.2023.2265306 ↑4

[15] Y. Xiaoju *et al.*, "Teaching practice of wind turbine practical training based on virtual teaching platform," *Comp. App. Eng. Ed.*, vol. 32, no. 2, art. e22710, 2024. https://doi.org/10.1002/cae.22710 ↑4

[16] A. Sánchez-López, J. Jáuregui_Jáuregui, N. García-Carrera, and Y. Perfecto-Avalos, "Evaluating effectiveness of immersive virtual reality in promoting students' learning and engagement: A case study of analytical biotechnology engineering course," *Front. Ed.*, vol. 9, art. 1287615, 2024. https://doi.org/10.3389/feduc.2024.1287615 ↑4

[17] V. Kolil and K. Achuthan, "Virtual labs in chemistry education: A novel approach for increasing student's laboratory educational consciousness and skills", *Ed. Info. Technol.*, vol. 29, pp. 25307-25331, 2024. https://doi.org/10.1007/s10639-024-12858-x ↑4

[18] S. Dong, F. Yu, and K. Wang, "A virtual simulation experiment platform of subway emergency ventilation system and study on its teaching effect", *Sci. Rep.*, vol. 12, art. 10787, 2022. https://doi.org/10.1038/s41598-022-14968-3 ↑4

[19] J. L. Calderón, "Aplicación de GeoGebra en la enseñanza de la cinemática de un mecanismo de cuatro barras," *Rev. Inst. GeoGebra São Paulo*, vol. 9, no. 2, pp. 3-19, 2020. http://dx.doi.org/10.23925/2237-9657.2020.v9i2p003-019 ↑4

[20] B. Stahre Wästberg *et al.*, "Design considerations for virtual laboratories: A comparative study of two virtual laboratories for learning about gas solubility and colour appearance," *Ed. Info. Technol.*, vol. 24, pp. 2059-2080, 2019. https://doi.org/10.1007/s10639-018-09857-0 ↑5

[21] V. M. Gándara, "Usos educativos de la computadora," in *El proceso de desarrollo y el diseño de interfaz al usuario*, J. M. Álvarez and A. M. Bañuelos, Eds., Ciudad de México, Mexico: UNAM, 1994. ↑5

[22] A. Joshi, M. Vinay, and P. Bhaskar, "Impact of coronavirus pandemic on the Indian education sector: perspectives of teachers on online teaching and assessments," *Inter. Technol. Smart Ed.*, vol. 18, no. 2, pp. 205-226, 2021. https://doi.org/10.1108/ITSE-06-2020-0087 ↑6

[23] A. Cova, X. Arrieta, and V. Riveros, "Análisis y comparación de diversos modelos de evaluación de software educativo," *Rev. Ven. Info. Tecnol. Conoc.*, vol. 5, no. 3, pp. 45-67, 2008. ↑13

[24] O. González, M. A. Aguilar, F. J. Aguilar, and M. L. Matheu, "Evaluación de entornos inmersivos 3D como herramienta de aprendizaje B-Learning," *Ed. XX1*, vol. 21, no. 2, pp. 417-440, 2018. https://doi.org/10.5944/educxx1.16204 ↑13

[25] Y. Arguelles, "Metodología para la evaluación de laboratorios virtuales en entornos 3D," M.S. thesis, Universidad de Moa, Cuba, 2021. ↑13

[26] J. Li and W. Liang, "Effectiveness of virtual laboratory in engineering education: A meta-analysis," *PLoS ONE*, vol. 19, no. 12, art. e0316269, 2024. https://doi.org/10.1371/journal.pone.0316269 ↑15, 16

Luz Adriana Mejía Calderón

She received her degree in Mechanical Engineering from Universidad Tecnológica de Pereira (1999). She holds an MSc in Automatic Production Systems from the same university (2005), as well as an MSc in Mechanical Engineering (2009) and a PhD in Engineering (2016) from the Polytechnic University of Valencia, Spain. She is currently a lecturer of the Faculty of Applied Mechanics at Universidad Tecnológica de Pereira. Her research interests include the modeling and simulation of multibody systems, machine and mechanism theory, and robotics.
Email: adriamec@utp.edu.co

Carlos Alberto Romero Piedrahita

He is an automotive engineer and holds an MSc from the Polytechnic Institute of Belarus (1986). He obtained a Specialist degree in Physical Instrumentation from Universidad Tecnológica de Pereira in 1992, an MSc in Internal Combustion Engines from the Thermal Engines Center of the Polytechnic University of Valencia (Spain) in 2007, and a PhD in Propulsion Systems for Transport Vehicles from the same university in 2009. He is currently a professor of the Faculty of Technology at Universidad Tecnológica de Pereira. His research areas include internal combustion engines, automotive design, and machine design.
Email: cromero@utp.edu.co

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.22458

Research

Study Comparing Empirical Data on Sensors used to Measure Obstacle Distance

Estudio comparativo de datos empíricos obtenidos de sensores utilizados para medir la distancia de obstáculos

Johan Leandro Téllez-Garzón[1]*, Jorge Saúl Fandiño-Pelayo[1], Antoine Bernard[2], and Giovanni Mazzini[2]

[1]Unidades Tecnológicas de Santander (Bucaramanga, Colombia)

[2]Polytech Aix-Marseille Université (Marseille, Francia)

Abstract

Context: Electronic sensors play a crucial role in different applications such as robotics or industrial or home automation. Sensors can measure essential environmental variables in order to feed digital signal processing algorithms and perform actions more efficiently. The sensors used for distance measurements follow different approaches. However, it is difficult to find a study with performance comparisons.

Method: An empirical study was performed to evaluate and compare the performance of ultrasonic and infrared sensors in frontal and lateral detection situations while considering distance and angle variations.

Results: The results show that the ultrasonic sensor detects the distance with good accuracy along its operational range. However, the distance measure is inaccurate when the obstacle is not orthogonal to the sensor. The ultrasonic sensor showed high accuracy in long-range, frontal obstacle detection, while the infrared sensor performed better at short distances with angled obstacles. Statistical analysis confirmed strong linear correlations, especially for the ultrasonic sensor, supporting the complementary use of both sensors in distance measurement applications.

Conclusions: An evaluation of ultrasonic and infrared sensors for distance measurement in applications involving robotics and the Internet of Things revealed that the former are more reliable for distant, orthogonal obstacles, while the latter perform better at short distances on angled surfaces, highlighting their complementary strengths and the need for future improvements to address environmental sensitivity and detection limitations.

Keywords: Internet of Things, robotics, obstacle detection, ultrasonic sensor, infrared sensor

Article history

Received: July 8th, 2024

Modified: December 2nd, 2024

Accepted: March 15th, 2025

Ing, vol. 30, no. 1, 2025, e22458

* **Correspondence:** jtellez@correo.uts.edu.co

Resumen

Contexto: Los sensores electrónicos tienen un papel crucial en diferentes aplicaciones como la robótica o la automatización industrial o doméstica. Los sensores pueden medir variables ambientales esenciales para alimentar algoritmos digitales de procesamiento de señales y realizar acciones de una manera más eficiente. Los sensores que se utilizan para mediciones de distancia siguen enfoques diferentes. Sin embargo, es difícil encontrar un estudio con comparaciones de desempeño.
Método: Se realizó un estudio empírico para evaluar y comparar el rendimiento de sensores ultrasónicos e infrarrojos en situaciones de detección frontal y lateral, considerando variaciones de ángulo y distancia.
Resultados: Los resultados muestran que el sensor ultrasónico detecta la distancia con buena precisión a lo largo de su rango operacional. Sin embargo, la medida de distancia se torna inexacta cuando el obstáculo no está ortogonal al sensor. El sensor ultrasónico mostró gran precisión en la detección de largo alcance de obstáculos frontales, mientras que el sensor infrarrojo se desempeñó mejor en distancias cortas con obstáculos en ángulo. Un análisis estadístico confirmó la presencia de correlaciones lineales fuertes, especialmente en el caso del sensor ultrasónico, lo que respalda el uso complementario de ambos sensores en aplicaciones de medición de distancias.
Conclusiones: La evaluación de sensores ultrasónicos e infrarrojos para la medición de distancias en aplicaciones que involucran robótica y el Internet de las cosas reveló que los primeros son más fiables para obstáculos distantes y ortogonales, mientras que los segundos funcionan mejor en distancias cortas y con superficies en ángulo, resaltando sus fortalezas complementarias y la necesidad de futuras mejoras para abordar la sensibilidad ambiental y otras limitaciones de detección.

Palabras clave: Internet de las Cosas, robótica, detección de obstáculos, sensor ultrasónico, sensor infrarrojo

Table of contents

1. Introduction

Scientific advances have allowed helping people in their daily life and work, as well as improving their safety. Advances in robotics and the development of associated sensors are fundamental to improving the efficiency of human activities in different contexts (*e.g.* industrial, commercial, and educational, among others). The use of sensors to detect obstacles or objects is key for some applications requiring autonomous vehicles (drones or ground robots) to perform risky tasks such as deactivating land mines or even rescuing living beings in earthquakes or landslides. In this vein, it is necessary to understand the accuracy of measurements carried out by ultrasonic and infrared sensors in order to adjust the operation of autonomous vehicles. Moreover, this type of sensors can be used to assist in and improve different tasks like obstacle avoidance for bipedal robots (1), route planning for agricultural inspection robots (2), unmanned aerial vehicle control (3), and environmental analysis in coffee crops (4). In the context of cognitive device operation, the accuracy of sensor measurements is a critical factor in decision-making processes (5).

Some work has been carried out to investigate the operation and accuracy of ultrasonic and infrared sensors, as detailed below.

The study by (6) presents a method that considers the infrared reflectance properties of a surface to estimate distance. However, it requires other sensing modalities to obtain information on the distance to the obstacle. Experiments indicated that this type of low-cost sensor can provide reliable distance measurements, and the results showed agreement between the Phong illumination model and real data from testing.

The work by (7) describes the use of wearable obstacle detection systems for safe human locomotion. The focus of this review is gait-assisting wearable sensors and hazard detection for pedestrians. This study analyzes the sensors used for detecting risks to pedestrians, highlighting the major advantages and disadvantages of each approach, analyzing the sensors and hardware components used, and explaining their potential for improving the quality of life of individuals with locomotor disabilities.

A performance study of ultrasonic and infrared sensors in a wireless obstacle detection system is conducted in (8). These sensors aim to enhance walking safety by identifying the risk of tripping and providing corrective feedback.

In (9), ultrasonic and infrared sensors are evaluated with the purpose of determining their operational performance regarding distance measurement. However, aspects such as detection angle and statistical correlation analysis are not addressed, as the focus is limited to validating their basic functionality.

In (10), the operation of infrared sensors is analyzed while considering the influence of light. The performance of ultrasonic sensors under the effects of temperature and humidity is also evaluated.

A comparison between ultrasonic and infrared sensors is made in (11) while considering obstacles with different materials. Obstacle distance detection performance is evaluated for cardboard, paper sheet, sponge, wood, plastic, rubber, and tile. This study is significant in that it aids in selecting the best combination of sensors to achieve good results in robotics applications.

In (12), an evaluation of ultrasonic and infrared sensors for mobile robot navigation is carried out. This work considers the wave interaction between the human body and metal surfaces. In terms of the root mean squared error, the results show complementary behaviors for the two sensors used, evidencing the need for data fusion strategies to minimize the positional error of obstacles for effective robotic navigation (12). In addition, the motion efficiency of a mobile robot equipped with an ultrasonic sensor decision-making algorithm is studied in (13), considering the time required to follow a predefined test path.

Ultrasonic and infrared sensors are very popular in some applications. For example, the design and reaction time required to find targets for a firefighter robot equipped with these types of systems are described in (14). Said robot can assist with activities that pose serious risks to human firefighters.

Most works have failed to present a detailed study on the performance of ultrasonic and infrared sensors for frontal or lateral situations (15). Although previous reports have compared these sensors under different conditions, they often focus on specific applications or general accuracy metrics without thoroughly analyzing their performance under varying obstacle orientations. This study differs by systematically evaluating these sensors in both frontal and lateral detection scenarios while considering the impact of angle and distance variations. Additionally, we incorporate a detailed correlation analysis between real and estimated distances, which is often overlooked in similar studies. This approach allows for a better understanding of each sensor's operational limits and its suitability for different applications. The findings not only complement existing research but also challenge certain assumptions regarding the accuracy and reliability of these systems under non-ideal conditions. In some cases, the literature is more focused on the study of the robotic system in general, without considering the effect of the sensors and the correlation of the measurements.

In this context, the proposed work provides a description of ultrasonic and infrared sensors which considers different experiments for detecting the distance of objects in frontal and lateral situations. A data analysis is performed to compute correlation coefficients, comparing the real distance against the estimated value. On the other hand, the HCSR04 ultrasonic sensor and the GP2Y0A021YK infrared sensor are compared with the purpose of determining the best option for specific use cases. This research aims to highlight the circumstances in which the sensors are effective while considering their use limits, as described in the existing data sheets. This work is organized into five sections, *i.e.*, 1) introduction, 2) methodology, 3) results, and 4) conclusion.

2. Methodology

2.1. Sensors and characteristics

Two sensor types (ultrasonic and infrared) were used in this work with the aim of identifying their performance in distance detection.

The ultrasonic sensor (HC-SR04), by Corporate Computer, has an operating range of 2 to 400 cm. It has two inputs and two outputs for VCC, TRIG, ECHO, and GND signals. The angle to the normal line of the obstacle surface cannot exceed 15°. Therefore, in this work, the obstacle had to be orthogonal to the surface of the sensor, *i.e.*, at a right angle (90°) to the sensing surface of the sensor. The ultrasonic sensor was fed with 3 V and was connected to ESP8266 device. The trigger input (TRIG) was used to send a series of pulses over eight cycles clocked at 40 kHz.

Fig. 1 shows the basic operation of the sensor. The ECHO output measures the time interval between sending the TRIG signal and receiving the ECHO signal, so the pulse width delivered by the output is used to determine the distance (16).

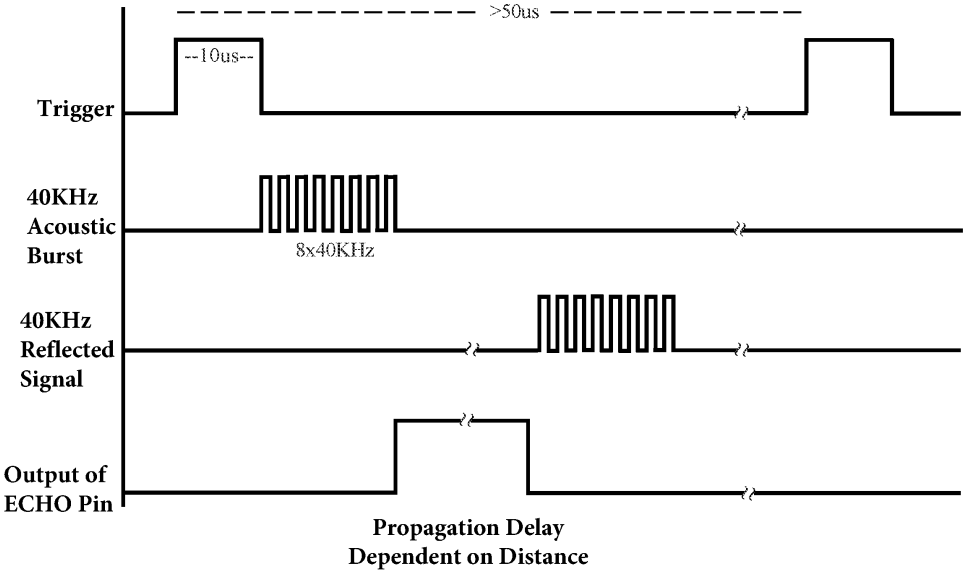


Figure 1. Timing diagram of signals in the ultrasonic sensor (17)

The ultrasonic sensor sends a TTL pulse with a duration of 10 us via the TRIG pin. A set of pulses clocked at 40 kHz are sent at the end of the TTL pulse (acoustic burst). Finally, the ECHO pin recovers the transmitted signal, and it sends a pulse with a specific width, which can take values from 150 to 25 ms. The distance d can be calculated using Eq. (1).

$$d = \frac{t}{59} \tag{1}$$

where, t is the time interval between sending the TRIG signal and receiving the ECHO signal (*i.e.*, the space between the acoustic bursts and the reflected signal). Two scenarios were analyzed for the

ultrasonic sensor: the first with the obstacle frontally aligned to the sensor (frontal) and the second with the obstacle non-frontally aligned (lateral), in order to obtain a map with the view field of sensor.

The GP2Y0A21YK0F infrared sensor by Sharp has an analog output and is controlled and fed by the ESP8266 device. The Vo output measures the time interval between sending the trigger signal and receiving it. Thus, the distance is calculated by the pulse width delivered by the output (18). The sensor performs various measurements, as shown in Fig. 2.

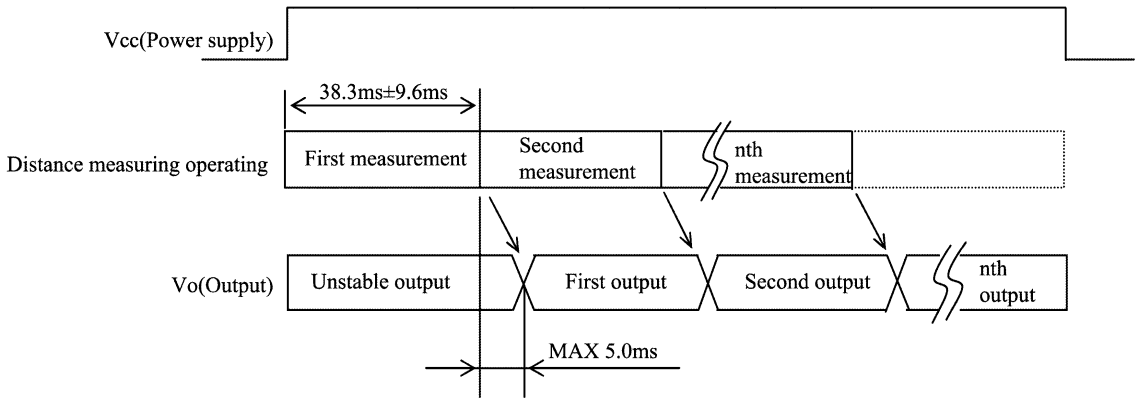


Figure 2. Timing diagram of signals in the infrared sensor (18)

To obtain the distance between sensor and obstacle, the equation to convert the ADC value into the estimated distance was determined. The theoretical curve provided by the manufacturer is shown in Fig. 3. It relates the distance to the object with the output voltage.

Furthermore, the behavior of the infrared sensor was determined. To this effect, a flat obstacle was positioned orthogonally to the sensor and then moved centimeter by centimeter. Fig. 4 shows the results of the measured distances vs. the ADC values for the experiment. These results provided the data curve and its representative Eq. (2).

$$d = 168869x^{-1,502} \tag{2}$$

This equation was used for the first study of the sensor. The numerical value given by the sensor was recorded during this research and compared against a theoretical distance. x represents the value generated by the ADC converter, which is denoted as $gp2y0a21Val$ in the program environment.

2.2. Experiments

The sensor's behavior was studied in two scenarios: frontal and lateral. Experiments were conducted in a controlled environment within an Internet of Things (IoT) laboratory, which has an approximate area of 6 × 6 m. No additional objects were present which could interfere with the measurements. The obstacle used was a 22 × 27 cm rectangle. In addition, for the infrared sensor, a flat white surface was used to optimize wave reflection.

To ensure the repeatability and precise control of distances and angles, a programmable system based on the ESP8266 was used. This system followed predefined trajectories, allowing for systematic

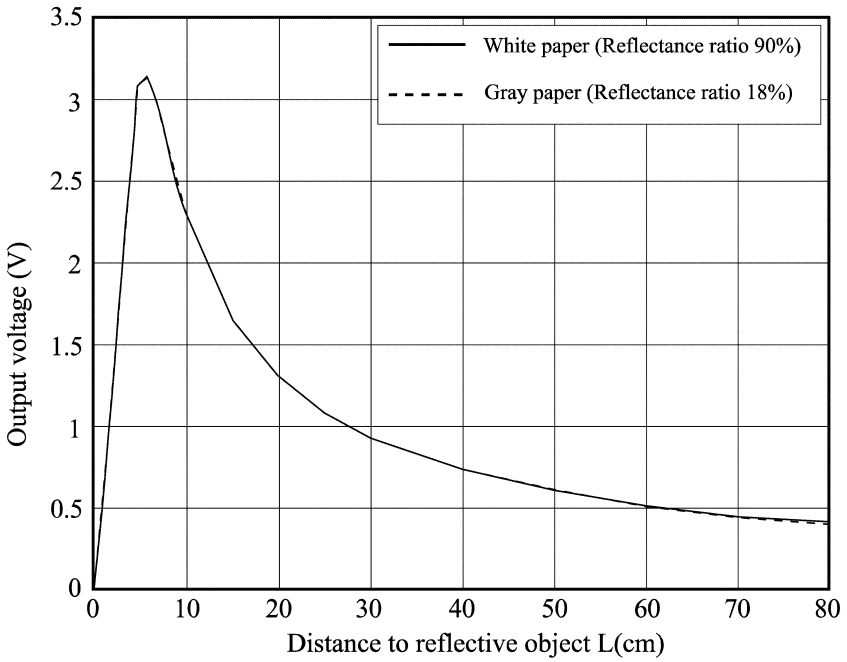


Figure 3. Timing diagram of signals in the ultrasonic sensor (18)

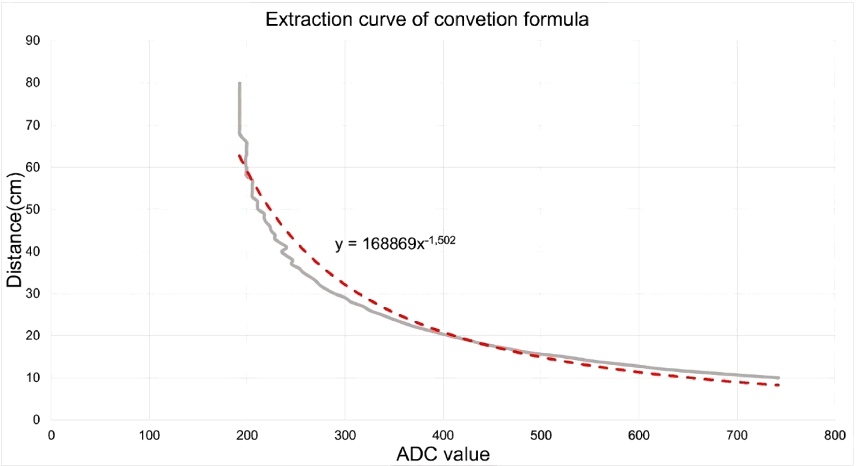


Figure 4. Computed and measured distances for different values of the ADC of sensor

measurements. Markers were placed on the ground at regular intervals to verify the actual distances between the sensor and the obstacle. Table I details the number of measurements taken for each point, with a total of 47 positions in 2 cm steps (0 to 92 cm) for the infrared sensor and 93 positions in 5 cm steps (0 to 460 cm) for the ultrasonic sensor. In this work, a test refers to an experiment involving measurements (10 per position point), and a *position point* is one of the possible locations of the obstacle relative to the sensor.

Table I. Details on the number of experiments

Type of sensor and situation	Infrared sensor, lateral	Infrared sensor, frontal	Ultrasonic sensor, lateral	Ultrasonic sensor, lateral
Number of tests performed	5	5	5	5
Number of measurements per point at each test	10	10	10	10
Total number of measurements per point	50	50	50	50
Total number of position points	47	47	93	93
Total number of measurements per scenario	**2350**	**2350**	**4650**	**4650**

2.2.1. Frontal situation

The first study focused on readings of the distance between the sensors and a frontally positioned obstacle. A class was created for each type of measurement and sensor. For each test, the program computed the average while considering 10 detected values to obtain a single measurement, with the purpose of reducing the impact of sensor fluctuations. For each measure in the ultrasonic sensor, the distance from the sensor to the obstacle was varied from 0 cm to 460 cm in 5 cm steps. The program recovered the propagation time in milliseconds, as well as the measured and actual distance in cm. For the infrared sensor, the distance between sensor and obstacle was varied from 0 cm to 92 cm. The maximum distance was that indicated by the manufacturer plus 15 % for both sensors. The program was used to extract the digital value of the sensor with the measured distance and compare it against the actual values. The values were stored in a dynamic array and presented as a text file in the following format: $time$; $d\_distance$; $d\_Realdistance$.

2.2.2. Lateral situation

To determine the detection response of sensors to non-frontal situations, the obstacle must be placed in the extension of the sensors and diagonally displaced to the right.

In this experiment, each obstacle is moved along the corresponding range, *i.e.*, from 0 to 460 cm (5 cm steps) for the ultrasonic sensor and from 0 cm to 92 cm (2 cm steps) for the infrared sensor. The program sends a notification when the obstacle is out of the measurement range and becomes invisible for sensor. Fig. 5 shows the positioning of the sensor and the obstacle. A real measurement of the distances is

performed for each point and fed to the program. This process is repeated for the subsequent distances. The program computes the sensor's angle with respect to the obstacle. Here, a right angle indicates whether the sensor is able to detect objects located along the diagonal, which is analogous to the left diagonal by the symmetry of sensors.

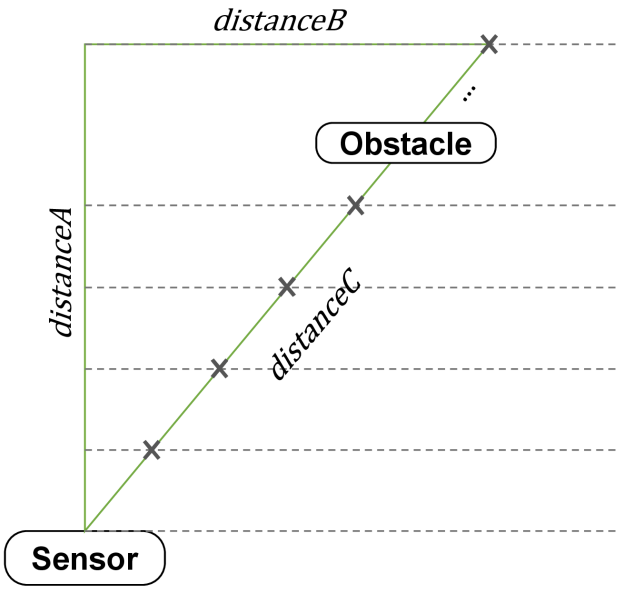


Figure 5. Lateral positioning of the obstacle

The values are stored in a dynamic array, and they are presented as a text file in the following format: d_time; $distanceA$; $distanceB$; $distanceC$; $angle$. The angle is obtained using Eq. (3).

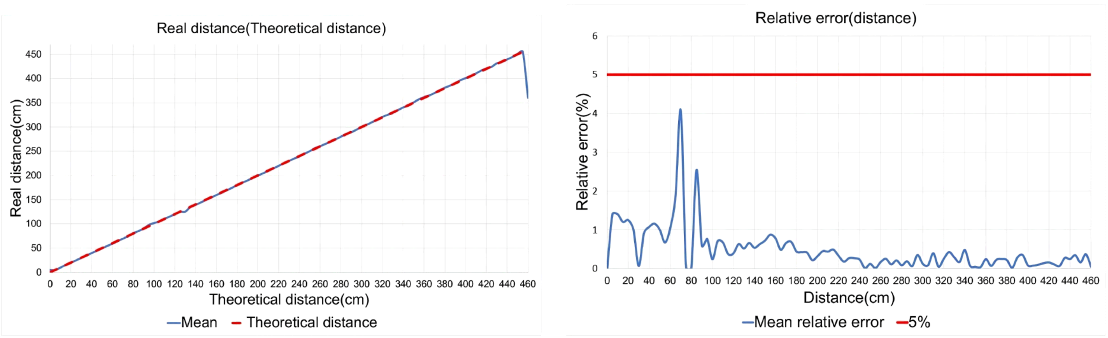
$$angle = \frac{1}{\pi} \tan^{-1} \left(\frac{distanceB}{distanceA} \right) \tag{3}$$

3. Results

3.1. Ultrasonic sensor

The HC-SR04 sensor can detect an obstacle if it is orthogonal to its field of view, with a theoretical angular tolerance of 15°. However, the sensor cannot identify an obstacle that is not orthogonal. The sensor is directional and provides accurate distance estimations in frontal scenarios, as shown in Fig. 6a. The distance measured in frontal situations (0° angle) shows excellent correspondence with the real distance, forming a linear relationship. Moreover, in most measurements (Fig. 6b), the error remains below 1 %, except for the estimation of shorter distances (0 to 90 cm), where it exceeds 1 % but stays below 5 %.

Studying the vision range allows understanding the sensor's ability to detect diagonal obstacles (Fig. 7). The results show better distance detection when the obstacle is located frontal to the sensor at



a) Distance measurement b) Relative error

Figure 6. Ultrasonic sensor in frontal situations

large distances close to the 400 cm limit. This has to do with the directivity of the ultrasonic sensor in radiating the power of signal, provided that it is not disturbed by its environment. Distance estimation can be performed at shorter distances, but only when the obstacle is diagonally positioned relative to the sensor. For example, it is possible to detect objects at 200 cm and 10°, or at 50-150 cm and 20°.

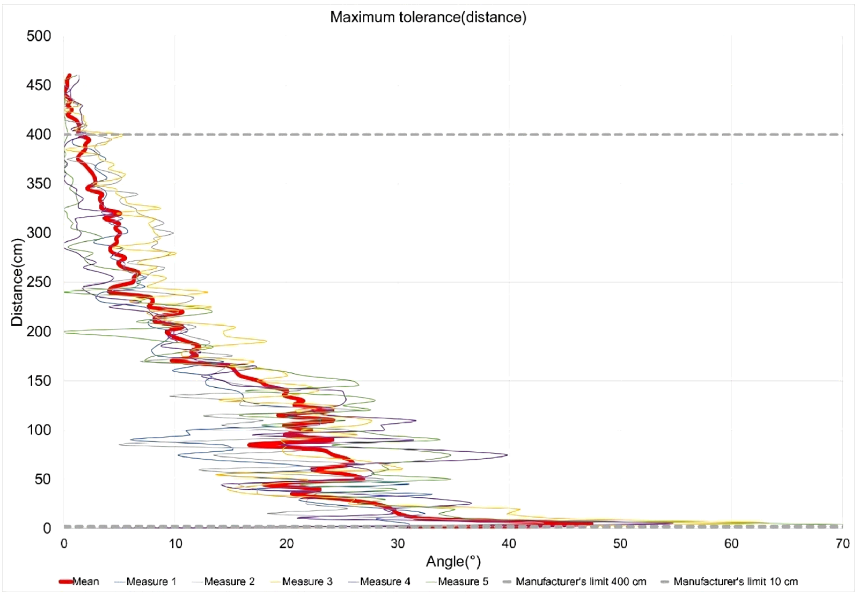


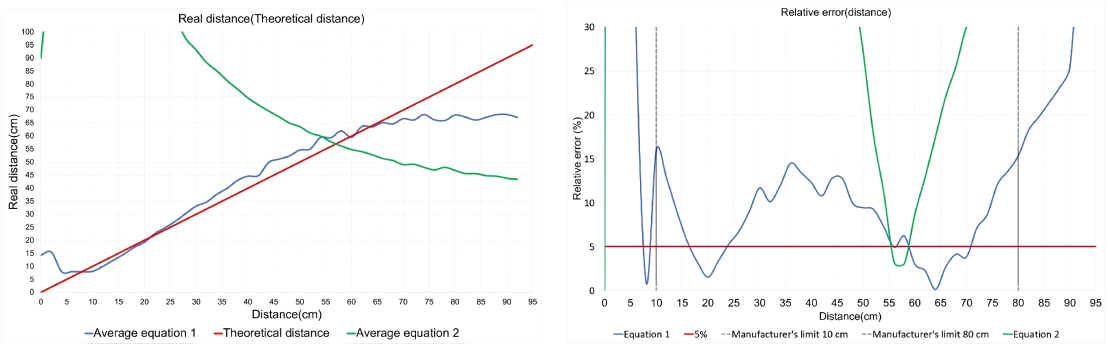
Figure 7. Ultrasonic sensor in lateral situations

The values show a defined tendency, without significant variations. Measurement accuracy is also a strength of this sensor; the error is low – almost negligible – along its range of use. The operating range specified by the manufacturer is therefore validated. However, the sensor is capable of operating beyond the theoretical limit.

Some advantages of the ultrasonic sensor include an operating range of 2 to 400 cm, good accuracy, high directivity, and ease of use. On the other hand, one of its drawbacks is its inability to detect obstacles that are significantly offset from its field of view, a phenomenon that becomes more pronounced as the distance increases.

3.2. Infrared sensor

The GP2Y0A021YK infrared sensor does not allow for obstacle detection at long distances when compared to its ultrasonic counterpart, since its detection range is limited to 80 cm according to the manufacturer's data. However, the sensor cannot be used throughout the recommended operating range. The experiment conducted in this study showed that the sensor can only detect distances between 10 and 60 cm. The use of Eq. (2) implies some error, which is more considerable for distances between 60 and 80 cm.



a) Distance measurement b) Relative error

Figure 8. Infrared sensor in frontal situations

Fig. 8 presents some tests carried out on the infrared sensor, altering the distance to the obstacle in frontal positions. Within the operating range (10 to 80 cm), estimation via Eq. (2) is adequate for applications requiring lower accuracy. However, in some tests (*e.g.*, 10, 37, 45, or 80 cm), the error is high (close to 15 %). As shown in the figure, distance measurement errors become excessively high outside the operating range.

On the other hand, Fig. 9 presents the sensor's performance in lateral scenarios. This sensor provides accurate distance estimations at angles of up to approximately 20° along most of its operating range. Even at short distances, the angle can reach nearly 30°. Different measures were taken, and a good correlation between them was observed. The red line in Fig. 9 shows the average for all measurements, representing the estimation tendency of the sensor.

The main advantage of the infrared sensor is its good detection of obstacles in lateral situations. However, it has some drawbacks, such as difficulties in obtaining a realistic equation for distance estimation, medium accuracy along the theoretical operational range, and low estimation distances (80 cm maximum limit).

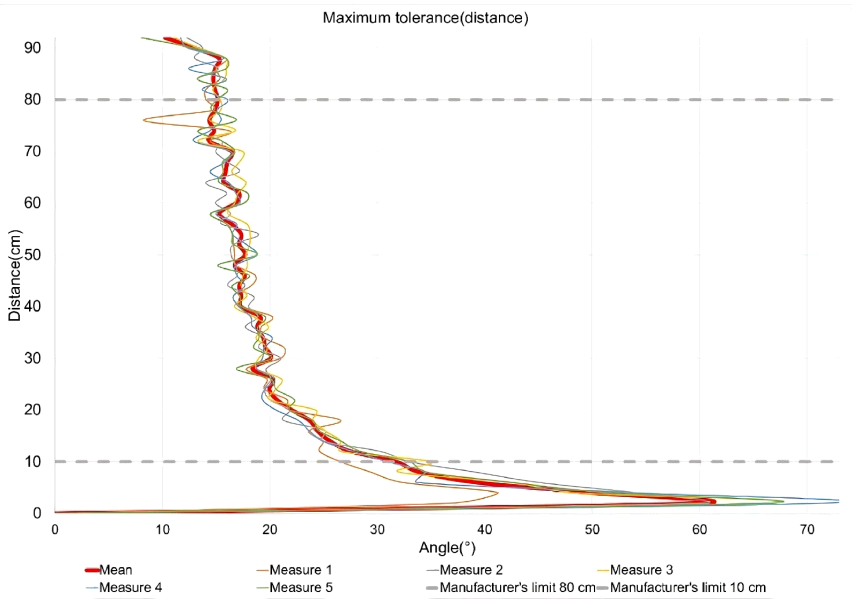


Figure 9. Infrared sensor in lateral situations

3.3. Statistical study

Using Python, several statistical studies were performed. A linear correlation was established between the two sensors in order to quantify their linear relationship. Fig. **10** presents a scatter diagram with linear adjustment, with the ultrasonic sensor on the horizontal axis and the infrared sensor on the vertical one.

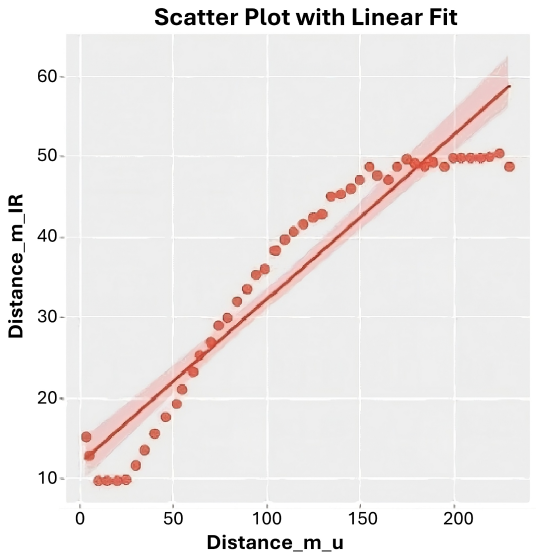


Figure 10. Scatter diagram with linear adjustment

The two mean distances increase simultaneously at a constant rate. The scatter plot in Fig. **10** shows a positive linear relation between the distance variables measured by the sensors. However, this is only the case for distances between 10 and 50 cm measured by the infrared sensor. Outside this range, the relation is not exactly linear anymore. In this vein, the different correlation coefficients were determined, and a distribution plot was elaborated for both sensors (Fig. **11**). This plot shows that the data distribution has a non-normal tendency and is slightly more uniform.

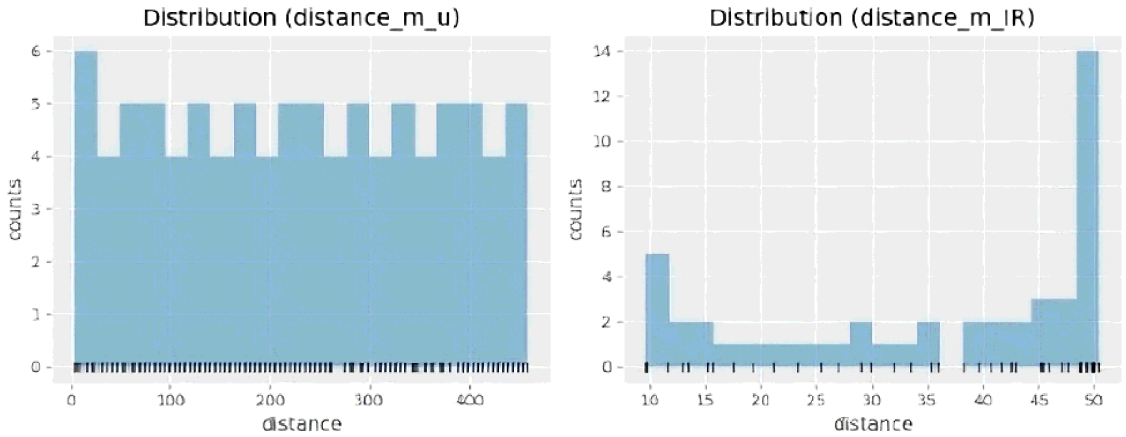


Figure 11. Distribution of values for both sensors

Fig. **12** presents the results of a study conducted to evaluate the distribution of the residuals, which are aligned along a straight line. However, very slight biases can be observed in the tails. The data lie close to the red line, indicating a linear correlation.

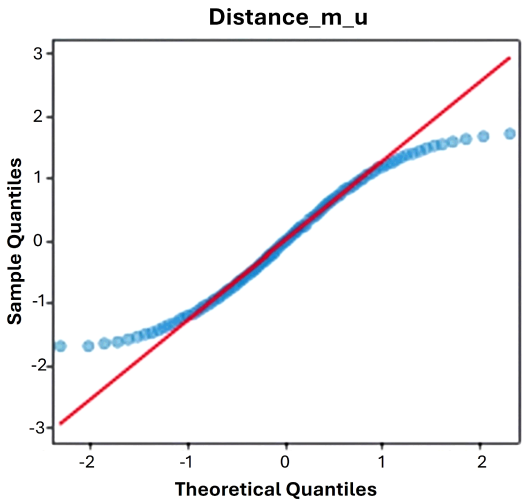


Figure 12. Quantile diagram for the ultrasonic sensor

An analysis was performed to qualify the results and compute the correlation coefficients with several methods: Pearson, Spearman, and Kendall. In all cases, a more sensitive result was observed

with regard to the extreme values, as well as a good correlation for most values.

For the ultrasonic sensor, the linear association is very high for all three methods, which confirms that the measured results are close to the theoretical values. The correlation coefficients computed for the ultrasonic sensor are as follows:

- Pearson correlation: 0.999
- Spearman correlation: 0.999
- Kendall correlation: 0.998

On the other hand, the infrared sensor exhibits a lower linear association than its ultrasonic counterpart, but it remains over 90 %, which is also a high value. The correlations for the infrared sensor are presented below:

- Pearson correlation: 0.952
- Spearman correlation: 0.982
- Kendall correlation: 0.926

Good results were obtained since the correlation for the infrared sensor was evaluated along its operating range, which shares many similarities with that of the ultrasonic sensor. It is possible to obtain lower correlations by comparing the infrared sensor used against other alternatives with different operating ranges.

4. Conclusions

In this work, the performance of ultrasonic and infrared sensors was evaluated in various situations, with the aim of understanding their operational characteristics and accuracy in measuring the distance to obstacles. This information is crucial for applications involving robotics and the Internet of Things, since precise obstacle distance measurements are fundamental for ensuring safe navigation, optimizing route planning, and enabling real-time decision-making in autonomous or semi-autonomous systems.

The results indicate that the ultrasonic sensor is the most effective for dealing with distant obstacles, provided that they are orthogonally oriented. However, when measuring shorter distances with obstacles that are not necessarily orthogonal, the infrared sensor exhibits a better performance. In other words, both sensors can complement each other to enhance the performance of devices such as robots or automated systems.

Nevertheless, it is important to highlight that infrared sensors are more susceptible to environmental factors, such as material reflectance variability, which can introduce measurement errors. On the other hand, ultrasonic sensors may struggle with detecting inclined or small surfaces. These limitations should be considered in future implementations and studies, exploring calibration methods or compensation algorithms to improve the accuracy of both devices.

The comparison between ultrasonic and infrared sensors underscores the need to consider application-specific trade-offs such as range, response time, and environmental conditions. The careful selection of either sensor – or indeed the combination of both – has been shown to substantially enhance system reliability and efficiency. A hybrid approach can often provide a more robust solution by leveraging the strengths of each technology in order to overcome their individual limitations.

5. Acknowledgements

The authors express their gratitude to Unidades Tecnológicas de Santander (UTS) and Polytech Aix-Marseille Université for their support during this research and within the framework of international academic collaboration between France and Colombia.

6. Author contributions

During the development of this work, Téllez-Garzón contributed through project administration, ensuring coordination and execution; supervision of the research process; and writing (original draft, review, and editing). Fandiño-Pelayo was responsible for formal analysis, applying analytical methods, as well as for the validation of the results and the visualization of the data. Bernard and Mazzini supported the study through data curation, organizing and maintaining the dataset; software, developing and implementing computational tools; and investigation, conducting the core research activities.

References

[1] N. Kalamain and M. Farrokhi, "Dynamic walking and stepping over large obstacles of biped robots: A Poincaré Map approach," *J. Eng. Sci. Tech. Rev.*, vol. 15, no. 6, pp. 186-200, 2022. https://doi.org/10.25103/jestr.156.23 ↑3

[2] K. Hu and C. Shidan, "Route planning of intelligent agricultural inspection robots based on improved ant colony algorithm," *J. Eng. Sci. Tech. Rev.*, vol. 16, no. 3, pp. 36-43, 2023. https://doi.org/10.25103/jestr.163.05 ↑3

[3] N. Amézquita et al., "Preliminary approach for UAV-based multi-sensor platforms for reconnaissance and surveillance applications," *Ing.*, vol. 28, no. 3, pp. 1-33, 2023. https://doi.org/10.14483/23448393.21035 ↑3

[4] M. William, E. Roberto, and J. Sánchez, "Application of a supervised learning model to analyze the behavior of environmental variables in a coffee crop," *Ing.*, vol. 25, no. 3, pp. 1-15, 2020. https://doi.org/10.14483/23448393.16898 ↑3

[5] J. L. Garzón, I. Muller, J. M. Winter, A. A. Salles, and C. E. Pereira, "Methodology of cooperative spectrum sensing and dynamic access to the channels in WSN," *Phys. Comm.*, vol. 31, pp. 28-39, 2018. https://doi.org/10.1016/j.phycom.2018.09.001 ↑3

[6] T. Mohammad, "Using ultrasonic and infrared sensors for distance measurement," in *World academy of science, engineering and technology*, Sep. 2009, vol. 51, pp. 273-278. ↑3

[7] A. M. Joseph, K. Azadeh and B. Rezaul, "State-of-the-art review on wearable obstacle detection systems developed for assistive technologies and footwear," *Sensors*, vol. 23, no. 5, pp. 1-31, 2023. https://doi.org/10.3390/s23052802 ↑3

[8] B. Mustapha, Z. Aladin, and K. B. Rezaul, " Ultrasonic and infrared sensors performance in a wireless obstacle detection system," in *IEEE 2013 1st Int. Conf. Art. Intell. Modell. Simul.*, Dec. 2013, pp. 487-492. ↑3

[9] A. Rahman, "Precision and accuracy of ultrasonic and infrared laser ToF IoT sensors," *J. Infor. Telecomm. Eng.*, vol. 8, no. 2, pp. 219-226, 2025. https://doi.org/10.31289/jite.v8i2.13406 ↑3

[10] K. Jia, "Comparative experimental study of infrared distance sensor and ultrasonic distance sensor," in *3rd Int. Symp. Sensor Tech. Control (ISSTC)*, Jan. 2024, pp. 348-353. ↑3

[11] S. Adarsh et al., "Performance comparison of Infrared and Ultrasonic sensors for obstacles of different materials in vehicle/robot navigation applications," in *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 149, art. 012141, Jul. 2016. ↑4

[12] J. Sankar, S. Adarsh, and K. I. Ramachandran, "Performance evaluation of ultrasonic and infrared waves on human body and metal surfaces for mobile robot navigation," *Mater. Today Proc.*, vol. 5, no. 8, pp. 16516-16525, 2018. https://doi.org/10.1016/j.matpr.2018.06.007 ↑4

[13] Z. Khaleel Hind and O. Bashra Kadhim, "Ultrasonic sensor decision-making algorithm for mobile robot motion in maze environment," *Bul. Elec. Eng. Infor.*, vol. 13, no. 1, pp. 109-116, 2024. https://doi.org/10.11591/eei.v13i1.6560 ↑4

[14] I. Prasojo, P. Nguyen, and S. Nishith, "Design of ultrasonic sensor and ultraviolet sensor implemented on a fire fighter robot using AT89S52," *J. Rob. Control*, vol. 1, no. 2, pp. 59-63, 2020. https://doi.org/10.18196/jrc.1212 ↑4

[15] U. Umiatin and A. F. Dendi, "Sharp IR GP2Y0A21 sensor calibration for prototyping application of smart anthropometric system," *J. Phys.*, vol. 2377, no. 1, pp. 1-6, 2022. https://doi.org/10.1088/1742-6596/2377/1/012026 ↑4

[16] C. Technologies, Product User's Manual-hc-sr04 Ultrasonic Sensor, Cytron , 2013. [Online]. Available: https://web.eece.maine.edu/~zhu/book/lab/HC-SR04%20User%20Manual.pdf ↑5

[17] C. Technologies, HC-SR04 User Guide, Cytron, 2020. Online]. Available: https://www.mpja.com/download/hc-sr04_ultrasonic_module_user_guidejohn.pdf ↑5

[18] C. Technologies, Data sheet GP2Y0A021YK, Cytron, 2020. Online]. Available: https://files.seeedstudio.com/wiki/Grove-80cm_Infrared_Proximity_Sensor/res/GP2Y0A21YK.pdf ↑6, 7

Johan Leandro Tellez Garzón

Telecommunications engineer, Universidad Militar Nueva Granada. Master of Electronics Engineering and Telecommunications, Universidade Federal do Rio Grande do Sul. PhD in Electrical Engineering, Universidade Federal do Rio Grande do Sul. Full profesor, Telecommunications program, Unidades Tecnológicas de Santander, Bucaramanga, Colombia.
Email: jtellez@correo.uts.edu.co

Jorge Saul Fandiño Pelayo

Telecommunications Engineer, Universidad Santo Tomas. Master in Telematics, Universidad Autónoma de Bucaramanga. Phd(C) and in Engineering, Universidad Autónoma de Bucaramanga. Full professor, Telecommunications program, Unidades Tecnológicas de Santander, Bucaramanga, Colombia.
Email: jfandino@correo.uts.edu.co

Bernard Antoine

Engineer in Microelectronics and Telecommunications, Aix-Marseille University, Polytech Marseille, France. Engineer in Digital Electronics, PYTHEAS TECHNOLOGY, La Ciotat, France.
Email: antoine.bernard.2@etu.univ-amu.fr

Mazzini Giovanni

Engineer in Microelectronics and Telecommunications, Aix-Marseille University, Polytech Marseille, France. Engineer in Analog Electronics, PYTHEAS TECHNOLOGY, La Ciotat, France.
Email: giovanni.mazzini@etu.univ-amu.fr

UNIVERSIDAD DISTRITAL
FRANCISCO JOSÉ DE CALDAS

Ingeniería

https://revistas.udistrital.edu.co/index.php/reving/issue/view/1266

DOI: https://doi.org/10.14483/23448393.22292

Research

Bibliometric Analysis and Overview of Matrix Product States in the Bose-Hubbard Model

Estudio bibliométrico y revisión del uso de estados producto de matrices en el modelo de Bose-Hubbard

[1]Juan Sebastián Gómez and [1]Karen Rodríguez

[1]Universidad del Valle Physics Department, Cali, Colombia

Abstract

Context: Quantum many-body systems have been a prominent topic over the past two decades, underpinning advancements in superconductors, ultracold atoms, and quantum computing, among other fields. This bibliometric analysis explores key concepts, influential authors, and the current significance of a powerful family of algorithms in computational physics, *i.e.*, density matrix renormalization group (DMRG) algorithms. Special emphasis is placed on the use of tensor product states in developing classical simulations of quantum systems.
Method: This paper presents a literature review sourced from the SCOPUS database. It analyzes trends and approaches related to uncertainty in numerical developments for quantum many-body systems, with a focus on the Bose-Hubbard Model, in order to better understand the imposition of additional constraints to ensure the validity of the results.
Results: The increasing number of publications on this topic over the last decade indicates a growing interest in solutions for many-body quantum systems, driven by promising advances in superconductive materials, quantum computing, and other impactful areas.
Conclusions: This work explored essential foundational works to help beginners understand a well-established technique that aims to overcome the limitations of classical computing. The use of matrix product states in DMRG algorithms is gaining significant traction in various fields, including quantum computing, machine learning, and statistical mechanics, with the purpose of addressing the challenges related to quantum many-body systems.
Acknowledgments: This work was supported by Universidad del Valle under internal project Cl. 71344.
Keywords: matrix product states, density matrix renormalization group, strongly correlated systems, Bose-Hubbard model, tensor networks

Article history

Received: 11th/June/2024

Modified: 30th/March/2025

Accepted: 4th/April/2025

Ing, vol. 30, no. 1, 2025. e22292

* **Correspondence:** juan.s.gomez@correounivalle.edu.co

Resumen

Contexto: Los sistemas cuánticos de muchos cuerpos han sido un tema prominente durante las últimas dos décadas, sustentando avances en superconductores, átomos ultrafríos y computación cuántica. Este análisis bibliométrico explora conceptos clave, autores influyentes y la significancia actual de una poderosa familia de algoritmos en física computacional, *i.e.*, los algoritmos del grupo de renormalización de matriz de densidad (DMRG). Se pone un énfasis especial en el uso de estados producto de tensor en el desarrollo de simulaciones clásicas de sistemas cuánticos.
Métodos: Este artículo presenta una revisión de la literatura obtenida de la base de datos de SCOPUS. Analiza tendencias y enfoques relacionados con el concepto de incertidumbre dentro del marco de desarrollos numéricos para sistemas cuánticos de muchos cuerpos, con énfasis en el modelo de Bose-Hubbard. Esto, con el objetivo de entender mejor la imposición de restricciones adicionales para asegurar la validez de los resultados.
Resultados: El número creciente de publicaciones sobre este tema en la última década indica un mayor interés en soluciones para sistemas cuánticos de muchos cuerpos, impulsadas por avances promisorios en materiales superconductores, computación cuántica y otras áreas de impacto.
Conclusiones: Este trabajo exploró los trabajos fundamentales para ayudar a los principiantes a comprender una técnica bien establecida que promete sobrepasar las limitaciones de la computación clásica. El uso de los estados de producto de matrices en los algoritmos DMRG está ganando una tracción significativa en numerosos campos, incluyendo la computación cuántica, el aprendizaje automático y la mecánica estadística, con el fin de abordar los desafíos asociados con los sistemas cuánticos de muchos cuerpos.

Acknowledgments: Este trabajo fue apoyado por la Universidad del Valle bajo el proyecto interno Cl. 71344.

Palabras clave: estados producto de matrices, grupo de renormalización de matriz de densidad, sistemas fuertemente correlacionados, modelo de Bose-Hubbard, redes de tensor

Table of contents

1. Introduction

Since the theoretical statistical-mechanical development by Bose and Einstein (1) and its experimental corroboration, ultracold quantum gases have been a topic of high interest for the scientific community.

Given their emergent physical properties, these types of systems attract great interest, bringing many fields together, such as condensed matter, quantum simulators, quantum information, and quantum error correction (2–4). Additionally, thanks to the development of tools for controlling cold atoms in atomic physics and quantum optics, it is possible to obtain highly controllable systems by leveraging phenomena such as radiation pressure and dipole force, which allow cooling and trapping low-energy atoms in periodic structures generated by laser light called *optical lattices* (5).

Among the many models that include interactions between atoms in optical lattices, this document focuses on the Bose-Hubbard (BH) model introduced in 1963 by Gersch and Knollman (6). This model describes bosons that can hop between the sites of a discrete lattice and interact when they coincide at the same site. It has undergone extensive theoretical and experimental scrutiny (7), and it has been studied using a wide range of techniques developed within the condensed matter community, from static mean-field approaches to the more complex dynamic mean-field theory (8), in addition to various diagrammatic approaches such as random phase approximation (9) and expansions based on spin and orbital angular momentum degeneracy (10). It has also been widely addressed using numerical methods like exact diagonalization (ED) and quantum Monte Carlo (QMC) (10).

However, describing a quantum system through the eigenstates of its Hamiltonian implies problems with the exponential growth of its Hilbert space. Therefore, in 1992, White developed an algorithm that extracted information from the system based on the eigenstates of the density matrix operator. This is known as the *density matrix renormalization group (DMRG) algorithm* (11) and enables the study of interacting many-body systems of low and high dimensionality with internal degrees of freedom.

The scope of traditional quantum mechanics is limited when investigating behavioral regimes different from the approximations made during theoretical calculations. Therefore, a few numerical methods have been developed to determine physical properties of high interest, such as observables and system correlations. The development of the DMRG has led to two decades of significant advances in addressing these problems; it became possible to relate the low-dimensional long-range phenomena of the BH model, which constitute the desired superfluidity and Mott insulator properties (12).

The success of the DMRG lies in its matrix product states (MPS), a type of tensor network (TN) that only deals with one spatial dimension (13). These are capable of handling the problem regarding the exponential growth of Hilbert spaces through truncations without losing system information or being limited to approximate systems. This paper seeks to explore how extensively MPS algorithms have been studied and assess their continued relevance based on recent research trends and output quality.

2. Methodology

The methodology followed in this work consists of several steps, ranging from defining the scope of the analysis to interpreting the results. A bibliometric approach was employed to examine the application of MPS in the BH model using data from SCOPUS. The process involved data collection, cleaning, visualization, and interpretation to identify key trends and contributions in the field. A summary of the methodology is presented in Table I.

Table I. Summary of the bibliometric analysis methodology

Step	Description
Definition of scope	This study focused on the application of MPS in the BH model and other TN methods in physics and engineering. The bibliometric analysis considered publications from 2013 to 2023.
Data collection	A SCOPUS query was formulated using relevant keywords. Search filters were applied to refine the results to articles discussing MPS and the BH model. Metadata, including titles, authors, publication years, journal/conference details, and citation counts, were extracted.
Data cleaning and processing	Duplicate entries and irrelevant publications were removed. Author names and affiliations were standardized where necessary, and consistency across bibliographic records was ensured.
Visualization and interpretation	Publication trends over the selected period were analyzed using VOSviewer. Co-authorship networks, journal impact, and keyword co-occurrence patterns were examined. Plots, charts, and network graphs were generated to illustrate trends.
Discussion of findings	The results were compared against those of prior studies to identify trends and research gaps. Key contributors and emerging themes in the field were highlighted.

The keywords used during the data collection stage were "tensor networks", "matrix product state", "many-body systems", "strongly correlated systems", ”ultra cold atoms", "Bose-Hubbard Hamiltonian",

"spin-1", and "bosons"for the decade between 2013 and 2023. Fig. 1 presents the research dynamics as evolving with the 479 scientific publications identified, where a pronounced increase in interest is observed from 2020 onward. The exact query entry used was as follows: *TITLE-ABS-KEY ("Tensor Networks" AND "Matrix Product State" OR "Many-body systems" OR "Strongly correlated systems" OR "Ultra Cold atoms" OR "Bose-Hubbard Hamiltonian" OR "Spin-1" OR "Bosons") AND PUBYEAR > 2012 AND PUBYEAR < 2024*.

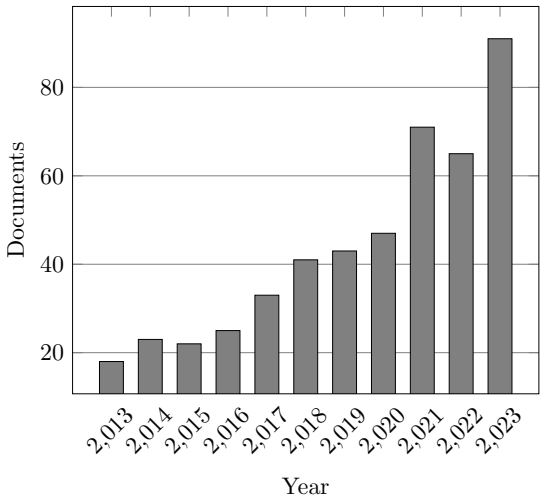


Figure 1. Publishing intensity regarding TNs in many-body systems over the last 10 years. **Source**: SCOPUS

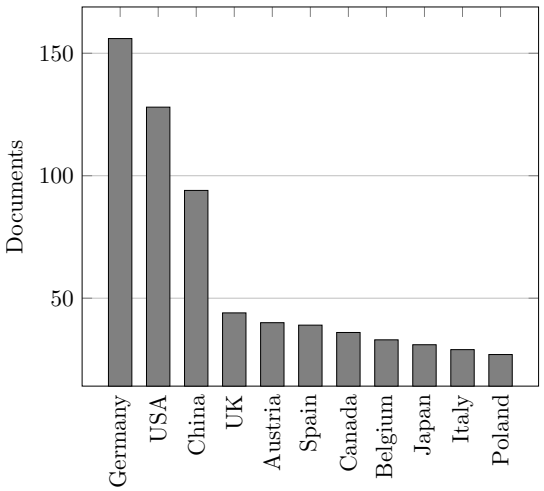


Figure 2. Documents per year by country. **Source**: SCOPUS, using the same query string as in Fig. 1

Furthermore, Fig. 3 illustrates the distribution of the research fields involved, highlighting "physics and astronomy"with 421 publications, followed by "materials science"with 155 documents and "mathematics"with 72.

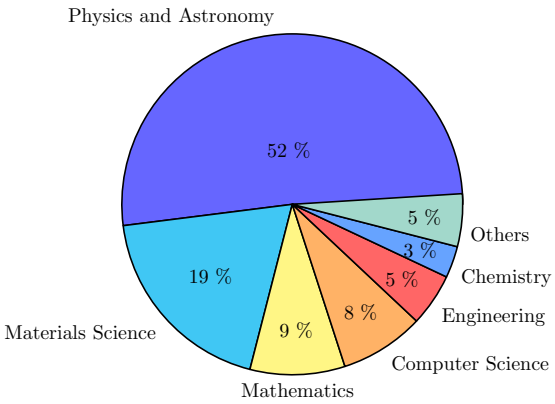


Figure 3. Distribution of fields of study with regard to scientific publications. **Source**: SCOPUS

As a result of this bibliometric analysis, an important aspect emerges: Fig. 4 shows that many countries are actively engaged in TN research in the context of many-body systems and ultra cold

atoms. The top three contributors are Germany (156 publications), United States (128), and China (94). Out of these publications, 12 papers have an author from Latin American countries, and four of them have a Colombian author (14–17). The topics addressed were quantum thermal machines and entanglement between single- and many-bodyness. The institutes responsible for most of the publications are consistent with Fig. 4: the first four of them are the Max Planck Institute, Universiteit Gent, Universität Wien, and the Munich Center for Quantum Sciences.

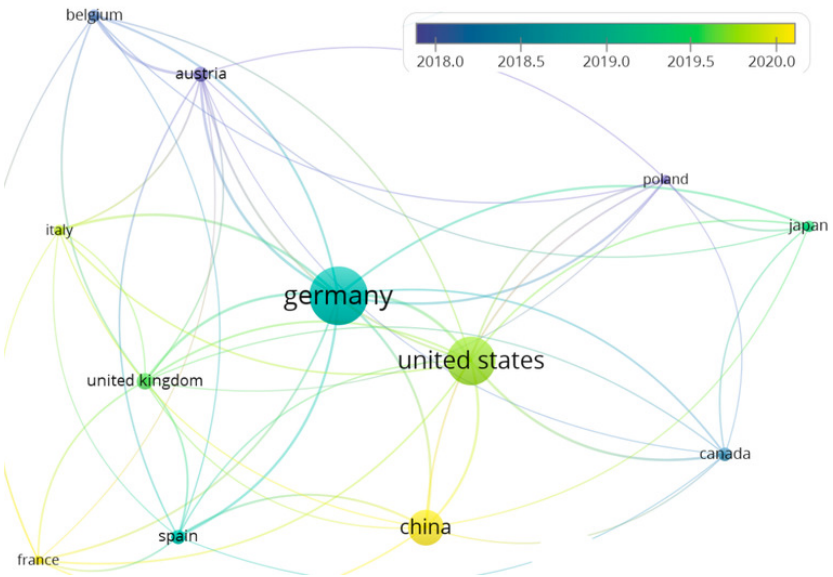


Figure 4. Countries, their collaborators, and highest average of documents published by year

The five journals that published the most articles on the topics over the past decade all belong to the Physical Review family, as illustrated in Fig.5.

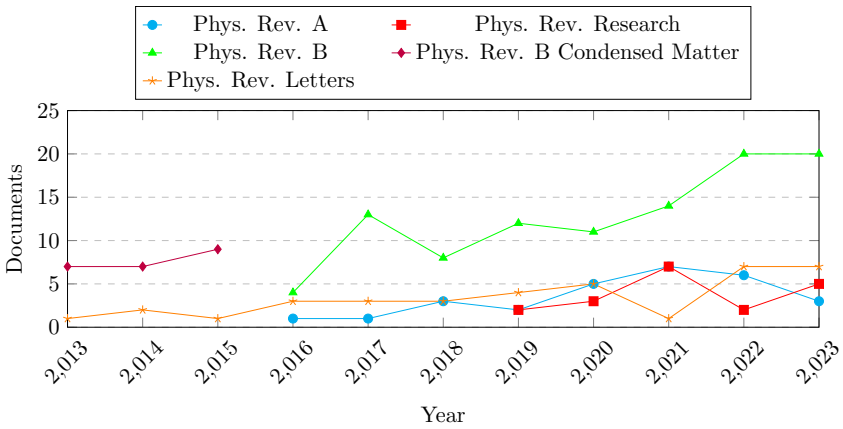


Figure 5. Documents per year by source. **Source**: SCOPUS

Finally, Fig. 6 shows the 26 most frequent keywords, each appearing more than 20 times in the selected articles. These reflect the main elements around which this topic revolves, including the determination of a ground state in quantum many-body systems, their phase diagrams, and certain fields of interest, *e.g.*, quantum optics, statistical mechanics, and machine learning.

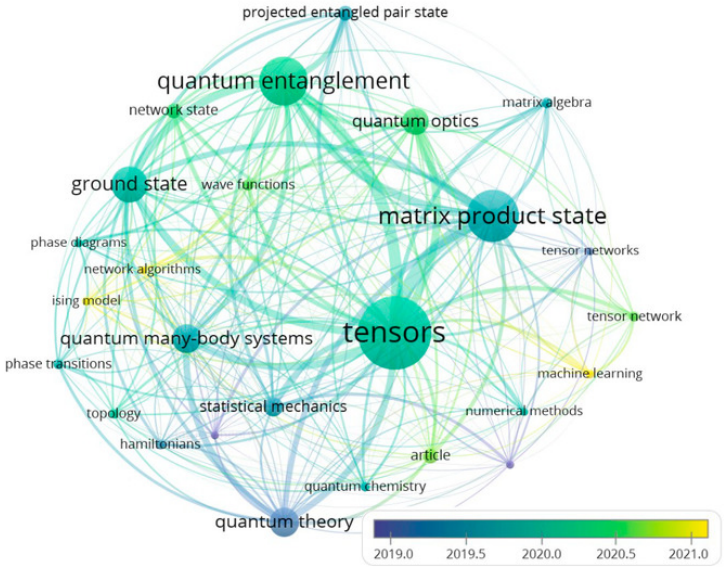


Figure 6. Most relevant keyword appearance by year

3. Results and discussion

Different logical operators were used in the query string to filter the initial 479 publications, focusing only on the topics most relevant to this research. By replacing the operator *OR* with *AND* before "spin-1ïn the query, we obtained 34 publications more directly related to our research topic.

The use of MPS and TNs to address many-body quantum systems has recently attracted substantial attention. With the growth of quantum computing and machine learning, MPS provide powerful solutions to problems where perturbative methods or exact diagonalization fail to accurately describe the complex dynamics of entangled quantum systems (18).

Fig. 4 indicates that research in this field is particularly prominent in Germany, where Frank Verstraete, JI Cirac, and MM Wolf developed foundational approaches in 2007 (19) and remain as the leading researchers. Institutions such as the Max Planck Institute of Quantum Optics, Universiteit Gent, Universität Wien, and the Munich Center for Quantum Science and Technology (MCQST) are among Europe's leading research centers in this area. Additionally, China and the United States have also contributed significantly to the field, even surpassing Germany in average annual publications. Three institutes that are ranked among the top 10 in these countries are the Chinese Academy of Sciences, the University of the Chinese Academy of Sciences, and the California Institute of Technology.

The dominance of certain institutions in the field of TN research is not merely a result of geographical biases, but rather a reflection of strategic investments in science and technology by both the public and private sectors (Fig.7). Governments in these regions prioritize funding for quantum technologies, artificial intelligence, and condensed matter physics, leading to well-financed research programs and cutting-edge facilities. Additionally, strong collaboration between academia and the industry foster innovation and accelerate publication rates. As shown in Fig. 7, the top 14 funding agencies supporting TN research highlight the crucial role of sustained financial backing in maintaining institutional leadership within this rapidly growing field.

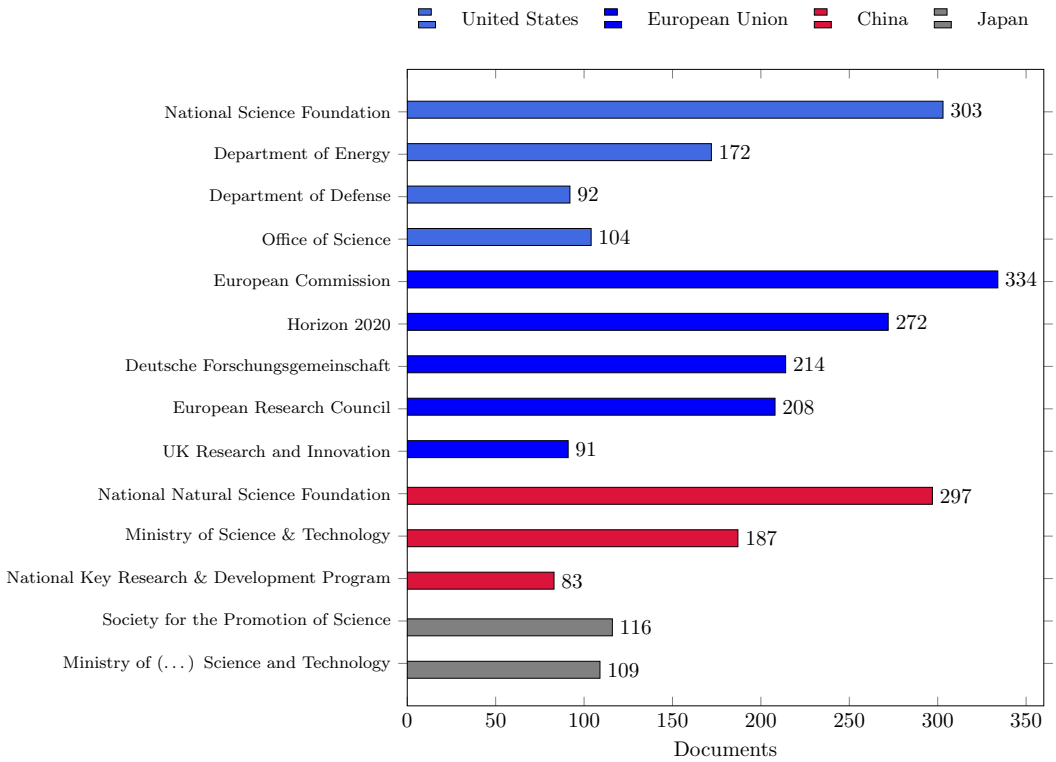


Figure 7. Top 14 funding sponsors by number of published documents on TNs from 2013 to 2023. The data highlight the leading institutions investing in TN research, with significant contributions from the United States, the European Union, China, and Japan

Source: SCOPUS

3.1. Bose-Hubbard model

Numerical developments flourish only when applied to complex systems, which is why quantum many-body models – such as the Fermi-Hubbard, Lieb-Liniger, and Bose-Hubbard models, among others – serve as essential benchmarks to evaluate their efficiency and precision (20). However, physics frequently involves complexities requiring the study of models with non-Hermitian Hamiltonians or quartic interaction terms, as is the case with the BH model, which describes a system of bosons in a 1D chain that interact and hop in a unidirectional fashion. The Hamiltonian that describes this model (20)

corresponds to Eq. (1):

$$\hat{H} = -t\left(\sum_{j=1}^{L-1} \hat{b}_j^{\dagger}\hat{b}_{j+1} + \hat{b}_j\hat{b}_{j+1}^{\dagger}\right) - \mu\sum_{j=1}\hat{n}_j + \frac{U}{2}\sum_{j=1}^{L}\hat{n}_j(\hat{n}_j - 1), \tag{1}$$

where $\hat{b}_j^{\dagger}$, $\hat{b}_j$, and $\hat{n}_j = \hat{b}_j^{\dagger}\hat{b}_j$ are bosonic creation, annihilation, and particle number operators applied to the j-th site; L refers to the total sites; U quantifies the on-site interaction strength; ϵ is set depending on boundary conditions; and t characterizes hopping between first neighbors. The model is indeed integrable from the perspective of the quantum inverse scattering method (20). Many exact solution approaches can be adopted using ansätze such as Bethe (20) or Gutzwiller (21).

3.2. Density matrix renormalization group

Investigating quantum many-body systems is crucial in condensed matter physics. Two main challenges include modeling complex interactions and solving these models computationally. In the latter, many computational issues also emerge, mainly due to the fact that the dimension of the Hilbert space of the whole system grows exponentially (d^n) with the amount of particles (n) involved and the internal degrees of freedom (d) which do not happen in classical mechanics, where the computational memory required is proportional to one times the other ($d\,n$) (22). To avoid this issue, based on the variational principle, Steven R. White developed the DMRG in 1992 (23) after realizing that the eigenstates of the density matrix provide a more suitable description for quantum systems than those of the Hamiltonian (24).

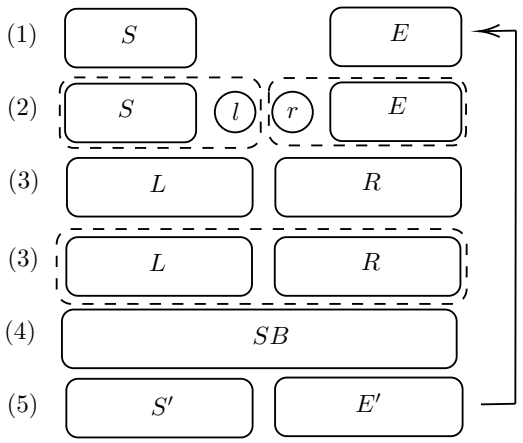


Figure 8. Schematic description of the infinite-system DMRG algorithm, as adapted from (22)

The procedure is as follows. First, we examine two blocks: system (S) and environment (E) (Fig. 8). These blocks are part of the overall system under analysis, so their labels are chosen arbitrarily. Next, we expand the system by adding a physical site to each block, resulting in the formation of blocks left (l) and right (r). We then create block SB (superblock) by combining blocks L and R. SB represents the entire system that we aim to describe at each iteration. It is important to note that aggregating these blocks requires consideration of both their individual Hamiltonians and their mutual interactions.

Finally, if we were to truncate blocks L and R using their corresponding low-energy states (5), we would create the new blocks S' and E' to be employed in the initial step of the subsequent iteration (22).

An alternative in the algorithm comes in step (5), where diagonalizing becomes viable using a well behaved eigenstate $|\psi\rangle$ (usually the ground state or among the lowest-energy eigenstates) to construct the density matrix. Considering multiple target states is not out of the picture, as long as the sum of the amplitudes of probability equals the identity, as shown in Eq.(2).

$$\rho = \sum_n c_n \left|\psi_n\right\rangle \left\langle\psi_n\right| \qquad\qquad \sum_n c_n = 1 \tag{2}$$

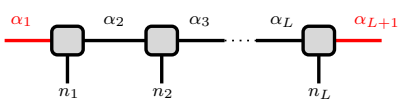
3.3. Matrix product states

Tensor network states (TNS) are extensively used within the DMRG community to estimate the ground states of Hamiltonians represented on lattices in d dimensions. The simplest type of TN, which we will briefly explain, consists of MPS where there is only a single spatial dimension (D=1). This is the core of the remarkable success of White's DMRG. Whether it be their dynamics with the time evolution block decimation algorithm (TEBD) or their variations, MPS have been very useful since they can accurately provide the properties of a ground state (13).

Constructing an MPS requires dealing with a system that can be split into two subsystems, *e.g*, a 1D chain of spins, as shown in Eq. (3). Starting from a physical lattice with L sites, a bipartite system is formed, where the second subsystem is taken as the lattice at site L and the first subsystem is the rest of the lattice. This is then represented as an entangled state, and singular value decomposition (SVD) is performed on the probability amplitude coefficients. After regrouping, the right site of the lattice is represented on an appropriate basis, with the left side representing the lattice from site 1 to $L-1$. This process is iterated until all sites are included (Fig. **9**) (25).

$$|\psi\rangle = \sum_{\{n\}}^{d} C_{n_1 n_2 \cdots n_L} \left|n_1 \cdots n_L\right\rangle \tag{3}$$

where the $A^{n_j}_{\alpha_j,\alpha_{j+1}}$ correspond to contractions of $\lambda_{\alpha_j}\Gamma^{n_j}_{\alpha_j,\alpha_{j+1}}$, *i.e.*, the Schmidt coefficient (λ) and the element from the transformation matrix (Γ), to the desired base (Fock's in this case).



$$|\psi\rangle = \sum_{\alpha_1 \cdots \alpha_L}^{\chi} \sum_{n_1 \cdots n_L}^{d} A^{n_1}_{\alpha_1 \alpha_2} A^{n_2}_{\alpha_2 \alpha_3} \cdots A^{n_L}_{\alpha_L \alpha_{L+1}} \left|n_1 \cdots n_L\right\rangle$$

Figure 9. Pictorial (left) and mathematical (right) representations of the MPS

3.4. Matrix product operators

In the same spirit, a quantum state can be represented in terms of products of matrices, and an analogous representation can be formulated for operators. This becomes really useful when intending to measure the expectation values of global operators that, rather than acting in a single site, act throughout

the entire chain, as is the case for most Hamiltonians (26–29). To this effect, the operators must be represented as a product of matrices, as shown in Eq. (5) for a general operator.

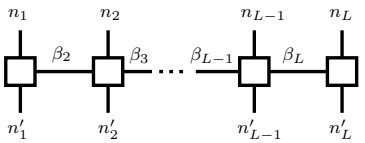
$$\hat{\mathcal{O}} = \sum_{\{n\},\{n'\}}^{d} C_{n_1 n_2 \cdots n_L}^{n_1' n_2' \cdots n_L'} \left|n_1 n_2 \cdots n_L\right\rangle \left\langle n_1' n_2' \cdots n_L'\right| \tag{4}$$

$$= \sum_{\{n\}\{n'\}}^{d} \sum_{\{\beta\}}^{w} W_{\beta_1} W_{\beta_1\beta_2}^{n_1 n_1'} W_{\beta_2\beta_3}^{n_2 n_2'} \cdots W_{\beta_{L-1}\beta_L}^{n_{L-1} n_{L-1}'} W_{\beta_L\beta_{L+1}}^{n_L n_L'} W_{\beta_{L+1}} \left|n_1 \cdots n_L\right\rangle \left\langle n_1' \cdots n_L'\right| \tag{5}$$

$$= \sum_{\{n\}\{n'\}}^{d} W^{[0]} W^{[1]} W^{[2]} \cdots W^{[L-1]} W^{[L]} W^{[L+1]} \left|n_1 \cdots n_L\right\rangle \left\langle n_1' \cdots n_L'\right|$$

In Eq. (5), the element $W_{\beta_k\beta_{k+1}}^{[k]}$ denotes a $w \times w$ matrix, where w is the local dimension of the matrix product operator (MPO) and $\{|n_k\rangle\}$ are the local basis vectors. Moreover, the matrices at the free ends of the chain will have a different structure. They are not square matrices; instead, they are column and row vectors of dimensions $1 \times w$ and $w \times 1$. Thus, the product between Ws yields a scalar, as shown in Eq. (6).

$$W^{[0]} = \begin{bmatrix} 0 & \cdots & 0 & \hat{\mathbb{I}} \end{bmatrix} \qquad W^{[L+1]} = \begin{bmatrix} \hat{\mathbb{I}} & 0 & \cdots & 0 \end{bmatrix}^T \tag{6}$$



$$\hat{\mathcal{O}} = \sum_{\{\beta\}}^{w} \sum_{\substack{\{n\} \\ \{n'\}}}^{d} W_{\beta_2}^{n_1' n_1} W_{\beta_2\beta_3}^{n_2' n_2} \cdots W_{\beta_L}^{n_L' n_L} \left|n_1 \cdots n_L\right\rangle \left\langle n_1' \cdots n_L'\right|$$

Figure 10. Graphic representation (left) and mathematical expression (right) of the MPO

The graphic representation of Eq. (5) is presented in Fig.10. The true wonders of the MPO representation lie in the fact that it allows expressing *exactly* any local Hamiltonian with (relatively) small matrices of dimension $w \times w$. Moreover, the structure of this representation is almost identical to that of a MPS; both of them are represented by a list of tensors at each site and, therefore, computing expectation values becomes an easier task (30). Taking into account that we are working with a Hamiltonian composed of a sum of finite-range interactions, we can construct the W matrix as entirely lower (or upper) triangular matrices by imposing that the top-left and bottom-right components be identity operators $\hat{\mathbb{I}}$. For example, consider an operator that is the sum of single-site local terms:

$$\hat{\mathcal{O}} = \sum_i \hat{X}_i = W^{[1]} W^{[2]} \cdots W^{[L-1]} W^{[L]} \quad \longrightarrow \quad W^{[k]} = \begin{bmatrix} \hat{\mathbb{I}} & 0 \\ \hat{X}_k & \hat{\mathbb{I}} \end{bmatrix}$$

For that specific operator, the internal dimension of the MPO is $w = 2$, and each operator (including that top-right 0) is $d \times d$ matrices. If we carry out the same procedure for a nearest-neighbor operator that also has single-site local terms, the following is obtained:

$$\hat{\mathcal{O}} = \sum_i \hat{X}_i \hat{Y}_{i+1} + \sum_j \hat{Z}_j = W^{[1]} W^{[2]} \cdots W^{[L-1]} W^{[L]} \quad \longrightarrow \quad W^{[k]} = \begin{bmatrix} \hat{\mathbb{I}} & 0 & 0 \\ \hat{Y}_k & 0 & 0 \\ \hat{Z}_k & \hat{X}_k & \hat{\mathbb{I}} \end{bmatrix}$$

Note that, in both the first and the second operator, the terms which act upon a single site always remain in the bottom-left corner. The BH Hamiltonian in Eq. (1) can also be expressed in this way, as seen in Eq. (7).

$$W^{[k]} = \begin{bmatrix} \hat{\mathbb{I}} & 0 & 0 & 0 \\ -t\hat{b}_k^\dagger & 0 & 0 & 0 \\ -t\hat{b}_k & 0 & 0 & 0 \\ -\mu\hat{n}_k + \frac{U}{2}\hat{n}_k(\hat{n}_k - 1) & \hat{b}_k & b_k^\dagger & \hat{\mathbb{I}} \end{bmatrix} \tag{7}$$

3.5. Other TN representations

Another interesting result from our SCOPUS analysis is the need for a higher-dimensional representation of states, as MPS are limited to one spatial dimension. When it comes to other physics fields such as quantum chemistry (31, 32), machine learning (33), quantum computation (34–37), and holography (2, 38–40), systems with higher dimensions ($d = 2$) are the norm rather than the exception. This is why other generalizations of MPS and MPO exist.

Among the most commonly used representations are projected entangled-pair states (PEPS) (41–45), the multiscale entanglement renormalization ansatz (MERA) (46), and tree tensor networks (TTNs) (47–50) (Fig. **11**). The first natural generalization corresponds to PEPS, where the linear MPS is replaced by a lattice that can take many geometries aside from the traditional squared lattice, as shown in Fig. **11a**. Translational invariance can once again be exploited for finite and infinite systems by choosing a unitary cell that is repeated over the whole lattice (13).

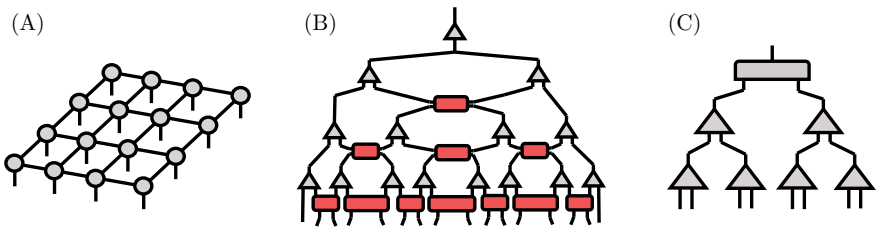


Figure 11. Depiction of other TN representations: a) projected entangled-pair states, b) multiscale entanglement renormalization, c) tree tensor networks

TNs have rapidly emerged as a powerful framework for tackling complex problems in quantum many-body physics, quantum information, and beyond, as showcased in Fig. **12**. Their ability to efficiently represent and manipulate high-dimensional data makes them invaluable for studying entanglement structures, simulating quantum systems, and optimizing variational methods. As research continues to uncover new applications, from enhancing quantum algorithms (51) to accelerating neural network training, TN methods (52) have become increasingly relevant for newcomers seeking to work at the forefront of computational and theoretical advancements. Given their expanding role in interdisciplinary research, mastering TNs is not just an asset but a necessity for those aiming to stay ahead in modern physics and computer science.

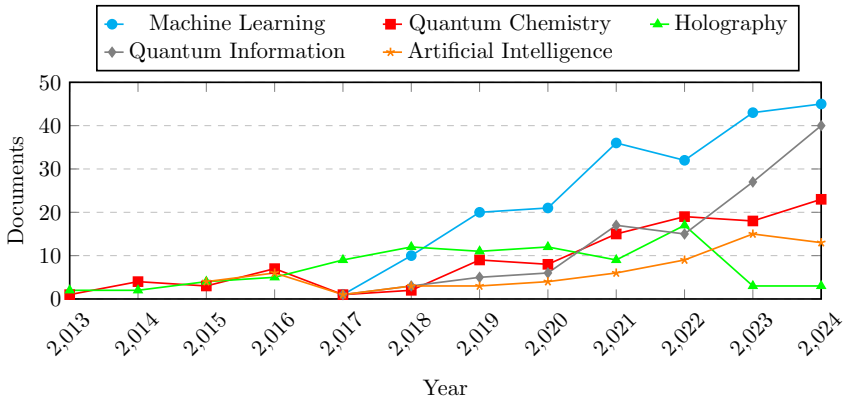


Figure 12. Number of documents published per year on TNs in the fields of machine learning, quantum chemistry, artificial intelligence, quantum information, and holography. **Source**: SCOPUS

4. Conclusions

- This paper reviews key foundational literature to introduce beginners to a well-established technique that aims to overcome the current limitations of classical computing.
- TNs are now becoming a trending technique for $d = 1, 2$ in many fields of physics and computer science, such as condensed matter physics, quantum chemistry, machine learning, artificial intelligence, and quantum computing, bridging theoretical insights with practical applications.
- Newer research directions are exploring physical systems beyond simplified toy models, where richer and more complex physics emerges. Examples include the spin-1 BH model, SU(N) Hubbard models, and topological matter, all of which push the boundaries of TN methods in capturing interesting correlations and exotic quantum phenomena.
- Future work may explore the integration of MPS with neural network-based quantum states, or the adaptation of MPS-inspired techniques to two-dimensional lattice models through PEPS or MERA frameworks.

5. Author contributions

JSG: methodology, data curation, formal analysis, investigation, writing (original draft).
KR: conceptualization, supervision, writing (review and editing).

References

[1] W. Wang, "Linear limit continuation: Theory and an application to two-dimensional bose–einstein condensates," *Chaos Sol. Frac.*, vol. 182, art. 114735, 2024. https://doi.org/10.1016/j.chaos.2024.114735 ↑ 3

[2] A. Jahn and J. Eisert, "Holographic tensor network models and quantum error correction: A topical review," *Quantum Sci. Technol.*, vol. 6, no. 3, art. 033002, 2021. https://doi.org/10.1088/2058-9565/ac0293 ↑ 3, 12

[3] A. J. Ferris and D. Poulin, "Tensor networks and quantum error correction," *Phys. Rev. Lett.*, vol. 113, art. 030501, 2014. https://doi.org/10.1103/PhysRevLett.113.030501 ↑ 3

[4] A. S. Darmawan and D. Poulin, "Tensor-network simulations of the surface code under realistic noise," *Phys. Rev. Lett.*, vol. 119, art. 040502, 2017. https://doi.org/10.1103/PhysRevLett.119.040502 ↑ 3

[5] S. F. Caballero-Benítez, "Materia cuántica en cavidades de alta reflectancia (many-body CQED)," 2022. https://doi.org/10.48550/arXiv.2201.06641 ↑ 3

[6] H. A. Gersch and G. C. Knollman, "Quantum cell model for bosons," *Phys. Rev.*, vol. 129, pp. 959–967, Jan 1963. https://doi.org/10.1103/PhysRev.129.959 ↑ 3

[7] Y. Nakamura, *et al*, "Experimental determination of bose-hubbard energies," *Phys. Rev. A*, vol. 99, no. 3, art. 033609, Mar 2019. http://dx.doi.org/10.1103/PhysRevA.99.033609 ↑ 3

[8] G. Sangiovanni, *et al*, "Static versus dynamical mean-field theory of mott antiferromagnets," *Phys. Rev. B*, vol. 73, no. 20, May 2006. http://dx.doi.org/10.1103/PhysRevB.73.205121 ↑ 3

[9] G. Co', "Introducing the random phase approximation theory," *Universe*, vol. 9, no. 3, art. 141, 2023. https://doi.org/10.3390/universe9030141 ↑ 3

[10] R. T. Scalettar, "An introduction to the Hubbard Hamiltonian," in *Quantum materials: Experiments and theory*, E. Pavarini, E. Koch, J. van den Brink, and G. Sawatzky, Eds. Jülich, Germany: Forschungszentrum Jülich, 2016, pp. 4.1–4.27. http://www.cond-mat.de/events/correl16 ↑ 3

[11] U. Schollwöck, "The density-matrix renormalization group," *Rev. Mod. Phys.*, vol. 77, pp. 259–315, Apr 2005. https://doi.org/10.1103/RevModPhys.77.259 ↑ 3

[12] M. Iskin, "Mean-field theory for the mott-insulator-paired-superfluid phase transition in the two-species bose-hubbard model," *Phys. Rev. A*, vol. 82, art. 055601, 2010. https://doi.org/10.1103/PhysRevA.82.055601 ↑ 3

[13] G. Evenbly and G. Vidal, "Tensor network states and geometry," *J. Stat. Phys.*, vol. 145, no. 4, art. 891–918, Jun 2011. http://dx.doi.org/10.1007/s10955-011-0237-4 ↑ 3, 10, 12

[14] H. Tang, J. L. Simancas-García, J. Mai, M. Cheng, I. Iqbal, and L. Y. Por, "Overview of digital quantum simulator: Applications and comparison with latest methods," *SPIN*, art. 2440004, 2024. https://doi.org/10.1142/S2010324724400046 ↑ 6

[15] M. Brenes, J. J. Mendoza-Arenas, A. Purkayastha, M. T. Mitchison, S. R. Clark, and J. Goold, "Tensor-network method to simulate strongly interacting quantum thermal machines," *Phys. Rev. X*, vol. 10, art. 031040, 2020. https://doi.org/10.1103/PhysRevX.10.031040 ↑ 6

[16] J. J. Mendoza-Arenas, "Dynamical quantum phase transitions in the one-dimensional extended fermi–hubbard model," *J. Stat. Mech.*, vol. 2022, art. 043101, 2022. https://doi.org/10.1088/1742-5468/ac6031 ↑ 6

[17] J. Reslen, "Entanglement at the interplay between single- and many-bodyness," *J. Phys. A*, vol. 56, art. 155302, 2023. https://doi.org/10.1088/1751-8121/acc291 ↑ 6

[18] D. Jaschke, M. L. Wall, and L. D. Carr, "Open source matrix product states: Opening ways to simulate entangled many-body quantum systems in one dimension," *Comput. Phys. Commun.*, vol. 225, pp. 59–91, 2018. https://doi.org/10.1016/j.cpc.2017.12.015 ↑ 7

[19] V. M. F. Verstraete and J. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," *Adv. Phys.*, vol. 57, no. 2, pp. 143–224, 2008. https://doi.org/10.1080/14789940801912366 ↑ 7

[20] M. Zheng, Y. Qiao, Y. Wang, J. Cao, and S. Chen, "Exact solution of the bose-hubbard model with unidirectional hopping," *Phys. Rev. Lett.*, vol. 132, art. 086502, Feb 2024. http://dx.doi.org/10.1103/PhysRevLett.132.086502 ↑ 8, 9

[21] V. Colussi, F. Caleffi, C. Menotti, and A. Recati, "Quantum gutzwiller approach for the two-component bose-hubbard model," *SciPost Phys.*, vol. 12, no. 3, art. 111, Mar 2022. http://dx.doi.org/10.21468/SciPostPhys.12.3.111 ↑ 9

[22] G. Catarina and B. Murta, "Density-matrix renormalization group: a pedagogical introduction," *Eur. Phys. J. B*, vol. 96, art. 111, Aug 2023. http://dx.doi.org/10.1140/epjb/s10051-023-00575-2 ↑ 9, 10

[23] I. Georgescu, S. Ashhab, and F. Nori, "Quantum simulation," *Rev. Mod. Phys.*, vol. 86, no. 1, art. 153–185, Mar 2014. http://dx.doi.org/10.1103/RevModPhys.86.153 ↑ 9

[24] S. R. White, "Density matrix formulation for quantum renormalization groups," *Phys. Rev. Lett.*, vol. 69, pp. 2863–2866, Nov 1992. https://doi.org/10.1103/PhysRevLett.69.2863 ↑ 9

[25] A. Haller, "Matrix product states: A variational approach to strongly correlated systems," Bachelor's thesis, Johannes Gutenberg University Mainz, 2014. https://www.tmqs.lu/images/haller_bsc.pdf ↑ 10

[26] S. Keller, M. Dolfi, M. Troyer, and M. Reiher, "An efficient matrix product operator representation of the quantum chemical hamiltonian," *J. Chem. Phys.*, vol. 143, art. 244118, 2015. https://doi.org/10.1063/1.4939000 ↑ 11

[27] J. Cui, J. I. Cirac, and M. C. Bañuls, "Variational matrix product operators for the steady state of dissipative quantum systems," *Phys. Rev. Lett.*, vol. 114, art. 220601, 2015. https://doi.org/10.1103/PhysRevLett.114.220601 ↑ 11

[28] A. Strathearn, P. Kirton, D. Kilda, J. Keeling, and B. Lovett, "Efficient non-markovian quantum dynamics using time-evolving matrix product operators," *Nat. Commun.*, vol. 9, art. 3322, 2018. https://doi.org/10.1038/s41467-018-05617-3 ↑ 11

[29] S. Moudgalya, N. Regnault, and B. A. Bernevig, "Entanglement of exact excited states of affleck-kennedy-lieb-tasaki models: Exact results, many-body scars, and violation of the strong eigenstate thermalization hypothesis," *Phys. Rev. B*, vol. 98, art. 235156, 2018. https://doi.org/10.1103/PhysRevB.98.235156 ↑ 11

[30] B. Pirvu, V. Murg, J. Cirac, and F. Verstraete, "Matrix product operator representations," *New J. Phys.*, vol. 12, art. 025012, 2010. https://doi.org/10.1088/1367-2630/12/2/025012 ↑ 11

[31] N. Nakatani and G. K.-L. Chan, "Efficient tree tensor network states (ttns) for quantum chemistry: Generalizations of the density matrix renormalization group algorithm," *J. Chem. Phys.*, vol. 138, art. 134113, 2013. https://doi.org/10.1063/1.4798639 ↑ 12

[32] S. Szalay *et al*, "Tensor product methods and entanglement optimization for ab initio quantum chemistry," *Int. J. Quantum Chem.*, vol. 115, no. 19, art. 1342 – 1391, 2015. https://doi.org/10.1002/qua.24898 ↑ 12

[33] D. Liu *et al*, "Machine learning by unitary tensor network of hierarchical tree structure," *New J. Phys.*, vol. 21, art. 073059, 2019. https://doi.org/10.1088/1367-2630/ab31ef ↑ 12

[34] M. Niedermeier, M. Nairn, C. Flindt, and J. L. Lado, "Quantum computing topological invariants of two-dimensional quantum matter," *Phys. Rev. Res.*, vol. 6, art. 043288, 2024. https://doi.org/10.1103/PhysRevResearch.6.043288 ↑ 12

[35] A. Wulff, B. Chen, M. Steinberg, Y. Tang, M. Möller, and S. Feld, "Quantum computing and tensor networks for laminate design: A novel approach to stacking sequence retrieval," *Comput. Methods Appl. Mech. Eng.*, vol. 432, part A, art. 117380, 2024. https://doi.org/10.1016/j.cma.2024.117380 ↑ 12

[36] A. Termanova, A. Melnikov, E. Mamenchikov, N. Belokonev, S. Dolgov, A. Berezutskii, R. Ellerbrock, C. Mansell, and M. Perelshtein, "Tensor quantum programming," *New J. Phys.*, vol. 26, art. 123019, 2024. https://doi.org/10.1088/1367-2630/ad985b ↑ 12

[37] A. Singh Bhatia and D. E. Bernal Neira, "Federated learning with tensor networks: a quantum ai framework for healthcare," *Mach. Learn. Sci. Technol.*, vol. 5, art. 045035, 2024. https://doi.org/10.1088/2632-2153/ad8c11 ↑ 12

[38] G. Evenbly, "Hyperinvariant tensor networks and holography," *Phys. Rev. Lett.*, vol. 119, art. 141602, 2017. https://doi.org/10.1103/PhysRevLett.119.141602 ↑ 12

[39] A. Peach and S. F. Ross, "Tensor network models of multiboundary wormholes," *Class. Quantum Gravity*, vol. 34, no. 10, art. 105011, 2017. https://doi.org/10.1088/1361-6382/aa6b0f ↑ 12

[40] A. Jahn, M. Gluza, F. Pastawski, and J. Eisert, "Holography and criticality in matchgate tensor networks," *Sci. Adv.*, vol. 5, no. 8, art. eaaw0092, 2019. https://doi.org/10.1126/sciadv.aaw0092 ↑ 12

[41] F. Verstraete, V. Murg, and J. Cirac, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," *Adv. Phys.*, vol. 57, no. 2, art. 143 – 224, 2008. https://doi.org/10.1080/14789940801912366 ↑ 12

[42] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, "Matrix product states and projected entangled pair states: Concepts, symmetries, theorems," *Rev. Mod. Phys.*, vol. 93, art. 045003, 2021. https://doi.org/10.1103/RevModPhys.93.045003 ↑ 12

[43] N. Schuch, D. Pérez-García, and I. Cirac, "Classifying quantum phases using matrix product states and projected entangled pair states," *Phys. Rev. B Cond. Mat. Mater. Phys.*, vol. 84, art. 165139, 2011. https://doi.org/10.1103/PhysRevB.84.165139 ↑ 12

[44] R. Orús, "A practical introduction to tensor networks: Matrix product states and projected entangled pair states," *Ann. Phys.*, vol. 349, pp. 117–158, 2014. https://doi.org/10.1016/j.aop.2014.06.013 ↑ 12

[45] J. Jordan, R. Orús, G. Vidal, F. Verstraete, and J. Cirac, "Classical simulation of infinite-size quantum lattice systems in two spatial dimensions," *Phys. Rev. Lett.*, vol. 101, art. 250602, 2008. https://doi.org/10.1103/PhysRevLett.101.250602 ↑ 12

[46] G. Vidal, "Class of quantum many-body states that can be efficiently simulated," *Phys. Rev. Lett.*, vol. 101, no. 11, 2008. https://doi.org/10.1103/PhysRevLett.101.110501 ↑ 12

[47] Y.-Y. Shi, L.-M. Duan, and G. Vidal, "Classical simulation of quantum many-body systems with a tree tensor network," *Phys. Rev. A*, vol. 74, art. 022320, 2006. https://doi.org/10.1103/PhysRevA.74.022320 ↑ 12

[48] S. Holtz, T. Rohwedder, and R. Schneider, "The alternating linear scheme for tensor optimization in the tensor train format," *SIAM J. Sci. Comput.*, vol. 34, no. 2, pp. A683–A713, 2012. https://doi.org/10.1137/100818893 ↑ 12

[49] V. Murg, F. Verstraete, O. Legeza, and R. Noack, "Simulating strongly correlated quantum systems with tree tensor networks," *Phys. Rev. B*, vol. 82, art. 205105, 2010. https://doi.org/10.1103/PhysRevB.82.205105 ↑ 12

[50] L. Tagliacozzo, G. Evenbly, and G. Vidal, "Simulation of two-dimensional quantum systems using a tree tensor network that exploits the entropic area law," *Phys. Rev. B*, vol. 80, art. 235127, 2009. https://doi.org/10.1103/PhysRevB.80.235127 ↑ 12

[51] M. L. Wall, A. Reilly, J. S. Van Dyke, C. Broholm, and P. Titum, "Quantum tensor network algorithms for evaluation of spectral functions on quantum computers," *Phys. Rev. B*, vol. 110, art. 214402, 2024. https://doi.org/10.1103/PhysRevB.110.214402 ↑ 12

[52] R. Huang, H. Yi, and P. H. Li, "Variational quantum recurrent neural networks for multi-feature forecasting," *Int. J. Quantum Inf.*, vol. 22, no. 7, art. 2450029, 2024. https://doi.org/10.1142/S0219749924500291 ↑ 12

Juan Sebastián Gómez

Undergraduate student of Physics at Universidad del Valle (Cali, Colombia). Member of the Solid State Theoretical Physics group as part of the research internship program, where he has spent a year working on the implementation of numerical methods based on tensor networks for the analysis of many-body systems. He participated in the International Physicist Tournament (IPT) 2024 held in Zürich, and his interests besides tensor networks include quantum computing and information.

Email: juan.s.gomez@correounivalle.edu.co

Karen C. Rodríguez

Dr. rer. nat., graduated in Physics from Universidad Nacional de Colombia. Her initial work focused on quantum simulators and cold atoms in the Quantum Optics group. She obtained her Master's degree in Theoretical Physics from Universität Stuttgart, working on DMRGs for one-dimensional many-body quantum systems at the intersection with condensed matter. She completed her PhD (*Dr. rer. nat*.) at Universität Hannover in Germany, continuing her work on numerical techniques for high-spin quantum simulators in the Quantum Optics group. She pursued a postdoctoral program at Universität zu Köln, where she extended her research on many-body quantum simulators. Currently, she is a faculty professor at Universidad del Valle, leading the Theoretical Solid State Physics group. Her interests include developing analytical and numerical tools for high-spin many-body quantum simulator systems, exploring novel superconductivity, and discovering new magnetic phases.
Email: karem.c.rodriguez@correounivalle.edu.co