

Proceedings of the

VIII Ibero-American Congress of Smart Cities

ICSC-CITIES 2025

Puebla, Mexico, November 10th to 12th

Hosted by Centro de Investigación en Computación and
Centro de Innovación e Integración de Tecnologías Avanzadas
of the Instituto Politécnico Nacional

CITIES

CIUDADES INTELIGENTES TOTALMENTE INTEGRALES, EFICIENTES Y SOSTENIBLES

Editors:

Pedro Moreno
Luis Hernández-Callejo
Sergio Nesmachnow
Diego Rossit
Carlos Torres-Aguilar
Ponciano J. Escamilla-Ambrosio

Editors

Pedro Moreno
Universidad Autónoma del Estado de Morelos, México

Luis Hernández-Callejo
Universidad de Valladolid, España

Sergio Nesmachnow
Universidad de la República, Uruguay

Diego Rossit
INMABB, DI, Universidad Nacional del Sur-CONICET, Argentina

Carlos Torres-Aguilar
Universidad Juárez Autónoma de Tabasco, México

Ponciano Jorge Escamilla-Ambrosio
Centro de Investigación en Computación del Instituto Politécnico Nacional, México

Elaborated by the Organizing Committee of the VIII Ibero-American Congress of Smart Cities (ICSC-CITIES 2025).

Disclaimer: The information in this book is true and complete to the best of the editor's knowledge. All recommendations are made without a guarantee on the part of the editors. The editors disclaim any liability in connection with the use of this information.

How to cite this book: Moreno P., Hernández-Callejo, L., Nesmachnow, S., Rossit, D., Torres-Aguilar, C. and Escamilla-Ambrosio, P.J. (Eds). (2025). Proceedings of the VIII Ibero-American Congress of Smart Cities. CITIES.

Cover: Carlos Torres-Aguilar

ISBN: 978-84-09-81656-9

This work is licensed under Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International. To view a copy of this license, visit https://creativecommons.org/licenses/by-nc-nd/4.0/.

Preface

The VIII Ibero-American Congress of Smart Cities (ICSC-CITIES 2025) was held from November 10 to 12, 2025, in Puebla, Mexico, in a hybrid format that combined in-person and virtual participation. This edition continues the trajectory of seven successful previous congresses organized in diverse locations: Soria, Spain, in 2018 and 2019; San Carlos, Costa Rica in 2020 and 2024; Cancún, Mexico, in 2021; Cuenca, Ecuador, in 2022; and Mexico City and Cuernavaca in 2023.

The primary objective of ICSC-CITIES 2025 was to provide a forum for researchers, scientists, educators, postgraduate students, industry professionals, and policymakers from Ibero-America and around the world to share their latest initiatives, research findings, and technological innovations to address the challenges of urbanization through smart and sustainable solutions. This proceedings book reflects the multidisciplinary nature of Smart City research, encompassing three major thematic areas: innovative computing solutions for smart cities; smart energy, energy efficiency, and sustainability; and innovative technologies for governance and citizenship.

The congress featured an extensive technical program that included 84 oral presentations, a poster session, and a panel discussion on inclusion in smart cities, emphasizing the social aspects of technological advancement. The Program Committee for ICSC-CITIES 2025 received a total of 136 manuscript submissions, of which 84 were accepted for oral presentation. Thirty-one articles, which specifically addressed computing, energy, and social applications in smart cities, were selected for publication in this proceedings book. This selection process followed a rigorous single-blind peer review conducted by at least three experts in the field. The accepted contributions highlight cutting-edge research and practical applications that enhance the understanding and implementation of smart city concepts.

ICSC-CITIES 2025 brought together more than 300 participants from 29 countries, solidifying its position as an international hub for knowledge exchange and collaboration. The diverse perspectives and expertise of attendees contributed to a dynamic and enriching dialogue on topics such as energy efficiency, the Internet of Things, artificial intelligence, urban informatics, big data analytics, governance, smart mobility, and sustainable development.

We extend our sincere gratitude to all the authors, reviewers, and members of the organizing and program committees for their dedication and professionalism, which were crucial to ensuring the academic quality of this proceedings book. We also appreciate the institutions that supported the congress and the participants whose engagement contributed to the success of the event. Finally, we thank the readers of this proceedings book; your interest and commitment will continue to drive innovation in Smart Cities.

December 2025

Editors of ICSC-CITIES 2025 Proceedings

Organization

General Chairs

Luis Hernández Callejo, Universidad de Valladolid, Spain

Sergio Nesmachnow, Universidad de la República, Uruguay

Pedro Moreno Bernal, Universidad Autónoma del Estado de Morelos, Mexico

Publication Chairs

Sergio Nesmachnow, Universidad de la República, Uruguay

Luis Hernández Callejo, Universidad de Valladolid, Spain, Spain

Pedro Moreno Bernal, Universidad Autónoma del Estado de Morelos, Mexico

Diego Rossit, Universidad Nacional del Sur, Argentina

Carlos Torres Aguilar, Universidad Juárez Autónoma de Tabasco, Mexico

Andrei Tchernykh, CICESE Research Center, Mexico

Submission and Conference Management Chair

Diego Rossit, Universidad Nacional del Sur, Argentina

Organization Committee

Luis Hernandez Callejo, Universidad de Valladolid, Spain

Sergio Nesmachnow, Universidad de la República, Uruguay

Pedro Moreno Bernal, Universidad Autónoma del Estado de Morelos, Mexico

Ponciano Jorge Escamilla-Ambrosio, Instituto Politécnico Nacional, Mexico

Jesús de la Casa Hernández, Universidad de Jaén, Spain

Local Organization Chair

Ponciano Jorge Escamilla-Ambrosio, Centro de Investigación en computación-IPN, Mexico

Local Organization Committee

Juan Humberto Sossa Azuela, Centro de Investigación en computación-IPN, Mexico

Ponciano J. Escamilla Ambrosio, Centro de Investigación en computación-IPN, Mexico

Tomás Garduño Pérez, Centro de Innovación e Integración de Tecnologías Avanzadas-IPN, Mexico

Ayesha Margarita Courrech Arias, Centro de Innovación e Integración de Tecnologías Avanzadas-IPN, Mexico

Program Committee Chairs

Luis Hernández Callejo, Universidad de Valladolid, Spain, Spain

Sergio Nesmachnow, Universidad de la República, Uruguay

Pedro Moreno Bernal, Universidad Autónoma del Estado de Morelos, Mexico

Diego Rossit, Universidad Nacional del Sur, Argentina

Carlos Torres Aguilar, Universidad Juárez Autónoma de Tabasco, Mexico

Track Chairs

Alberto Redondo Plaza, Universidad de Valladolid, Spain

Carlos Torres-Aguilar, Universidad Juárez Autónoma de Tabasco, Mexico

Carlos Baier, Universidad de Talca, Chile

Diego Rossit, Universidad Nacional del Sur, Argentina

Héctor Matteo Romero, Universidad de Valladolid, Spain

Francisco José Sánchez Sutil, Universidad de Jaén, Spain

Jesús Armando Aguilar, Universidad Autónoma de Baja California, Mexico

Jesús De La Casa Hernández, Universidad de Jaén, Spain

Jesús Del Carmen Peralta Abarca, Universidad Autónoma del Estado de Morelos, Mexico

Joao Soares, Instituto Politécnico do Porto, Portugal

José Alberto Hernández Aguilar, Universidad Autónoma del Estado de Morelos, Mexico

Lilian Johanna Obregon, Universidad de Valladolid, Spain

Luis Hernández-Callejo, Universidad de Valladolid, Spain

Oscar Izquierdo, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Spain

Pedro Moreno Bernal, Universidad Autónoma del Estado de Morelos, Mexico

Rodrigo Porteiro, Universidad de la República, Uruguay

Sergio Nesmachnow, Universidad de la República, Uruguay

Susana Gómez Redondo, Universidad de Valladolid, Spain

Vanessa de Almeida Guimaraes, Universidade Federal do Rio de Janeiro, Brazil

Víctor Alonso Gómez, Universidad de Valladolid, Spain

Publicity Chair

Nadia Lara Ruiz, Universidad Autónoma del Estado de Morelos, Mexico

Program Committee

Adrian Toncovich, Universidad Nacional del Sur, Argentina

Agustín Barrera Sánchez, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Spain

Agustina Anselmino, Universidad Tecnológica Nacional-CONICET, Argentina

Alberto Redondo Plaza, Universidad de Valladolid, Spain

Alberto José Cháves Chacón, Universidad Técnica Nacional, Costa Rica

Alicia Martinez Rebollar, Tecnológico Nacional de México CENIDET, Mexico

Amy-Zulema Velasco-Bonilla, Universidad de Valladolid, Spain

Ana Carolina Olivera, Universidad Nacional de Cuyo, Argentina

Ana-Kristhel Esteban-López, Universidad Juárez Autónoma de Tabasco, Mexico

Ana-Lilia Sánchez-Brito, Centro Nacional de Investigación y Desarrollo Tecnológico, Mexico

Andrea Insfran-Rivarola, Universidad Nacional de Asunción, Paraguay

Andrei Tchernykh, CICESE Research Center, Mexico

Andrés Adolfo Navarro Newball, Pontificia Universidad Javeriana, Colombia

Andres Felipe Fuentes Vasquez, Pontificia Universidad Javeriana, Colombia

Ángel Hernández Jiménez, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Spain

Ángel-L Zorita-Lamadrid, Universidad de Valladolid, Spain

Angela Ferreira, Polytechnic University of Bragança, Portugal

Antonio Cano Ortega, University of Jaén, Spain

Antonio Gusmão, Universidade Federal Fluminense, Brazil

Antonio Mauttone, Universidad de la República, Uruguay

Antonio Muñoz, University of Malaga, Spain

Araceli Ávila, Instituto Tecnológico Superior de Perote, Mexico

Armando Huicochea, Universidad Autónoma del Estado de Morelos, Mexico

Beatriz Martínez Bahena, Universidad Autónoma del Estado de Morelos, Mexico

Begoña González Landín, Universidad de Las Palmas de Gran Canaria, Spain

Belisario Morales Morales, Instituto Tecnológico Superior de Centla, Mexico

Benjamín Alvarez-Alor, Instituto Tecnológico Superior de Coatzacoalcos, Mexico

Bouras Abdelkarim, Badji Mokhtar University, Algeria

Brandon Cortés-Caicedo, Universidad Distrital Francisco José de Caldas, Colombia

Brandon Cortés-Caicedo, Universidad Distrital Francisco José de Caldas, Colombia

Bruno Dias, Universidade Federal de Juiz de Fora, Brazil

Bruno França, Universidade Federal Fluminense, Brazil

C. Alberto Ochoa-Zezatti, Universidad Autónoma de Ciudad Juárez, Mexico

Carlos Baier, Universidad de Talca, Chile

Carlos Barroso-Moreno, Universidad Rey Juan Carlos, Mexico

Carlos Grande, Universidad Centroamericana José Simeón Cañas, El Salvador

Carlos Muñoz, Universidad de Talca, Chile

Carlos Torres, Universidad Juárez Autónoma de Tabasco, Mexico

Carlos-Miguel Jiménez-Xaman, Universidad Autónoma de Campeche, Mexico

Carmen Vasquez, Universidad Nacional Experimental Politécnica Antonio José de Sucre , Venezuela

César Varela, Tecnológico Nacional de México CENIDET, Mexico

Christian Cintrano, Universidad de Málaga, Spain

Christian Medina, Universidad Nacional de Asunción , Paraguay

Claudio Paz, Universidad Tecnológica Nacional, Argentina

Claudio Risso, Universidad de la República, Uruguay

Conrado Garcia-Gonzalez, Universidad Autónoma de Baja California, Mexico

Cristina Sáez Blázquez, Universidad de Salamanca, Spain

Daniel Dias, Universidade Federal Fluminense, Brazil

Daniel Rossit, Universidad Nacional del Sur-CONICET, Argentina

Daniel Sanin-Villa, Universidad EAFIT, Colombia

Daniel H. Stolfi, University of Luxembourg, Luxembourg

David Balladares de la Cruz, Universidad Juárez Autónoma de Tabasco, Mexico

David Lorenzi Junior, Universidade Federal de Santa Maria, Brazil

David Peña Morales, University of Cadiz, Spain

Deyslen Mariano, Instituto Tecnológico de Santo Domingo, Dominican Republic

Diego Arcos-Aviles, Universidad de las Fuerzas Armadas ESPE, Ecuador

Diego Rossit, Universidad Nacional del Sur-CONICET, Argentina

Diego Alberto Godoy, Universidad Gastón Dachary, Argentina

Edgar Vicente Macias Melo, Universidad Juárez Autónoma de Tabasco, Mexico

Eduardo Espinosa, Universidad Católica de la Santísima Concepción, Chile

Eduardo Fernández, Universidad de la República, Uruguay

Eduardo Redondo, Universidad Nacional de Asunción, Paraguay

Edwin Gerardo Acuña Acuña, Universidad Latinoamericana de Ciencia y Tecnología, Costa Rica

Efrain Simá-Moo, Tecnológico Nacional de México CENIDET, Mexico

Elkin Henao, Instituto Tecnológico Metropolitano, Colombia

Elmer Iván Sánchez-Rabadán, Universidad Autónoma del Estado de Morelos, Mexico

Emmanuel Millan, Universidad Nacional de Cuyo-CONICET, Argentina

Enedino Arroyo-Martínez, Instituto Tecnológico de Zacatepec., Mexico

Enrique Fernández-Vilas, Universidad de Valladolid, Spain

Enrique González, Universidad de Salamanca, Spain

Erika Avila Melgar, Universidad Autónoma del Estado de Morelos, Mexico

Esteban I. Marciel, Universidad de Talca, Chile

Eyko Medeiros-Rios, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Brazil

Fabian Castillo Peña, Universidad Libre, Colombia

Fábio Castro, Instituto Politécnico do Porto, Portugal

Federico Alonso Pecina, ITESM Campus Cuernavaca, Mexico

Fernanda Villarreal, Universidad Nacional del Sur, Argentina

Fernando Lezama, Instituto Politécnico do Porto, Portugal

Fernando-Adrian Mezzano, Universidad Nacional de San Luis, Argentina

Filiberto Cruz Flores, Centro Nacional de Investigación y Desarrollo Tecnológico, Mexico

Florencia Caro, Universidad de la República, Uruguay

Franciele Weschenfelder, Universidad Tecnológica del Uruguay, Uruguay

Francisco Beltrán-Carbajal, Universidad Autónoma Metropolitana, Mexico

Francisco Esteban, Universidad Nacional de San Luis, Argentina

Francisco Magaña-Hernández, Universidad Juárez Autónoma de Tabasco, Mexico

Francisco Valbuena, Universidad de Valladolid, Spain

Francisco José Sánchez Sutil, Universidad de Jaén, Spain

Gilberto Martinez, Instituto Politécnico Nacional, Mexico

Gonzalo Bojorquez-Morales, Universidad Autónoma de Baja California, Mexico

Guillermo Santamaría-Bonfil, BBVA de Mexico, Mexico

Gustavo Medina Angel, Universidad Autónoma del Estado de Morelos, Mexico

Héctor Chicchiarini, Universidad Nacional del Sur-CONICET, Argentina

Hector Felipe Mateo Romero, Universidad de Valladolid, Spain

Homero Alonso-Jiménez, Universidad Autónoma del Estado de Morelos, Mexico

Hortensia Amaris, University Carlos III de Madrid, Spain

Ignacio Turias, Universidad de Cádiz, Spain

Irma Hernández-Báez, Universidad Politécnica del Estado de Morelos, Mexico

Irvin López, Universidad Autónoma Metropolitana, Mexico

Irving Hernández-López, Universidad de Sonora, Mexico

Ivonne Yazmín Arce García, Universidad Autónoma del Estado de Morelos, Mexico

Jakub Jurasz, Wrocław University of Science and Technology, Poland

Jamal Toutouh, Universidad de Málaga , Spain

Javier Rocher, Universitat Politècnica de València, Spain

Javier Uriarte, Centro Nacional de Metrología, Mexico

Jesus Arce, Tecnológico Nacional de México CENIDET, Mexico

Jesus Lopez Gomez, Universidad Juárez Autónoma de Tabasco, Mexico

Jesús Soto, Instituto Tecnológico de Nuevo León, Mexico

Jesús Vegas, Universidad de Valladolid, Spain

Jesús Armando Aguilar, Universidad Autónoma de Baja California, Mexico

Jesus C. Hernandez, Universidad de Jaén, Spain

Jesús Del Carmen Peralta Abarca, Universidad Politécnica del Estado de Morelos, Mexico

Joao Soares, Instituto Politécnico do Porto, Portugal

João Coelho, Instituto Politécnico de Bragança, Portugal

João Coelho, Instituto Politécnico de Bragança, Portugal

Jonathan Muñoz Tabora, Universidad Nacional Autónoma de Honduras, Honduras

Jonathan Muraña, Universidad de la República , Uruguay

Jorge López-Rebollo, University of Salamanca, Spain

Jorge Esteban Rodas Benítez, Universidad Nacional de Asunción, Paraguay

Jorge-Luis Mirez-Tarrillo, Universidad Nacional de Ingeniería, Peru

José Almeida, Instituto Politécnico do Porto, Portugal

José Alberto Hernández, Universidad Autónoma del Estado de Morelos, Mexico

José Daniel Gámez Avilez, Universidad Autónoma de Baja California, Mexico

José Eduardo Hernández Hernández, Instituto Politécnico Nacional, Mexico

Jose Ignacio Morales Aragonés, Ziguratt Sci., Spain

José Mauricio Rodríguez Valencia, Universidad Juárez Autónoma de Tabasco, Mexico

José Ramón Aira, Universidad Politécnica de Madrid, Spain

José-Alejandro Suástegui, Universidad Autónoma de Baja California, Mexico

Jose-Isidro Hernández-Vega, Instituto Tecnológico de Nuevo León, Mexico

José-Ramón Ayala, Universidad Autónoma de Baja California, Mexico

Joseph K. Sasu, University of Ghana, Ghana

Juan Espinoza, Universidad de Cuenca, Ecuador

Juan Giraldo, Universidad EAFIT, Colombia

Juan Pavón, Universidad Complutense de Madrid, Spain

Juan Romay-Coca, Universidad de Valladolid, Spain

Juan Carlos Martínez Jiménez, Universidad Juárez Autónoma de Tabasco, Mexico

Juan Humberto Sossa Azuela, Instituto Politécnico Nacional, Mexico

Juan Manuel Ramírez Alcaraz, Universidad de Colima, Mexico

Juan Matías Novas, Universidad Nacional de Córdoba-CONICET, Argentina

Juan Miguel Durán, Universidad Autónoma de Campeche, Mexico

Juan-Manuel Hurtado, Instituto de Biotecnología - UNAM, Mexico

Julian Patiño, Institución Universitaria Pascual Bravo, Colombia

Juliana Torres Morales, Tecnológico de Antioquia Institución Universitaria, Colombia

Karla María Aguilar Castro, Universidad Juárez Autónoma de Tabasco, Mexico

Kevin Pavez, Universidad Nacional de San Luis, Argentina

Laene Oliveira Soares, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Brazil

Larizza Delorme, Universidad Nacional de Asunción, Paraguay

Lenin Ramos, Universidad Juárez Autónoma de Tabasco, Mexico

Leonardo Cardinale-Villalobos, Instituto Tecnológico de Costa Rica, Costa Rica

Leonardo Suarez Matarrita, Universidad Técnica Nacional, Costa Rica

Leticia Neria Monroy, Universidad Autónoma del Estado de Morelos, Mexico

Lilian Johanna Obregón, Universidad de Valladolid, Spain

Luis Angeles, Universidad Juárez Autónoma de Tabasco, Mexico

Luis Hernández-Callejo, Universidad de Valladolid, Spain

Luis Torres, Universidad Nacional de Río Cuarto, Argentina

Luis Aarón Martínez Figueroa, Universidad Centroamericana "José Simeón Cañas" (El Salvador) , El Salvador

Luis Antonio Gonzalez Uribe, Universidad Autónoma de Baja California, Mexico

Luis Fernando Grisales Noreña, Universidad del Valle, Colombia

Luis G. Montané Jiménez, Universidad Veracruzana, Mexico

Luis Manuel Navas Gracia, Universidad de Valladolid, Spain

Luis-Enrique Angeles-Montero, Universidad Juarez Autónoma de Tabasco, Mexico

Magdalena Irurtia, Administración Nacional de Usinas y Trasmisiones Eléctricas, Uruguay

Mailyn Moreno-Espino, Universidad Complutense de Madrid, Spain

Manuel Olan Ramos, Instituto Tecnológico Superior de Comalcalco, Mexico

Manuel Jesus Rodriguez-Perez, Universidad Autónoma de Campeche, Mexico

Marcin Seredynski, E-Bus Competence Center, Luxemburg

Marcos Coronado, Universidad Autónoma de Baja California, Mexico

Margherita Valle Pilia, Universidad Latinoamericana de Ciencia y Tecnología, Costa Rica

María Sánchez-Aparicio, Universidad Politécnica de Madrid, Spain

Maria Teresa Cepero, Universidad Veracruzana, Mexico

Marian Acero, Universidad de Zaragoza, Spain

Mariana Coccola, Universidad Tecnológica Nacional-CONICET, Argentina

Mariana Funes, Universidad Nacional de Córdoba, Argentina

Mariano Frutos, Universidad Nacional del Sur-CONICET, Argentina

Mario Carbono de la Rosa, Universidad Nacional Autónoma de México, Mexico

Miguel Aybar, Instituto Tecnológico de Santo Domingo, Dominican Republic

Miguel Chen Austin, Universidad Tecnologica de Panama, Panama

Miguel Davila, Universidad Politécnica Salesiana, Ecuador

Miguel Angel Che Pan, Instituto Tecnológico Superior de Perote, Mexico

Miguel Ángel Juárez Merino, Universidad Nacional Autónoma de México, Mexico

Miguel-Ángel Muñoz-García, Universidad Politécnica de Madrid, Spain

Miriam Navarrete, Universidad Autónoma del Estado de Morelos, Mexico

Monica Alonso, University Carlos III de Madrid, Spain

Mónica Montoya Giraldo, CIDET, Colombia

Nadia Lara Ruiz, Universidad Autónoma del Estado de Morelos, Mexico

Ney Balderramo, Universidad Técnica de Manabí, Ecuador

Nicolás Plaza Gómez, Universidad de Salamanca, Spain

Noelia Uribe-Perez, European Software Institute TECNALIA, Spain

Noelia Soledad Pinto, Universidad Tecnológica Nacional, Argentina

Olalla García Pérez, Universidad Europea de Madrid, Spain

Oscar Duque-Perez, Universidad de Valladolid, Spain

Oscar Izquierdo, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Spain

Oscar Danilo Montoya Giraldo, Universidad Tecnológica de Pereira, Colombia

Osvaldo González, Universidad Nacional de Asunción, Paraguay

Pablo Daniel Godoy, Universidad Nacional de Cuyo, Argentina

Paula Guerra, Kennesaw State University, United States

Paula Peña-Carro, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Spain

Pedro Moreno, Universidad Autónoma del Estado de Morelos, Mexico

Pedro Piñeyro, Universidad de la República, Uruguay

Pedro Henrique González, Universidade Federal do Rio de Janeiro, Brazil

Ponciano Jorge Escamilla-Ambrosio, Instituto Politécnico Nacional, Mexico

Rafael Campos-Amezcua, Tecnológico Nacional de México CENIDET, Mexico

Ramiro Martins, Polytechnic Institute of Bragança, Portugal

Raquel Lozano Blasco, Universidad de Zaragoza, Spain

Raúl Alberto López Meraz, Universidad Veracruzana, Mexico

Regina León-Ovelar, Universidad Nacional de Asunción, Paraguay

Renzo Massobrio, University of Antwerp, Belgium

Ricardo Beltran-Chacon, Centro de Investigación en Materiales Avanzados, Mexico

Roberto Flores, Universidad Autónoma del Estado de Morelos, Mexico

Roberto Quintal Palomo, Universidad Autónoma de Yucatán, Mexico

Roberto Ramírez Basterrechea, , Germany

Roberto Villafafila, Universitat Politècnica de Catalunya, Spain

Roberto-Carlos Cárdenas, Tecnológico Nacional de México, Mexico

Rodrigo Alonso-Suárez, Universidad de la República, Uruguay

Rodrigo Porteiro, Universidad de la República, Uruguay

Rodrigo Tóffano, Universidade Federal de Goiás, Brazil

Rogelio Vargas Lopez, Universidad Tecnológica de Jalisco and Universidad Autónoma de Guadalajara, Mexico

Roger Castillo Palomera, Universidad Juárez Autónoma de Tabasco, Mexico

Rosenberg Romero, Universidad Autónoma del Estado de Morelos, Mexico

Rubén-Iván Bolaños, Instituto Tecnológico Metropolitano, Colombia

Ruitong Yang, Northeast Petroleum University, China

Ryszard Edward Rozga Luter, Universidad Autónoma Metropolitana, Mexico

Samanta López Salazar, Northeast Petroleum University, Mexico

Saúl Islas-Pereda, Universidad Autónoma de Baja California, Mexico

Sebastián Pereira, Universidad Tecnológica Nacional-CONICET, Argentina

Sergio Nesmachnow, Universidad de la República, Uruguay

Sérgio Ramos, Instituto Politécnico do Porto, Portugal

Silvia Soutullo, Centro de Investigaciones Energéticas, Medioambientales y Tecnológicas, Spain

Susana Gómez Redondo, Universidad de Valladolid, Spain

Teresa Batista, Universidade de Évora, Portugal

Ulises Cruz Jacobo, Universidad Autónoma del Estado de Morelos, Mexico

Ulises Cruz Jacobo, Universidad Autónoma del Estado de Morelos, Mexico

Valentín Cardeñoso-Payo, Universidad de Valladolid, Spain

Vanessa Machado, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Brazil

Vanessa de Almeida Guimarães, Universidade Federal do Rio de Janeiro, Brazil

Vicente Leite, Instituto Politécnico de Bragança, Portugal

Victor Garrido, Universidad de Pamplona, Spain

Victor Viveros González, Instituto Politécnico Nacional, Mexico

Victor Zezatti, Universidad Autónoma del Estado de Morelos, Mexico

Víctor Alonso-Gómez, Universidad de Valladolid, Spain

Víctor Hugo Pacheco Valencia, Universidad Autónoma del Estado de Morelos, Mexico

Vitervo López, Tecnológico Nacional de México CENIDET, Mexico

Wallice Souza, Centro Federal de Educação Tecnológica Celso Suckow da Fonseca, Brazil

Walter Gil, Universidad Tecnológica de Pereira, Colombia

Walter Gil, Universidad Tecnológica de Pereira, Colombia

Yamila Soledad Grassi, Universidad Nacional del Sur-CONICET, Argentina

Contents

Innovative computing solutions for smart cities

Method for calibration of MQ family sensors for use in the Arduino IDE ……..…..…1

Dariel Gustavo Hernández Montalvo, Abelardo Rodriguez León, Guillermo Efren Ovando Chacón, Olaya Pirene Castellanos Onorio and Miguel Angel Perez Cabada

Sustainable IoT-Based Architecture for Real-Time water Quality Monitoring in Smart Cities……16

Nicole Michelle Leyva Munguía and Rocio Maciel

Classification of avocado leaves into healthy and unhealthy by computer vision and deep learning ……..…….…….....30

Liana Leiva Romero, José Alberto Hernández, Outmane Oubram, Pedro Moreno, Gustavo Medina-Angel, Ponciano Jorge Escamilla-Ambrosio and Jesús Del Carmen Peralta Abarca

Data Strategy as a Value Enabler in Sustainable, Smart, and Circular Cities……...……43

Augusto Velasquez Mendez, Tito Pablo Neira Avila, Jorge de Jesús Lozoya Santos and José Fernando Jiménez Vargas

Análisis comparativo de métodos predictivos para la completitud de datos en series de producción fotovoltaica: regresión lineal y Gradient Boosting..……58

Marta Lafuente-Cacho, Oscar Izquierdo-Monge, Paula Peña-Carro, Ángel Hernández-Jiménez and Carlos Barrera-del Amo

Artificial Neural Networks for Temperature Prediction in Urban Gardens Using Data from Grey Urban Spaces in Valladolid, Spain…….71

Francisco Tomatis, Francisco Javier Diez, Adriana Correa-Guimaraes and Luis Manuel Navas Gracia

Multi-objective vehicle routing problem with time windows and service restrictions……87

Luis Rementería, Andrei Tchernykh and Sergio Nesmachnow

Smart energy, energy efficiency, and sustainability

Systematic review of methods for flow optimization in photovoltaic solar pumping systems under conditions of irradiance variability ……..……102

Manuel González, Félix Santos, Deyslen Mariano-Hernández, Martin Melendez, Armando J Taveras-Cruz, Miguel Aybar-Mejía, Fabio Andrade Rengifo and Marcel Castro-Sitiriche

Analysis of the Impact of the Penetration of Variable Renewable Energies on the Inertia of the National Intercon-nected Electric System (SENI) of the Dominican Republic and Mitigation Proposals ……..……116

Miguel Aybar-Mejía, Onésimo Ramirez, Rony Montero, Deyslen Mariano-Hernández, Roberto Eduardo Quintal Palomo, Armando J Tavera-Cruz and Máximo Alberto Domínguez Garabitos

Agrivoltaic evaluation of horticultural crops for Balearic tourism: Comparative analysis of lettuce, tomato and zucchini……...131

Alicia Nadal-Salas, Marta Pons-Nieto, Esperanza Dalmau, Antoni Capellà-Ferretjans and Susana Simal

Estimation of heat transfer coefficients in a dual falling film condenser……..146

Ariadna Huerta Ortiz and Laura Irene Morales Gomez

Urban decarbonization through PV installations on building facades: the real-life case of a university building in southern Spain…….161

Carlos Gilabert-Torres, Giuseppe M. Tina, Gabino Jiménez-Castillo, Francisco José Muñoz-Rodríguez and Catalina Rus-Casas

Regeneration of absorbent solution in a desorption column for post-combustion CO2 capture at laboratory scale……..176

Karla Anahi Altamirano Tlali, Laura Lília Castro Gómez, Miriam Navarrete Procopio, Víctor Manuel Zezatti Flores and Juan Carlos Garcia Castrejon

Optimizing Solar Panel Allocation in Smart-City Buildings Using Genetic Algorithms……..188

Ricardo Alcaraz Fraga and Jorge Ponciano Escamilla Ambrosio

Quantifying Environmental Benefits of Energy Refurbishment: From Fossil Fuel to Biomass Heating in University Buildings - A Case Study……199

Diana Alexandra Murcia-Velasco, Adriana Correa-Guimaraes and Luis Manuel Navas-Gracia

Operational Lessons from a Commercial PV and Battery System in Jamaica: Curtailment, Frequency, and Demand Charge Management……...207

Toni Ann Hendricks, Jayaka Campbell, Dudley Williams, A. Anthony Chen, Kevin Mills, Randy Koon Koon and Aston Stephens

Comparative Analysis of Artificial Intelligence and Classical Recursive Estimation Methods for AMI Monitoring in Low-Voltage Distribution Networks ……..…… 222

Marlon Geovanny Roa Serpa, Wilson Andrés Ocampo Cardozo and Victor Garrido Arevalo

Smart Industry Using Industry 4.0 and 5.0 to Improve Manufacturing Processes in Small and Medium-Sized Businesses in México ……..……..……233

Jose-Isidro Hernández-Vega, Luis-Alejandro Reynoso-Guajardo, Jesús-Enrique Soto-Soto, María-Ernestina Macías-Arias, Mario-Carlos Gallardo-Morales and Jorge-Eduardo Ortega-Lopez

Observación Territorial integral: Análisis de Pendiente, Radiación y Confort térmico como Determinantes de la Caminabilidad en Entornos Urbanos ……..……..……246

Cinthya Butrón-Revilla, Liliana Valverde-Caballero and Yolanda Pamo Diaz

Innovative technologies for governance and citizenship

Data Literacy for Smart Governance: An Educational-Analytical Framework for Ibero-American Cities……261

Mailyn Moreno-Espino, Jorge Alberto Pachecho-Sernard, Yahima Hadfeg-Fernández, Gerardo Acevedo-Sánchez, Nayma Cepero-Pérez and Cornelio Yáñez-Márquez

Implementing data governance programs in a Colombian public institution ……...…271

Gina Maestre-Góngora, Diana López, Fernando Osorio Giraldo, Osvaldo Cifuentes Restrepo, Juan Aragon-Mendoza and Myriam Patricia Cifuentes

Urban agriculture in Morocco – Systematic Review…….……284

Ouiam Fatiha Boukharta, Loïc Sauvé, Francisco Tomatis, Leticia Chico-Santamarta and Luis Manuel Navas Gracia

Artificial Intelligence in Digital Public Policies in Latin America: Tensions, Gaps, and Transferable Lessons…….…...294

Felipe Bonilla-Sánchez, Adrian Cabral Martínez and Elmer Iván Sánchez Rabadan

From Digital Government to Smart and Sustainable Cities in Latin America……..…….310

Miguel Ángel Juárez Merino

Integra Hub: Territorial Educational Innovation for Inclusive Smart Cities A University. Submodel for Technological and Social Articulation in Regional Contexts……325

Alejandro Vergara Berra and José David Ortega Cerecedo

Assessing the Impacts of Mega Projects on Urban Thermal Dynamics in Riyadh, Saudi Arabia…..338

Ayman Imam

Reframing Urban Mobility: Assessing Land Use, Accessibility, and Equity through the Riyadh Metro……...….353

Abdulrhman Gbban

Urban Agriculture with an IAR-FAO Approach During COVID-19: A Socio-Educational Strategy from Cuenca, Ecuador……368

Fredi Portilla Farfán, Paola Duque-Sarango, María José Quizhpi Pintado and Tania Alexandra Sarango Guamán

El marco de las Ciudades Inteligentes como un reforzamiento de la Cohesión Urbana Inclusiva…380

Miroslaba Taide Bautista Flores, Ryszard Edward Rozga Luter and Miroslaba Taide Bautista Flores

Identificación y análisis de factores relevantes del desempeño en matemáticas en México con datos PISA 2012……395

Esteban Paul Louis Nicolas, Gilberto Lorenzo Martínez Luna and Adolfo Guzmán Arenas

Landscapes of memory and sustainable education in rural contexts: a STEAM learning project around hydraulic mills ……..……..410

Susana Gómez, Raúl Maján Navalón and Lilian Johanna Obregón

Posters

The energy transition in Cuba: process characteristics ……..……420

Miguel Castro Fernández, Miriam Vilaragut Llanes and Roberto Eduardo Quintal Palomo

Preliminary Study on Convolutional Neural Network-Based Burn Diagnosis: Advancing Smart Health in Urban Environments ……421

Jose de Jesus Alvarez Ramirez, Rocio Maciel Arellano and Víctor Larios Rosillo

Photovoltaic Potential in a City like Elche: Solar Plants and Agriculture ……422

D. López, S. Valero, C. Senabre and Jesus C. Hernandez

Hybrid solar lighting system with optical fiber and LED with continuous lighting control using PW ……..……..….423

Carlos Henriques Oliveira and Marcio Fortes

Aplicación de algoritmos genéticos para la optimización de la generación fotovoltaica en las redes de distribución de Puerto Rico mediante el software PowerFactory ..……424

Daniel Aragones Vargas, Marcel Castro-Sitiriche, Fabio Andrade and Jesús Hernández

Implementation and Validation of an FPGA-Based System for Harmonic and Supraharmonic Analysis in Single-Phase Power Grids ……425

Roger Challco, Francisco José Sánchez Sutil, José Carlos U. Peña, Antonio Cano Ortega and Jesús de la Casa Hernández

Position Control of an 18 DoF Humanoid Robot Using a Saturated Control Law ...……426

Miguel Angel Ortega Palacios, Fernando Reyes Cortés and Amparo Palomino Merino

NSN, AIoT en organizaciones civiles de base y redes comunitarias ……427

Carlos Astroza and Adolfo Mackay

Urban energy modelling and data flow: review and application cases for decarbonization and resilience in Spanish cities ……428

Pablo de Agustín Camacho and David Garcia

Eficiencia del extractor de noticias fitosanitarias (FREM) ...……429

Rodrigo Olarte Astudillo, Gilberto Lorenzo Martínez Luna, Adolfo Guzmán Arenas, Manolo Muñiz Merino, Sagrario Hernández Mata and Yadira Juliet Pascualito Jimenez

A machine-learning aid model for estimating lean manufacturing implementation in the Latin-American industry …..……430

Diego Rossit, Jesús del Carmen Peralta-Abarca, Andrea Insfran-Rivarola, Pedro Moreno, Ana Pamela Arévalos, Daniel Rossit and Sergio Nesmachnow

Agent-Based Model Simulation as a Decision Support Tool for Renewable Energy Communities Diffusion ……..……..……431

Eduard Falcó Artieda, Roberto Villafafila-Robles, Marina Codina Escolar, Alessandro Sebastiano Carrus, Ivan Blečić and Giuseppe Desogus

OptIoT-ResourSense: Agriculture Optimization with IoT of Resources through Sensing……...…432

Enrique González González, Miguel Ángel Sánchez Herreros, Alonso Garrido Limiñana, Ana del Campo Sánchez, Pedro Antonio Plaza Álvarez and Juan Isidoro Martínez Gutiérrez

Energy-Saving Potential of Natural Ventilation in Temperate Climates: A Case Study of Spain…433

Carolina Aparicio-Fernández, José-Luis Vivancos, Ali Rafat Gigasari and Mar Cañada Soriano

Implementación de Medidas para Mejorar la Eficiencia Energética y los Indicadores Claves de Rendimiento (ICRs) y de Medio ambiente (IAs), en la Producción de Vapor en un Caldero Pirotubular de una nueva PyME - Pequeña y Mediana Empresa ..……434

Félix Rustan Roca Subirana

A System Dynamics Model for the Development of Innovation Districts ……..……435

Jorge Alberto Durán Encalada, Carina A. Rappeti and Josep M. Piqué

Comparative Analysis of Power Demand Models for Electric Traction Systems in Colombian Public Transport ……..……..……436

Harrynson Ramírez Murillo, Juan Sebastián Martínez Donoso, Carlos Andrés Torres Pinzón, Óscar Danilo Montoya Giraldo and Jesús C. Hernández

Towards the Adaptability of Information Visualization for Smart City Technologies……..……437

Irma Elizabeth Romero Arrioja, Luis Gerardo Montane Jimenez and Luis Alfredo Rojano Ruiz

Desafíos metodológicos en la recolección de datos de movilidad urbana basados en encuestas con monitoreo personal. Caso: Comunidad Universitaria de Arequipa y Lima ……..……439

Cinthya Lady Butron Revilla and Jesús Santiago Ramos

The role of third-party logistic providers in Brazilian e-waste collection: A Case Study with Repense ……..……440

Gabriela Rodrigues, Thiago Teixeira and Vanessa de Almeida Guimarães

Articles selected for publication in Volume 2742 of the Communications in Computer and Information Science series by Springer-Nature

Enhancing Solar Energy Research: A Database Approach for Photovoltaic Laboratories
Hector Felipe Mateo Romero, Luis Hernández Callejo, Miguel Angel González Rebollo, Oscar Martinez Sacristan, Alberto Redondo Plaza and Víctor Alonso Gómez

Smart meter based on demand forecasting for real-time architecture
Dario Benavides, Paul Arévalo, Alberto Ríos, Leonardo Torres and Danny Ochoa-Correa

IoT-Driven Smart Mobility: Enhancing Urban Transportation Efficiency with AI and Big Data
Edwin Gerardo Acuña Acuña, Christopher Hernández López, José David Leal Caro, Faider José Solano Herrera, Carlos Vega Quirós and Kenneth Andrés López Cerdas

Quantum-resilient smart cities: Cryptographic strategies and implementation challenge
Gabriel Silva

Inspection of Land Surface Temperature Estimation Methods contrasted to in-situ Measurements: A Case Study in Panama
Miguel Chen Austin, Mileyka Bustamante, Dafni Mora and Félix Henríquez

Clustering techniques to improve the prediction of the THD index for poorly correlated predictor variables
Manuel R. Vasquez Castellanos, Andrés Ortiz Salazar, Francisco J. Sánchez Sutil, Jesús De la Casa Hernández and Antonio Cano Ortega

Evolutionary algorithms for the optimization of irrigation systems
Nicolás Buero, Joaquín Tomás, Diego Rossit and Sergio Nesmachnow

IoT platform for detecting pavement irregularities on urban roads
Orlando Reyes, Pedro Moreno, Gustavo Medina-Ángel, Felipe Bonilla-Sánchez and Juan Manuel Hurtado-Ramírez

A QR-based traceability system for avocado production in Mexico
Jesús Del Carmen Peralta-Abarca, Pedro Moreno, José Alberto Hernández-Aguilar, Felipe Bonilla-Sánchez and Diego Rossit

Flood analysis using remote sensing and cloud computing: a WebGIS tool for urban flood management
Junior A. Calvo Montañez, Jorge López-Rebollo and Susana Del Pozo

Optimal Dispatch in a Multinodal AC Microgrid Model with Battery Energy Storage System (BESS) and Solid-State Transformer (SST)
Emerson Madrid-Lorca, Yamisleidi Salgueiro-Sicilia, Esteban I. Marciel, Carlos A. Munoz, Luis Grisales-Noreña, Oscar Danilo Montoya and Carlos R. Baier

Comparative Forecasting of Hospital Cardiovascular Emergency Demand in the Dominican Republic Using Time Series and Machine Learning Models
Deyslen Mariano-Hernández, Navila Leon-Tejeda, Edwin Sánchez Camilo, Armando J. Taveras-Cruz, Miguel Aybar-Mejia and Iván Jiménez-Durán

A bi-objective model for the bins allocation and collection routing problem in waste management
Diego Rossit, Begoña González Landín, Mariano Frutos and Máximo Méndez Babey

Neural networks for modeling and predicting pollution in urban environments
Josefina Cardozo and Sergio Nesmachnow

A multiobjective evolutionary algorithm for the bus stops location problem: optimizing demand coverage, travel times, and operational costs
Santiago Gestal, Santiago López de Haro, Diego Rossit and Sergio Nesmachnow

Eddies as indicators of local sargassum inundation risks on the south coast of Dominican Republic
Iván Jiménez-Durán, Lodriz Lorenzo Rodríguez, Miguel Aybar-Mejía, Deyslen Mariano-Hernández and Carlos Sanlley

A Neural Network Approach for Super-resolution of a Digital Elevation Model
Facundo Locatelli, Sergio Nesmachnow and Carlos López

A bi-objective model for task assignment with multi-skilled workers and customer deadlines in service companies
Daniel Viera, Diego Rossit, Sergio Nesmachnow and Claudio Risso

Intelligent Operation of BESS for Active and Reactive Power Management in AC Microgrids Using the Crow Search Optimizer
Daniel Sanin-Villa, Luis Fernando Grisales Noreña and Rubén Iván Bolaños

Estimation model for arrival times in urban public transportation of Cuernavaca using artificial neural networks
Edelmira Tapia, Pedro Moreno, Sergio Nesmachnow, Carlos Torres-Aguilar, Jesús del Carmen Peralta-Abarca and Jose Alberto Hernández-Aguilar

Traffic sign detection for autonomous driving at 1:10 scale a YOLO approach
Luis Antonio Marquez-Carlos, Erika P. Sanchez-Femat, Rafael Reveles-Martinez, Huitzilopoztli Luna-Garcia, Jose Guadalupe Arceo-Olage, Juvenal Villanueva-Maldonado, Juan Ruben Delgado-Contreras, Manuel Alejandro Soto-Murillo, Jorge Isaac Galvan-Tejada and Jose Maria Celaya-Padilla

A Q-learning approach for long-term hydrothermal dispatch
Florencia Roque, Matías Rama, Ignacio Salas, Mónica Carle and Rodrigo Porteiro

Electric Vehicles–based Microgrid for Demand Response Studies using a Stochastic Algorithm
Emmanuel Hernández-Mayoral, Gregorio Martínez-Reyes and Iridian Guadalupe Ortíz-Jiménez

Design and Implementation of a Low-Cost IoT Air-Quality Monitoring Network at Queen Mary University of London Mile End Campus
Andrés Aharhel Mercado-Velázquez, Stefan Poslad and Ponciano Jorge Escamilla-Ambrosio

Towards timely detection of student stress and anxiety
Evelyn Scarlett Angeles Calleja, Ponciano Jorge Escamilla Ambrosio, Gilberto Lorenzo Martínez Luna, Abril Valeria Uriarte Arcia and Enrique Garcia Ceja

Clustering Algorithm for Crack Detection on 2D Surfaces: Analysis of Its Dependence on Simulated Surface Size
Gustavo Medina-Ángel, Gennadiy Burlak, Pedro Moreno, Erika Martínez-Sánchez, Mayra Medina-Ángel and Dafne Daniela Manzano-Fuentes

An smart industry study approach in photovoltaic panels and FEM design
David Balladares-de-la-Cruz, Carlos-Enrique Torres-Aguilar, Luis-Enrique Angeles-Montero, Luis Manuel Lopez-Manrique, Julio-Cesar Ramírez-Hernández and Justino Alavez Ramirez

Implementation of a Metaheuristic based on K-means and Ant Colony Optimization
Sinuhe Gines Palestino, Eduardo Roldán Reyes, Marcela Quiroz Castellanos and Adilene Palma Asunción

Brazilian Bus Transport Efficiency: A Simar -Wilson DEA with Beta-Tobit Analysis
Ana Beatriz Fernandes Fangueiro, Jorge Junio Moreira Antunes and Vanessa de Almeida Guimarães.

Articles selected for publication in the Springer Proceedings of Energy series by Springer-Nature

Economic impact of thermal comfort: A climate zone and electricity tariff subsidy approach
J. A. Aguilar-Jiménez, J. Rios-Arriola, N. Velázquez-Limón, S. Islas-Pereda, J. A. Suástegui-Macías, and M. Zamora-Machado

Lightweight energy consumption models for high performance computing
Jonathan Muraña and Sergio Nesmachnow

Optimizing solar chimney configuration considering cost, quality of service and carbon footprint
Carlos Torres-Aguilar, Diego Rossit, Pedro Moreno, Sergio Nesmachnow, Edgar Vicente Macias Melo and Karla María Aguilar Castro

Economic–environmental dispatch of thermal power plants using the Newton-Raphson optimization algorithm
David Alejandro Tohapanta-Quiranza, Jairo Andres Sedano-Duque, Carlos Adrián Correa-Florez,Oscar Danilo Montoya, and Rubén-Iván Bolaños

Electric Vehicle Bidirectional Charging Infrastructure: A Review of its Impact on the Grid
Miguel Aybar-Mejía, Anyel Guzmán-De los Santos, Cristhian Medina-Peña, Oscar Pérez-Otáñez, Yendry Estévez-Feliz, Armando J Taveras-Cruz, and Deyslen Mariano-Hernández

Methodological analysis and techno-economic simulation of the bioenergy potential of agricultural waste in Manabí, Ecuador
Cristopher E. Barberán Torres, Ney R. Balderramo Vélez, María G. Intriago Cedeño, Gabriel E. Pico Mera, and Miguel Castro Fernández

A Mixed Combinatorial Optimization Model to Manage the Hydrothermal Dispatch for the Uruguay River Hydroelectric Plant
Claudio Risso, Sergio Nesmachnow, Rodrigo Porteiro, J. Mario Vignolo, Emiliano Sierra, and Mario Ibarburu

Assessing PCM Melting Range for Passive Cooling in Tropical Climate: An Infrared Thermography Approach Using Jamaican Climate Data
Dudley Williams, Jean-françois Dorville, Jayaka Campbell, and Marhourn Ferhat

Design of a control strategy to reduce DC offset and maximum HFT current of a DAB converter under open circuit fault in the first active bridge
K. Pavez, F. Esteban, F. Serra, C. Baier, E. Espinosa, B. França, and M. Aybar-Mejia

Evaluating Non-Network Solutions for Distribution Network Reinforcement Deferral: A 30-Year Scenario-Based Analysis
Fabio Castro, Fernando Lezama, João Soares, Bruno Canizes, Jesús C. Hernández, and Zita Vale

Design and Comparative of a Single-Stage Amplifier Using Radiation-Hardened Annular MOSFETs vs. Conventional MOSFETs
Jacobo Patiño Bucheli, Fabio Andrade Rengifo, and Federico Serra

Analysis of Classic Multilevel Inverters for Electric Vehicle Applications
Eduardo E. Espinosa, Ricardo Lizana F., Guillermo Ramírez A., Pedro Melin, Carlos Baier, Federico M. Serra, Osvaldo Gonzalez, and Jesús C. Hernández

Design and Evaluation of a Hybrid Electronic Speed Controller for Brushless DC Motors for Smart Mobility Applications
David Marcos-Andrade, Francisco Beltran-Carbajal, Ivan Rivas-Cambero, Daniel Guillen, and Jesús C. Hernández

Incorporating Voltage-Controlled Nodes into the MatrixBased Successive Power Flow Method
Oscar Danilo Montoya, Carlos R. Baier, and Luis Fernando Grisales-Noreña

FCS-MPC multi-objective in NPC with LCL filter in GCPVS under irradiance transients
Adolfo Gonzalez, Gianyacomo Zucchini, Jairo De Oliveira, Matias Aguilar, Osvaldo González, Raúl Gregor, Julio Pacher and Eduardo Espinosa

Adaptive Proportional-Resonant Current Control for Three-Phase Induction Motors
Néstor Villamayor, Magno Ayala, Osvaldo Gonzalez, Gustavo Ojeda, Gianyacomo Zucchini, and Franciele Weschenfelder

Sliding Mode Control of Three-Level NPC Converter-Based STATCOM for Reactive Power Compensation
Jairo De Oliveira, Matías Aguilar, Juan Insfran, Fernando Antúnez, Adolfo González, Raúl Gregor, Leonardo Comparatore and Osvaldo González

DISO Interleaved Buck Converter with PV–Wind Sources for 2nd-Life Li-Ion Battery Charging
Mauricio Mendes da Silva, Conrado Fleck dos Santos, and Claudio Yorsi de los Santos

Application of the ACT5S Numerical Method for Solving Power Flow Equations in DC Monopolar Systems
Oscar Danilo Montoya, Juan Diego Pulgarín Rivera, and Walter Gil-González

Electrical characterization of a photovoltaic thermal system using I-V curves: an experimental study in Soria, Spain
Amy Zulema Velasco-Bonilla, Rachel Josephine Aniny Okango, Luis Hernández-Callejo, Alberto Redondo-Plaza, Víctor Alonso Gómez, Carlos Enrique Torres-Aguilar, Angel Zorita Lamadrid, Alejandro del Amo Sancho, Ali Bassam, and Ghada Amami

Electroluminescence Imaging Comparison for Crystalline Silicon Photovoltaic Modules
Alberto Redondo-Plaza, Amy Zulema Velasco-Bonilla, Vanessa de Almeida Guimarães, Ghada Amami, Ángel L. Zorita-Lamadrid, Víctor Alonso-Gómez and Luis Hernández Callejo .

Renewables to save the island from islanding: the Cozumel study case
Roberto Eduardo Quintal Palomo, Ezequiel Cedillos Panti, Francisco Javier Mendoza Vizcaino , Mauricio Escalante Soberanis, Miguel Aybar-Mejía

Improving AI imaging models for a Multimodal Photovoltaic fault detection system
Leonardo Cardinale-Villalobos, Efrén Jiménez-Delgado, and Magaly Segura-Arias

Reinforcement Learning for short-term battery dispatch in renewable-thermal power systems
Magdalena Irurtia, Santiago Irurtia and Rodrigo Porteiro

Electromobility Markets and Policies in Ibero-America: Progress and Challenges
Carmen Luisa Vásquez Stanescu, Luis Aarón Martínez and Renato Andara

Policies towards zero-emission maritime port infrastructure and buildings
Carmen Luisa Vásquez, Rodrigo Ramírez-Pisco, Luís Manuel Navas, Teresa Batista

Ultra-Fast EV Charging Stations in Distribution Grids: Impacts and Hosting Capacity Assessment
Caio Cesar Amorim Silva; Walquíria Nascimento Silva; Bruno Henriques Dias; Jesús de la Casa Hernández

Geothermal heritage data into digital repositories: the IGME-CSIC case study
J.A. Rodríguez-Rama, J. Mediato, R. Marrero, J.R. Hernández-Manchado, A. Bretones, E. Carrión-Ocaña, A. Perez, I. Moreno, J. García-Crespo and T. Sánchez- García

Geothermal Energy and Data Centres in Spain: An Exploratory Framework for Sustainable and Efficient Digital Infrastructures
J.A. Rodríguez-Rama, J. Mediato, E. Carrión-Ocaña, A. Pérez, I. Moreno, J. García-Crespo, P. Fernández-Canteli and T. Sánchez-García

Design, construction, and experimental characterization of a low-cost test bench for Peltier modules for thermoelectric power generation
José Daniel Gamez Avilez, Saúl Islas Pereda, Jesús Armando Aguilar Jiménez, and Nicolas Velázquez Limón

Harnessing Low-Enthalpy Geothermal Energy to Enhance Road Infrastructure Resilience: Insights from GEO-ROAD Project
Cristina Sáez Blázquez, Sergio Alejandro Camargo Vargas, Daniel Herranz Herranz and Miguel Ángel Maté-González

Can offshore wind make green ammonia competitive? Insights from a Spanish case study
Armando Díaz-Motta, and Francisco Díaz-González

The impact of data quality on the viability of wind energy projects
Franciele Weschenfelder, Priscila Silveira Ebert, Vittorio Valetta Rodríguez, and Osvaldo González

Methodology for energy modeling of cities or small countries with a view to decarbonization and climate resilience
Allison Vásquez, Allisson Zelaya, Luis Aaron Martinez, Miguel Chen Austin, and Jesús C. Hernández

Agrivoltaic Systems: An Integrated Solution for Energy, Food, and Environmental Sustainability
Bárbara Teruel, Byron Alejandro Acuña Acurio, João Soares, and Luiz Carlos Pereira da Silva

Innovative computing solutions for smart cities

Method for calibration of MQ family sensors for use in the Arduino IDE.

Dariel G. Hernández-Montalvo[1[0000-0002-3136-0442]], Abelardo Rodríguez-León[1[0000-0002-2179-4837]], Guillermo E. Ovando-Chacón[1[0000-0003-2441-3021]], Olaya P. Castellanos-Onorio[1[0000-0003-3510-2640]] y Miguel Ángel Pérez Cabada[1[0009-0007-5044-0601]]

[1] TECNM Campus Veracruz, Calzada Miguel Ángel de Quevedo 2779, Veracruz, México
darielhernandezmontalvo99@gmail.com

Abstract. This paper proposes a methodology for calibrating MQ family sensors to obtain reliable readings when measuring various pollutant gases that affect human health, and experimentally validates the results by comparing them with those of a commercial sensor.

Keywords: Methodology, calibration, MQ sensors.

1 Introduction

Air pollution, both outdoors and indoors, refers to the presence of chemical, physical, or biological agents that alter the natural characteristics of the atmosphere. Some familiar sources of pollution include domestic combustion appliances, motor vehicles, and industrial facilities. According to the World Health Organization (WHO), air pollution is currently one of the primary environmental issues that directly impacts people's health. It is estimated that each year, exposure to air pollution causes approximately 6.7 million premature deaths [1] in addition to reducing the quality of life. According to WHO data, almost the entire world population (99%) breathes air that exceeds the limits recommended by the WHO and contains high levels of pollutants [2]. For this reason, monitoring pollutant gases is of great importance, and the choice of sensors used for this purpose is even more crucial. Technology is becoming increasingly sophisticated, with compact and inexpensive devices replacing existing equipment in different areas. Since 2013, low-cost sensors have been a viable alternative to current air quality monitoring technologies [3]. The definition of low-cost sensors varies slightly among authors. However, they are characterized by the fact that their price must be below the average market value.

1.1 Problem

The advantage of having low-cost pollutant gas sensors (such as those in the MQ family), which generally offer high measurement reliability, has led us to develop mathematical equations that correct for significant variations in measurement caused by environmental conditions (which can be as high as 60% depending on the manufacturer) with respect to the actual concentration levels of the gas being sensed.

1.2 State of the Art

There are currently a large number of sensors available for gas monitoring, ranging from professional measuring equipment, such as the System Environmental Monitoring PM2.5 PM10 from Yaoan (Shenzhen) Electronic (see Figure 1), to sensors that can be implemented in electronics projects, such as the TB200B-ES1/ES4-NO2-1000 Nitrogen Dioxide Gas sensor from Ningxia Maiya (see Figure 2). It should be noted that, although both have high accuracy, only the former is certified equipment. Both are expensive, limiting their accessibility to low-budget projects. In this sense, the use of MQ sensors is of great interest.



Fig. 1. Professional measurement team.

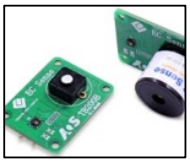


Fig. 2. NO_2 sensor

The MQ family of sensors is a device designed to detect the presence of different chemical components in the air. Each model is designed to detect one or more substances, such as flammable gases, air quality, or alcohol in breathable air. MQ sensors are widely used in research and electronic projects due to their ease of implementation and affordability.

Some studies, such as [4], utilize the MQ131 and MQ135 to monitor O_3, NO_2, NO, and CO_2 present in the air. The study reported by [5] utilizes the MQ2 to measure CO, propane, and hydrogen, the MQ135 to measure NH_3, NOx, alcohol, benzene, and CO_2, the MQ131 for O_3, and the MQ136 for measuring SO_2, to create a low-cost device for homes. Other research utilizes various chemical compounds, monitored by different sensors from the MQ family, including MQ135 (for CO_2), MQ9 (for CO_2), and MQ2 (for hydrocarbons). Specifically, in [6], motorcycle emissions are monitored, and in [7], they are utilized in an air filtration system. In [8], the MQ137 sensor is used to monitor NH_3in fish tanks. Finally, some other studies, such as [9], utilize the MQ7 and MQ2 to detect butane (C_4H_{10}), propane (C_3H_8), methane (CH_4), and CO_2, while [10] uses the MQ135 to detect CO_2 in a portable system.

The versatility of use and the large number of reported studies on this family of sensors provide a certain degree of reliability in the measurements made with them. In addition, studies on the reliability of these sensors have shown that MQ sensors have a high degree of accuracy in their response to the gases they can monitor. In [11], three sensors (MQ137, FECS44-100, and Mix2801) for measuring ammonia are evaluated in comparison with standard equipment (IPMU) under atmospheric conditions in an open compost barn during winter in Brazil. The results show that the MQ137 sensor obtained the best response compared to the other two. Similarly, [12] evaluates four electrochemical sensors for ozone measurement (SensoriC O3 3E 1, SM50 (OZU), SP3-61-00, MQ131) against ozone concentration data from the MLU 400 analyzer (Messtechnik für Luft und Umwelt, Essen, Germany) as reference values, obtaining good accuracy in the readings of the MQ131 sensor. Finally [13] compared the MQ6 and MQ7

sensors for measuring LPG and CO, respectively, against the MX6-IBRID equipment and an MGT-5 gas meter, obtaining absolute errors of less than 10% in both cases.

The gas sensors of the MQ family are classified as chemical sensors, which, according to the International Union of Pure and Applied Chemistry (IUPAC), are "devices that transform chemical information about a sample into an analytically useful signal." These "have a sensitive layer of alloy elements consisting of a metal oxide (for example, for the MQ135 it is tin dioxide S_nO_2 that reacts in the presence of a gas. The metal oxide causes the gas to dissociate into ions, resulting in electron transfer and a change in the conductivity of the layer" [14]. This sensitive layer exhibits lower conductivity in clean air, and when a contaminating gas is present, the sensor's conductivity increases in direct proportion to the gas concentration. As this process depends on temperature, all MQ models have a heating filament to maintain the working temperature and ensure that their materials acquire the required sensitivity. "The heater in most MQ sensors operates at 5V and has a resistance of 3Ω±5%. Until the heater reaches operating temperature, the sensor reading will not be reliable" [15]. Figure 3 illustrates a diagram, and its components are described in Table 1. MQ sensors are devices with high inertia, meaning that the response takes a long time to stabilize after a change in gas concentration. MQ sensors assembled into modules have both analog and digital outputs (see Figure 4). Their analog output is via a voltage divider, as shown in Figure 5. The generic operation of MQs can be more easily understood as follows: "The sensitive resistance R_S of the sensor (measured between terminals A and B) and R_L (load resistance) form a voltage divider. The H terminals are used to supply current to the sensor heater." [14].

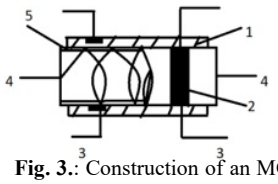


Fig. 3.: Construction of an MQ sensor. Credits HANWEI ELECTRONICS



Fig. 4: MQ sensor module.

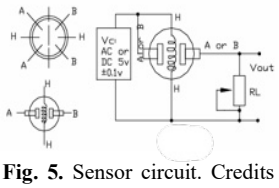


Fig. 5. Sensor circuit. Credits HANWEI ELECTRONICS

Table 1. Parts of the sensor

	Part
1.-	Gas detection layer (Material depends on MQ sensor)
2.-	Electrode
3.-	Electrode line
4.-	Ni-Cr alloy heater filament
5.-	Ceramic tubular Al_2O_3

1.3 Objetive

This study hypothesizes that a mathematical-computational correction model, derived from the manufacturer's sensitivity curves and based on environmental conditions, can improve the accuracy and reliability of MQ sensor measurements. The aim is to obtain more accurate data in the implementation of these sensors. Therefore, our study will focus on this particular family of sensors, as they have the same construction and operating characteristics.

2 Methodology

To start up the MQ sensors, refer to the manufacturer's data sheets to determine their operating conditions. Figure 6 shows the sensitivity curve on a logarithmic scale for the MQ135 sensor. Where R_0 is the resistance of the sensor in clean air and R_s is the sensitive resistance or resistance of the sensor in the presence of a contaminating gas. It can be observed that the sensor responds differently to different gases.

2.1 Sensor operation

To obtain a reliable reading, we first need to understand how the sensor behaves as a voltage divider. The value of R_s is calculated from the voltage divider formed by the load resistance (R_L) and the electronic interface of the sensor, shown in Figure 7 [16]. By Ohm's law

$$V = I \times R$$
$$V_{in} = I \times (R_s + R_L)$$
$$I = V_{in}/(R_s + R_L)$$
$$V_{out} = I \times R_L$$

$$V_{out} = \frac{V_{in}}{(R_s + R_L)} \times R_L$$

$$R_s = R_L \left[\frac{V_{in}}{V_{out}} - 1\right] \quad (1)$$

Where V_{in} is the sensor supply voltage (usually 5V), R_L is the load resistance, which may vary depending on the module manufacturer (typically ranging from 100 to 100K Ohms), R_s is the sensor resistance in the presence of a gas, and V_{out} the voltage read at the analog output.

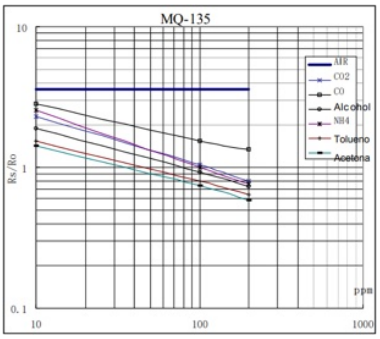


Fig. 6. Response curve for the MQ135. Credits HANWEI ELECTRONICS

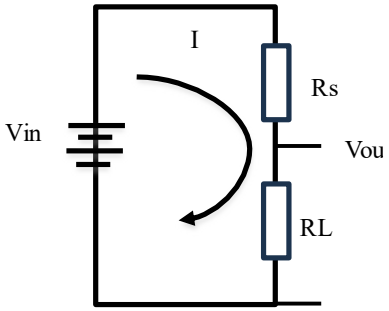


Fig. 7. Voltage divider diagram of the sensor.

2.2 Obtaining the sensor response equation.

Figure 8 shows the general methodology for implementing MQ sensors. Next, we will explain the process, highlighted in blue, for obtaining the sensor sensitivity equation.

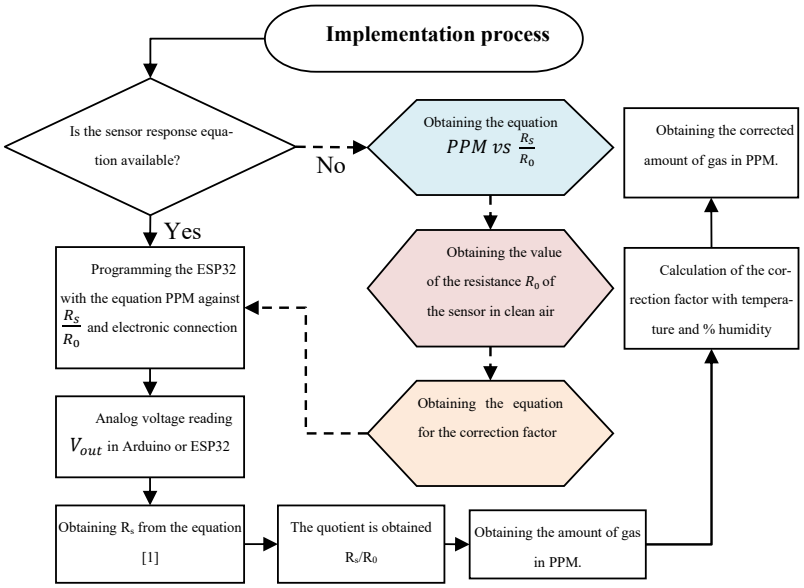
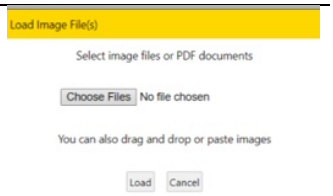
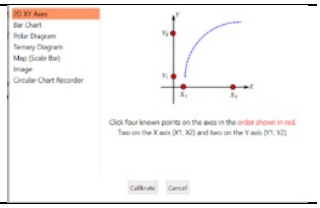
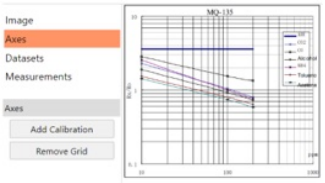
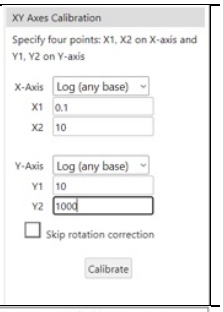
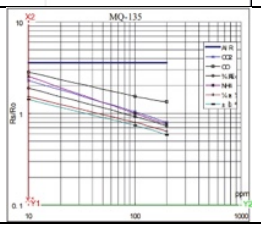


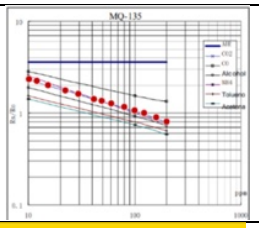
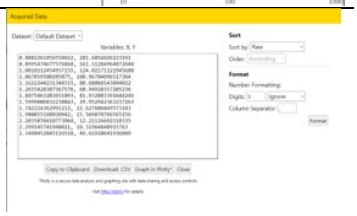
Fig. 8. Implementation methodology of MQ sensors.

To obtain the sensor response equation, we first assume that R_0 is available; how this is obtained will be explained later. First, numerical values are obtained from the sensitivity curve in the data sheet. In [17], WebPlotDigitizer software is used, which is an

online program that helps analyze graph images by plotting points on the curve. The analysis process is not fully detailed in [17], so the following steps are deduced to obtain the equation. The description of the complete process is in Table 2.

Table 2. Obtaining sensor response curve values.

Description	Example
1.- We load the image from our directory.	
2.- The following window appears where we select "2D XY Axes" and click on "Calibrate".	
3.- Select “Axes”	
4.- The following box will appear on the right. Select the X and Y axes for the logarithmic scale and modify the upper and lower limits. (As we want to obtain PPM, our variable "Y" will be the horizontal axis. Before clicking on "Calibrate", we move to the next step.	
5.- On the image, we place markers to draw the axes in the following order: X1-X2-Y1-Y1-Y2, where the "X" axis in this case is vertical, and the "Y" axis is horizontal, as shown in the image. Click on "Calibrate" in the previous step.	

6.- Next, we begin plotting points on the curve of interest. For example, we use the CO2 curve.	
7.- Once we have entered the desired number of points, we click on the left side where it says "View Data", and the following window will appear. In it, we click on "Download.CVS" to be able to make use of the data.	

Once the above process has been completed, the values obtained are loaded into an XLSX spreadsheet file, where a "scatter plot" graph of the "potential" subtype is created. The equation is presented in the graph. As shown in Figure 9, the values on the Y-axis are the PPM concentration of CO_2, and on the X-axis, R_s/R_0.

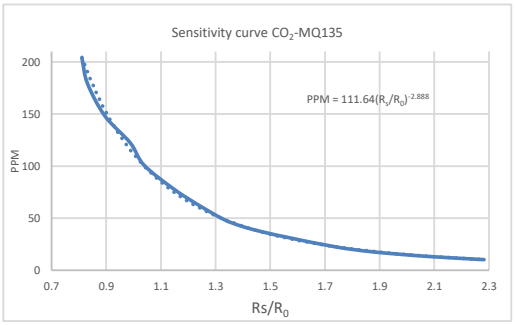


Fig. 9. Sensitivity curve of the MQ 135 sensor for CO2

Now we have an equation that, as shown in Figure 9, is quite close to the plotted points. The equation for gas concentration in PPM for CO_2 generated by the spreadsheet is

$$PPM_{CO2} = 111.64\left(\frac{R_S}{R_0}\right)^{-2.88} \quad (2)$$

We can apply the equation as long as we have the values of the sensitive resistance, R_s, and the sensor resistance in clean air, R_0.

2.3 Calibration of R_0

The process of obtaining the sensor's resistance in clean air or air free of abnormal levels of the gas to be measured is also known as sensor calibration. The data sheet indicates that, before continuous use, the sensor must be preheated for 24 hours. After this period, the sensitive resistance value R_s is read, and using data provided by the National Monitoring System on the current ambient gas concentration (or a commercial module), a value for the gas concentration in ppm is proposed. With this, R_0 is solved in equation 2. We repeat this process as many times as necessary and calculate the average. The value obtained is the one that will be used for subsequent measurements. The process described above must be performed on each MQ sensor and for each of the gases to be monitored with it.

2.4 Correction model due to environmental values

Environmental variables influence the sensor reading at the location where it is located (relative humidity and temperature), so it is not always the same. Manufacturers include the sensor response curve in the data sheet. In this case, Figure 10 shows the relative humidity curve ranging from 33% to 85% for MQ135. The y-axis represents the correction factor for R_s/R_0, and the x-axis represents the ambient temperature in degrees Celsius.

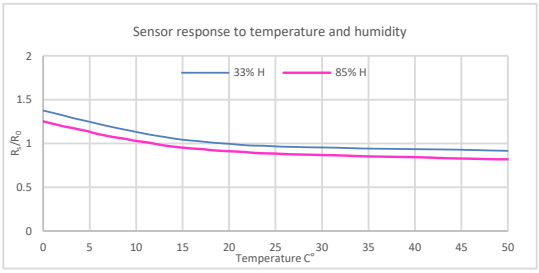


Fig. 10. Response to relative humidity and temperature.

To obtain an equation that takes into account the three variables involved (temperature, relative humidity, and correction factor for R_s/R_0) and provide the most accurate correction factor possible, it is necessary to consider an estimate of relative humidity. This paper will present two methods for achieving this, emphasizing that the result is an estimate rather than an exact answer, but one that brings us much closer to a more realistic measurement. First, the equation is calculated using the same steps as those previously employed for the response curve, as outlined in Table 2. The relative humidity and temperature values are obtained from a DHT11 sensor, which will not be discussed in further detail in this paper (see Figure 11).



Fig. 11. DHT11 sensor

The first option is as follows: You can choose to use one of the two curves (see Figure 10) according to a specific humidity range, for example.

1. Blue curve for humidity percentage range <85%
2. Pink curve for humidity percentage range >85%

However, this option is not the most accurate, so a second option is proposed, which takes into account the change in humidity percentage. To this end, an average curve is proposed (red line in Figure 12), considering that both curves have a similar trend line and are shifted on the y-axis in proportion to the change in relative humidity, as shown in Figure 12. The curve obtained is shown in Figure 13, along with the polynomial mathematical equation that represents it (equation 3).

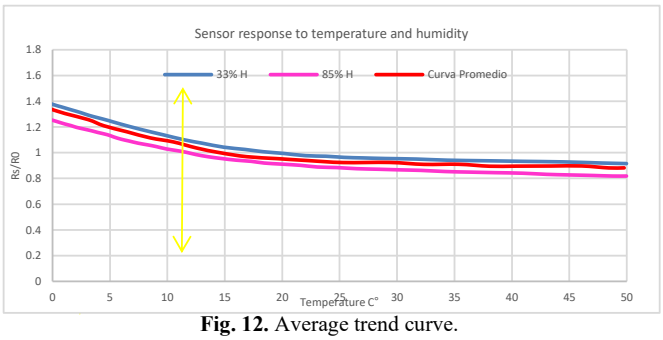


Fig. 12. Average trend curve.

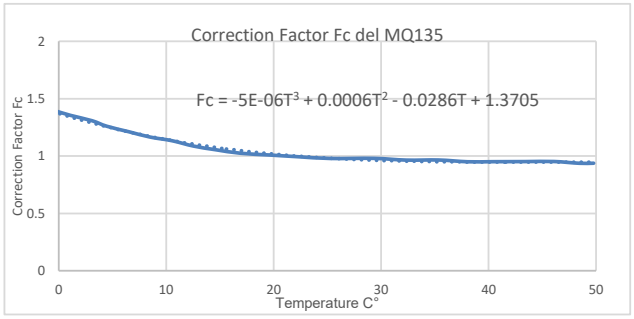


Fig. 13. Trend curve for correction factor.

$$F_c = -5 \times 10^{-6} T^3 + 0.0006 T^2 - 0.0286 T + 1.3705 \quad (3)$$

Considering relative humidity, the curve shifts on the vertical axis, depending on the relative humidity, taking into account the difference between 33% and 85% humidity, and is then normalized to 33% humidity. As a result, the equation is as shown in equation 4.

$$F_c = -5 \times 10^{-6} T^3 + 0.0006 T^2 - 0.0286 T + 1.3705 - 0.002380502(H - 33) \quad (4)$$

Where F_c is the correction factor that multiplies R_s/R_0 before applying equation 2, T is the temperature in degrees Celsius, and H is the relative humidity percentage of the medium. The calculation of the last term of equation 4 is solved with the average shift of the correction curve for each increase in relative humidity.

3 Results

The equations obtained above are validated by implementing an MQ135 sensor (Figure 4) with an ESP32, in addition to using a DHT11 sensor (Figure 11) to monitor the environmental temperature and relative humidity. The connection diagram is shown in Figure 14, and the physical prototype is depicted in Figure 15.

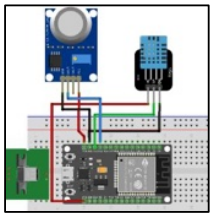


Fig.14. Connection diagram

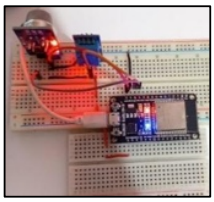


Fig. 15. Real prototype

The program for reading the MQ135 sensor was developed on the ESP32, utilizing equations 2 and 4 to determine the CO_2 concentration at a specific location. First, the sensor was calibrated and left to preheat for 24 hours. After this period, readings of the sensor's resistance in ambient air were taken. The global average value for CO_2 concentration in clean air, according to [18] and [19], is currently 426.91 PPM. Currently, the National Air Quality Information System (SINAICA) in Mexico lacks up-to-date information on the. Table 3 shows the R_0 readings at 10-second intervals between readings. The R_0 value used in the PPM equation is the average of the readings in Table 3, which is R_0 = 23.52 KΩ. Once we have determined the resistance value of the sensor in clean air, performance tests are carried out.

Table 3. R_0 average resistance

Reading time	Value of R_0 in KΩ	11:31:30 am	23.64
11:30:00 am	23.15	11:31:40 am	23.77
11:30:10 am	23.83	11:31:50 am	24.07
11:30:20 am	23.59	11:32:00 am	23.40
11:30:30 am	23.56	11:32:10 am	23.59
11:30:40 am	23.45	11:32:20 am	23.27
11:30:50 am	23.59	11:32:30 am	23.37
11:31:00 am	23.75	11:32:40 am	23.56
11:31:10 am	22.77	11:32:50 am	23.51
11:31:20 am	23.56	11:33:00 am	23.56

TOTAL:	446.99	AVERAGE:	23.52

Direct exhalation test in an open space. We performed the test in an open space at the following coordinates: 19.201316, -96.188733. The test was conducted on February 19, 2025. Figure 16 shows the readings taken over a 4-minute period, with a 2-second sampling interval, at 11:00 a.m. The temperature recorded by the DHT11 was 35.2 °C, and the relative humidity was 53.8166%. For the test, the author kept the sensor in fresh air and, at certain times, also exhaled air directly onto it; these moments corresponded to very high reading peaks. The first and third peaks correspond to a slight exhalation, and the second peak corresponds to a sustained and repeated exhalation, which causes the readings to spike. Once the exhaled air dissipates, the CO_2 levels stabilize. Figure 17 shows the period from 0 to 80 seconds, during which the sensor is not affected. The values read are very close to the average concentration of CO_2 in the environment.

Exhalation test against a commercial sensor. Another exhalation test was performed, where the MQ135 and the commercial sensor of the TTAKA7 brand were placed at the same point, then exhaled onto both sensors. The data from the TTAKA7 sensor was extracted visually. Figure 18 illustrates the data collected by both sensors, and Figure 19 displays the relationship between the measurements.

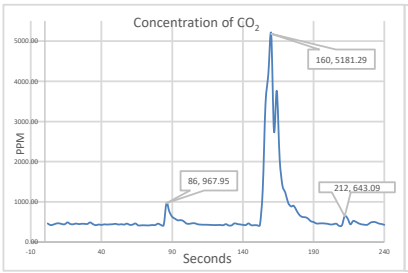


Fig. 16. CO_2 reading. 19/02/2025

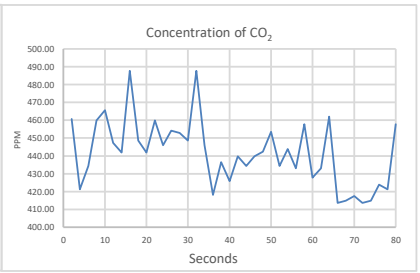


Fig. 17. Stabilized sensor values.

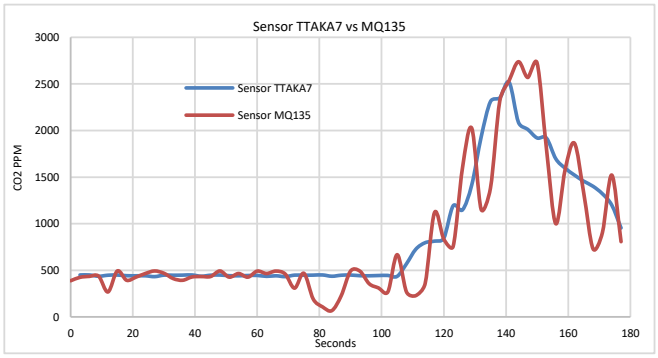


Fig.18. Comparison with Commercial Sensor.

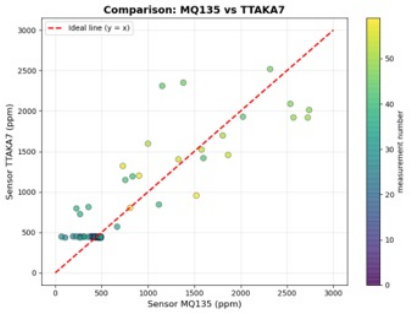


Fig. 19 Relationship between sensors

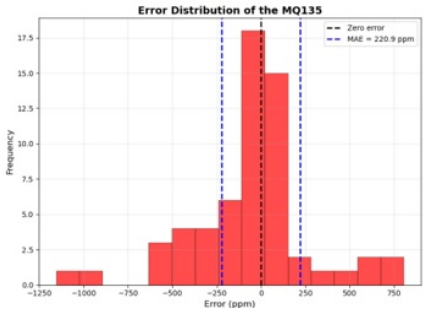


Fig. 20 Error distribution

Finally, Figure 20 illustrates the distribution of errors, and Table 4 presents the calculated error values, along with the level of reliability validated by comparing the MQ135 with the TTAKA7 sensor.

Table 4 Calculated error values

Metric	Result
Mean absolute error (MAE)	220.86
Root mean square error (RMSE)	343.27
Mean absolute percentage error (MAPE)	22.95
Reliability percentage	77.05

Dynamic sensor test in a busy area. The second test was carried out at the TECNM Veracruz Campus facilities on February 25 at 11:00 a.m. Figure 21 shows the route taken from point A to point B. The results shown in Figures 22 and 23 are analyzed below. Figure 22 shows the complete reading taken over 14 minutes and 30 seconds. We turned on the ESP32 and placed it at point A of building E, in the shade. The DHT11 recorded a temperature of 28 °C and a humidity of 53%. Before starting the readings, the MQ sensor was allowed to warm up until the measurements stabilized. From second 0 to approximately 350, we took readings at point A, where there were about 25 people. However, as it was an open space, CO_2 levels remained normal. Figure 23 shows this data more clearly. From approximately seconds 350 to 560, we moved the sensor to point B, where it was now directly exposed to the sun's rays. In this new position, the humidity recorded by the DHT11 decreased to 52%, and the temperature rose to 29.3 degrees, resulting in a slight increase in the readings, which registered around 500 PPM.

As we approached point B, the sensor was brought into a closed building. Once inside, the readings skyrocketed as the sensor detected a sudden change in environmental conditions, primarily in humidity, reaching a value of 31.3%. The temperature remained virtually unchanged. There were about six people in the building's lobby. It is worth noting that this area of the building lacks adequate ventilation systems. After approximately 3 minutes, the readings stabilized again, indicating that there were high levels of CO_2due to inadequate ventilation in the building. An average of up to 1200 PPM was recorded, as shown in Figure 22.



Fig. 21. Sketch of the TECNM Veracruz Campus.

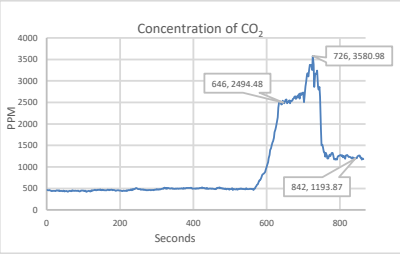


Fig. 22. CO_2 reading 25/02/2025

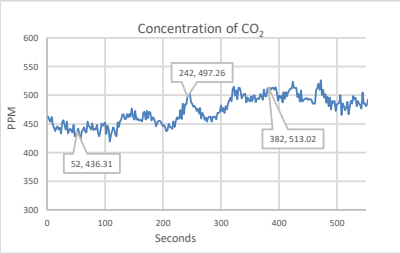


Fig. 23. Concentration of CO_2; seconds 0-550

Sensor test in a closed classroom. The last test was conducted in a classroom measuring 4.0 by 5.0 meters during a one-hour class at 3 p.m. on March 10, 2025. There were a total of 30 students and a teacher in the classroom. The classroom is air-conditioned. The results show that, at the beginning of the sampling period, CO_2 levels did not deviate from the standard average value of 430 ppm (Figure 24). However, as the session progressed, CO_2 levels reached very high values, demonstrating the need for an efficient ventilation system in this space.

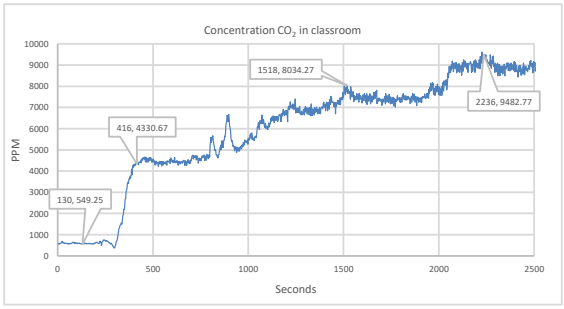


Fig. 24. CO_2 reading10/03/2025

4 Conclusions

The study presents a mathematical-computational method for calibrating MQ family sensors, utilizing the MQ135 model, which can be applied to any model in this sensor

family, resulting in improved accuracy and reliability of the measurements of pollutant gases detected by this sensor family. The results obtained demonstrated that the method described, which is based on the manufacturer's response curves adjusted for environmental factors such as temperature and humidity, allows for measurements closer to the actual values. In addition, the following main conclusions stand out:

1. Effectiveness of the calibration method: The implementation of equations derived from the sensor's sensitivity curves, combined with correction for temperature and humidity, proved highly effective. Tests conducted in various environments (open spaces, busy areas, and closed classrooms) confirmed that the readings from the adjusted MQ135 sensor are significantly closer to the expected values, such as the global average CO_2 level (422.7 ppm).
2. Influence of environmental conditions: The results obtained showed that ambient temperature and humidity affect sensor readings. Including a polynomial correction factor (equation 4) in the calculations minimized these effects and improved the accuracy of sensor measurements under variable conditions.
3. Versatility and applicability: The proposed method is not only applicable to the MQ135 sensor but can also be applied to other sensors in the MQ family due to their similar construction and operation, which makes it an accessible method for projects that use this type of sensor.
4. Practical validation: Tests comparing the results with the commercial sensor (TTAKA7) yielded similar results, supporting the reliability of the method. Additionally, tests conducted in real-world environments demonstrate the sensor's effectiveness in identifying ventilation and CO_2 concentration issues.
5. Limitations and recommendations: The method significantly improves accuracy, but it was observed that initial calibration (calculation of R_0) is essential to achieve this. However, this calibration must be performed under the most ideal conditions possible (e.g., outdoors), with CO_2 levels close to those reported by [18] for 2024 of 422.7 PPM.

In conclusion, this research presents a calibration method for MQ family sensors, enhancing the accuracy of measurements obtained from them. This improvement is easy to replicate and, together with its high availability and low cost on the commercial market, encourages its use in academic, commercial, and even industrial projects. The results obtained support the potential of these sensors as reliable tools for assessing air quality in different environments, provided they are correctly calibrated.

References

1. Pure Energy UK: The Air We Breathe is Killing Us... (2025). https://www.pureenergyuk.com/2023/08/30/the-air-we-breathe-is-killing-us/
2. Abbas, F.N., Mis.Intisar, M.S., Khyioon-Abdalrdha, Z. and Nassir-Abud, E.: Capable of gas sensor mq-135 to monitor the air quality with arduino uno. International Journal of Engineering Research and Technology, 13(10), 2955-2959 (2020)
3. Anitha, M. and Kumar, L.S.: Development of an IoT-enabled air pollution monitoring and air purifier system. MAPAN-Journal of Metrology Society of India. 38(3), 669–688 (2023).

4. Badura, M., Batog, P., Drzeniecka-Osiadacz, A. and Modzel, P.: Low and Medium-Cost Sensors for Tropospheric Ozone Monitoring-Results of an Evaluation Study in Wrocław, Poland. Atmosphere, 13(4), 542 (2022)
5. Balasubramanian, S., Sneha, T., Vinushiya, B. and Saraswathi, S.: Air Pollution Monitoring and Prediction using IOT and Machine Learning. International Journal of Computer Science and Information Technologies, 12 (2), 60-65 (2021)
6. Bhosale, M.C., Kulkarni, N.M., Shaligram, A.D. and Pustake, S.: Measurement of co2 gas concentration using mq135 sensor for air pollution monitoring. (Paper presented at the International Conference on "Recent trends in Science and Technology"). (2023, January)
7. De-Oliveira, V.C., de-F.F.-Tinôco, I., da-Silva, L.F., Oliveira, C.E.A., Damasceno, F.A. and Saraz, J.A.O.: Evaluation of ammonia sensor modules in a compost barn system during winter in brazil. Engenharia Agrícola, 43(4). (2023)
8. De Pablo-Gozalo, V. de: Design and implementation of a portable CO2 emissions measurement system. Final Degree Project / Final Degree Project, E.T.S.I. Telecomunicación (UPM). (2020)
9. Erazo-Moreta, O., Santana-Sornoza, A. R., Vera-Alarcón, M. J., Oviedo-Bayas, B. W.: Indoor pollutant gas monitoring using WSNs for preventive action. University and Society, 11(4), 116–122 (2020)
10. Fernando-Ilustre, L., Claire-Mortos, Y., Enriquez, J., Mindoro, J., and Arjay-Malbog, M.: See carbon: Real-time iot monitoring system for motorcycle vehicle carbon emissions. E3S Web of Conferences, 488(03014) (2024).
11. García-Bravo, E. I., Guatapi-Yachimba, A. M.: Design and implementation of a prototype for a system for measuring, analyzing, and purifying air polluting gases using Arduino and Ubidots IoT. Dissertation. Universidad Politecnica Salesiana, (2023). http://dspace.ups.edu.ec/handle/123456789/24088.
12. García-Navarrete, G., and Rico-Soto, K. G.: Low-cost sensors for air quality monitoring. EPISTEMUS, 13(27), 30–37 (2020) https://doi.org/10.36790/epistemus.v13i27.108
13. Lindsey, R. (2025, May 21). *Climate change: atmospheric carbon dioxide*. NOAA Climate.gov. https://www.climate.gov/news-features/understanding-climate/climate-change-atmospheric-carbon-dioxide
14. Mazin-Talib, A. and Nsaif-Jasim, M.: Geolocation based air pollution mobile monitoring system. Indonesian Journal of Electrical Engineering and Computer Science, 23(1), 162-170 (2021)
15. Nagy, A. S., Polanco-Risquet, A., Martínez-de-la-Cotera, O. L., & Carralero-Ibargollen, O.: Simultaneous gas measurement with MQ sensors. Electronic, Automation and Communications Engineering, 41(1), 34-43 (2020). Retrieved from http://scielo.sld.cu/scielo.php?script=sci_arttext&pid=S1815-59282020000100034&lng=es&tlng=es.
16. Pulungan, I.H., Nasution, T.I., Ningsih, H.W., Tarigan, K. and Fathurrahman.: An automatic and realtime control of ammonia concentration in catfish pond water based on mq137 sensor. Journal of Physics: Conference Series, 1428(1):012054, 1-5 (2020)
17. Segura-Cruz, A. P.: Implementation of wireless sensor networks in a telemetry system to detect and control LPG and carbon monoxide leaks in domestic applications. Escuela Superior Politécnica de Chimborazo, Riobamba, (2017). http://dspace.espoch.edu.ec/handle/123456789/7548.
18. Tiseo I.: "Average carbon dioxide (CO_2) levels in the atmosphere worldwide from 1959 to 2024", Jan 15, 2025 https://www.statista.com/statistics/1091926/atmospheric-concentration-of-co2-historic/
19. WHO: Ambient (outdoor) air quality and health, (2024). https://www.who.int/es/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health.

Sustainable IoT-Based Architecture for Real-Time water Quality Monitoring in Smart Cities

Nicole Michelle Leyva Munguia [1][0000-1111-2222-3333] and Rocio Maciel [2][0000-0002-5548-2073]

[1] Universidad de Guadalajara, Zapopan, Jalisco 45100, Mexico
[2] Universidad de Guadalajara, Zapopan Jalisco 45100, Mexico
nicole.leyva@alumnos.udg.mx

Abstract. Ensuring safe access to drinking water demands continuous and near-real-time assessment methods capable of operating beyond the economic and temporal constraints of conventional laboratory analysis. This study introduces and validates a compact Internet of Things (IoT) framework conceived for water quality monitoring in dynamic environments. The prototype combines low-cost sensors for pH, turbidity, and total dissolved solids (TDS) and evaluates their performance through a statistical examination of data collected under simulated continuous-flow conditions. A purpose-built hydraulic mock-up was designed to reproduce pipeline dynamics, enabling the observation of sensor response under realistic flow dynamics. The results confirmed the stability of pH readings, the expected dispersion of TDS, and the high responsiveness of turbidity measurements to flow-induced variations. Altogether, the findings demonstrate the framework's reliability in dynamic environments representative of urban water distribution systems, establishing a validated foundation for large-scale deployment in decentralized, data-driven smart-city infrastructures.

Keywords: Internet of Things, Water Quality, Statistical Analysis

1 Introduction

Rapid urbanization and demographic expansion are placing unprecedented pressure on existing water infrastructure [1]. In growing urban centers, ensuring a reliable supply of safe drinking water has become a critical challenge, recognized globally as a fundamental human right by the United Nations General Assembly. Efficient management of this resource is essential not only for disease prevention but also for fostering resilient economic growth and reducing urban poverty [2]. However, global water demand continues to outpace population growth, driven by demographic expansion and the increasing needs of the agricultural, industrial, and energy sectors. According to the Sustainable Development Goals Report 2023, water stress already affects nearly half of the world's population for at least one month each year [3]. Furthermore, the United Nations Secretary-General's SDG 2025 Report warns that despite 73 million new annual domestic water service connections, population growth still outpaces these advances, undermining international goals for equitable and sustainable access [4].

The conventional paradigm for water quality monitoring relies on systematic sampling and periodic laboratory analysis of water conditions. Although highly accurate, this approach involves considerable operational costs, extended processing times, and a structural inability to deliver real-time information for timely decision-making [5]. According to López-Ramírez and Aragón-Zavala [6], traditional methods, such as laboratory-based monitoring (TMLB) and in-situ monitoring (TMIS), require specialized personnel, reducing sampling frequency and potentially leading to data loss between campaigns. These persistent limitations have accelerated the shift toward new technological solutions that enable continuous, automated, and low-cost monitoring, particularly those based on the Internet of Things (IoT) and wireless sensor networks.

The proliferation of low-cost sensing hardware and advances in wireless communication technologies have driven the rapid expansion of the Internet of Things (IoT), a domain projected to encompass nearly 27 billion connected devices by 2025 [7, 8, 9]. This technological evolution has established IoT as a key enabler for water quality monitoring (WQM), offering new possibilities for decentralized environmental assessment. IoT-based architectures enable real-time, remote, and cost-effective measurement of critical parameters such as pH, turbidity, and total dissolved solids (TDS). Beyond continuous data acquisition, these systems are designed to transform raw measurements into actionable information that supports efficient, adaptive, and sustainable water management practices [10]. Moreover, IoT platforms provide a foundation for modular and scalable implementations, strengthening early anomaly detection, facilitating predictive maintenance, and ultimately contributing to more transparent and resilient water governance frameworks [11, 12].

While many IoT prototypes for WQM exist, their validation is often limited to static, proof-of-concept laboratory tests. This approach fails to account for the dynamic hydraulic conditions such as continuous flow, pressure fluctuations, and flow induced micro vibrations present in real world distribution networks. This study addresses this critical gap by presenting a rigorous experimental validation of an IoT prototype under simulated, continuous-flow conditions. This work provides a technical and statistical analysis of sensor performance in this dynamic environment, assessing the stability of pH, the dispersion of TDS, and critically, using the turbidity sensor's response as a key indicator of the system's sensitivity to hydraulic movement. These findings provide a necessary foundation for assessing the real-world reliability of low-cost IoT solutions, acting as an intermediate benchmark toward deployment.

2 Related Work

Research on water quality monitoring systems has advanced significantly with the integration of the Internet of Things (IoT), which has emerged as a viable solution to overcome the limitations of traditional laboratory sampling methods. Recent literature reflects a growing interest in the development of prototypes and platforms that leverage low-cost technologies for water surveillance.

2.1 Empirical IoT Systems for Water Quality Monitoring

The development of Internet of Things (IoT) architectures for water quality monitoring has evolved from early proof-of-concept prototypes to more robust and application-oriented implementations. Most recent studies have sought to demonstrate feasibility through low-cost and modular designs, focusing on domestic, urban, or remote contexts with varying degrees of validation.

Early studies such as those by Lakshmikantha et al. [13] and Hakimi and Jamil [14] focused on developing low-cost monitoring solutions for basic water quality parameters, including pH, turbidity, temperature, and electrical conductivity. Lakshmikantha et al. [13] proposed an IoT-based prototype integrating multiple sensors and conceptually outlined cloud connectivity, although validation was limited to local, static tests. In contrast, Hakimi and Jamil [14] conducted short-term laboratory calibration of turbidity and temperature sensors, verifying linearity and stability under controlled conditions. While both studies contributed to establishing the technical feasibility of low-cost monitoring, their validation remained restricted to static laboratory settings without evaluating performance under dynamic or real-time conditions.

Building upon these preliminary efforts, Nanyanzi et al. [15] implemented a microcontroller-based monitoring system, demonstrating deployment feasibility but still lacking extensive operational validation. Rokhman et al. [16] and Sakib et al. [17] investigated sensor reliability and network behavior in open systems, while Abrajano et al. [18] introduced a solar-powered IoT prototype tailored for isolated rural communities. These works collectively highlight the versatility of IoT platforms but also reveal challenges in achieving long-term stability, calibration consistency, and reliable communication under environmental variability.

More recently, research efforts have expanded towards the integration of IoT within smart water and urban management systems. Abdelmoez et al. [19] developed a standalone sensing and actuation IoT solution for water management and leakage localization, combining flow and quality sensors with automated control of filtration units. This study emphasizes the relevance of IoT as an enabler for intelligent water distribution and conservation; however, they primarily address infrastructure-level automation and predictive analytics rather than empirical validation of sensor performance under dynamic hydraulic conditions.

Despite this progress, most research remains centered on feasibility demonstrations rather than full experimental validation. Although Abdelmoez et al. [19] developed an IoT-based prototype in simulated dynamic conditions, their work focused on leak detection and system actuation rather than a statistical validation of sensor performance. Furthermore, a persistent methodological gap exists regarding the analysis of sensor performance under controlled, continuous-flow conditions mimicking pipelines. The work by López et al. [20], for instance, identified sensor fluctuations in a lagoon, but correctly attributed them to "wave motion" in a static deployment. To provide a clearer comparison of these empirical studies, Table 1 summarizes their main approaches, contributions, and limitations, highlighting this validation gap.

Table 1. Comparative overview of experimental IoT implementations and validation strategies in water quality monitoring.

Authors	Validation Context and Method	Contribution	Limitations
Lakshmikantha et al. [13];	Designed and validated an IoT-based prototype integrating pH, turbidity, temperature, and EC sensors.	Demonstrated the technical feasibility of a low-cost multi-sensor IoT prototype for water quality monitoring.	Cloud connectivity remained theoretical; no experimental validation of data transmission or performance under dynamic flow.
Hakimi & Jamil [14]	Calibrated turbidity and temperature sensors under static laboratory conditions. Compared sensor outputs with reference instruments during short-term tests.	Verified the linearity of turbidity response and the stability of temperature readings, confirming basic sensor reliability.	Evaluations were short and static; no long-term testing, dynamic flow analysis, or IoT integration was performed.
Nanyanzi et al. [15]	Prototypes for domestic and urban environments in Uganda, using microcontrollers cloud-based data visualization.	Proposed a viable solution for urban and domestic contexts.	Proof-of-concept stage only
Rokhman et al. [16] Sakib et al. [17]	Implemented monitoring systems for surface water bodies and validated sensor reliability, respectively.	Provided context-specific IoT applications in Indonesia and India.	Limited to surface water
Abrajano et al. [18]	Solar-powered IoT solution for rural communities.	Demonstrated the feasibility of energy-autonomous monitoring systems in remote areas.	Context-specific for rural communities; no exhaustive validation mentioned.
Abdelmoez et al. [19]	developed and tested an IoT-based prototype for leakage detection and flow control in a simulated water distribution network.	Demonstrated real-time leak detection, localization, and actuation capabilities within a smart water distribution framework.	Focused on automation and network management rather than quantitative validation of water-quality parameters.
López et al. [20]	Real-time lagoon monitoring prototype in Mexico, recording pH, turbidity, temperature, and total dissolved solids.	Achieved results comparable to dedicated instruments.	Fluctuations due to wave motion. Limited to a single water body.

2.2 Reviews on IoT and ML

Complementary to empirical works, several reviews synthesize the role of IoT and ML in water quality monitoring. Essamlali et al. [10] emphasize ML applications for prediction, anomaly detection, and classification in WQM platforms. López-Ramírez and Aragón-Zavala [6] review wireless sensor networks (WSN) as enablers of IoT deployments, highlighting both the inefficiency of traditional laboratory methods and the opportunities of IoT solutions. Zulkifli et al. [21] report challenges related to energy consumption and network reliability, while Obunga et al. [22] extend this by identifying cybersecurity and sustainability as overlooked aspects in current deployments.

Despite these advances, three recurring gaps are evident: reliance on closed datasets, limiting reproducibility, lack of field validation, as most systems are tested only under controlled laboratory conditions and insufficient strategies for long-term sustainability and energy efficiency.

2.3 Contribution of This Work

The present study directly addresses the methodological and experimental gaps identified in the existing literature. In contrast to previous IoT-based water monitoring initiatives that relied on static tests, this work provides one of the first empirical validations of an IoT architecture evaluated under simulated continuous-flow conditions. This paper presents the foundational statistical validation of test data derived from a broader research project, which aims to develop and validate a complete IoT solution for water quality monitoring. The specific contribution of this manuscript is the rigorous statistical characterization of sensor reliability under dynamic flow, which demonstrates the system's operational consistency and bridges the gap between conceptual prototypes and field-ready implementations. This analysis represents the foundational statistical validation of the hardware, an essential initial step to assess the prototype's viability for future scaling into a robust, multi-unit monitoring solution.

By emphasizing experimental validation and analyzing the interaction between water flow and turbidity response, this research establishes a framework for assessing system sensitivity to hydraulic disturbances. This aspect has been largely overlooked in previous studies. The results provide a foundation for the future integration of computational agent algorithms aimed at automatic reading compensation and adaptive calibration. These advancements will ultimately enhance the precision and interpretability of urban water quality monitoring systems.

Furthermore, this study contributes a comprehensive perspective that merges technical robustness, scalability considerations, and data security aspects. This validated architecture provides the groundwork for developing these systems beyond simple prototypes. The work supports the advancement of resilient smart-city infrastructures by focusing on the foundational challenge of ensuring reliable data collection under real-world operational dynamics.

3 Methodology

This study adopted a quantitative, experimental research methodology designed to empirically validate a low-cost, IoT-based water quality monitoring system. The core of this methodology is the system's evaluation under simulated, continuous-flow conditions, which departs from common static laboratory tests. This approach is essential for evaluating the prototype's reliability and responsiveness under conditions representative of real-world hydraulic networks.The project's lifecycle was adapted from Design Thinking principles [23] and structured into a five-stage process, focusing on technical validation rather than user-facing features.

3.1 Design and Project Stages

The project used a five-stage process to create a scalable IoT water quality monitoring system. This methodical approach provided a clear workflow for development and testing, ensuring the system's reliability for both local use and future smart city applications. The process is shown in Figure 1.

1. Component Selection and *Calibration*: Low-cost, open-hardware sensing elements were chosen to measure pH, turbidity, and total dissolved solids (TDS). Each module was independently calibrated with certified reference solutions to align response curves before integration with the control hardware.
2. Hardware Integration and Prototyping: The calibrated sensors were connected to an ESP32 microcontroller responsible for data collection and wireless communication. A protective enclosure produced through additive manufacturing safeguarded the electronic components and stabilized sensor placement during immersion tests.
3. Backend and Data Pipeline: A single-board computer was deployed to host both the MySQL database and the MQTT broker responsible for message handling. This lightweight, local infrastructure enabled high-frequency data capture and secure storage while maintaining low power consumption.
4. Hydraulic Mock-up Construction: To reproduce plumbing dynamics conditions, a recirculating hydraulic mock-up was constructed. The closed-loop system allowed the control of flow rate and pressure, creating a consistent environment for testing the sensors' operational stability under continuous circulation.
5. Experimental System Validation: The complete prototype was installed in the hydraulic mock-up for continuous flow testing. This final stage focused on collecting a robust dataset to statistically analyze the system's performance, stability, and responsiveness under dynamic, non-static conditions.

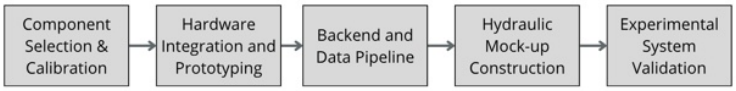


Fig. 1. Project Development Workflow.

3.2 System Architecture

The system architecture follows a multilayer design aimed at enabling reliable, real-time data acquisition for statistical validation. It integrates a custom hardware prototype with a two-tier software and communication backend, providing a low-cost and scalable framework for continuous monitoring. This configuration addresses the limitations of traditional non-real-time approaches by enabling high-frequency data transmission from sensors to a central database, as illustrated in Figure 2.

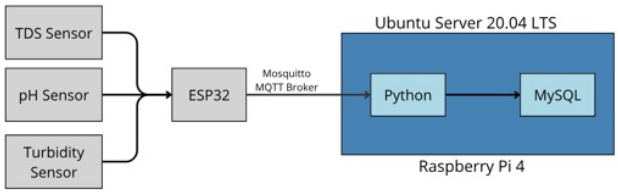


Fig. 2. Experimental architecture of the IoT-based water quality monitoring prototype.

Data Acquisition and Sensing Layer. To conduct this statistical validation, the experiment required a method to capture real-time data from widely available low-cost sensors. An ESP32-based microcontroller was selected as the core of the experimental acquisition node. The critical factor in this choice was its integrated processing and Wi-Fi capabilities, which allowed it to function as a standalone logger, autonomously capturing and transmitting the pH, turbidity, and TDS measurements from the hydraulic circuit without complex external interfaces.

Communication and Transmission Layer. The primary methodological requirement for this validation was the integrity of the collected time-series data. A stable, continuous, and high-frequency data stream was essential for the subsequent statistical analysis. The MQTT protocol was therefore implemented as the transport mechanism. Its lightweight and efficient standard was the deciding factor, as it ensured low-latency, reliable data transfer and minimized network overhead, thereby preserving the fidelity of the raw sensor data for the validation study.

Backend and Data Storage Layer. A robust and simple data-logging tool was needed to receive and persist the continuous data stream generated by the experiment. A Raspberry Pi was selected to serve this limited function, chosen for its minimal energy footprint and ability to autonomously run server tasks. MySQL was employed specifically for its capability to structure and log the incoming time-series data reliably. This entire

backend setup was designed not as a scalable product, but as the necessary laboratory infrastructure to produce the validated, structured log required for the offline statistical analysis.

3.3 System Validation

Hardware Calibration and Preparation. To ensure the subsequent statistical analysis was founded on reliable data, it was essential to first establish an operational baseline for the sensor array. This involved a standardization process where each sensor was benchmarked against certified reference points: the pH sensor was normalized using a 7.00 pH buffer, the turbidity sensor was zeroed against a 0 NTU reference, and the TDS sensor's accuracy was verified with a 1000 ppm solution. This process validated that all sensors were operating within acceptable error margins, confirming their suitability for the dynamic experiment. Concurrently, a custom enclosure was fabricated to provide a stable immersion platform for the sensors, thereby minimizing physical instability as a potential experimental variable.

Hydraulic Mockup and Flow simulation. A core methodological contribution of this study was the shift from conventional static tests to validation under dynamic flow. To create this environment, a custom-built, closed-loop hydraulic mock-up was designed specifically to function as the experimental apparatus. To generate continuous recirculation that mimics hydraulic plumbing, a pump with a 2050 L/h rated capacity was selected. This component successfully achieved stable, continuous flow through the sensor chamber in under 5 seconds. The express purpose of this mock-up was not to create a final product, but to serve as the controlled, non-static environment required for this paper's primary research goal: validating sensor performance under realistic, dynamic conditions.

3.4 Data Integrity and Analysis Workflow

The validity of the subsequent statistical analysis was critically dependent on the integrity of the raw data collected from the IoT system. Therefore, before analysis, the entire data acquisition pipeline was verified to act as an experimental control against transmission or parsing errors. It was confirmed that the MQTT data transmission from the ESP32 node remained stable. Furthermore, a dedicated Python validation script was implemented to monitor incoming JSON messages, confirming correct formatting and guaranteeing accurate ingestion into the MySQL database. This software validation step was essential, as it ensured that the final dataset used for the study was complete, reliable, and free from any system-induced artifacts, thus validating it as a sound basis for the statistical analysis

The data acquisition protocol was refined to operate with a fixed sampling interval of 10 seconds, ensuring a uniform data stream suitable for rigorous time-series analysis.

Measurements were collected continuously at this interval, resulting in 628 observations over a 104-minute period. This extended high-frequency dataset, distinct from the preliminary validation runs, was subsequently exported from the MySQL database in CSV format for offline statistical processing.

The dataset was processed in Python 3.11 using the pandas, NumPy, and matplotlib libraries. The analysis workflow was structured in two stages to fully characterize the sensors' behavior under dynamic flow. First, measures of central tendency (mean) and dispersion (variance, standard deviation, coefficient of variation) were calculated. Second, key distribution parameters, specifically skewness and kurtosis, were computed to characterize the data asymmetry and the concentration of values around the mean, respectively.

Finally, the analysis generated graphical representations to visualize the data distribution. Histograms were constructed for each parameter, which were then complemented with kernel density estimation (KDE) curves and fitted normal distribution curves. This visual analysis is essential for interpreting the statistical findings and understanding the sensor's response to the dynamic flow conditions. A summary of the main statistical indicators is presented in Table 2.

Table 2. Statistical indicators for IoT sensor measurements under continuous-flow conditions

Sensor	Mean	Variance	Standard Deviation	Coefficient of Variation (%)	Skewness	Kurtosis
pH	7.302	0.385	0.621	8.50%	0.313	2.304
Turbidity *(NTU)*	2.525	0.006	0.024	0.97%	0.954	6.137
Total Dissolved Solids *(ppm)*	91.33	1984.96	44.58	48.78%	0.203	1.740

4 Results and Discussion

4.1 Statistical Analysis

The pH values exhibited a mean of 7.302, with a range of 3.12, spanning from 6.49 to 9.61. The median was 7.39, which is close to the central tendency, while the mode was 6.60. Variability within the dataset was low, as reflected by a variance of 0.385, a standard deviation of 0.6212, and a coefficient of variation of 8.50%. The distribution presented a slight positive skew (skewness = 0.313) and a slightly platykurtic profile (kurtosis = 2.304). In Fig. 3, the histogram shows this distribution, where the KDE line highlights deviations from the fitted normal curve.

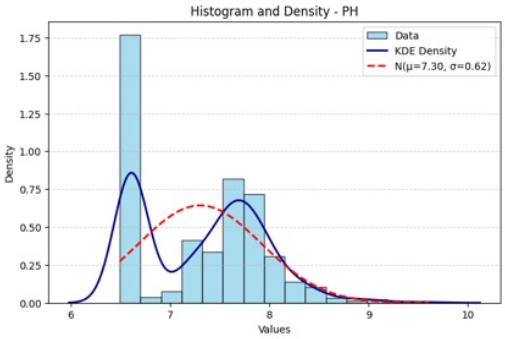


Fig. 3. Histogram and Density of pH measurements.

The turbidity parameter served as a key indicator of the system's responsiveness to flow-induced variations. Measurements averaged 2.525 NTU, with real-time fluctuations observed within a narrow range (2.48-2.66 NTU). The dataset showed minimal dispersion (CV = 0.97%) but was positively skewed (skewness = 0.954) and strongly leptokurtic (kurtosis = 6.137). This statistical profile indicates that, while the baseline remained stable, the sensor effectively captured subtle, high-frequency variations generated by hydraulic movement. As illustrated in Figure 4, the sharply peaked distribution confirms this behavior, showing a consistent baseline punctuated by rapid, minor deviations, validating the sensor's sensitivity under dynamic conditions.

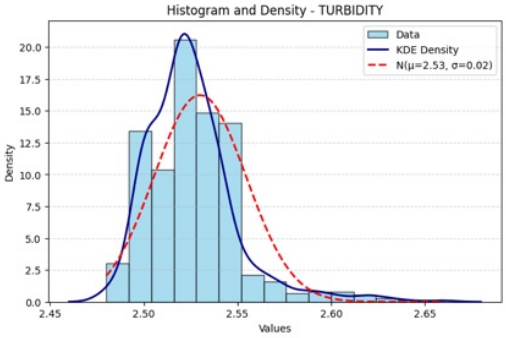


Fig. 4. Histogram and Density of Turbidity measurements.

TDS values had the highest variability among the parameters, with a mean of 91.3317 ppm and a wide range of 174.72 ppm (5.81–180.53 ppm). The median was 94.35 ppm, whereas the mode was 58.28 ppm. The variance reached 1984.96, with a standard deviation of 44.5885 and a coefficient of variation of 48.78%, confirming a high degree of dispersion. The distribution was nearly symmetric (skewness = 0.204) and platykurtic (kurtosis = 1.740). In Fig. 5, the histogram exhibits multiple peaks, indicating this heterogeneity, with the KDE diverging notably from the normal fit.

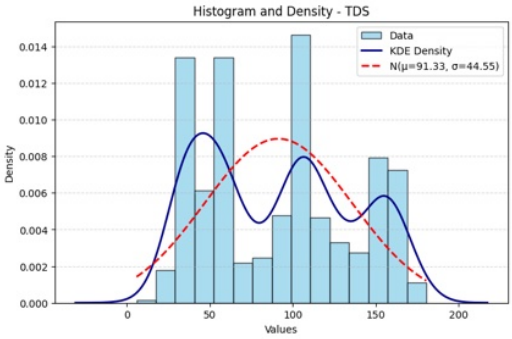


Fig. 5. Histogram and Density of TDS measurements.

4.2 Technical Discussion

The statistical analysis of the sensor data under continuous flow provides critical validation of the prototype's performance in a dynamic environment. The pH measurements demonstrated robust and reliable behavior throughout the 104-minute experiment. The low coefficient of variation (8.5%) and stable distribution confirm that the hydraulic flow within the mock-up did not introduce significant electrical noise or measurement deviations. This observation aligns with related field studies, such as López et al. [20], who also reported pH stability under environmental motion, reinforcing the general resilience of this parameter to physical disturbances.

The turbidity analysis provided the primary evidence of the system's responsiveness to hydraulic dynamics. Although the average value remained stable (2.53 NTU), the strongly leptokurtic distribution (kurtosis = 6.137) constitutes the key finding. This sharply peaked distribution (see Figure 4) indicates that the sensor was highly sensitive, capturing subtle, high-frequency fluctuations generated by flow-induced movement. This outcome clearly differentiates the present work from previous studies [13–18, 20], which validated IoT-based water quality monitoring systems under static laboratory or uncontrolled environmental conditions, and from Abdelmoez et al. [19], whose research focused on flow control and leak detection rather than on sensor performance evaluation.

In contrast, the TDS readings exhibited substantial variability, as reflected by the high coefficient of variation (48.78%). This behavior is consistent with the parameter's intrinsic nature in a non-static, closed-loop system. The platykurtic and multimodal distribution (see Figure 5) suggests that the transport and accumulation of dissolved solids were non-uniform, highlighting the heterogeneous behavior of particulate matter within the hydraulic circuit. These findings corroborate previous observations from non-controlled environments, such as those reported by Lupi et al. [5] and López et al. [20], who also noted pronounced TDS fluctuations resulting from environmental dynamics.

5 Conclusions and Future Work

This study presented the experimental validation of an IoT-based water quality monitoring system, focusing on the statistical characterization of pH, turbidity, and TDS parameters under simulated continuous-flow conditions. The results confirmed the stability of pH readings and the expected variability of TDS, demonstrating that the proposed setup provides reliable data under dynamic scenarios. By incorporating statistical indicators such as skewness and kurtosis, the study offered a rigorous evaluation of sensor performance. The key finding was the validation of the system's responsiveness, evidenced by the turbidity sensor's leptokurtic distribution, which demonstrated sensitivity to flow-induced hydraulic variations.

These findings emphasize the importance of employing flow-based hydraulic mockups as an essential intermediate step in sensor validation. This experimental approach bridges the gap between idealized, static laboratory tests and the complexities of real-world field implementations. The methodological contribution presented here is fundamental for accurately assessing the reliability and responsiveness of low-cost sensors before their deployment in buildings or urban plumbing systems.

Future work will focus on enhancing the system's technical robustness, autonomy, and scalability. The next development phase involves the integration of autonomous computational agents to manage core operational tasks. These agents will implement adaptive sampling strategies, such as reducing acquisition frequency during stable readings, to optimize energy efficiency and extend sensor lifespan. Additionally, they will incorporate secure data-handling mechanisms and localized data processing to minimize transmission overhead.

Building upon the statistical baseline established in this study, these computational agents will further integrate machine learning algorithms for noise reduction and real-time anomaly detection. A primary objective will be to train the system to differentiate between benign, flow-induced statistical noise (as observed in the turbidity readings) and genuine contamination events. Classification algorithms will also be employed to categorize water quality states (e.g., potable, marginal, contaminated) based on multi-parameter sensor fusion, supporting automated decision-making.

Future developments envision scaling these autonomous agents into a distributed, inter-municipal intelligent sensor network. By correlating spatial-temporal data from devices across multiple locations, this agent-based framework could identify and trace contamination events as they propagate through the water distribution system, enabling automatic generation of real-time risk maps for specific zones. This multi-layered approach will enhance system reliability, ensure data integrity, and lay the groundwork for scalable, resilient, and scientifically validated IoT solutions for water monitoring within smart-city frameworks, fostering sustainable and data-driven management of urban water resources.

Acknowledgments. The authors did not receive any specific grant or financial support for this research.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. World Bank: Water Overview. World Bank, Washington, D.C. (2025), https://www.worldbank.org/en/topic/water/overview, last accessed 2025/09/15
2. World Health Organization: Drinking Water. WHO, Geneva (2023), https://www.who.int/news-room/fact-sheets/detail/drinking-water, last accessed 2025/09/15
3. United Nations: The Sustainable Development Goals Report 2023. United Nations, New York (2023), https://sdgs.un.org/sites/default/files/2023-07/The-Sustainable-Development-Goals-Report-2023_0.pdf, last accessed 2025/09/15
4. United Nations: Progress towards the Sustainable Development Goals. Report of the Secretary-General (A/80/81-E/2025/62). United Nations, New York (2025), https://unstats.un.org/sdgs/files/report/2025/secretary-general-sdg-report-2025--EN.pdf, last accessed 2025/09/15
5. Lupi, O., Zaradnik, I., Canziani, M.: State of the art of water quality monitoring systems. ReDDi 5(2), 1–8 (2020), http://repositoriocyt.unlam.edu.ar/handle/123456789/1221, last accessed 2025/09/15
6. López-Ramírez, G.A., Aragón-Zavala, A.: Wireless Sensor Networks for Water Quality Monitoring: A Comprehensive Review. IEEE Access 11, 95120–95142 (2023). https://doi.org/10.1109/ACCESS.2023.3308905
7. IoT Analytics: Insights Release – Number of connected IoT devices (Sep 4, 2024), https://iot-analytics.com/wp-content/uploads/2024/09/INSIGHTS-RELEASE-Number-of-connected-IoT-devices-vf.pdf, last accessed 2025/09/15
8. Singhvi, H.: The Internet of Things in 2025: Trends, Business Models, and Future Directions for a Connected World. Int. J. Internet Things 3(1), 17–24 (2025). https://doi.org/10.34218/ijiot_03_01_003
9. Transforma Insights: Global IoT connections forecast above 40 billion by 2034. Transforma Insights Reports (2024), https://transformainsights.com/news/global-iot-connections-forecast-above-40-billion-2034, last accessed 2025/09/15
10. Essamlali, I., Nhaila, H., El Khaili, M.: Advances in machine learning and IoT for water quality monitoring: A comprehensive review. Heliyon 10(6), e27920 (2024). https://doi.org/10.1016/j.heliyon.2024.e27920
11. Forhad, H.M., Uddin, M.R., Chakrovorty, R.S., Ruhul, A.M., Faruk, H.M., Kamruzzaman, S., Sharmin, N., Jamal, A.H.M.S.I., Haque, M.U., Morshed, A.K.M.M.M.: IoT-based real-time water quality monitoring system in water treatment plants (WTPs). Heliyon 10(24), e40746 (2024). https://doi.org/10.1016/j.heliyon.2024.e40746
12. Hamel, P., Ding, N., Cherqui, F., Zhu, Q., Walcker, N., Bertrand-Krajewski, J., Champrasert, P., Fletcher, T.D., McCarthy, D.T., Navratil, O., Shi, B.: Low-cost monitoring systems for urban water management: Lessons from the field. Water Research X 22, 100212 (2024). https://doi.org/10.1016/j.wroa.2024.100212
13. Lakshmikantha, V., Hiriyannagowda, A., Manjunath, A., Patted, A., Basavaiah, J., Anthony, A.A.: IoT-based Smart Water Quality Monitoring System. Global Transitions Proceedings 2, 181–186 (2021). https://doi.org/10.1016/j.gltp.2021.08.062

14. Hakimi, I.M., Jamil, Z.: Development of Water Quality Monitoring Device Using Arduino UNO. IOP Conf. Ser.: Mater. Sci. Eng. 1144, 012064 (2021). https://doi.org/10.1088/1757-899X/1144/1/012064
15. Nanyanzi, D.R., Ocen, G.G., Omara, T., Bwire, F., Matovu, D., Semwogerere, T.: Design and assembly of a domestic water temperature, pH and turbidity monitoring system. BMC Res. Notes 14(1), 161 (2021). https://doi.org/10.1186/s13104-021-05578-9
16. Rokhman, F., Arini, F.Y., Hidayat, A., Kurniawan, T.: Water quality monitoring system for temperature, pH, turbidity, DO, BOD, and COD parameters based on Internet of Things in Garang watershed. IOP Conf. Ser.: Earth Environ. Sci. 1395(1), 012004 (2024). https://doi.org/10.12912/27197050/174412
17. Sakib, M.R., Haque, M., Ahmed, M., Rahman, M., Bain, R.: A Design and Implementation of IoT-Based Water Quality Monitoring System. Bachelor's thesis, Department of Computer Science and Engineering, Sonargaon University, Dhaka (2024), https://www.researchgate.net/publication/379992549_A_Design_and_Implementation_of_IoT_Based_Water_Quality_Monitoring_System, last accessed 2025/10/27
18. Abrajano, J.V., Apanay, J., Botangen, K.A., Nabua, J., Peña, C.F.: IoT-based water quality monitoring system in Philippine off-grid communities. arXiv preprint arXiv:2410.14619v2 (2024). https://doi.org/10.48550/arXiv.2410.14619 last accessed 2025/10/27
19. Abdelmoez, M.N., Ibrahim, K., Ali, A.S., Heshmat, M.: A standalone sensing and actuation IoT solution for water management, leakage detection, and localization problems. Water Conserv. Sci. Eng. 9, 16 (2024). https://doi.org/10.1007/s41101-024-00248-w
20. López, M.A., Torrealba, R., Arriaga, C.: Water quality monitoring system with an IoT orientation. In: International Congress on Mechatronics, Control and Artificial Intelligence, pp. 123–127 (2023), https://virtual.cuautitlan.unam.mx/intar/memoriascimcia/wp-content/uploads/sites/20/2024/01/123-Sistema-de-monitoreo-de-la-calidad-del-agua-con.pdf, last accessed 2025/09/15
21. Zulkifli, C.Z., Garfan, S., Talal, M., Alamoodi, A.H., Alamleh, A., Ahmaro, I.Y.Y., Sulaiman, S., Ibrahim, A.B., Zaidan, B.B., Ismail, A.R., Albahri, O.S., Albahri, A.S., Soon, C.F., Harun, N.H., Chiang, H.H.: IoT-Based Water Monitoring Systems: A Systematic Review. Water 14(22), 3621 (2022). https://doi.org/10.3390/w14223621`
22. Obunga, P.O., Rwanga, S.S., Dinka, M.O.: Review of the emerging technologies in the water sector with a focus on the deployment of Internet of Things solutions. npj Clean Water 8(1), 53 (2025). https://doi.org/10.1038/s41545-025-00487-x
23. Beltrán, S.: Design thinking and agile methodologies (SCRUM) as an enhancing effect in project management. Universidad Militar Nueva Granada (2022), https://repository.umng.edu.co/items/2ea53e73-2091-48c1-9cd1-a07fd3755b81, last accessed 2025/09/19

Classification of avocado leaves into healthy and unhealthy by computer vision and deep learning

Liana Leiva-Romero[1[0009-0006-5696-9461]] , José Alberto Hernández-Aguilar [1[0000-0002-5184-0005]], Outmane Oubram [1[0000-0003-2783-7364]], Pedro Moreno-Bernal[1[0000-0002-2811-5331]], Gustavo Medina-Ángel [1[0000-0002-0279-3492]], Ponciano Jorge Escamilla-Ambrosio [2[0000-0003-3772-3651]] and Jesús del Carmen Peralta-Abarca[1[0000-0003-2995-9277]]

[1] Universidad Autónoma del Estado de Morelos, Av. Universidad No. 1001, Col. Chamilpa, Cuernavaca, Morelos 62209, México
[2] Centro de Investigación en Computación, Instituto Politécnico Nacional, Av. Juan de Dios Bátiz, esq. Miguel Othón de Mendizábal,
Col. Nueva Industrial Vallejo, Alcaldía Gustavo A. Madero, CDMX 07738, México

lianaleiva6@gmail.com, jose_hernandez@uaem.mx,outmane@uaem.mx, pmoreno@uaem.mx,pescamilla@cic.ipn.mx,gustavo.medina@uaem.edu.mx,carmen.peralta@uaem.mx

Abstract. The article presents a study on classifying avocado leaves as healthy and unhealthy using computer vision and deep learning techniques. The primary objective is to develop an automated system for detecting diseases in avocado leaves at an early stage, thereby enhancing the efficiency of crop management. For this purpose, a data repository was created using images of healthy and unhealthy avocado leaves from two datasets: NZDLPlantDisease-v1 and the Avocado Leaf Disease Dataset - Kaithala, which were used to train the neural network. MobileNetV2 was used to extract features, and its output was passed to a CNN consisting of a flattening layer, a dense layer, a normalization layer, with a dropout of 0.5, and a learning rate of 0.00001. The loss function used was binary cross-entropy, and the optimizer was the Adam algorithm. The model was executed for 50, 150, and 250 epochs. For training, an accuracy of 99.21% and a loss of 0.0518 were obtained; for validation, an accuracy of 98.35% and a loss of 0.0636; and in the test data, an accuracy of 97.1% and a loss of 0.111.

Keywords: Convolutional Neural Network (CNN), avocado leaves, disease, binary classification.

1 Introduction

In agricultural production, crops are susceptible to several diseases that require prompt identification and treatment. These diseases can significantly affect the quality and

yield of crops, which in turn impacts the agricultural economy and food security. The existing method for detecting diseases in crops is through visual observation by agricultural experts. This process requires time and is dependent on the availability and high cost of expert consultants, which can be a significant constraint, especially in rural areas or developing countries [1].

It is therefore essential to have automatic methods for the detection and classification of diseases in crops. These methods utilize advanced technologies, such as computer vision and deep learning, to analyze images of their leaves and quickly and accurately detect signs of disease. [2]. The implementation of these automatic systems would not only reduce dependence on human experts but also enable early and timely detection, which is crucial for effective disease management and the proper development of crops. The existing methods include neural networks and computer vision. The detection of plant diseases based on deep learning has gained significant attention from the scientific community [1].

Mexico has exponentially increased the planting, production, and export of avocados in the last 60 years; in turn, it is the world's leading producer and exporter [3].Therefore, the use of modern techniques, such as precision agriculture and artificial intelligence, is essential to maintaining competitiveness. Given the importance of this crop, which may be susceptible to certain diseases, this project focuses on classifying avocado leaves into healthy and unhealthy categories to detect diseases in their early stages. The hypothesis to be validated is: H1. A deep learning model with digital image processing will be able to identify distinctive characteristics that allow the classification of healthy and unhealthy leaves of avocado plants.

Among the scope and limitations of the project is the database, which was created with images of avocado leaves from two datasets reported in the literature. One is the Avocado Leaf Disease Dataset - Kaithala [9], and the other is the dataset _ NZDLPlantDisease-v1 [1].

On the other hand, by utilizing in-depth learning techniques, it is expected to develop an app that can be used by farmers and producers to continuously and accurately monitor the health of their avocado crops.

This document is organized as follows: The second section provides a summary of the literature review, highlighting articles related to the topic. The third section will then discuss the methodology to reach the proposed model. The fourth section shows the results and their discussion. Finally, the conclusions and future work will be set out.

2 Review of Literature

In Marimuthu et al. (2020) [5], a method is proposed to detect diseases in tomato plants using a Convolutional Neural Network (CNN) architecture. The database used in this investigation was PlantVillage. In total, 7,166 images of tomato leaves are available in the dataset, which are divided into training and testing sets at a ratio of 80% to 20%, respectively. The proposed classifier is designed to detect 10 different diseases (bacterial stain, early slate, late slate, leaf mold, septoria foliar stain, two-spot mite, target

spot, mosaic virus, leaf yellow curling virus). It was tested with four different variants of CNN architecture, varying the filter size in 3 x 3, 5 x 5, 7 x 7, and 11 x 11. All variants feature three convolutional layers matched with three maximum pooling layers, utilizing ReLU as the activation function. The learning rate was set to 0.001, and a dropout rate of 0.5 was used. The fully connected layer is implemented using a Dense layer with 10 classes and a SoftMax activation function. The highest accuracy obtained was 90.86%, achieved with a 3 x 3 filter.

In Hassan et al. (2021) [4], it is proposed to replace the standard convolution with the in-depth separable convolution, which reduces the number of parameters and the computational cost. InceptionV3, InceptionResNetV2, MobileNetV2, and EfficientNetB0 models were used. The models implemented were trained on an open dataset called PlantVillage, which comprises 14 different plant species and 38 distinct categorical classes of diseases and healthy plant leaves. Three different representations (color, grayscale, and segmented) were used, with the color representation yielding better results. The implemented models achieved precision rates in disease classification of 98.42%, 99.11%, 97.02%, and 99.56%, using InceptionV3, InceptionResNetV2, MobileNetV2, and EfficientNetB0, respectively.

In Mishra et al. (2022) [6], a model is proposed that utilizes the MobileNetV2 convolutional neural network and the MSCA-PSO hybrid optimization algorithm, proven to be able to classify avocado fruits into healthy and unhealthy. The proposed model utilizes the MobileNetV2 convolutional neural network and the hybrid MSCA-PSO optimization algorithm, demonstrating superiority in terms of convergence. Public datasets available on the Fruit-360 dataset and images collected from different regions of Ethiopia were combined. The proposed model, MSCA PSO Mobile-NetV2, achieved an accuracy of 98.42%.

In Saleem et al. (2022) [1], a new dataset, NZDLPlantDisease-v1, is generated. In this work, a Deep Learning approach is developed to detect diseases in five important horticultural crops in New Zealand. The methodology includes stratified cross-validation and increasing techniques, such as five-fold cross-validation and tests on external data. An optimized version of Completely Conventional Networks Based on Regions (RFCN) was proposed, achieving a mean average precision (mAP) of 93.80%. This study suggested the use of automated systems to apply disease control techniques and explore new data enhancement techniques, as well as the use of Deep Learning models based on segmentation.

In Abd Algani et al. (2023) [7], an innovative technique, known as Ant Colony Optimization with the Convolutional Neural Network (ACO-CNN), is proposed for deep learning-based disease detection and classification. The collection of data is based on healthy leaves and those infected with the following sets: green, chancre, melanosis, and black spots. The proposed model was used to classify infused leaves from healthy leaves. The ACO-CNN model has shown better accuracy, recall, and F1 score compared to the C-GAN, CNN, and SGD models. In C-GAN, CNN, and SGD, the accuracy rates are 99.6%, 99.97% and 85%, respectively.

In (Abera and Avula, 2023) [8], a deep learning-based avocado leaf disease detection and classification system is proposed. The considered diseases include algal spot, leaf

burn, persea mite damage, and powdery mildew. This research has undergone the following steps: image acquisition, image preprocessing, feature extraction, model training, and detection and classification. The proposed method was classified into three phases by changing the values of hyperparameters. All phases had passed the above processing steps. Avocado leaves affected by four diseases and healthy leaves were trained and classified into their disease class. To test the model's performance, the dataset was split into a training set and a validation set in an 80:20 ratio, respectively. A Convolutional Neural Network (CNN) model is used for detection and classification purposes. CNN has demonstrated high performance in training, disease prediction, and classification, achieving 100% training accuracy and 98.8% testing accuracy, with an overall accuracy of 99%.

3 Methodology

The methodology used in this research is as follows, shown in Figure 1:

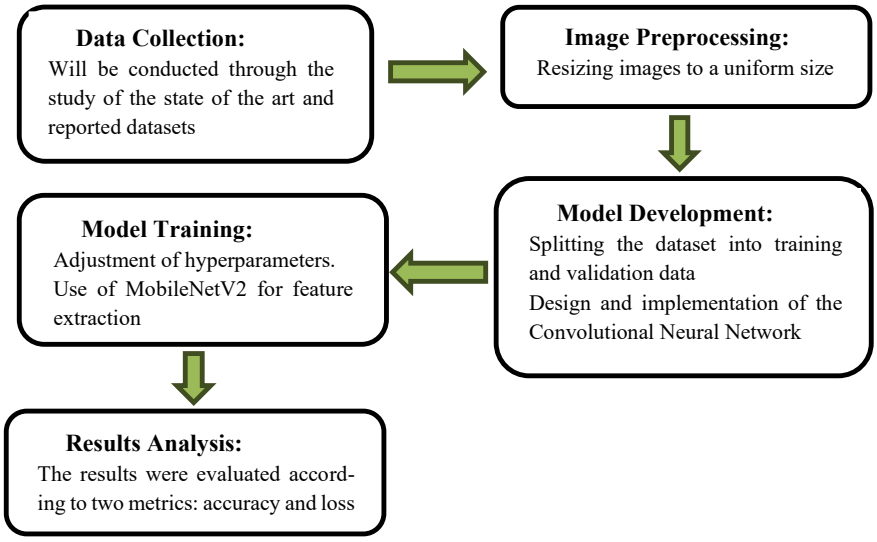


Fig. 1. Proposed Methodology.

3.1 Acquisition of images

The images used were collected from two databases reported in the revised literature: K-Kotagiri (NRB) Avocado Leaf Dataset [9] and dataset_NZDLPlantDisease-v1 [1].

The K-Kotagiri (NRB) Avocado Leaf Dataset [9] consists of 435 raw images, showcasing both healthy and disease-infected avocados. Specifically, 216 images depict avocados affected by multiple diseases, while 219 images represent healthy specimens.

The Dataset_NZDLPlantDisease-v1 [1] contains plant disease in five different crops in New Zealand, including kiwifruit, apple, pear, avocado, and grapevine. The images were captured using a Samsung Galaxy S10 Plus, featuring a 12 MP f/1.5-2.4 (wide), 12 MP f/2.4 (telephoto), and 16 MP f/2.2 (ultrawide) camera. The images were taken at a working distance of 200-300 mm. From this dataset, only the original avocado images were used, avoiding duplicated images, of which 304 belong to the Disease class and 226 to the Healthy class.

The resulting dataset comprised the binding of all images from the K-Kotagiri (NRB) Avocado Leaf Dataset [9] and the original avocado images of the NZDLPlantDisease-v1 [1]. A total of 445 images were allocated to the Healthy class and 520 to the Disease class.

3.2 Preprocessing of images and model development

The images were divided into training (80%), validation (10%), and test (10%). The split dataset function takes all the images from the *Healthy* and *Disease* folders, randomly shuffles them, and then copies a defined percentage of those images into the train, validation, and test folders for each category. This results in 772 images for training, 96 images for validation, and 97 images for testing. These images were resized to a resolution of 224x224 and used in color. Data augmentation techniques (horizontal flipping and 90-degree rotation) were applied only to the training data to increase the diversity of the avocado leaf images and improve model generalization.

3.3 In-depth learning model training

First, the pre-trained MobileNetV2 was used for feature extraction with the following architecture:

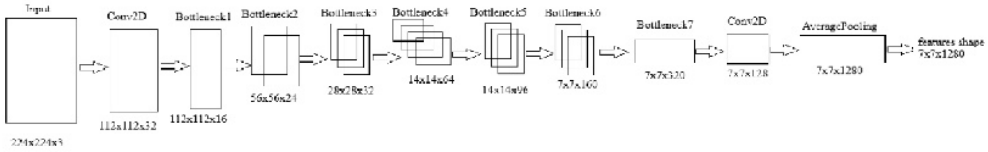


Fig. 2. Architecture of MobileNetV2.

MobileNetV2 is an efficient, lightweight neural network architecture primarily designed for mobile devices and resource-constrained applications. Introduced by Google in 2018, it is based on:

- Bottleneck Depth-wise Separable Convolutions. Divide standard convolutions into two more efficient steps.
- Inverted Residuals with Linear Bottlenecks. Expand dimensionality with a 1x1 convolution, applies depth-wise convolution, and then reduces dimensionality

with another linear 1x1 convolution without ReLU activation, preserving critical information.

These blocks are connected via skip connections when the input and output dimensions match [11].

The architecture of this network comprises 2,257,984 total parameters, consisting of 2,223,872 trainable parameters and 34,112 non-trainable parameters.

Figure 4 illustrates the shapes of the outputs for training, validation, and testing.



Fig. 3. Shape of the output data for training, validation, and testing.

These features were used as input to the following CNN architecture:

```
inputs = keras.Input(shape=(7, 7, 1280))
x = layers.Flatten()(inputs)
x = layers.Dense(256)(x)
x = layers.BatchNormalization()(x)
x = layers.Activation("relu")(x)
x = layers.Dropout(0.5)(x)
outputs = layers.Dense(1, activation="sigmoid")(x)
optimizer = keras.optimizers.Adam(learning_rate=0.00001)
model = keras.Model(inputs, outputs)
model.compile(loss="binary_crossentropy",
              optimizer=optimizer,
              metrics=["accuracy"])
```

The input is a tensor with shape (7, 7, 1280). First, this tensor is flattened and then passed through a dense layer with 256 units. Batch Normalization is applied to stabilize training, followed by a ReLU activation that introduces non-linearity. Next, a Dropout layer with a rate of 0.5 is used to prevent overfitting. Finally, the output is obtained

through a dense layer with a single unit and sigmoid activation, which is ideal for binary classification.

The model is compiled using the Adam optimizer (with a learning rate of 0.00001), the binary cross-entropy loss function, and accuracy as the evaluation metric, making it suitable for fine-tuning tasks in two-class classification. The model is then trained. A callback is defined to save only the best version based on the validation loss. The model was trained for 50, 150, and 250 epochs. In the code below, the number of epochs is set to 50.

```
callbacks = [
    keras.callbacks.ModelCheckpoint(
      filepath="feature_extraction.keras",
      save_best_only=True,
      monitor="val_loss")
]
history = model.fit(
    train_features, train_labels,
    epochs=50,
    validation_data=(val_features, val_labels),
    callbacks=callbacks)
```

3.4 Analysis of results

The results were evaluated using two metrics: accuracy and loss, considering 50, 150, and 250 epochs, for which the graphs and figures are presented below.

Results for 50 epochs. In Figure 5, the graph shows that the training accuracy (dots) quickly reaches values close to 0.94, while the validation accuracy (line) remains stable at around 0.92. This suggests that the model is learning well.

In Figure 6, the training loss steadily decreases to 0.2, and the validation loss also decreases steadily to 0.2, indicating that the model maintains a solid overall performance without significant signs of overfitting.

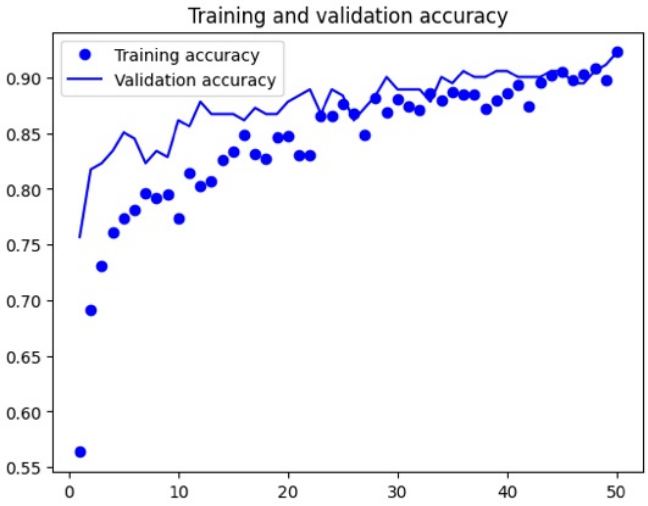


Fig. 4. Accuracy for Training and Validation with 50 Epochs

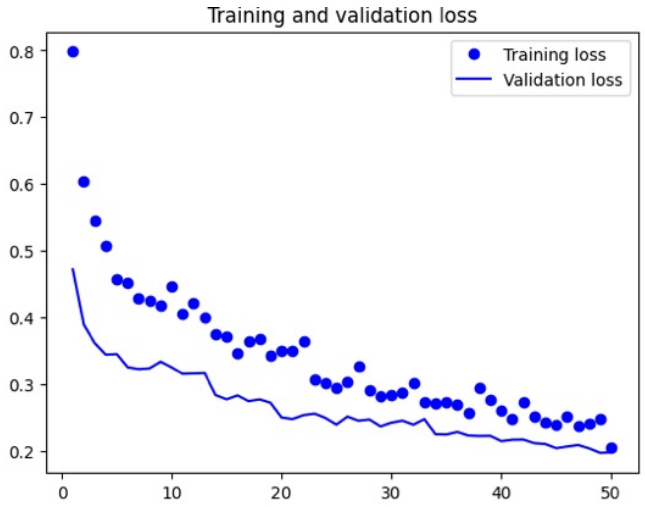


Fig.5. Loss for Training and Validation with 50 Epochs

Results for 150 epochs. As shown in Figure 7, the training accuracy (dots) reaches values close to 0.97, while the validation accuracy (line) converges to 0.94, indicating that the model is learning effectively.

In Figure 8, the training loss steadily decreases to very low values, while the validation loss also decreases consistently at first.

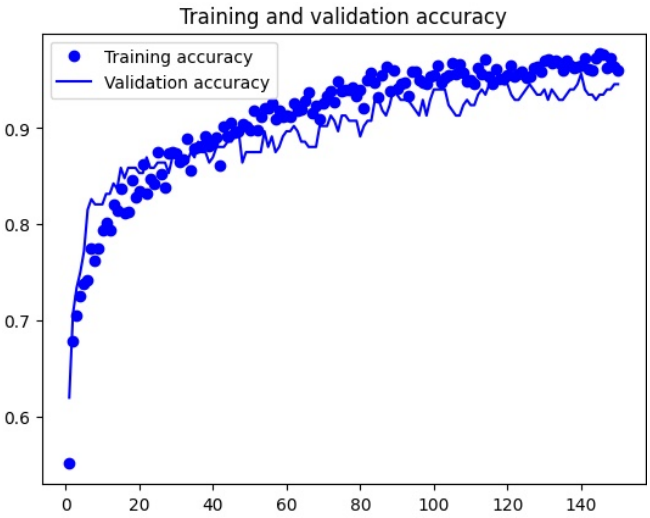


Fig. 6. Accuracy for Training and Validation with 150 Epochs

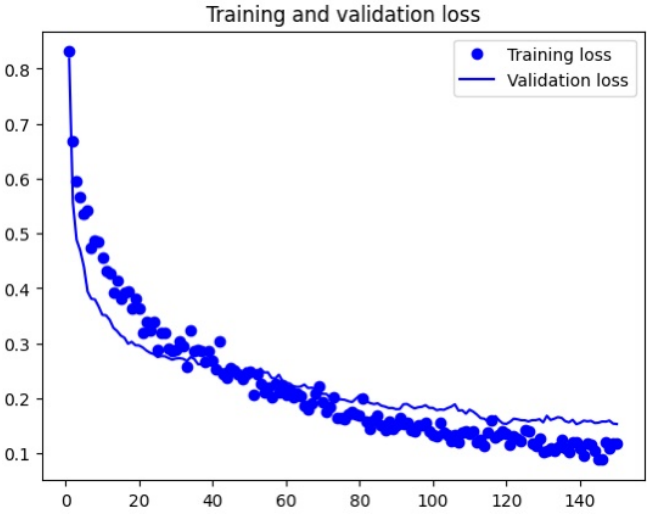


Fig. 7. Loss for Training and Validation with 50 Epochs

Results for 250 epochs. As shown in Figure 9, the training accuracy (dots) reaches values close to 0.99, while the validation accuracy (line) also converges to 0.97, indicating that the model is learning effectively. In Figure 10, the training loss steadily decreases to very low values, while the validation loss also initially decreases consistently.

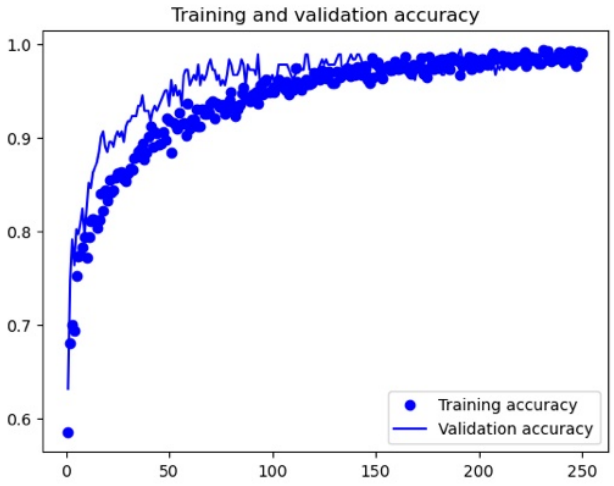


Fig. 8. Accuracy for Training and Validation with 250 Epochs

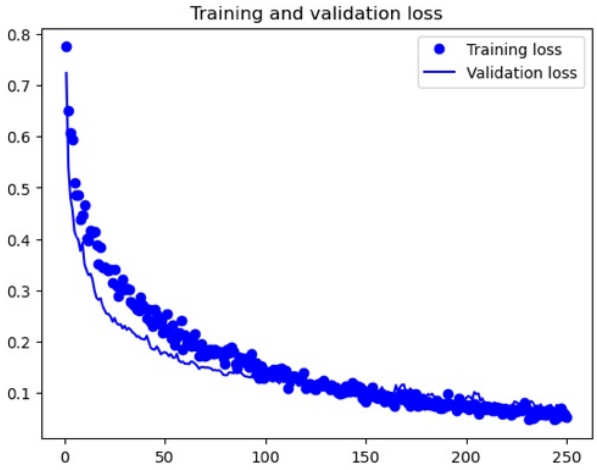


Fig. 9. Loss for Training and Validation with 250 Epochs

Table 1 summarizes the accuracy and loss results for training, validation, and testing with 50, 150, and 250 epochs.

Table 1. Accuracy and Loss Results for Training, Validation, and Testing Stages with 30, 50, and 100 Epochs

Epochs	Training	Validation	Testing
50 epochs	accuracy: 92.92% loss: 0.1854	accuracy: 92.27% loss: 0.1983	accuracy: 92.1% loss: 0.2011
150 epochs	accuracy: 97.76% loss: 0.19	accuracy: 95.78% loss: 0.23	accuracy: 95.4% loss: 0.18
250 epochs	accuracy: 99.21% loss: 0.0518	accuracy: 98.35% loss: 0.0636	accuracy: 97.1% loss: 0.111

The results obtained exceed 90% accuracy, and for 250 epochs, an accuracy of 97.1% is achieved on the test data, which is comparable to results reported in [8] where a Convolutional Neural Network model was used for classifying diseases in avocado leaves, achieving a test accuracy of 98.8%.

The primary contributions of this research are centered on the development of an efficient and accurate system for classifying avocado leaves. First, we propose a computer vision and deep learning approach to distinguish between healthy and unhealthy avocado leaves, providing support for automated crop monitoring. To achieve this, the MobileNetV2 architecture was implemented and validated, providing a lightweight and efficient model that can be deployed on mobile devices, thereby enabling farmers and agricultural technicians to utilize the model directly in the field. The model was trained and optimized over 250 epochs, achieving a maximum accuracy of 97.1% on the test data, which highlights its robustness and effectiveness in disease detection and classification. Additionally, a processed dataset of avocado leaves was generated, incorporating preprocessing techniques to enhance image quality, thereby contributing to reproducibility and future research efforts. Overall, this work contributes to the development of accessible technological solutions for precision agriculture by leveraging computationally efficient models that allow scalability and practical application in resource-limited environments.

4 Conclusions and Future Work

The use of public datasets such as the Avocado Leaf Disease Dataset – Kaithala [9] and NZDLPlantDisease-v1 [4] has been fundamental for the development of the model. These datasets have provided a solid foundation for training and validating the model, enabling accurate classification of avocado leaves.

The deep learning model based on MobileNetV2 has proven to be highly efficient for classifying avocado leaves as healthy or unhealthy. The best results resulted in 99.21% accuracy on the training set, 98.35% on the validation set, and 97.1% on the test data. This indicates a high capacity of the model to generalize to unseen data.

The implementation of this classification system can have a significant impact on agriculture, especially in the early detection of diseases in avocado crops. Farmers and producers can utilize this model to continuously and accurately monitor the health of

their crops, thereby reducing their reliance on human experts and enabling timely intervention.

The hypothesis proposed for this project is H0. A deep learning model with digital image processing will be able to identify distinctive features that allow for the classification of healthy and unhealthy avocado plant leaves. The results obtained throughout training, validation, and testing support the hypothesis proposed in this project.

A deep learning model was successfully developed, capable of classifying avocado plant leaves as healthy or unhealthy with 99.98% accuracy, using digital image processing techniques. The model was able to identify distinctive features associated with each class.

Limitations: Although the results are promising, the model has certain limitations, such as dependence on the quality and diversity of the training data.

Future Work: This research demonstrates that the proposed approach is practical for addressing diagnostic problems in crops and lays the foundation for future practical implementations using mobile computing in the automated monitoring of the phytosanitary status of avocado plants. Future research could focus on collecting more varied data, particularly from the Latin American region, and on implementing data augmentation techniques to enhance the model's robustness. Different models could also be tested with this data to compare performance. In future studies, we will evaluate these models.

References

[1] M. H. Saleem, J. Potgieter, and K. M. Arif, "A Performance-Optimized Deep Learning-Based Plant Disease Detection Approach for Horticultural Crops of New Zealand", *IEEE Access*, vol. 10, pp. 89798–89822, 2022, doi: 10.1109/ACCESS.2022.3201104.

[2] A. Upadhyay *et al.*, "Deep learning and computer vision in plant disease detection: a comprehensive review of techniques, models, and trends in precision agriculture", *Artif Intell Rev*, vol. 58, Num. 3, Mar. 2025, doi: 10.1007/s10462-024-11100-x.

[3] R. Díaz Castellanos, "mercado mundial de aguacate: 60 años del liderazgo de México y su impacto en la próxima década", *The Anáhuac Journal*, vol. 21, núm. 2, pp. 12–49, 2021, doi: 10.36105/theanahuacjour.2021v21n2.01.

[4] S. M. Hassan, A. K. Maji, M. Jasiński, Z. Leonowicz, and E. Jasińska, "Identification of plant-leaf diseases using CNN and transfer-learning approach", *Electronics (Switzerland)*, vol. 10, Num. 12, jun. 2021, doi: 10.3390/electronics10121388.

[5] S. Marimuthu, J. Bhuvana and T.T. Mirnalinee, "Disease Detection in Tomato Plants Using Deep Learning", 2020. doi: 10.3233/apc200140.

[6] S. Mishra, T. H. Ayane, V. Ellappan, D. S. Rathee, and H. Kalla, "Avocado fruit disease detection and classification using modified SCA–PSO algorithm-

based MobileNetV2 convolutional neural network", *Iran Journal of Computer Science*, vol. 5, Num. 4, pp. 345–358, Dec. 2022, doi: 10.1007/s42044-022-00116-7.

[7] Y. M. Abd Algani, O. J. Marquez Caro, L. M. Robladillo Bravo, C. Kaur, M. S. Al Ansari, and B. Kiran Bala, "Leaf disease identification and classification using optimized deep learning", Measurement: Sensors, vol. 25, Feb. 2023, doi: 10.1016/j.measen.2022.100643.

[8] H. Abera and V. A. Avula, "Avocado Leaf Disease Classification Using Convolutional Neural Network", *International Journal of Science and Research (IJSR)*, vol. 12, Num. 9, pp. 780–785, Sep. 2023, doi: 10.21275/sr23908191606.

[9] T. Rajasekaran, P. Naveen, and V. Balasubramaniam, "K-Kotagiri (NRB) Avocado Leaf DataSet", 2024. doi: 10.17632/6zy6wxhf2v.1.

[10] S. Liu *et al.*, "An apple leaf disease identification model for safeguarding apple food safety", *Food Science and Technology (Brazil)*, vol. 43, 2023, doi: 10.1590/fst.104322.

[11] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks", ene. 2018, Available on: http://arxiv.org/abs/1801.04381

Data Strategy as a Value Enabler in Sustainable, Smart, and Circular Cities

Augusto Velasquez Mendez[1*], Tito Pablo Neira Avila[1], Jorge de Jesús Lozoya Santos[2], José Fernando Jiménez Vargas[1]

[1]Universidad de Los Andes, Bogotá, Colombia
[2]Instituto Tecnológico de Monterrey, Monterrey, México

*a.velasquem2@uniandes.edu.co, tp.neira@uniandes.edu.co, jorge.lozoya@tec.mx, fjimenez@uniandes.edu.co

Abstract. The move toward Sustainable, Smart, and Circular Cities (SSCCs) requires technological integration and a clear data strategy that promotes urban intelligence, stakeholder engagement, and value creation. This paper examines how data strategies create operational, social, environmental, and economic benefits in SSCCs by enabling four key mechanisms: informed decision-making, improved urban operations, product and service innovation, and business model transformation. Building on ideas from corporate and urban fields, we define a city data strategy as a cross-cutting planning tool that manages the collection, integration, governance, and use of urban data to support sustainability and circularity goals. Insights from the Fenicia Urban Living Lab (ULL) in Bogotá illustrate how local data ecosystems can activate these mechanisms. The first case involves a waste management project that combines AI, georeferenced images, and a digital twin to detect and respond to waste hotspots in near real-time. The second features a digitally enhanced urban garden that uses IoT sensors to monitor soil and environmental conditions, allowing for optimized irrigation and community-driven agricultural experiments. When integrated within participatory innovation infrastructures like ULLs, these interventions demonstrate how strong data strategies can improve service efficiency, empower citizens, and promote circular resource flows. Our findings indicate that data strategies in SSCC act as socio-technical infrastructures for coordinating value across multiple areas. We conclude by proposing a conceptual framework for assessing urban data strategies and discussing implications for data governance, real-time responsiveness, and innovation in complex urban systems.

Keywords: data strategy, smart cities, circular economy, urban living labs, optimized operations, informed decision-making, real-time management, digital twin, citizen participation.

1 Introduction

Cities worldwide face transformation from urbanization, climate change, and socio-economic factors. Sustainable, Smart, and Circular Cities (SSCCs) combine digital innovation, sustainability, and circular economy to build resilient, inclusive urban areas [1-4]. They aim to integrate intelligence and sustainability into urban governance, services, and infrastructure, moving beyond mere technological upgrades.

A key driver of this paradigm shift is the digital transformation of urban systems, which changes city operations, public service delivery, and resident interactions. This is fueled by more urban data from Internet of Things (IoT), remote sensing, citizen devices, and platform-based interactions [1,3,4]. When well-organized, these data enable real-time monitoring, predictive modeling, and system-wide optimization [2,5].

In this context, the strategic use of data becomes an essential part of SSCCs. A clear data strategy helps cities extract actionable insights from large data streams, coordinate stakeholders across sectors, and implement targeted actions aligned with sustainability and circularity goals [1,2,6]. As several scholars have highlighted, such strategies can assist cities in achieving optimized urban operations, improved decision-making, and real-time management, thereby creating value across operational, social, and environmental domains [1,2,5,7].

Despite its growing significance, the concept of a data strategy for SSCCs remains underexplored in academic research, often divided into technical, governance, and sustainability perspectives. To address this gap, this paper investigates the following research question:

How does a data strategy generate value in sustainable, smart, and circular cities?

The article develops a conceptual framework for understanding data strategies as value enablers in SSCCs to answer this question. Drawing on a comparative analysis of enterprise and urban contexts, the paper examines how data strategies support city functions. It also analyzes two empirical cases from the Fenicia Urban Living Lab (ULL) in Bogotá: (i) a waste management experiment using georeferenced data and AI for hotspot detection, and (ii) a digitally enhanced urban garden with IoT sensors for environmental monitoring and community engagement.

This study demonstrates how integrating AI, IoT, and digital twin technologies within participatory frameworks can develop data-driven infrastructures that support sustainable urban transformation. Building on this foundation, the paper conceptualizes data strategies as socio-technical infrastructures that align technological capabilities, institutional processes, and community participation to promote value creation in sustainable and circular urban systems.

While previous studies have examined data strategies in smart or sustainable cities, few have explored their combined role across sustainability, circularity, and governance dimensions. This study fills that gap by framing data strategy as a systemic, value-enabling mechanism, validated through empirical experiments conducted in the Fenicia ULL. The paper is organized as follows: Section 2 reviews the evolution of data strategies. Section 3 discusses value creation in SSCCs. Section 4 compares urban and enterprise data frameworks. Section 5 shares empirical insights from Fenicia ULL. Section 6 offers final remarks and suggestions for future research.

2 Theoretical and Conceptual Foundations

In urban innovation, data now actively shapes decision-making, planning, and operations. As cities focus on sustainability, circularity, and tech, a data strategy is vital for explaining data collection, governance, analysis, and value creation. Although common in corporate settings, 'data strategy' in urban systems remains fragmented and evolving.

This section examines the data strategy within SSCC by linking enterprise digital transformation with urban social and technical realities. It traces the concept's evolution from business origins (2.1), explores digital transformation in cities and how data infrastructures reshape operations (2.2), discusses the evolving citizen role in urban data ecosystems, including participation and data justice issues (2.3), and identifies obstacles to value creation in smart city systems (2.4).

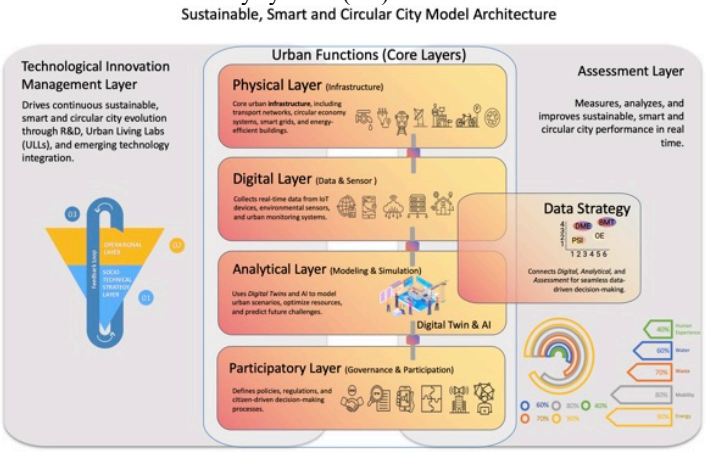


Fig. 1. Sustainable, Smart and Circular City – SSCC Model Architecture, highlighting the systemic integration of the data strategy.

This exploration uses the SSCC model, which organizes urban functions into physical, digital, analytical, and participatory layers [8]. These are connected with the Technological Innovation Model (TIM) and Assessment layers [9,10]. The Data Strategy Layer ensures coherence and governance across sensing, simulation, and evaluation processes. Figure 1 shows this architecture and the integration of the data strategy within the SSCC model.

2.1 Defining Data Strategy: From Corporate to Urban Settings

In corporate settings, data strategy is a plan aligning business goals with data resources, analytics, and governance to create value [11]. It covers data collection, quality, access, analysis, and monetization. According to Neira and Vesga [12,13], data strategy is vital for digital transformation, boosting efficiency, decision-making, and innovation in products and models [12,13].

Applying this concept to urban settings, a data strategy for SSCC must navigate a complex environment involving multiple actors and scales. Cities prioritize public value, sustainability, equity, and resilience over profit. Therefore, urban data strategies should enable cross-sector coordination, transparency, inclusion, and sustainability goals. According to Bibri [1,2] and Bibri & Krogstie [6], such a strategy should connect sensing infrastructure, modeling systems like digital twins, and governance platforms to inform planning, optimize resources, and support circularity.

2.2 Digital Transformation in SSCC: More Than Just Smart Services

Urban digital transformation often focuses on smart tech or digitizing services. In SSCC, it signifies a systemic shift in city management, experiences, and co-production, involving technical upgrades and redesigning institutional arrangements, citizen interactions, and value models [3,4,6].

Authors like Bibri [2,5] and Omrany [14] emphasize that transforming cities involves integrating ICTs, AI, IoT, and digital twins to optimize operations, inform decision-making, and enable real-time responses. These technologies must also be incorporated into governance frameworks that support inclusivity, accountability, and strategic alignment. In SSCC, digital transformation serves as a tool to advance climate action, promote circular resource flows, and strengthen social cohesion. A data strategy should coordinate these capabilities, adjusting to local contexts, capacities, and regulations.

2.3 The Role of the Citizen in the Urban Data Ecosystem

While smart city stories depict citizens as passive data sources, the SSCC paradigm views them as co-producers and active contributors to urban change. Cardullo [15] notes urban data ecosystems should move from technocratic control to civic-centered models emphasizing participation, transparency, and rights.

The transformation is not only digital but also civic and cultural, with citizens creating and interacting with data through various platforms. However, this introduces risks such as inequality, surveillance, and exclusion, which require ethical guidelines and governance [5,16,17]. Urban digital citizenship grants citizens the right to access services, question data, and influence how it is used. Therefore, SSCC's data strategy should include ways to empower citizens, safeguard their rights, and promote participation in urban data systems.

2.4 A Data Strategy Perspective on SSCC Bottlenecks

Despite efforts to implement data-driven SSCC, challenges remain in transforming urban data into insights. Fang [18] cites three main bottlenecks: (i) inefficient data collection, (ii) fragmented processing, and (iii) weak data application frameworks. These highlight the need to rethink data infrastructures as backbones aligned with urban sustainability. A comprehensive data strategy is vital for bridging sensing and smart governance.

To address challenges, we recommend examining the SSCC data lifecycle through three aspects: acquisition, processing, and application as shown in Figure 2. A

comprehensive strategy must encompass these areas, integrating strategic planning, governance, and socio-technical factors.

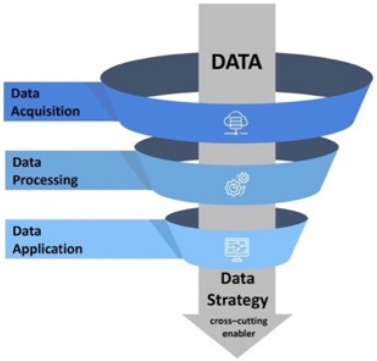


Fig. 2. Framing SSCC Data Bottlenecks Through a Strategic Lens

Data Acquisition: Smart cities rely on diverse data sources like IoT devices, mobile platforms, Unmanned Aerial Vehicles (UAVs), citizen sensors, and institutional data to model urban dynamics in real time. However, Fang [18] notes that current systems often struggle with data collection and integration, resulting in inconsistent coverage and low granularity. A strong data strategy can overcome these issues by establishing standards for interoperability, calibration, and synchronization, and by incorporating participatory sensing to involve communities in urban decision-making.

Data processing: Turns urban data into insights through aggregation, cleansing, normalization, and semantic alignment. Many smart city platforms struggle with siloed systems and limited hardware, limiting real-time and multimodal analytics. A data strategy with a scalable architecture, like the edge-network-cloud model, can improve this. Using AI pipelines to detect anomalies, simulate urban events, and predict resource needs enhances sustainability and circularity.

Data Application: Data strategies turn information into insights but often overlook the full data–information–knowledge chain, limiting analytics in urban planning and resilience. A good data strategy aligns sectors like waste, mobility, and energy with decision-making and policy, enabling evidence-based actions, feedback, and co-created public value.

A data strategy unifies sensing, analysis, and decision-making, ensuring coherence. In the SSCC model, this integration is a cross-cutting mechanism linking the Digital, Analytical (via digital twins and AI), and Assessment layers for dynamic urban management.

This perspective also underscores the importance of citizen participation, ethical data governance, and contextual relevance as critical components of urban data strategy development. Furthermore, by addressing the structural weaknesses identified in empirical and theoretical research, data strategies become essential infrastructures for fostering resilience, inclusivity, and sustainability in contemporary urban systems.

3 Methodology

This study employs a comparative, case-based methodology drawing from both corporate and urban sectors to examine how data strategies generate value in SSCC. The focus is on identifying how data strategies create value across operational, economic, social, and environmental dimensions. Specifically, the methodology compares a well-documented enterprise case of data strategy implementation—from Neira's study of a Latin American financial group—with use cases and frameworks developed by leading scholars in urban studies, such as Bibri, Bokolo, Munonye, and Omrany [1,2,5,6,7,11,12,14]. By comparing these cases, the analysis aims to identify common value creation mechanisms and assess how they can be applied and adapted to complex urban systems.

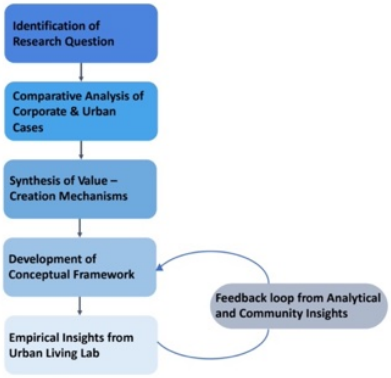


Fig. 3. Research Methodology.

Figure 3 presents the study's methodological structure, which combines a comparative, case-based approach—drawing from both corporate and urban fields—with an empirical validation loop in the Fenicia ULL. The process consists of five stages: (1) identifying the research question and conceptual gap, (2) comparing corporate and urban data strategy cases, (3) synthesizing common value-creation mechanisms, (4) developing a conceptual framework, and (5) interpreting empirical insights from the ULL.

The framework uses comparative analysis to show how data strategies add value in sustainable, smart, and circular cities across operational, economic, social, and environmental aspects. Its iterative approach incorporates analytical findings and community feedback to stay grounded in real-world socio-technical dynamics.

This approach focuses on understanding context and socio-technical patterns. The cases explore data strategies, the value they create, and the support conditions for their success.

3.1. Value Aspects in Urban Data Strategies

In digital transformation, data has become essential for creating value and driving innovation in urban systems. A data strategy (DS) is best described as an organizational

plan that combines business goals, data assets, technologies, analytical methods, and stakeholder experiences to generate value by enhancing existing business models or developing new, data-driven solutions [12, 19].

Urban data strategies are vital for creating sustainable smart cities. Their value comes from intentional arrangements that align data collection and use with urban goals like sustainability and resilience [6,20].

A data strategy (DS) in urban settings can be described as:

"*a systemic plan that guides how urban organizations collect, manage, analyze, and use data to support strategic goals such as sustainability, circularity, social equity, and innovation*".

This definition builds on core contributions from Davenport [21], Mazzei & Noble [22], and recent empirical work by Neira & Vesga [12,13], highlighting the importance of integrating technological capabilities, human skills, and governance structures in a unified way.

3.2. Mechanisms of Value Creation

In data-driven smart cities, value creation involves systemic transformation beyond technical optimization. Effective data strategies are dynamic enablers turning digital capabilities into urban advantages. These benefits result from structured, intentional processes that let data create value within city systems and stakeholder setups.

Building on Neira and Vesga's review [12], four core value creation mechanisms emerge, showing how data strategies produce outcomes aligned with urban sustainability, resilience, and inclusion.

1) **Decision-Making Enhancement (DME)**: Facilitating quicker, more informed, and context-aware policy choices [23]. Data strategies help cities make faster, evidence-based, and context-sensitive decisions. This involves using real-time data streams, predictive analytics, and scenario modeling tools to guide public policy, urban planning, and crisis response. Integrating urban digital twins and AI systems further boosts this process by enabling anticipatory governance [14].
2) **Operational Efficiency (OE)**: Improving urban services and lowering resource use through data-driven insights [24]. Operational optimization enhances city services and infrastructure performance with real-time data, predictive analytics, and integrated management systems [18]. By combining IoT, geospatial analytics, and automated control systems, urban data strategies help optimize resource flows and service delivery. Examples in waste management, mobility, and energy distribution show how cities can cut costs, reduce environmental impact, and enhance service quality [20].
3) **Product and Service Innovation (PSI)**: Using data to create new digital services, especially in mobility, energy, and waste management [25]. Data strategies help develop new urban services, particularly in emerging areas like micromobility, decentralized energy systems, and circular economy solutions. By building innovation ecosystems that connect public institutions, startups,

citizens, and research organizations, cities can use data for experimentation and value co-creation.

4) **Business Model Transformation (BMT)**: Creating new governance and economic models supported by real-time, interoperable data ecosystems [26]. The most fundamental mechanisms involve reconfiguring institutional and economic logics through interoperable data infrastructures. Examples include open data marketplaces, data trusts, and performance-based procurement schemes that shift value flows and governance models. These changes are significant in sustainable smart city models that combine circular economy and digital commons principles [18].

These four interconnected mechanisms work together: improved decision-making boosts efficiency, service innovation develops new models, and transformed governance enables inclusive, adaptable data practices. In the SSCC architecture, these are coordinated through layers like the Digital layer with sensor data, the Analytical layer with Digital Twin simulations, and the Innovation Management layer driven by ULLs, ensuring localized and systemic value creation. Figure 4 illustrates these connections.

Understanding that value depends on context, not all data strategies deliver the same results. Success depends on factors like data quality, organizational maturity, leadership, and governance, as outlined by Neira and Vesga [12,13].

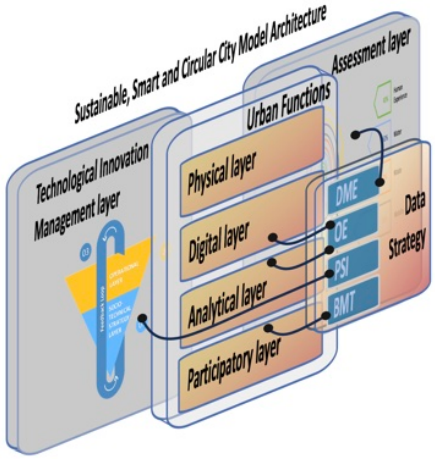


Fig. 4. Connections between Data Strategy mechanisms of value creation and SSCC Model Architecture's layers.

A comprehensive data strategy must include flexible mechanisms to assess and improve its ability to generate value in urban systems.

3.3. Critical Success Factors

Implementing data strategies in urban settings requires more than just infrastructure or formal plans—it relies on key technical and organizational enablers influencing their systemic impact. According to Neira and Vesga [13], four crucial factors emerge:

1) Data and Technology Management: Effective platforms and skilled teams are crucial for ensuring interoperability, cybersecurity, and system integration [27].
2) Data Quality and Maturity: Reliable, well-governed data supports analytics and decision-making; lacking standards can compromise system functionality [28,29].
3) Organizational Culture: Building a culture of experimentation, transparency, and evidence-based practice is essential across institutions [30].
4) Leadership and Collaboration: Effective, adaptable leadership and stakeholder coordination are vital to overcome inertia and maintain strategic consistency [13, 31].

These factors are interconnected and change over time, especially in cities working toward inclusive and sustainable growth.

3.4. From Strategy to Systemic Impact

For data strategies to enable systemic transformation, they must be incorporated into institutional routines, regulatory frameworks, and long-term urban agendas. Neira and Vesga [13] propose three main mechanisms to support this change: (1) experimentation cycles, often through ULLs, which validate the effectiveness of data-driven interventions in real-world settings [32]; (2) impact measurement, including both operational metrics and intangible factors like public trust and participation; and (3) policy feedback loops, which enable ongoing adjustments based on insights generated from data.

These mechanisms extend beyond isolated projects and establish data strategies as cross-cutting infrastructures within the SSCC model—integrating sensing, modeling, participation, and evaluation [33]. Achieving systemic impact ultimately depends on institutional learning, scalable practices, and governance models that adapt to evolving socio-technical dynamics.

4 Comparative Insights: From Corporative to Smart Cities

While the concept of data strategy originally developed in corporate settings as a structured plan to align business goals with digital capabilities and data assets [21,34], adapting it to urban systems—especially in Sustainable, Smart, and Circular Cities (SSCC)—requires a more extensive reconfiguration. This adaptation addresses the inherently multi-actor, multi-scalar, and value-diverse nature of cities, which aim not for profit maximization but for creating public value, resilience, and equity [6, 12].

Drawing on cross-sector insights from enterprise and urban research, Table 1 compares data strategy features across 11 key dimensions. These dimensions are identified by combining conceptual developments from both fields, including organizational purpose, value orientation, stakeholder involvement, governance structure, data infrastructure, and impact logic.

Corporate data strategies usually operate within hierarchical governance models, with clearly defined internal stakeholders and proprietary infrastructures optimized for

short- to medium-term financial performance [19]. In contrast, urban data strategies must orchestrate a larger network of actors—including public agencies, civil society, academia, and private entities—under regulatory frameworks and participatory approaches emphasizing transparency, accountability, and inclusion [13,15].

Additionally, while enterprises focus on product and service innovation to gain a competitive edge, cities engage in socio-technical experimentation to develop public services and sustainable urban solutions collaboratively. This involves integrating data strategies into long-term urban plans, utilizing tools such as digital twins, interoperable standards, and ULLs as innovation infrastructure [5,14].

Furthermore, impact measurement varies significantly across different areas. While corporate strategies rely on KPIs and return-on-investment (ROI) metrics, SSCCs need to evaluate value in multiple dimensions: economic, environmental, social, and participatory. This broader perspective on impact supports the core principles of sustainability and circularity embedded in the SSCC model architecture.

By presenting these comparative insights, Table 1 below highlights the challenges of translating corporate data strategy models into the urban domain and uncovers the unique institutional, infrastructural, and cultural conditions that need to be considered when designing effective data strategies for cities.

Table 1. Comparative Table: Enterprise Vs Urban Data Strategy.

Dimension	Enterprise Data Strategy	Urban Data Strategy
Primary Objective	Profit maximization and operational efficiency	Sustainability, resilience, and social equity
Value Orientation	Shareholder and customer value	Public value and collective wellbeing
Stakeholders	Internal departments, partners, customers	Government, citizens, civil society, private sector, academia
Governance Model	Hierarchical, centrally governed	Multi-actor, participatory, regulatory integration
Data Sources	Transactional systems, CRM, ERP, IoT	IoT, open data, citizen-generated data, satellite imagery
Innovation Approach	Product/service innovation for market advantage	Co-creation and socio-technical experimentation
Time Horizon	Short to medium term (quarterly, annual)	Medium to long term (multi-annual, generational)
Risk Management	Security, compliance, ROI-focused	Ethics, inclusion, resilience
Impact Measurement	KPIs, ROI, customer metrics	Multidimensional: environmental, social, economic indicators
Infrastructure	Proprietary platforms, cloud services	Open platforms, interoperable standards, digital twins
Cultural Context	Business-driven, competitive environment	Governance-driven, civic and institutional

5 Results

This section explains how the conceptual framework was developed, which is the main result of the comparative analysis. The framework combines the value creation mechanisms identified in both corporate and urban data strategy cases—such as operational efficiency, economic performance, social inclusion, and environmental sustainability—into a single model. This model shows how data strategies create value within the context of SSCCs. By linking data governance, technological innovation, and circular economy principles, the framework illustrates how data are turned into actionable knowledge and collective benefits. This structure is then examined using empirical insights from the Fenicia ULL, serving as the testing ground to validate and refine the proposed model.

It highlights two urban experimentation initiatives: (1) the participatory Cyber-Physical System (CPS) for waste detection and management using a digital twin platform, and (2) the initial deployment of an IoT sensor-enabled urban garden as a testbed for circularity and community-driven environmental monitoring. Data sources included sensor readings, participatory observations, and georeferenced images collected through open-source platforms. Community participants and researchers jointly validated these datasets during the demonstration and evaluation phases, ensuring reliability and contextual accuracy. Both interventions are structured within the SSCC model architecture and address the three strategic objectives outlined in the innovation management model: Quality of Life**,** Productivity**,** and ICT-based Sustainability**.**

The results are analyzed using the Key Performance Indicators (KPIs) established in the strategic layer of the TIMM and validated through the operational mechanisms deployed in each experiment, as shown in Table 2. Additionally, each finding is interpreted in relation to the four value creation mechanisms identified in Chapter 3—Decision-Making Enhancement**,** Operational Efficiency**,** Product and Service Innovation**,** and Business Model Transformation—to evaluate the transformative potential of the data strategy implemented at the neighborhood scale.

The model clearly outlines a pathway for turning strategic goals into practical operational and governance results by activating the data strategy through real-world experimentation. Specifically, the integration of digital tools, participatory mechanisms, and analytical infrastructures enables the generation of actionable data, supports the validation of co-created solutions, and strengthens institutional learning processes.

Implementing the data strategy within the ULL demonstrates its potential as a catalyst for systemic change, fostering the development of more inclusive, efficient, and sustainable urban innovation processes.

Table 2. Data Strategy: Activated Value Mechanism(s) for Fenicia ULL experiments.

Experiment	Strategic Objective	Key Performance Indicator (KPI)	Empirical Evidence	Activated Value Mechanism(s)
Waste CPS	Quality of Life	Citizen engagement rate	Community workshops, Epicollect5 platform data submission.	Decision-Making Enhancement (DME), Business Model Transformation (BMT)

Waste CPS	Productivity	Operational efficiency of digital twin	Efficiency: 81.1% recall, 100% precision, <10 min latency.	Operational Efficiency (OE)
Waste CPS	ICT-based Sustainability	Use of non-proprietary tech, participatory sensing.	Open-source tools, community-driven data flow.	Product and Service Innovation (PSI), Business Model Transformation (BMT)
Urban Garden	Quality of Life	Community participation in environmental monitoring	IoT Sensor installation, community-led planting events.	Decision-Making Enhancement (DME), Product and Service Innovation (PSI)
Urban Garden	Productivity	Data usability for adaptive irrigation	Soil sensor readings integrated into digital log.	Operational Efficiency (OE)
Urban Garden	ICT-based Sustainability	Integration of low-cost IoT prototypes	Co-design with ORUS Technology, botanical support, FIWARE open-source integration.	Product and Service Innovation (PSI), Business Model Transformation (BMT)

6 Discussion and Contributions, and Conclusions

Implementing the TIMM in the Fenicia ULL provides empirical evidence that a structured, participatory, and data-driven approach can effectively translate strategic goals into measurable results across different urban contexts. The findings from the two experimental projects—waste management through a cyber-physical system (CPS) and digital twin, and the initial deployment of an IoT-integrated urban garden—demonstrate that value-creation mechanisms can be activated even in environments with technological and institutional limitations.

A key insight from the results is the reinforcing role of socio-technical integration: combining community-based data collection, AI-powered analysis, and digital visualization improved local decision-making and service responsiveness. Additionally, the use of open-source tools and participatory sensing demonstrates that sustainable and inclusive digital infrastructures can be created with limited financial resources—an especially relevant finding for cities in the Global South.

Beyond the empirical findings, this paper offers three significant contributions to academic and professional discussions on smart and sustainable cities. First, it affirms the theoretical framework linking data strategy with four mechanisms of urban value creation—Decision-Making Enhancement, Operational Efficiency, Product and Service Innovation, and Business Model Transformation—and demonstrates how these mechanisms can be operationalized through real-world experiments. Second, it confirms the dual-layer structure of the SSCC model (strategic–operational), which enables dynamic interaction between high-level goals and ground-level actions. This interaction creates feedback loops that support iterative learning and cross-sector collaboration.

Third, it provides a comparative view of corporate and urban data strategies, emphasizing both their similarities and differences in goals, governance, data sources, and value frameworks. While corporate strategies usually focus on internal efficiency and competitiveness, urban data strategies need to consider public value, systemic effects, and inclusive governance. Using enterprise data logic in urban settings helps reframe smart city projects through the perspective of institutional diversity and public interest.

Compared to data-driven smart-sustainable city literature, our results expand Bibri's instrumentation–datafication–computerization program by implementing value pathways and KPI-based evidence in a ULL setting [1–3,6]. Unlike Fang's data-driven city stack, we clarify where value is generated across acquisition–processing–application stages and connect these stages to decision and service outcomes [18]. Unlike guidance-oriented roadmaps, our framework incorporates participatory sensing and governance feedback loops as essential for circular urban value [17,34].

Overall, this study enhances the understanding of data strategy as a systemic, value-creating mechanism that functions through the interaction of technological, organizational, and social layers in sustainable and circular urban systems.

7 Conclusions

The experiments carried out in the Fenicia ULL demonstrate that a data strategy based on systemic thinking, community participation, and technological flexibility can promote sustainable urban transformation. The study shows that data strategies generate value in sustainable, smart, and circular cities by activating four interconnected mechanisms—improving decision-making, increasing operational efficiency, fostering product and service innovation, and transforming business models—through integrated technological and social experimentation.

The results confirm that ULLs serve as effective testing grounds for innovation and as platforms for integrating data strategies into broader governance frameworks. Future research should examine the long-term impacts and scalability of such strategies, as well as their integration with digital governance and policy ecosystems in various urban settings.

References

1. Bibri, S. (2019). The anatomy of the data-driven smart sustainable city: instrumentation, datafication, computerization, and related applications. *Journal of Big Data*, 6. https://doi.org/10.1186/s40537-019-0221-4.
2. Bibri, S. (2021). Data-driven smart sustainable cities of the future: urban computing and intelligence for strategic, short-term, and joined-up planning. *Computational Urban Science*, 1. https://doi.org/10.1007/s43762-021-00008-9.
3. Bibri, S. (2019). On the sustainability of smart and smarter cities in the era of big data: an interdisciplinary and transdisciplinary literature review. *Journal of Big Data*, 6. https://doi.org/10.1186/s40537-019-0182-7.

4. Bibri, S. (2021). The underlying components of data-driven smart sustainable cities of the future: a case study approach to an applied theoretical framework. *European Journal of Futures Research*, 9. https://doi.org/10.1186/s40309-021-00182-3.
5. Bibri, S., Huang, J., Jagatheesaperumal, S., & Krogstie, J. (2024). The synergistic interplay of artificial intelligence and digital twin in environmentally planning sustainable smart cities: A comprehensive systematic review. *Environmental Science and Ecotechnology*, 20. https://doi.org/10.1016/j.ese.2024.100433.
6. Bibri, S., & Krogstie, J. (2021). Data-driven smart sustainable cities of the future: An evidence synthesis approach to a comprehensive state-of-the-art literature review. *Sustainable Futures*. https://doi.org/10.1016/J.SFTR.2021.100047.
7. Munonye, W., & Ajonye, G. (2024). Redesigning Urban Infrastructure for Circularity: The Role of Smart Cities in Reducing Waste. *Journal of Environmental Science Studies*. https://doi.org/10.20849/jess.v7i2.1468.
8. Velasquez-Mendez, A.; Lozoya-Santos, J.d.J.; Jiménez-Vargas, J.F. (2025). Validating a Sustainable, Smart, and Circular City Architecture Through Urban Living Lab Experiments. *Urban Sci. 9*, 377. https://doi.org/10.3390/urbansci9090377.
9. Velasquez-Mendez, A.; Lozoya-Santos, J.; Jimenez, J.F. (2025). Technological Innovation Management Model for Urban Living Labs in Smart Cities: Strategies for Urban Problem-Solving. In Proceedings of the 23rd LACCEI International Multi-Conference for Engineering, Education, and Technology: "Engineering, Artificial Intelligence, and Sustainable Technologies in service of society", Mexico City, Mexico, 16–18 July 2025.
10. Velasquez-Mendez, A.; Lozoya-Santos, J.; Jimenez, J.F. (2024). Assessing Smart Cities in Developing Countries: Multidimensional Approaches. In Proceedings of the VII Ibero-American Congress of Smart Cities ICSC-CITIES 2024, San Carlos, Costa Rica, 12–14 November 2024.
11. Bokolo, A. (2023). Data driven approaches for smart city planning and design: a case scenario on urban data management. *Digital Policy, Regulation and Governance*. https://doi.org/10.1108/dprg-03-2022-0023.
12. Neira, T. Vesga, R. (2024). Factores que afectan la generación de valor en la estrategia de datos de una organización. RISTI. *Iberian Journal of Information Systems and Technologies.*
13. Neira, T. Vesga, R. (2024). Factores determinantes en la eficacia de las estrategias de datos: Evidencia del sector bancario colombiano. RISTI. *Iberian Journal of Information Systems and Technologies.*
14. Omrany, H., Mehdipour, A., Oteng, D. *et al.* The uptake of urban digital twins in the built environment: a pathway to resilient and sustainable cities. *Comput.Urban Sci.* **5**, 20 (2025). https://doi.org/10.1007/s43762-025-00177-x
15. Cardullo, P., Kitchin, R. Being a 'citizen' in the smart city: up and down the scaffold of smart citizen participation in Dublin, Ireland. *GeoJournal* 84, 1–13 (2019). https://doi.org/10.1007/s10708-018-9845-8
16. Lee, J., Babcock, J., Pham, T., Bui, T., & Kang, M. (2022). Smart city as a social transition towards inclusive development through technology: a tale of four smart cities. *International Journal of Urban Sciences*, 27, 75 - 100. https://doi.org/10.1080/12265934.2022.2074076.
17. Bieser, J. Charles, A. Hussein, H. Tong, K. (2023). Data for the City of Tomorrow: Developing the Capabilities and Capacity to Guide Better Urban Futures. World Economic Forum. Insight Report. access: https://www.weforum.org/publications/data-for-the-city-of-tomorrow-developing-the-capabilities-and-capacity-to-guide-better-urban-futures/

18. Fang, Y. Shan, Z. and Wang, W. (2021). Modeling and Key Technologies of a Data-Driven Smart City System. in *IEEE Access*, vol. 9, pp. 91244-91258, 2021, doi: 10.1109/ACCESS.2021.3091716.
19. Ylijoki, O. & Porras, J. (2018). A recipe for big data value creation. Business Process Management Journal. 25. 10.1108/BPMJ-03-2018-0082.
20. Costa, D., Bittencourt, J., Oliveira, F., Peixoto, J., & Jesus, T. (2024). Achieving Sustainable Smart Cities through Geospatial Data-Driven Approaches. *Sustainability*. https://doi.org/10.3390/su16020640.
21. Davenport, T. Barth, P. and Bean, R. (2012). "How 'big Data' Is Different." MIT Sloan Management Review 54(1).
22. Mazzei, M. Noble, D. (2017). Big data dreams: A framework for corporate strategy, Business Horizons, Volume 60, Issue 3, Pages 405-414, ISSN 0007-6813, https://doi.org/10.1016/j.bushor.2017.01.010.
23. Sharma, R. & Mithas, S. & Kankanhalli, A. (2014). Transforming decision-making processes: A research agenda for understanding the impact of business analytics on organisations. European Journal of Information Systems. 23. 433-441. 10.1057/ejis.2014.17.
24. Roden, S., Nucciarelli, A., Li, F., & Graham, G. (2017). Big data and the transformation of operations models: a framework and a new research agenda. *Production Planning & Control*, *28*(11–12), 929–944. https://doi.org/10.1080/09537287.2017.1336792
25. Duan, Y. Cao, G. Edwards, J. (2020). Understanding the impact of business analytics on innovation, European Journal of Operational Research, Volume 281, Issue 3, Pages 673-686, ISSN 0377-2217, https://doi.org/10.1016/j.ejor.2018.06.021.
26. Chen, Y. & Kreulen, J. & Campbell, M. & Abrams, C. (2011). Analytics Ecosystem Transformation: A Force for Business Model Innovation. Proceedings - 2011 Annual SRII Global Conference, SRII 2011. 11 - 20. 10.1109/SRII.2011.12.
27. Alharthi, A. Krotov, V. Bowman, M. (2017). Addressing barriers to big data, Business Horizons, Volume 60, Issue 3, Pages 285-292, ISSN 0007-6813, https://doi.org/10.1016/j.bushor.2017.01.002.
28. Merino, J. Caballero, I. Rivas, B. Serrano, M. and Piattini, M. (2016). "A Data Quality in Use Model for Big Data." Future Generation Computer Systems 63. doi: 10.1016/j.future.2015.11.024.
29. Comuzzi, M. and Patel, A. (2016). "How Organisations Leverage: Big Data: A Maturity Model." Industrial Management and Data Systems 116(8). doi: 10.1108/ IMDS-12-2015-0495
30. Sandberg, B. and Aarikka-Stenroos, L. (2014). "What Makes It so Difficult? A Systematic Review on Barriers to Radical Innovation." Industrial Marketing Management 43(8). doi: 10.1016/j.indmarman.2014.08.003.
31. Flores Arocupita, J. P. Manrique, M.A.L. Serna, G.J. and Aybar. I.E. (2021). "Liderazgo En Tiempos de 4ta Revolución Industrial." *Revista Venezolana de Gerencia* 26(96). doi: 10.52080/rvgluz.26.96.7.
32. Thomke, S. and Euchner. J. (2020). "High Velocity Business Experiments: An Interview with Stefan Thomke." *Research Technology Management* 63(4).
33. Walentek, D. (2021). Datafication Process in the Concept of Smart Cities. Energies 2021, 14, 4861. https://doi.org/10.3390/en14164861
34. Townsend, A., & Zambrano-Barragán, P. (2019). *Big Urban Data: A Strategic Guide for Cities*. https://doi.org/10.18235/0001918

Análisis comparativo de métodos predictivos para la completitud de datos en series de producción fotovoltaica: regresión lineal y Gradient Boosting

Marta Lafuente-Cacho[1], Oscar Izquierdo-Monge[1], Paula Peña-Carro[1], Ángel Hernández-Jiménez[1], Carlos Barrera del Amo[2]

[1] CEDER-CIEMAT, Autovía de Navarra A15 salida 56, 42290 Lubia (Soria), España
M.L.C.: marta.lafuente@ciemat.es (correspondence author); O.I.M.: oscar.izquierdo@ciemat.es; P.P.C.: paula.pena@ciemat.es; A.H.J.: angel.hernandez@ciemat.es

[2] Föra, Campus Universitario Duques de Soria, edificio I+D+i, 42004 Soria, España
C.B.D.A.: carlos.barrera@fora.es

Resumen: La calidad y completitud de los datos son fundamentales para el análisis y la predicción del rendimiento en sistemas fotovoltaicos. Sin embargo, las series temporales suelen presentar vacíos que limitan la validez de los estudios. Este trabajo aborda la reconstrucción de los registros de radiación solar y producción de la planta experimental PEPA II (2022–2024), combinando información procedente de estaciones meteorológicas cercanas y técnicas de modelización estadística y de aprendizaje automático. En una primera etapa, se aplicó un procedimiento iterativo de completitud de radiación solar basado en seis estaciones de la red Wunderground. Esta estrategia permitió reducir el número de valores ausentes de 6.349 a 130, logrando una cobertura del 99,9 % de la serie. En una segunda fase, se estimaron los huecos en la producción fotovoltaica comparando dos enfoques: un modelo de regresión lineal y un regresor Gradient Boosting. Los resultados mostraron que, aunque ambos modelos alcanzaron un ajuste satisfactorio, el Gradient Boosting ofreció un mejor desempeño, con un R^2 de 0,738 y un RMSE de 2.852,17, frente al R^2 de 0,713 y el RMSE de 2.982,58 de la regresión lineal. El enfoque propuesto demuestra la eficacia de combinar fuentes externas con técnicas de machine learning para garantizar series de datos completas y coherentes, sentando las bases para aplicaciones posteriores en monitorización, predicción y optimización de sistemas solares.

Palabras clave: Energía fotovoltaica, series temporales, datos faltantes, regresión lineal, Gradient Boosting, web scraping.

1. Introducción

La energía solar fotovoltaica se ha consolidado como una de las tecnologías renovables más prometedoras en el marco de la transición energética y la descarbonización del sistema eléctrico [1]. Su carácter limpio, modular y en constante reducción de costes la sitúa como un pilar estratégico para alcanzar los compromisos internacionales de reducción de emisiones [2]. No obstante, la fiabilidad y el rendimiento de las plantas fotovoltaicas dependen en gran medida de la calidad de los datos registrados, que incluyen series históricas de radiación solar, temperatura ambiente y producción eléctrica [3][4]. La disponibilidad de registros completos y consistentes resulta esencial no solo para el seguimiento operativo de las instalaciones, sino también para la calibración de modelos predictivos capaces de optimizar la gestión y la planificación de la generación renovable [5].

En la práctica, los sistemas de medición suelen verse afectados por fallos en los sensores, interrupciones en la transmisión de datos o errores en el almacenamiento digital [6][7]. Estas incidencias originan vacíos en las series temporales que, aun cuando representan un porcentaje reducido, pueden comprometer la fiabilidad de indicadores de eficiencia, la validación de modelos de predicción o la estimación de curvas de producción [8]. El tratamiento de estos valores ausentes constituye, por tanto, un desafío metodológico que requiere estrategias específicas de reconstrucción de series.

Las técnicas más sencillas incluyen la sustitución por valores medios, interpolaciones o modelos de regresión lineal [9]. Aunque estas aproximaciones son fáciles de implementar y de bajo coste computacional, presentan limitaciones cuando las relaciones entre variables son no lineales o dependen de factores contextuales complejos [10][11][12]. En contraste, los métodos de aprendizaje automático han mostrado mayor capacidad para capturar patrones no lineales y gestionar la incertidumbre asociada a datos incompletos [13][14]. Entre ellos destacan los algoritmos basados en árboles de decisión, como Gradient Boosting [15], ampliamente utilizados en predicción energética y en la reconstrucción de series debido a su robustez y flexibilidad [16][17].

Además, la incorporación de fuentes externas de información puede enriquecer las estrategias de reconstrucción [18]. En el caso de la radiación solar, los valores ausentes pueden ser completados mediante estaciones meteorológicas cercanas cuyos registros están disponibles en línea [19][20][21]. El uso de técnicas de web scraping permite integrar estas series externas de forma coherente con la variabilidad climática local, garantizando así una mayor consistencia de los datos.

En este trabajo se propone un procedimiento de reconstrucción en dos etapas: en primer lugar, la recuperación de huecos en la radiación solar mediante la integración de datos de estaciones meteorológicas cercanas; y en segundo lugar, la estimación de la producción fotovoltaica a través de la comparación entre un modelo estadístico clásico (regresión lineal) y un enfoque avanzado de aprendizaje automático (Gradient Boosting). Este contraste metodológico permite evaluar ventajas y limitaciones de cada

paradigma, al tiempo que proporciona una base más sólida para la gestión y explotación de instalaciones fotovoltaicas.

2. Metodología

El trabajo se estructuró en cuatro etapas principales: recopilación y descripción de los datos, análisis de la calidad de la serie temporal, reconstrucción de valores ausentes mediante un procedimiento en dos fases y, finalmente, evaluación cuantitativa de los modelos empleados para estimar la producción fotovoltaica a partir de la radiación solar.

2.1 Datos

La base de datos principal procede de la planta experimental PEPA II, gestionada por el CEDER (Centro de Desarrollo de Energías Renovables) situado en Lubia (Soria), e incluye registros de radiación solar incidente y de producción fotovoltaica con una frecuencia de muestreo de 15 minutos entre los años 2022 y 2024. Estos registros se mantuvieron en formato bruto, sin aplicar procesos adicionales de depuración, salvo la exclusión de aquellos valores negativos imposibles en radiación y potencia. A partir de julio de 2024, la planta operó con limitación de potencia, lo que explica que la producción registrada resultara inferior a la capacidad real de generación.

2.2 Análisis de huecos y valores nulos

La inspección preliminar de la serie permitió identificar la existencia de valores ausentes y nulos tanto en la radiación solar como en la producción fotovoltaica. Los vacíos se distribuyen de forma irregular a lo largo del periodo de estudio, con mayor incidencia en determinados meses debido a incidencias técnicas en la adquisición de datos. La detección de huecos se realizó mediante el recuento de entradas vacías y la visualización de discontinuidades temporales, lo que ofreció un diagnóstico inicial sobre la calidad de la base de datos y constituyó el punto de partida para diseñar la estrategia de completitud.

2.3 Estrategia de completitud en dos etapas

La reconstrucción de la serie se planteó en dos fases diferenciadas. En la primera, se abordaron los vacíos de radiación solar mediante la integración de información procedente de estaciones meteorológicas cercanas pertenecientes a la red Wunderground. Se seleccionaron aquellas estaciones situadas en un radio aproximado

de 15–20 km de la planta PEPA II, priorizando la proximidad geográfica, la representatividad climática y la disponibilidad de registros horarios de radiación. De acuerdo con estos criterios, se trabajó con las estaciones ICUBOD2, ILOSRB1, IALMAZ2, IQUINT63, IGOLMA6 e ISORIA35. Estos datos externos se obtuvieron mediante técnicas de *web scraping*, se sometieron a un proceso de limpieza y estandarización, y se alinearon temporalmente con la serie original aplicando un criterio de vecino más próximo con tolerancia de ±15 minutos. Este procedimiento permitió recuperar la mayor parte de los valores ausentes en la variable de radiación.

En la segunda fase se abordó la completitud de la producción fotovoltaica a partir de la radiación ya reconstruida. Para este fin se calibraron dos modelos representativos de paradigmas distintos. Por un lado, se utilizó la regresión lineal simple, elegida como referencia debido a su simplicidad, facilidad de implementación y capacidad de interpretación. Por otro, se aplicó el algoritmo Gradient Boosting, una técnica de aprendizaje automático basada en ensambles de árboles de decisión que ha demostrado un desempeño robusto en problemas de predicción energética y que resulta especialmente adecuada para capturar relaciones no lineales entre variables. La comparación de ambos enfoques permitió evaluar sus fortalezas y limitaciones en la reconstrucción de valores de producción eléctrica.

La validación de los modelos se llevó a cabo utilizando exclusivamente observaciones completas, es decir, registros sin vacíos simultáneos en radiación ni en producción. Como indicadores de desempeño se calcularon el error cuadrático medio (RMSE), que mide la magnitud media de los errores de predicción, y el coeficiente de determinación (R^2), que refleja la proporción de varianza explicada por el modelo. El análisis se implementó en Python (versión 3.10) empleando las librerías scikit-learn, pandas y numpy.

3. Resultados

En esta sección se presentan los principales resultados obtenidos tras la aplicación de la metodología propuesta. En primer lugar, se realiza un análisis de la calidad de los datos registrados y de las incidencias detectadas en las series temporales. A continuación, se evalúa el grado de completitud de los registros de radiación solar, para finalmente examinar la consistencia y cobertura de la producción fotovoltaica.

3.1 Calidad y análisis de los datos

El conjunto de datos correspondiente a la planta experimental PEPA II abarcó un total de 125.760 registros con frecuencia de 15 minutos para el periodo 2022–2024. En esta serie se identificaron valores ausentes tanto en la radiación solar como en la producción

fotovoltaica. En concreto, la radiación presentó 6.349 entradas nulas (5,05 % del total) (*Figura 1*), mientras que la producción registró 7.589 valores nulos (6,03 %) (*Figura 2*). Aunque estos porcentajes son moderados, resultan suficientes para comprometer la validez de análisis posteriores si no se adoptan estrategias de completitud.

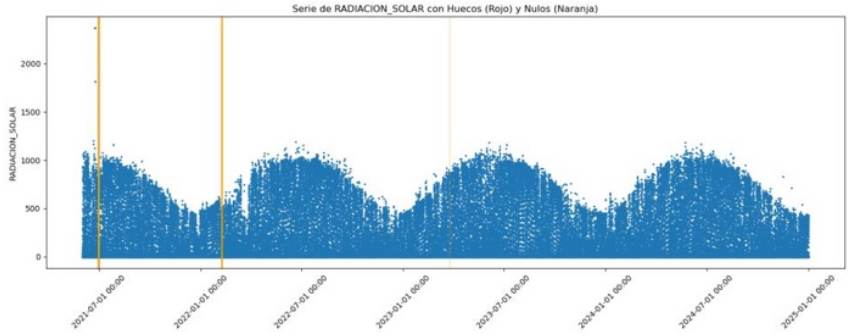


***Figura 1**: Serie temporal de radiación solar.*

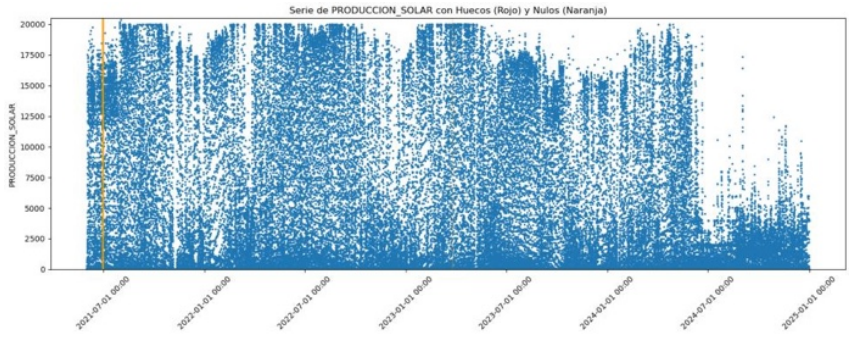


***Figura 2**: Serie temporal de producción solar.*

El estudio temporal de los nulos reveló que no se distribuyen de manera uniforme a lo largo del periodo. El mapa de calor (*Figura 3*) permite visualizar la localización de estos valores faltantes, destacando que no se trata únicamente de registros aislados, sino que aparecen también en bloques continuos de varias horas e incluso días. Este patrón sugiere la coexistencia de distintas causas: desde fallos puntuales en los sensores hasta interrupciones prolongadas en el sistema de adquisición.



***Figura 3**: Mapa de calor de datos faltantes.*

Un aspecto visible en el mapa de calor es que los vacíos no siempre se produjeron de forma simultánea en radiación y en producción, lo que apunta a fallos independientes en los dispositivos de medida. En contraste, la estructura temporal se mantuvo íntegra: todas las marcas horarias previstas en intervalos de 15 minutos estuvieron presentes, de modo que no existieron huecos en la secuencia temporal. Esto facilita la posterior reconstrucción, ya que los problemas se limitan a la ausencia de valores y no a discontinuidades en el calendario de muestreo.

Por último, la verificación de coherencia física detectó algunos valores negativos de radiación y tramos planos en la producción, probablemente vinculados a errores de calibración o bloqueos en la adquisición. Aunque se trata de casos minoritarios, refuerzan la necesidad de un preprocesamiento cuidadoso antes de la reconstrucción de la serie.

3.2 Completitud de radiación solar

La estrategia de completitud aplicada permitió reducir de manera sustancial el número de vacíos en la variable de radiación solar. Inicialmente, se identificaron 80 días con registros ausentes en la serie de la planta PEPA II.

Para abordar este problema, se recurrió a los datos horarios de seis estaciones meteorológicas de la red Wunderground consideradas cercanas, con condiciones climáticas representativas de la zona y con suficiente disponibilidad de registros (*Figura 4*). El procedimiento se desarrolló de forma iterativa, estableciendo una jerarquía de prioridad en función de la proximidad y la disponibilidad histórica de cada estación. Se comenzó con ICUBOD2, seguida de ILOSRB1, IALMAZ2, IQUINT63, IGOLMA6 e ISORIA35. En los casos en que una estación no ofrecía registros válidos en determinadas fechas, el proceso continuaba con la siguiente en la lista. De este modo, ICUBOD2, IALMAZ2 e IQUINT63 aportaron datos válidos que permitieron cubrir la mayoría de los vacíos, mientras que ILOSRB1 e IGOLMA6 no proporcionaron información utilizable.



***Figura 4**: Localización de estaciones meteorológicas.*

Para garantizar la coherencia, se verificó previamente la correlación entre los valores de radiación registrados por las estaciones y los datos locales en los periodos en los que la planta sí disponía de información. Esta comprobación confirmó la representatividad de las estaciones seleccionadas y su idoneidad para servir como fuente de completitud.

Como resultado, el número de días con vacíos se redujo de 80 a únicamente 4. En términos absolutos, de los 6.349 valores nulos iniciales en la variable de radiación solar, solo permanecieron 130 sin poder reconstruirse, lo que equivale a una cobertura del 99,9 % de la serie total. Estos huecos residuales corresponden a intervalos muy concretos en los que ninguna estación cercana disponía de datos válidos.

Si bien el procedimiento no permite capturar microvariaciones locales debidas a la nubosidad sobre el emplazamiento específico de la planta, la magnitud de los huecos remanentes fue mínima y no compromete los análisis posteriores. El conjunto final de radiación solar reconstruida se consolidó en un archivo completo, que constituye la base para la estimación de la producción fotovoltaica en la fase siguiente del análisis.

3.3 Completitud de producción fotovoltaica

Una vez consolidada la serie de radiación solar, el siguiente paso consistió en la reconstrucción de los valores ausentes de la producción fotovoltaica. Para este fin se implementaron y compararon dos enfoques complementarios: un modelo de regresión lineal (LR), ampliamente utilizado como referencia en estudios energéticos por su simplicidad y facilidad de interpretación, y un Gradient Boosting Regressor (GB), técnica de ensamble basada en árboles de decisión que ha demostrado gran eficacia en la captura de relaciones no lineales y en la mejora de la precisión predictiva en sistemas complejos.

Ambos modelos fueron entrenados utilizando únicamente los registros completos de radiación y producción, lo que permitió ajustar sus parámetros a la dinámica real observada en la planta. Posteriormente, se evaluaron mediante dos métricas estándar: el error cuadrático medio (RMSE), que cuantifica la magnitud promedio de los errores de predicción, y el coeficiente de determinación (R^2), que mide la proporción de la variabilidad de la producción explicada por la radiación.

La *Figura 5* sintetiza esta comparación.

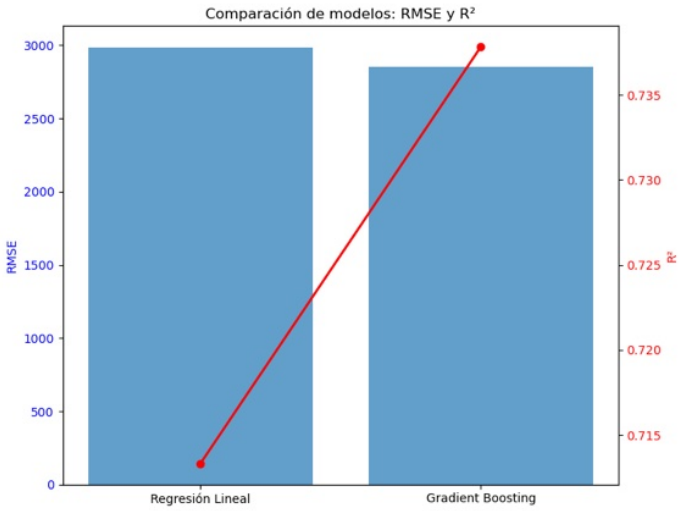


***Figura 5**: Comparación métrica de error.*

El modelo lineal obtuvo un RMSE de 2.982,58 y un R^2 de 0,713, lo que implica que explicó aproximadamente un 71 % de la variabilidad de la producción. El Gradient Boosting, en cambio, redujo el error a 2.852,17 y elevó el R^2 a 0,738, mejorando así en un 4,4 % la precisión y capturando 2,5 puntos porcentuales adicionales de la variabilidad. Si bien la diferencia numérica puede parecer moderada, desde el punto de vista aplicado resulta significativa, ya que un error menor y un ajuste más estrecho incrementan la fiabilidad de los datos reconstruidos y reducen la propagación de incertidumbre en análisis posteriores.

Más allá de las métricas globales, la *Figura 6* permite apreciar la naturaleza de la relación entre radiación y producción. En la dispersión de los datos reales, la regresión lineal se traduce en una tendencia recta que tiende a infraestimar la producción en condiciones de alta irradiancia, al no poder capturar la saturación que caracteriza a los sistemas fotovoltaicos cuando se aproximan a su potencia nominal. El Gradient Boosting, en cambio, traza una curva más flexible que reproduce con mayor realismo esta dinámica, adaptándose tanto a los niveles bajos como a los máximos de irradiancia.

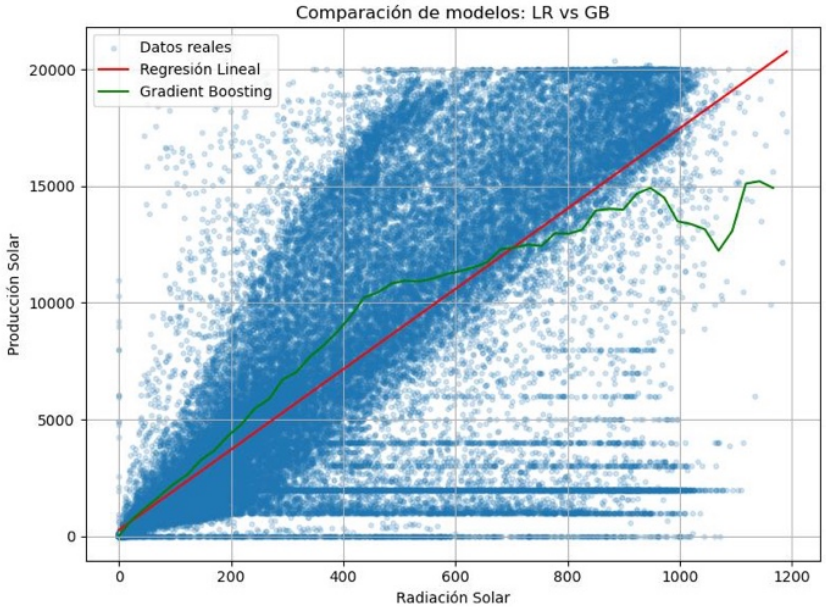


***Figura 6**: Dispersión radiación-producción.*

La ventaja del modelo no lineal se confirma también en la Figura 7, que muestra la distribución de errores de ambos enfoques. Mientras que los errores del modelo lineal presentan una dispersión mayor y un sesgo hacia valores negativos (indicando su

tendencia a infraestimar), los del Gradient Boosting se concentran más estrechamente en torno a cero, reflejando predicciones más equilibradas y consistentes a lo largo de todo el rango de radiación.

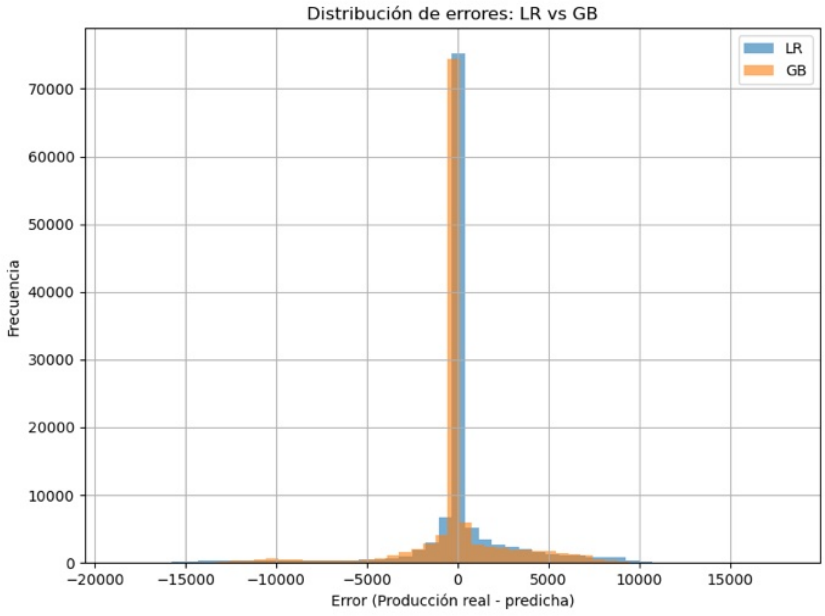


***Figura 7**: Distribución de errores.*

En conjunto, estos resultados justifican la adopción del Gradient Boosting como modelo de referencia para completar los registros faltantes de producción. Su capacidad de capturar relaciones no lineales y reducir la dispersión de errores lo convierte en la alternativa más adecuada para asegurar la calidad de la serie temporal.

4. Conclusiones

El presente trabajo ha demostrado la viabilidad de reconstruir series temporales incompletas de variables críticas para el análisis del rendimiento fotovoltaico mediante una combinación de fuentes externas y técnicas de modelización.

En primer lugar, la estrategia de completitud aplicada a la radiación solar permitió reducir drásticamente la proporción de datos ausentes: de 6.349 valores nulos iniciales se pasó a solo 130, alcanzando una cobertura del 99,9 % de la serie. El uso de estaciones

meteorológicas cercanas resultó determinante, confirmándose su representatividad gracias a la alta correlación encontrada con los registros originales de la planta. Esta aproximación, aunque no capta microvariaciones locales, garantiza una base suficientemente robusta para los análisis posteriores.

En segundo lugar, la comparación de dos enfoques de modelado para la estimación de la producción fotovoltaica —regresión lineal frente a Gradient Boosting— puso de manifiesto las ventajas de recurrir a técnicas más flexibles. Mientras que la regresión lineal alcanzó un R^2 de 0,713 y un RMSE de 2.982,58, el Gradient Boosting mejoró los resultados con un R^2 de 0,738 y un RMSE de 2.852,17, reflejando con mayor precisión la curvatura de la relación entre radiación y producción y reduciendo los sesgos en condiciones de alta irradiancia.

De esta manera, el modelo de Gradient Boosting se consolidó como la opción más adecuada para completar los vacíos de producción, garantizando al mismo tiempo coherencia física al imponer producción nula en ausencia de radiación. El producto final fue una serie completa y consistente, lista para ser utilizada en estudios de eficiencia, predicción y planificación energética.

En conjunto, los resultados refuerzan la importancia de aplicar metodologías híbridas que combinen datos medidos, fuentes externas y técnicas de aprendizaje automático para mejorar la calidad de la información en el sector energético. Este tipo de enfoques no solo aumenta la fiabilidad de los análisis históricos, sino que también abre la puerta a aplicaciones futuras, como la detección temprana de anomalías o la estimación de producción en emplazamientos sin monitorización directa.

5. Referencias

[1] Breyer, C., Bogdanov, D., Aghahosseini, A., Gulagi, A., Child, M., Oyewo, A. S., ... & Vainikka, P. (2018). Solar photovoltaics demand for the global energy transition in the power sector. *Progress in Photovoltaics: Research and Applications*, *26*(8), 505-523.

[2] Victoria, M., Haegel, N., Peters, I. M., Sinton, R., Jäger-Waldau, A., Del Cañizo, C., ... & Smets, A. (2021). Solar photovoltaics is ready to power a sustainable future. *Joule*, *5*(5), 1041-1056.

[3] Zhang, P., Li, W., Li, S., Wang, Y., & Xiao, W. (2013). Reliability assessment of photovoltaic power systems: Review of current status and future perspectives. *Applied energy*, *104*, 822-833.

[4] Sharma, V., & Chandel, S. S. (2013). Performance and degradation analysis for long term reliability of solar photovoltaic systems: A review. *Renewable and sustainable energy reviews*, *27*, 753-767.

[5] Ukoba, K., Olatunji, K. O., Adeoye, E., Jen, T. C., & Madyira, D. M. (2024). Optimizing renewable energy systems through artificial intelligence: Review and future prospects. *Energy & Environment*, *35*(7), 3833-3879.

[6] Mahmood, M. A., Seah, W. K., & Welch, I. (2015). Reliability in wireless sensor networks: A survey and challenges ahead. *Computer networks*, *79*, 166-187.

[7] Raghunath, K. K., & Rengarajan, N. (2013). Investigation of faults, errors and failures in wireless sensor network: A systematical survey. *International Journal of Advanced Computer Research*, *3*(3), 151.

[8] García, F. P., Pedregal, D. J., & Roberts, C. (2010). Time series methods applied to failure prediction and detection. *Reliability Engineering & System Safety*, *95*(6), 698-703.

[9] Lam, N. S. N. (1983). Spatial interpolation methods: a review. *The American Cartographer*, *10*(2), 129-150.

[10] Samanta, S., Pal, D. K., Lohar, D., & Pal, B. (2012). Interpolation of climate variables and temperature modeling. *Theoretical and Applied Climatology*, *107*(1), 35-45.

[11] Hocking, R. R. (2013). *Methods and applications of linear models: regression and the analysis of variance*. John Wiley & Sons.

[12] Kravtsov, S., Kondrashov, D., & Ghil, M. (2005). Multilevel regression modeling of nonlinear processes: Derivation and applications to climatic variability. *Journal of Climate*, *18*(21), 4404-4424.

[13] Solomatine, D. P., & Shrestha, D. L. (2009). A novel method to estimate model uncertainty using machine learning techniques. *Water Resources Research*, *45*(12).

[14] Cheng, S., Quilodrán-Casas, C., Ouala, S., Farchi, A., Liu, C., Tandeo, P., ... & Arcucci, R. (2023). Machine learning with data assimilation and uncertainty quantification for dynamical systems: a review. *IEEE/CAA Journal of Automatica Sinica*, *10*(6), 1361-1387.

[15] Singh, U., Rizwan, M., Alaraj, M., & Alsaidan, I. (2021). A machine learning-based gradient boosting regression approach for wind power production forecasting: A step towards smart grid environments. *Energies*, *14*(16), 5196.

[16] Charbuty, B., & Abdulazeez, A. (2021). Classification based on decision tree algorithm for machine learning. *Journal of applied science and technology trends*, *2*(01), 20-28.

[17] Hassan, M. A., Khalil, A., Kaseb, S., & Kassem, M. A. (2017). Exploring the potential of tree-based ensemble methods in solar radiation modeling. *Applied Energy*, *203*, 897-916.

[18] Chang, Y., Wilkinson, S., Brunsdon, D., Seville, E., & Potangaroa, R. (2011). An integrated approach: managing resources for post-disaster reconstruction. *Disasters*, *35*(4), 739-765.

[19] Demirhan, H., & Renwick, Z. (2018). Missing value imputation for short to mid-term horizontal solar irradiance data. *Applied Energy*, *225*, 998-1012.

[20] Turrado, C. C., López, M. D. C. M., Lasheras, F. S., Gómez, B. A. R., Rollé, J. L. C., & de Cos Juez, F. J. (2014). Missing data imputation of solar radiation data under different atmospheric conditions. *Sensors*, *14*(11), 20382-20399.

[21] Qing, X., & Niu, Y. (2018). Hourly day-ahead solar irradiance prediction using weather forecasts by LSTM. *Energy*, *148*, 461-468.

Artificial Neural Networks for Temperature Prediction in Urban Gardens Using Data from Grey Urban Spaces in Valladolid, Spain.

Francisco Tomatis[1][https://orcid.org/0000-0002-5075-2310], Francisco Javier Diez[1][https://orcid.org/0000-0002-1391-2688], and Luis Manuel Navas Gracia[1][https://orcid.org/0000-0002-7895-925X]

[1] TADRUS Research Group, Department of Agricultural and Forestry Engineering, University of Valladolid, 34004 Palencia, Spain

Abstract. Urban green spaces play a critical role in mitigating the impacts of climate change by reducing temperatures and improving thermal comfort in cities. However, continuous monitoring of temperature variations across diverse green infrastructures remains a challenge. This study explores the application of Artificial Neural Networks (ANNs) for estimating ambient temperatures in urban green areas such as urban allotment gardens and urban parks using data collected from grey urban spaces. Using a 6-Y-1 network architecture trained on data from Valladolid, Spain, we tested model performance across three different seasonal periods -41 days, 31 days for the training phase and 10 days for testing. Results indicate that ANNs can accurately predict temperature variations, with the best performance observed (RMSE = 0.44 °C) for the urban forest park during October to November 2022. While ANNs do not replace traditional monitoring methods, their ability to generate reliable estimates with limited input data highlights their potential as a low-cost, flexible support tool in urban climate analysis and the design of climate-resilient green infrastructure and smart cities.

Keywords: urban climate; urban microclimate; urban green spaces; urban gardens; temperature; artificial neural networks.

1 Introduction

Two of the most pressing challenges of the 21st century are climate change and the growing global population increasingly concentrated in urban areas —estimated to exceed 70% by 2050 [1,2]. These challenges are interconnected, as climate change and the rise in global temperature threaten human health, economic activities, and even urban ecosystems [3,4]. Moreover, rising temperatures caused by climate change are associated by complex urban structures that store and re-radiate large amounts of solar radiation, mainly due to construction materials [5], contributing to the Urban Heat Island (UHI) effect [3-6]. UHI can be defined as the development of higher temperatures in city centers compared to adjacent rural areas [6,7].

In this context, the role of urban green spaces in mitigating urban temperatures is increasingly recognized [8-11]. These spaces are considered Nature-Based Solutions (NbS) and are valued for the Ecosystem Services (ES) they provide, such as temperature regulation, nature conservation, water infiltration, nutrient cycling, and biodiversity support [8-10,12]. NbS offer integrated, multifunctional solutions to many urban challenges by leveraging natural processes [13]. Urban green spaces include parks, green corridors, urban forests, and even urban agriculture initiatives [14,15]. Urban agriculture offers similar ecosystem services to green spaces, with the added benefit of food production [16,17]. One definition of urban agriculture is "the production of food, flowers, fiber, feed, and herbs on land in urban and peri-urban areas" [12], encompassing initiatives such as urban gardens, community gardens, allotment gardens, family gardens, rooftop farms, among others [18].

Recognizing that urban areas are not sufficiently adapted to global warming [19], international recommendations call on cities to become sustainable, resilient, adaptive, healthy, smart, and edible [2,20]. Adaptation, in human systems, refers to “the process of adjusting to actual or expected climate and its effects, in order to moderate harm or exploit beneficial opportunities” [21]. In natural systems, it “is the process of adjustment to actual climate and its effects” [21]. These recommendations challenge urban planners, local governments, and researchers, who must contribute to achieving the desired urban transformation. In this regard, research on rising urban temperatures and UHI has increased significantly over recent decades [22,23], as have studies on the cooling potential of urban green spaces as a climate adaptation strategy [24-28]. In this context, urban green infrastructure is generally regarded as an effective strategy for urban cooling [29].

To study the cooling effects of green spaces, a variety of methods and techniques are used [28,29]. The choice of method depends on the research goals, scale, and the type of green infrastructure studied [30]. The main methodological approaches include observational or empirical methods, statistical modeling and simulation, or experimental techniques —applied either separately or in combination [30]. Traditionally, *in situ* observations are conducted using fixed or remote environmental stations, enabling temperature comparisons between different spaces [24,27,30-34]. Other methodologies include remote sensing techniques using satellite or aircraft sensors, modeling through atmospheric simulations, numerical calculators, or computational fluid dynamics models [33,35-37]. In this methodological context, some of the most important advances are concentrated in the field of modeling, where various artificial intelligence techniques, including Artificial Neural Networks (ANNs), have been applied [38,39]. ANNs are innovative tools because of their capacity to uncover hidden or non-obvious patterns in climate data that may not be directly extractable through traditional methods [38,40]. ANNs can be developed as predictive models for urban temperatures using inputs from various data sources. Studies have found that ANNs can effectively relate urban temperatures to green spaces, such as urban parks [41], urban gardens [42], and green roofs [43]. However, despite their potential, there is a lack of scientific evidence applying ANNs to predict urban temperatures with a specific focus on the influence of

green spaces [42]. Furthermore, there is a scarcity of studies that develop ANN models to generate data in places where such information is unavailable, enabling it to be used as input for other models [44].

In this study, we developed ANN models to estimate urban temperatures in urban allotment gardens and in an urban forested park using data collected from grey urban spaces in Valladolid, Spain. The objective of this research is to seek new and improved results that build upon previous similar study [42]. In both cases, the ANNs were trained over a period of 41 days, with 31 days for the training phase and 10 days for testing. However, what differentiates this research from the previous one [42] is the testing period and the input data used. Additionally, this research builds on prior lessons learned and focuses exclusively on a single ANN architecture (6-Y-1), deliberately avoiding architectures that had previously yielded poorer performance [42]. The results contribute to optimizing the development of ANNs as innovative tools for studying urban temperatures and support local climate governance efforts toward making cities more adaptive and smart [45].

2 Materials and Methods

2.1 Temperature Data Collection

The study was conducted in the city of Valladolid, Spain, located in the northwestern quadrant of the Iberian Peninsula [46]. Its climate is classified as Mediterranean-continental, with an average annual temperature of approximately 12°C and an annual precipitation of 400 mm [46]. Valladolid has implemented various initiatives aimed at advancing its urban transformation toward a more sustainable, smart, and renatured city [45,46, 47,48]. Within this framework, the local government manages four urban allotment gardens, primarily intended for unemployed residents [49]. Allotment gardens are defined as spaces where land is subdivided into individual plots, each cultivated by a household or individual [50]. The four allotment gardens in Valladolid are:

- *Jardín Botánico* (total area: 2,600 m²; 33 plots of 50 m² each)
- *Valle de Arán* (4,620 m²; 50 plots of 50 m², plus one 800 m² community plot)
- *Parque Alameda* (3,300 m²; 48 plots of 50 m², 300 m² community plot)
- *Los Santos-Pilarica* (5,200 m²; 50 plots of 50 m², 300 m² community plot)

Field studies conducted between 2022 and 2024 identified tomatoes (*Solanum spp*.), peppers (*Capsicum spp*.), and onions (*Allium spp*.) as the most common crops. Other crops observed included lettuce *(Lactuca spp*.), garlic (*Allium spp*.), cabbage (*Brassica spp.)*, strawberries (*Fragaria spp.)*, beans (*Phaseolus spp.*), raspberries (*Rubus spp.*), among others [24]. No significant differences were identified among the allotment gardens regarding the crop species [24].

To study temperatures across different areas of interest in Valladolid, nineteen environmental stations were installed *in situ* throughout the city. Twelve stations were placed

in the four allotment gardens (three *per* garden), one station was installed in a forested urban park called "*Campo Grande*"— with a surface area of 101,376 m² and the highest tree density in the city [51], and the remaining five stations were installed in urban grey spaces, named at *Puente Poniente*, *San Miguel*, *Catedral*, *Don Sancho*, and *Dos de Mayo* **(Figure 1)**. The data collection started on June 2022 and ended in May 2024.

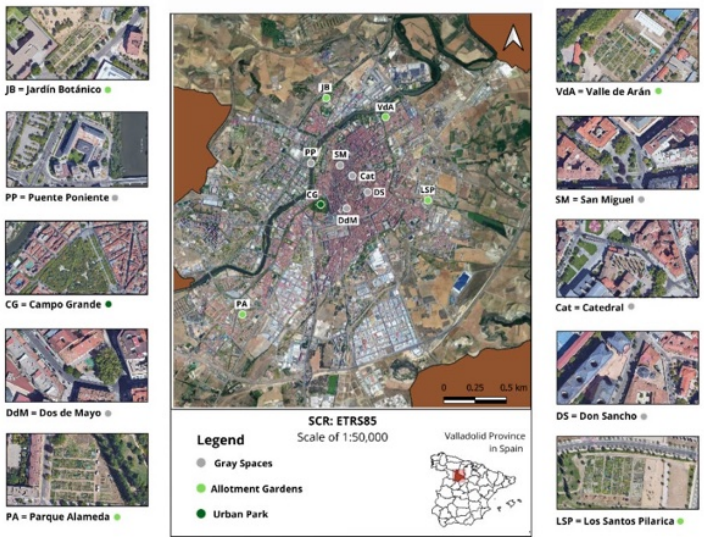


Fig. 1. Spatial orientation of the research sites in Valladolid, Spain. The locations of the four urban gardens are identified in light green, *Campo Grande* (park) in green, and the gray spaces in gray.

The environmental stations used were Onset HOBO UA-002-64 loggers (5.8 cm × 3.3 cm × 2.3 cm), with an operating range of −20 to 70 °C, and an accuracy of ±0.53 °C between 0 and 50 °C. Devices were placed 1.5 meters above the ground and protected with plastic shields (Figure 2). The data collected consisted of the hourly average temperature (°C), calculated from individual temperature measurements taken every hour.



Fig. 2. Environmental stations in *Jardín Botánico, Parque Alameda*, and *Campo Grande*.

2.2 ANNs Models

ANN were constructed using different combinations of inputs (X), hidden layers (Y), and outputs (Z), following the practical guideline of "keeping networks simple to obtain better results". Based on results obtained from similar previous studies [42], the selected ANN architecture was 6-Y-1 **(Figure 3)**.

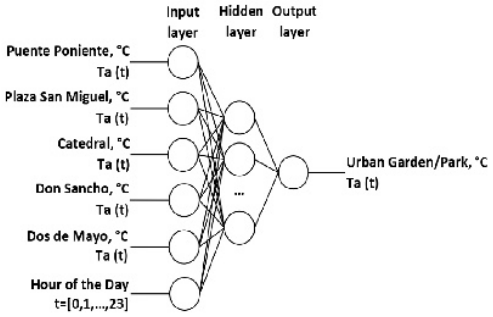


Fig. 2. 6-Y-1 ANN architecture.

The training period for each ANN spanned 41 days, with 31 days allocated for training and 10 days for testing. The input data (X) consisted of hourly temperature records and the hour of the day collected from the five environmental stations located in the urban grey spaces. The number of hidden layers (Y) varied between 2, 3, …, 10. The output layer (Z) have one single neuron for the temperature prediction from for the four urban allotment gardens (*Jardín Botánico, Valle de Arán, Parque Alameda,* and *Los Santos Pilarica*) and *Campo Grande*. A total of three Test Blocks were conducted using 6-Y-1 architecture, each differing in the dates used for training and testing phases:

- **Test Block 1:**
  Training from August 20, 2022 to September 20, 2022. Testing from September 21 to September 30, 2022.
- **Test Block 2:** Training from October 6, 2022 to November 6, 2022. Testing from November 7 to November 19, 2022 (excluding November 16th, 17th, and 18th).
- **Test Block 3:**
  Training from April 5, 2023 to May 11, 2023. Testing from May 12 to May 21, 2023.

This three-Test Block strategy, covering different seasonal periods, is crucial for assessing the model's robustness and generalization capacity. A single train-test experiment could lead to overfitting to the specific conditions of that period, limiting the validity of the findings.

The ANN configuration followed the principle of Occam's razor, which advocates simplicity in model design: the simplest solution is often the most likely, though not necessarily the correct one [52]. This heuristic is commonly applied in scientific modeling to avoid overfitting and enhance generalization. The ANN models were implemented using MATLAB's *feedforwardnet* function. The dimensionality of the input and output data vectors determined the respective layer sizes, resulting in a Multilayer Perceptron (MLP) with a single hidden layer. The activation function used between neurons in the hidden layer was the hyperbolic tangent sigmoid (tansig), while the transfer function for the output layer neurons was linear (purelin). The Levenberg–Marquardt backpropagation algorithm (BP-LM) was applied for fast optimization (trainlm) [52,53]. Although other types of ANNs may improve interpolation performance, the current study focused on identifying the most suitable (X-Y-Z) architecture with minimal data requirements. The models were trained using the *train* function in MATLAB, based on input–output data matrices covering 41 days. After training, model performance was evaluated using the *sim* function.

2.3 ANNs Evaluation

The accuracy of the ANNs in the validation stage was achieved by analyzing statistics values, which include: the Root Mean Square Error (RMSE, temperature °C) using Equation (1) and the Coefficient of Determination (R^2), as an indicator of the level of model fit, using Equation (2).

$$RMSE\ (°C) = \sqrt{\frac{\Sigma_{i=1}^{n}(\ Y_i - \hat{Y}_i\ )^2}{n}} \qquad \text{Equation (1)}$$

$$R^2 = 1 - \frac{\Sigma_{i=1}^{n}\ (\ Y_i - \hat{Y}_i\ )^2}{\Sigma_{i=1}^{n}\ (\ Y_i - \bar{Y}\ )^2} \qquad \text{Equation (2)}$$

3 Results

This section presents the results obtained using the 6-Y-1 ANN architecture for predicting daily ambient temperature and estimating hourly temperatures at four urban allotment gardens (*Jardín Botánico, Valle de Arán, Parque Alameda,* and *Los Santos Pilaric*a) and at *Campo Grande*. The results are according to each Test Block.

- For **Test Block 1 (Table 1)**, the best performance was achieved using the 6-3-1 architecture, with an output for *Valle de Arán* achieving an RMSE of 0.93 °C **(Figure 3)**. The worst result in this Test Block was obtained with the 6-2-1 architecture for *Campo Grande*, with an RMSE of 1.87 °C.

- For **Test Block 2 (Table 2)**, the best result was obtained with the 6-8-1 architecture for *Campo Grande*, with an RMSE of 0.44 °C **(Figure 4)**. The worst performance was recorded for the 6-5-1 architecture, also for *Campo Grande*, with an RMSE of 1.20 °C.
- In **Test Block 3 (Table 3)**, the most accurate prediction was provided by the 6-7-1 architecture for *Parque Alameda*, with an RMSE of 1.30 °C **(Figure 5)**. The lowest performance in this block was seen with the 6-5-1 architecture for *Jardín Botánico*, which produced an RMSE of 1.55 °C.

Table 1. Test Block 1. Prediction models for average hourly temperatures (°C) in urban gardens and forested urban park, according to 6-Y-1ANN architecture (X-Y-Z), changing the number of neurons in the hidden layers Y = (2, 3, …,10). The inputs are X = [Ta(t) *Puente Poniente*, Ta(t) *Plaza San Miguel*, Ta(t) *Catedral*, Ta(t) *Don Sancho*, Ta(t) *Dos de Mayo*, Hour of day (t = 0, 1, …, 23)], and the output is Z = [Ta(t) *Jardín Botánico*], Z = [Ta(t) *Valle de Arán*], Z = [Ta(t) *Los Santos-Pilarica*], Z = [Ta(t) *Parque Alameda*], and Z = [Ta(t) Campo Grande] in each case.

Outputs for Campo Grande									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.87	1.49	1.24	1.27	1.28	1.11	1.55	1.39	1.58
R^2	0.856	0.909	0.937	0.934	0.932	0.950	0.901	0.920	0.898
Outputs for Jardín Botánico									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.17	1.22	1.13	1.08	1.12	1.08	1.37	1.11	1.27
R^2	0.960	0.956	0.962	0.966	0.963	0.966	0.944	0.964	0.952
Outputs for Los Santos Pilarica									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	0.99	1.21	1.44	1.56	1.20	1.28	1.44	1.66	1.52
R^2	0.970	0.955	0.936	0.926	0.956	0.950	0.937	0.916	0.929
Outputs for Parque Alameda									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.21	1.11	1.03	1.05	1.03	0.99	1.31	1.12	1.15
R^2	0.954	0.961	0.967	0.966	0.967	0.969	0.946	0.960	0.959
Outputs for Valle de Arán									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.11	0.93	0.94	0.96	0.99	1.15	1.13	1.16	1.27
R^2	0.961	0.974	0.972	0.971	0.969	0.958	0.960	0.957	0.949

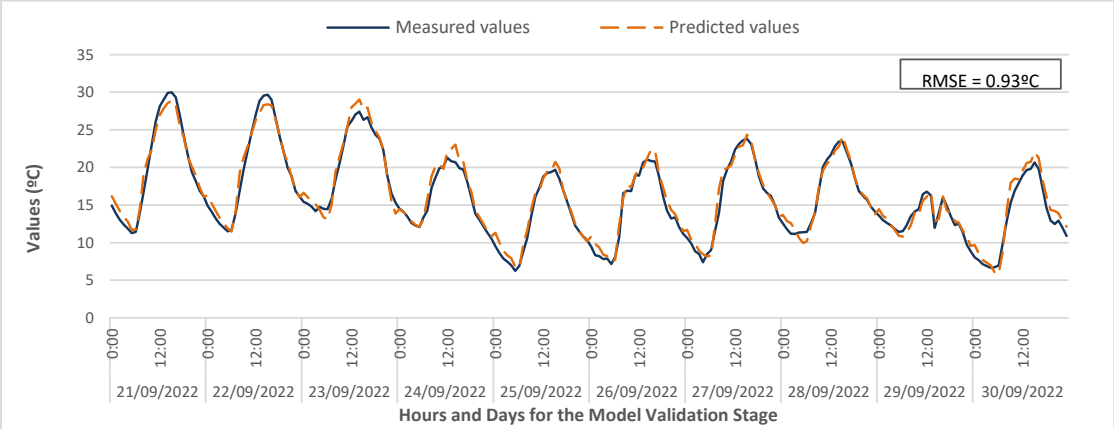


Figure 3. Test Block 1. Measured data and predictions in *Valle de Arán* for the validation stage, obtained with the 6-3-1 ANN architecture.

Table 2. Test Block 2. Prediction models for average hourly temperatures (°C) in urban gardens and forested urban park, according to 6-Y-1ANN architecture (X-Y-Z), changing the number of neurons in the hidden layers Y = (2, 3, …,10). The inputs are X = [Ta(t) *Puente Poniente*, Ta(t) *Plaza San Miguel*, Ta(t) *Catedral*, Ta(t) *Don Sancho*, Ta(t) *Dos de Mayo*, Hour of day (t = 0, 1, …, 23)], and the output is Z = [Ta(t) *Jardín Botánico*], Z = [Ta(t) *Valle de Arán*], Z = [Ta(t) *Los Santos-Pilarica*], Z = [Ta(t) *Parque Alameda*], and Z = [Ta(t) Campo Grande] in each case.

Outputs for Campo Grande									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	0.63	0.52	0.53	0.53	0.49	0.46	0.44	0.50	0.50
R^2	0.957	0.971	0.970	0.970	0.975	0.977	0.980	0.973	0.974
Outputs for Jardín Botánico									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.02	1.11	1.11	1.20	1.03	1.13	0.97	1.01	0.98
R^2	0.928	0.915	0.915	0.901	0.928	0.913	0.935	0.930	0.933
Outputs for Los Santos Pilarica									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	0.73	0.79	0.75	0.82	0.81	0.73	0.82	0.73	0.69
R^2	0.963	0.956	0.961	0.953	0.955	0.962	0.953	0.963	0.967
Outputs for Parque Alameda									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	0.77	0.79	0.77	0.75	0.78	0.71	0.78	0.73	0.71
R^2	0.956	0.953	0.955	0.958	0.954	0.962	0.955	0.960	0.962
Outputs for Valle de Arán									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	0.84	0.84	0.83	0.84	0.78	0.82	0.81	0.78	0.81
R^2	0.951	0.950	0.951	0.950	0.957	0.952	0.953	0.957	0.953

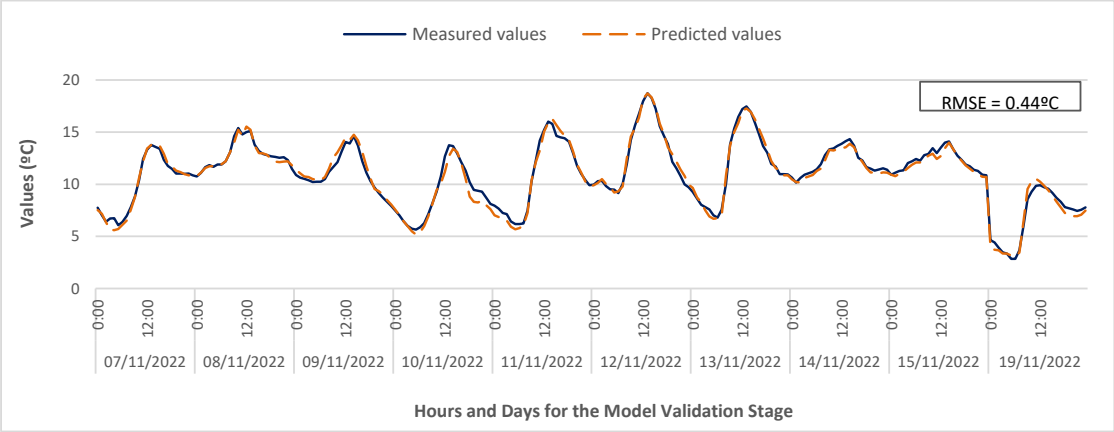


Figure 4. Test Block 2. Measured data and predictions in *Campo Grande* for the validation stage, obtained with the 6-8-1 ANN architecture.

Table 3. Test Block 3. Prediction models for average hourly temperatures (°C) in urban gardens and forested urban park, according to 6-Y-1 ANN architecture (X-Y-Z), changing the number of neurons in the hidden layers Y = (2, 3, …,10). The inputs are X = [Ta(t) *Puente Poniente*, Ta(t) *Plaza San Miguel*, Ta(t) *Catedral*, Ta(t) *Don Sancho*, Ta(t) *Dos de Mayo*, Hour of day (t = 0, 1, …, 23)], and the output is Z = [Ta(t) *Jardín Botánico*], Z = [Ta(t) *Valle de Arán*], Z = [Ta(t) *Los Santos-Pilarica*], Z = [Ta(t) *Parque Alameda*], and Z = [Ta(t) Campo Grande] in each case.

Outputs for Campo Grande									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.39	1.31	1.33	1.33	1.33	1.35	1.34	1.38	1.33
R^2	0.878	0.891	0.888	0.888	0.888	0.884	0.887	0.880	0.887
Outputs for Jardín Botánico									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.47	1.51	1.55	1.55	1.48	1.51	1.50	1.53	1.49
R^2	0.905	0.900	0.895	0.894	0.904	0.900	0.901	0.897	0.902
Outputs for Los Santos Pilarica									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.49	1.46	1.54	1.54	1.49	1.42	1.49	1.48	1.51
R^2	0.898	0.902	0.891	0.891	0.898	0.907	0.898	0.900	0.895
Outputs for Parque Alameda									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.43	1.39	1.38	1.38	1.33	1.30	1.35	1.38	1.34
R^2	0.907	0.913	0.914	0.915	0.921	0.924	0.917	0.914	0.918
Outputs for Valle de Arán									
ANN	**6-2-1**	**6-3-1**	**6-4-1**	**6-5-1**	**6-6-1**	**6-7-1**	**6-8-1**	**6-9-1**	**6-10-1**
RMSE (ºC)	1.42	1.38	1.37	1.37	1.40	1.44	1.42	1.42	1.41
R^2	0.911	0.916	0.917	0.917	0.913	0.908	0.911	0.911	0.912

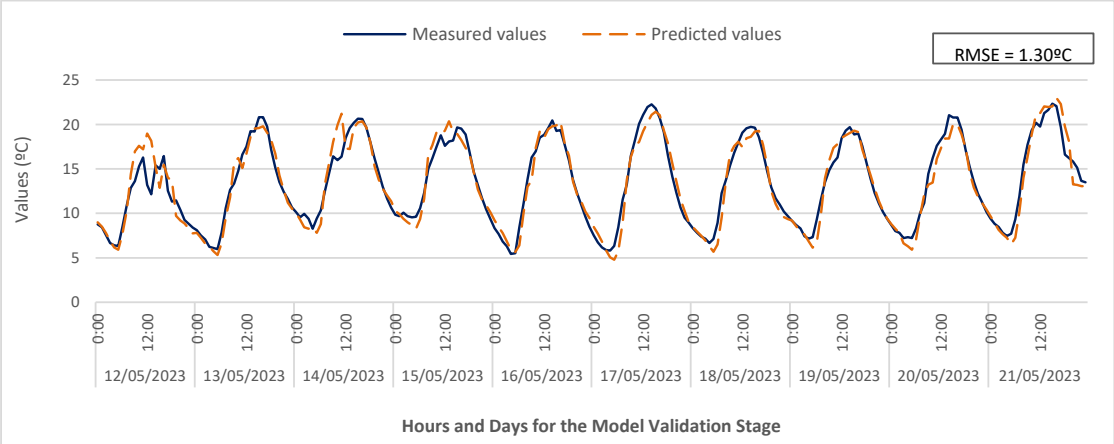


Figure 5. Test Block 3. Measured data and predictions in *Parque Alameda* for the validation stage, obtained with the 6-7-1 ANN architecture.

Comparing the best results obtained across each test block, we found that Test Block 2 (RMSE = 0.44 °C) produced more accurate temperature predictions and estimations than Test Block 1 (RMSE = 0.93 °C) and Test Block 3 (RMSE = 1.30 °C). The differences between the best-performing results were 0.49 °C and 0.86 °C, respectively. Furthermore, while Test Block 2 yielded its best result for *Campo Grande* (a forested urban park), the best results for Test Block 1 and Test Block 3 were associated with urban allotment gardens, specifically *Valle de Arán* and *Parque Alameda*.

4 Discussion

This research is a complementary extension of a previous study titled "Prediction of Daily Ambient Temperature and Its Hourly Estimation Using Artificial Neural Networks in Urban Allotment Gardens and an Urban Park in Valladolid, Castilla y León, Spain" [42]. In that publication, we developed and tested multiple ANN architectures in search of the best performance, following the practical rule of *"making networks simple to obtain better results"*. The architectures tested in the previous study included: 7-Y-5 (Y = 6, 7, …, 14), 6-Y-5 (Y = 6, 7, …, 14), 7-Y-1 (Y = 2, 3, …, 8), 6-Y-1 (Y = 2, 3, …, 8), 4-Y-1 (Y = 1, 2, …, 7), 3-Y-1 (Y = 1, 2, …, 7), and 2-Y-1 (Y = 2, 3, …, 8) all using a 41-day dataset [42]. Based on the conclusions of Tomatis F., et al, 2023[42], the present research focused on the best-performing architecture (6-Y-1) and applied it again over a 41-day period (31 days for training and 10 days for testing), but with different test dates. While the previous study used data from June–July 2022 [42], this research tested the same architecture across three different time periods: Test Block 1 (August–September 2022), Test Block 2 (October–November 2022), and Test Block 3 (April–May 2023). Comparing the results of both studies, the best performance in the current research was achieved in Test Block 2, with an RMSE of 0.44 °C for *Campo*

Grande. In contrast, the previous study reported a slightly better result, with an RMSE of 0.42 °C for *Valle de Arán* [42].

Moreover, it is worth highlighting that in the present study, the best result was obtained for the forested urban park *(Campo Grande)*, whereas in the previous study, the best performance was achieved for an urban allotment garden *(Valle de Arán)* [42]. This encourages us to continue exploring the development and application of ANNs across different types of urban green spaces. In particular, further research is needed to better understand the relationship between ANN performance and the specific vegetation characteristics present in the urban green areas under study.

It is acknowledged that the three isolated test blocks, while spanning different seasons (summer, autumn, and spring), do not cover a full year of seasonal variability and therefore do not provide a comprehensive assessment for long-term prediction (e.g., beyond 10 days). However, the primary objective of this study was to evaluate the performance of the artificial neural network (ANN) in short-term estimation (10-day test phase) and its adaptability to different short-term thermal dynamics, rather than to create a robust year-round predictive system, especially considering data availability.

Under the climate change scenario, the increasing frequency of extreme heat events means that historical data quickly become less representative. Therefore, maintaining the ANN's predictive accuracy requires constant recalibration and periodic retraining with the most recent data. This iterative process ensures the model incorporates the observable warming trend and remains a reliable support tool for climate adaptation strategies in cities.

Although the present study did not surpass the best result obtained in the previous work [42], it contributes significantly to the evaluation, development, and optimization of ANN integration for studying and predicting urban temperatures, especially in connection with urban green spaces. This is particularly relevant, as ANN represent a novel methodological approach that can complement more traditional techniques commonly used in this field, such as fixed or mobile *in situ* sensors, satellite imagery analysis, thermal cameras, or drone-based studies [24,27,54-56].

5 Limitations

While the results of this study are promising, several limitations must be acknowledged. First, the models were trained and evaluated using data from a single city (Valladolid, Spain), which may restrict the generalizability of the findings to other urban contexts with different climatic, geographic, or socio-ecological conditions. Second, although the selected ANN architecture (6-Y-1) demonstrated good performance, only one architecture was tested, leaving room for exploring more complex or adaptive models. Third, the analysis focused exclusively ambient temperature and did not incorporate

other potentially relevant environmental variables such as humidity, solar radiation, wind speed, or vegetation indices, which could enhance predictive capabilities. Regarding vegetation analysis, more specific data from both urban gardens and *Campo Grande* are needed to better correlate environmental variables and improve the interpretation of their influence on temperature and ANN model performance. Additionally, although *in situ* temperature data were collected over two years, only three Test Blocks were tested, suggesting that multiple additional tests covering different periods within the full two-year span could be conducted. Specifically, the lack of full annual coverage means that the model's performance for periods longer than the 10-day testing phase, particularly during extreme seasonal transitions (e.g., peak winter or summer), remains untested. Future research should address these limitations by expanding the spatial and temporal scope of data and integrating additional environmental and vegetative parameters.

6 Conclusions

This study demonstrates the usefulness of ANNs as a complementary methodological tool for predicting and estimating ambient temperatures in urban green spaces. Using the 6-Y-1 ANN architecture over three different testing periods, the best performance was observed in Test Block 2 (October–November 2022), with an RMSE of 0.44 °C for the forested urban park *Campo Grande* —slightly below the best result obtained in a previous study (RMSE = 0.42 °C for the urban allotment garden *Valle de Arán*). These findings underscore the importance of continuing to develop and test ANN models across different seasons and types of green infrastructure. Although ANNs may not yet replace traditional temperature monitoring methods (e.g., *in situ* sensors, satellite imagery, thermal cameras, or drones), their flexibility, low data requirements, and predictive capacity make them a valuable addition to urban climate research. Overall, this work contributes to the growing body of knowledge supporting data-driven approaches for designing more sustainable, resilient, smart, and climate-adaptive cities.

Acknowledgments. Authors would like to express our sincere gratitude for the support provided by the Ibero-American Program of Science and Technology for Development (CYTED) through the TRAPECIO Network. Francisco Tomatis has been financed under the call for University of Valladolid 2020 predoctoral contracts, co-financed by Banco Santander.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. United Nations: World Urbanization Prospects. In: Demographic Research, Vol. 12 (2019). Available online at: https://population.un.org/wup/
2. UN-Habitat: Envisaging the Future of Cities. World Cities Report 2022 (2022). Available online at: https://unhabitat.org/sites/default/files/2022/06/wcr_2022.pdf

3. Ickovics, J.R., Astbury, K., Campbell, M. et al.: Indicadores de The Lancet Countdown on Health and Climate Change: Perspectivas y experiencia de líderes municipales de 118 ciudades. Journal of Urban Health (2025). https://doi.org/10.1007/s11524-024-00952-x
4. IPCC: Resumen para responsables de políticas. In: Masson-Delmotte, V. et al. (eds.) Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the IPCC. Cambridge University Press (2021)
5. Rizwan, A.M., Dennis, L.Y., Liu, C.: A review on the generation, determination and mitigation of Urban Heat Island. Journal of Environmental Sciences 20(1), 120–128 (2008) https://doi.org/10.1016/S1001-0742(08)60019-4
6. Oke, T.R.: The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society 108(455), 1–24 (1982)
7. Fernández García, F., Montávez Gómez, J.P., González Rouco, J.F., Valero Rodríguez, F.: Relación entre la estructura espacial de la isla térmica y la morfología urbana de Madrid. (2004) http://hdl.handle.net/20.500.11765/9070
8. Dubbeling, M.: Integrating urban agriculture and forestry into climate change action plans. (2015) https://cdkn.org/sites/default/files/files/SriLanka-Argentina-Background-Paper.pdf
9. United Nations Human Settlements Programme: Planning Sustainable Cities: Global Report on Human Settlements 2009. Earthscan, London (2009) https://doi.org/10.1080/17535069.2010.481379
10. Heidt, V., Neef, M.: Benefits of Urban Green Space for Improving Urban Climate. In: Carreiro, M.M., Song, Y.C., Wu, J. (eds.) Ecology, Planning, and Management of Urban Forests. International Perspectives, pp. 84–96. Springer, New York (2008) https://doi.org/10.1007/978-0-387-71425-7_6
11. Kabisch, N., Korn, H., Stadler, J., Bonn, A. (eds.): Nature-Based Solutions to Climate Change Adaptation in Urban Areas. Linkages Between Science, Policy and Practice. Springer, Cham (2017)
12. Kingsley, J., Egerer, M., Nuttman, S., Keniger, L., Pettitt, P., Frantzeskaki, N., et al.: Urban agriculture as a nature-based solution to address socio-ecological challenges in Australian cities. Urban Forestry & Urban Greening 60, 127059 (2021) https://doi.org/10.1016/j.ufug.2021.127059
13. Dumitru, A., Wendling, L.: Evaluating the Impact of Nature-Based Solutions: A Handbook for Practitioners. European Commission, Luxembourg (2021). https://data.europa.eu/doi/10.2777/244577
14. Bowler, D.E., Buyung-Ali, L., Knight, T.M., Pullin, A.S.: Urban greening to cool towns and cities: A systematic review of the empirical evidence. Landscape and Urban Planning 97, 147–155 (2010) https://doi.org/10.1016/j.landurbplan.2010.05.006
15. Cao, S.Y., Yin, W.D., Su, J.Y., et al.: Spatial and Temporal Evolution of Multi-scale Green Space Environments and Urban Heat Islands: A Case Study of Beijing Sub-center. Sensors and Materials 35, 589–606 (2023) https://doi.org/10.18494/SAM4189
16. Lucertini, G., Di Giustino, G.: Urban and peri-urban agriculture as a tool for food security and climate change mitigation and adaptation: The case of Mestre. Sustainability 13, 5999 (2021) https://doi.org/10.3390/su13115999
17. Jabareen, Y.: Planning the resilient city: Concepts and strategies for coping with climate change and environmental risk. Cities 31, 220–229 (2013) https://doi.org/10.1016/j.cities.2012.05.004
18. Tomatis, F., Egerer, M., Correa-Guimaraes, A., Navas-Gracia, L. M.: Urban Gardening in a Changing Climate: A Review of Effects, Responses and Adaptation Capacities for Cities. Agriculture 13, 502 (2023) https://doi.org/10.3390/agriculture13020502

19. Mancebo, F.: Gardening the city: Addressing sustainability and adapting to global warming through urban agriculture. Environments 5, 38 (2018) https://doi.org/10.3390/environments5030038
20. UN-Habitat: Urbanization and Development–Emerging Futures. World Cities Report 2016. Available online: https://unhabitat.org/sites/default/files/download-manager-files/WCR-2016-WEB.pdf
21. IPCC: Contribution of Working Group II to the Sixth Assessment Report of the IPCC. In: Pörtner, H.-O., Roberts, D., Tignor, M., Poloczanska, E., Mintenbeck, K., et al. (eds.) Climate Change 2022: Impacts, Adaptation and Vulnerability, p. 3056. Cambridge University Press, Cambridge/New York (2022)
22. Deilami, K., Kamruzzaman, M., Liu, Y.: Urban heat island effect: A systematic review of spatio-temporal factors, data, methods, and mitigation measures. Int. J. Appl. Earth Obs. Geoinf. 67, 30–42 (2018) https://doi.org/10.1016/j.jag.2017.12.009
23. Almeida, C. R. D., Teodoro, A. C., Gonçalves, A.: Study of the urban heat island (UHI) using remote sensing techniques: A systematic review. Environments 8, 105 (2021) https://doi.org/10.3390/environments8100105
24. Tomatis, F., Egerer, M., Navas-Gracia, L. M.: Evaluating Food Production, Cooling Potential, and Gardener Perspectives in Urban Allotment Gardens of Valladolid, Spain. Horticulturae 10, 1232 (2024) https://doi.org/10.3390/horticulturae10111232
25. Mancebo, F.: Gardening the city: Addressing sustainability and adapting to global warming through urban agriculture. Environments 5, 38 (2018) https://doi.org/10.3390/environments5030038
26. Lwasa, S., Mugagga, F., Wahab, B., Simon, D., Connors, J., Griffith, C.: Urban and peri-urban agriculture and forestry: Transcending poverty alleviation to climate change mitigation and adaptation. Urban Climate 7, 92–106 (2014) https://doi.org/10.1016/j.uclim.2013.10.007
27. Arzberger, S., Egerer, M., Suda, M., Annighöfer, P.: Thermal regulation potential of urban green spaces in a changing climate: winter insights. Urban Forestry & Urban Greening, 128488 (2024) https://doi.org/10.1016/j.ufug.2024.128488
28. Hami, A., Abdi, B., Zarehaghi, D., Maulan, S. B.: Assessing the thermal comfort effects of green spaces: a systematic review of methods, parameters, and plant attributes. Sustainable Cities and Society 49, 101634 (2019) https://doi.org/10.1016/j.scs.2019.101634
29. Saaroni, H., Amorim, J. H., Hiemstra, J. A., Pearlmutter, D.: Urban Green Infrastructure as a tool for urban heat mitigation: Survey across climatic regions. Urban Climate 24, 94–110 (2018) https://doi.org/10.1016/j.uclim.2018.02.001
30. Koc, C. B., Osmond, P., Peters, A.: Evaluating the cooling effects of green infrastructure (2018) https://doi.org/10.1016/j.solener.2018.03.008
31. Egerer, M., Lin, B. B., Kendal, D.: Temperature variability differs in urban agroecosystems across two metropolitan regions. Climate 7, 1–18 (2019) https://doi.org/10.3390/cli7040050
32. Skoulika, F., Santamouris, M., Kolokotsa, D., Boemi, N.: On the thermal characteristics and the mitigation potential of a medium-size urban park in Athens, Greece. Landscape Urban Plan. 123, 73–86 (2014) https://doi.org/10.1016/j.landurbplan.2013.11.002
33. Feyisa, G. L., Dons, K., Meilby, H.: Efficiency of parks in mitigating urban heat island effect: Addis Ababa example. Landscape Urban Plan. 123, 87–95 (2014) https://doi.org/10.1016/j.landurbplan.2013.12.008
34. Georgi, J. N., Dimitriou, D.: The contribution of urban green spaces to environmental improvement: Case study of Chania, Greece. Build. Environ. 45, 1401–1414 (2010) https://doi.org/10.1016/j.buildenv.2009.12.003

35. Gupta, A., De, B.: Enhancing the city-level thermal environment through strategic utilization of urban green spaces employing geospatial techniques. Intl. J. Biometeorol., 1–19 (2024) https://doi.org/10.1007/s00484-024-02733-2
36. Pritipadmaja, Garg, R. D., Sharma, A. K.: Assessing the cooling effect of blue-green spaces: implications for UHI mitigation. Water 15, 2983 (2023) https://doi.org/10.3390/w15162983
37. Wavrek, M. T., Carr, E., Jean-Philippe, S., McKinney, M. L.: Drone remote sensing in urban forest management: A case study. Urban Forestry & Urban Greening 86, 127978 (2023) https://doi.org/10.1016/j.ufug.2023.127978
38. Zhang, Z.: Multivariate Time Series Analysis in Climate and Environmental Research. Springer, Cham (2018) https://doi.org/10.1007/978-3-319-67340-0
39. Blocken, B.: Computational Fluid Dynamics for urban physics: Importance, scales, possibilities, limitations and ten tips and tricks towards accurate and reliable simulations. Build. Environ. 91, 219–245 (2015) https://doi.org/10.1016/j.buildenv.2015.02.015
40. Diez, F. J., Correa-Guimaraes, A., Chico-Santamarta, L., Martínez-Rodríguez, A., Murcia-Velasco, D. A., Andara, R., Navas-Gracia, L. M.: Prediction of Daily Ambient Temperature and Its Hourly Estimation Using Artificial Neural Networks in an Agrometeorological Station in Castile and León, Spain. Sensors 22, 4850 (2022) https://doi.org/10.3390/s22134850
41. Chan, S. Y., Chau, C. K.: Development of artificial neural network models for predicting thermal comfort evaluation in urban parks in summer and winter. Build. Environ. 164, 106364 (2019) https://doi.org/10.1016/j.buildenv.2019.106364
42. Tomatis, F., Diez, F. J., Wilhelm, M. S., Navas-Gracia, L. M.: Prediction of Daily Ambient Temperature and Its Hourly Estimation Using Artificial Neural Networks in Urban Allotment Gardens and an Urban Park in Valladolid, Castilla y León, Spain. Agronomy 14, 60 (2023) https://doi.org/10.3390/agronomy14010060
43. Anahita, A., Arefi, H., Fathipoor, H.: Simulation of green roofs and their potential mitigating effects on the urban heat island using an artificial neural network: A case study in Austin, Texas. Adv. Space Res. 66, 1846–1862 (2020) https://doi.org/10.1016/j.asr.2020.06.039
44. Franco, B. M., Hernández-Callejo, L., Navas-Gracia, L. M.: Virtual weather stations for meteorological data estimations. Neural Comput. Appl. 32, 12801–12812 (2020) https://doi.org/10.1007/s00521-020-04727-8
45. Ayuntamiento de Valladolid: SmartVA! Strategic Plan for Digital Transformation and Applied Innovation in the City of Valladolid—Innovation and Smart City Plan (In Spanish). Available online: https://www.ideva.es/sites/default/files/Innovation_Plan.pdf (accessed on 16 December 2023)
46. GEOCyL & CARTIF: Adaptation Strategy to Climate Change of the Municipality of Valladolid. Spain-Portugal Cross-Border Cooperation Program POCTEP 2014–2020. (2020)
47. CARTIF: URBAN GreenUP. New strategy for Renaturing Cities through Nature Based Solutions. EU Horizon 2020 project 730426. (In Spanish) (2017)
48. Ayuntamiento de Valladolid: Agenda Urbana de Valladolid 2030. (2021). Available online: https://auva2030.es/wp-content/uploads/2022/07/agenda-urbana-valladolid-2030-3.pdf (accessed on 4 January 2024)
49. Ayuntamiento de Valladolid: Bases para la convocatoria de huertos ecológicos 2024-2025 para desempleados y huertos ecológicos de explotación comunitaria. (2024) Available online: https://www.valladolid.es/es/temas/hacemos/convocatoria-huertos-ecologicos-2024-2025-desempleados-huer.ficheros/ (accessed on 11 February 2024)
50. Burchard-Dziubińska, M.: The role of urban gardening in building city resilience to climate change. Ekonomia i Środowisko – Economics and Environment 3(3), 32–43 (2021) 10.34659/2021/3/19

51. Serrano Stampa, J.: Análisis y Cuantificación del Carbono Almacenado en los Parques y Jardines de la Ciudad de Valladolid. Master's thesis, Universidad de Valladolid (2016) http://uvadoc.uva.es/handle/10324/18785
52. Patz, J. A., Campbell-Lendrum, D., Holloway, T., Foley, J. A.: Impact of regional climate change on human health. Nature 438, 310–317 (2005) https://doi.org/10.1038/nature04188
53. AghaKouchak, A., Cheng, L., Mazdiyasni, O., Farahmand, A.: Global warming and changes in risk of concurrent climate extremes: Insights from the 2014 California drought. Geophys. Res. Lett. 41, 8847–8852 (2014) https://doi.org/10.1002/2014GL062308
54. Egerer, M. H., Wagner, B., Lin, B. B., Kendal, D., Zhu, K.: New methods of spatial analysis in urban gardens inform future vegetation surveying. Landscape Ecology 35, 761–778 (2020) https://doi.org/10.1007/s10980-020-00974-1
55. Arabi Aliabad, F., Ghafarian Malamiri, H., Sarsangi, A., Sekertekin, A., Ghaderpour, E.: Identifying and monitoring gardens in urban areas using aerial and satellite imagery. Remote Sensing 15, 4053 (2023) https://doi.org/10.3390/rs15164053
56. Jato-Espino, D., Tomatis, F., Forestieri, G., Peña, M.: Street-level surface and air temperatures in the Urban Center of Málaga, Spain. Urban Climate 61, 102490 (2025) https://doi.org/10.1016/j.uclim.2025.102490

Multi-objective vehicle routing problem with time-windows and service restrictions

Luis Rementería[1] [0009-0009-1619-4405], Andrei Tchernykh[1] [0000-0001-5029-5212], Sergio Nesmachnow[2] [0000-0002-8146-4012]

[1] CICESE Research Center, Ensenada, México, [2] Universidad de la República, Montevideo, Uruguay.

rementeria@cicese.edu.mx, chernykh@cicese.mx, sergion@fing.edu.uy

Abstract. The increasing population in recent decades has brought with it various problems related to the needs of citizens in large cities. There are several cities in the world with high population densities, which causes a number of problems, such as scarcity of natural resources, environmental pollution, health problems, and traffic congestion, among others. An efficient response to this is the concept of the smart city, which seeks to reduce the ecological footprint, boost the economy, and improve the quality of life by integrating information and communication technologies. Service Routing Problems are optimization problems within the context of logistics, tourism, and operations, looking to determine one or more routes for a group of vehicles to visit a given set of points of interest under certain operational constraints. This work proposes a novel variant of this problem with three objective functions: maximum route time, maximum waiting time, and total user satisfaction, and parameters that include not only the service time windows and service duration, but also the number of simultaneous services that can be provided at a point of interest. We propose a two-step evaluation approach. On the first step, we focus on calibration of algorithms and obtaining the reference solutions for small problems. The second step is to solve large-scale problems with a more complex graph where no exact algorithms are feasible, by multi-objective metaheuristics.

Keywords: service routing problem, multi-objective optimization, time window, simultaneous services.

1 Introduction

Vehicle Routing Problems (VRPs) were first addressed in 1959 by Dantzig & Ramser [8]. These are combinatorial optimization problems that aim to design a set of routes for a fleet of vehicles that depart from one or more depots and must meet customer demands under certain operational constraints. In 1964, Clarke and Wright developed a heuristic method to solve a variant of a VRP [5], and in 1981, Lenstra and Kan [15] demonstrated that VRPs are NP-hard problems. They are considered a generalization

of the well-known Traveling Salesman Problem (TSP), since the TSP consists of routing a vehicle to visit all customers once without considering operational constraints.

Formally and very generally, a VRP addresses the following approach:

Given an undirected graph $G = (V, E)$ *such that* $V = \{0,1,2,\dots,n\}$*, with the node* 0 *representing a distribution center, a homogeneous set K of vehicles, and a cost* c_{ij} *associated with each edge* $(i,j) \in E$*, a solution to the optimization problem is a collection of* K *sequences of vertices, where each sequence represents a tour for one of the vehicles, the union of the collection of sequences includes all vertices, and also minimizes the objective function:*

$$\sum_{k \in K} \left(\sum_{(i,j) \in E} c_{ij}^{(k)} \cdot x_{ij}^{(k)} \right)$$

With $x_{ij}^{(k)} \in \{0,1\}$ *as a decision variable that is equal to* 1 *if route k passes through edge* (i,j) *and is equal to* 0 *in the opposite case.*

This approach can solve problems related to courier or personnel transport companies. Other factors that can be taken into account within the model are the load capacity of a vehicle, time-windows in which a node can be visited, the use of multiple distribution centers, battery recharging in the case of hybrid or electric vehicles, and even the use of policies such as FIFO (First In, First Out), LIFO (Last In, First Out), RR (Round Robin), SJF (Shortest Job First), among others.

We can also find many variants of VRP's in the literature, but the importance of each of these variants lies in how well they reflect the problems that have to be faced in real life. In this article, we consider a variant we call Service Routing Problem (SRP), that is based on a tourism environment with the following characteristics:

- Time windows are set for visiting points of interest (POIs). (for instance, open hours for museums and restaurants).
- A POI can only be visited by a limited number of groups at a time, either due to the limit of space, tourist guides, parking slots, other operational limitations, or government regulations of the POIs. A waiting time can be generated when a group attend to a full POI.
- When visiting a POI, service time, for instance, the time limits of guided tours, can be set.
- Each route time must be less than a certain time limit that represents the maximum labor hours for drivers, or a group time budget.
- Every route begin and end on the same warehouse.
- Each group can visit every POI at most once.

It is not difficult to imagine a problem with these characteristics also in logistics and supply chains, with a variety of optimization criteria. In our case, we start from the minimization of maximum route time, maximum service waiting time, and maximization of the total user satisfaction.

Waiting times are generated when a service is not provided just after the group arrives at a POI. User satisfaction is associated with their preferences regarding the order in which they visit POIs. This complex formulation poses a significant methodological challenge, making the search for the Pareto Front (PF) unsolvable for the large-scale instances observed in operational reality. The need to solve complex and extensive problems without compromising the quality of solutions confirms the relevance of this study.

Considering the challenge of this SRP, we seek to answer: What is the most promising methodological approach for establishing a benchmarking framework for solving large-scale instances relevant to the tourism and logistics sectors? This article proposes a two-step methodology for the analysis of metaheuristics using the exact solutions as a reference. The following sections will outline the problem formulation and the proposed experimental design.

2 Background

There are many practical problems where SRP's are applicable, such as goods distribution, supply chains, public and/or private transportation, and even other less direct problems, such as irrigation systems, electrical grid distribution, and data traffic, among others. This means that the diversity in the taxonomy can be quite broad, where each of the possible variants of a SRP can find a real-life problem to which it can be applied. In recent years, it has been home healthcare (HHC) that has attracted the most attention, although it is not the only one that has been the subject of study. One of the advantages of taking HHC problems as a subject of study is the great diversity they offer. For example, Kandakoglu [11] developed a mathematical model for the purpose of creating routes for nurses who need to visit patients at their homes and perform dialysis treatments. This model creates an itinerary with two objectives: to minimize the distances and travel times for nurses, but with the condition of visiting all patients. This model supports distance and travel time between nodes, time windows, service times, and rest times. In addition, its optimization objectives are focused on minimizing travel time, waiting times, and route operating costs. Additionally, Euchi et al. use this model and adapt it to consider synchronized visits. That is, each patient must be visited by two nurses at the same time, regardless of how the nurse combinations are made [9].

In recent literature, it is most common to find mathematical models that focus on optimizing three objective functions together with different constraints that depend on the problem at hand. To provide a clear reference on the state of the art, we show relevant articles that are closely related to the scope of this problem, comparing them with the most important characteristics proposed by our research, as seen in *Table 1*.

From this review, we find that only four (4) articles address techniques based on MOCell, while ten (10) deal with other types of evolutionary algorithms (NSGA-II being the most widely implemented among them), four (4) deal with MIP (Mixed-Integer Programming), and a few others with NS (Neighborhood Search), SA (Simulated Annealing), and even exact algorithms such as Dijkstra. On the other hand, we can note that none of them have all the features that our model addresses, although the work by

Peña [27] comes closest. To complement this, Masmoudi's work addresses the constraints for simultaneous visits by applying genetic algorithms [20]. Adding service times to the model is not complicated, as it is a constant value that must be added to the route time every time one of the POIs is visited.

#	Methods	Benchmark	Route time	Route cost	Waiting time	Service time	Time-window	User satisfaction	Simultaneous visits	Pollution	Multiple Depot	Capacity	Heterogeneous Fleet
[1]	NS			✓	✓		✓				✓	✓	✓
[2]	SA, GA		✓	✓			✓						
[3]	-		✓	✓	✓		✓					✓	
[6]	MOEA/D						✓					✓	✓
[9]	AI, ACO		✓	✓	✓		✓		✓				
[11]	MIP		✓	✓	✓		✓						
[13]	MOEA	✓		✓			✓						
[14]	MOEA/D		✓	✓	✓		✓						
[16]	NS, TS, GA						✓		✓				
[17]	NSGA-III	✓		✓			✓			✓			
[18]	MIP, ACO			✓			✓		✓	✓			
[19]	MIP, NS						✓		✓	✓			✓
[20]	SA, NS, GA	✓	✓				✓		✓				
[22]	Dijkstra		✓	✓		✓	✓						
[24]	MOCell, GA		✓	✓								✓	
[25]	MCFEA						✓					✓	
[26]	MOCell		✓	✓			✓					✓	✓
[27]	MOCell	✓	✓	✓	✓		✓			✓		✓	✓
[28]	MS	✓	✓				✓		✓				
[30]	MOCell, GA		✓	✓						✓			
[33]	MIP		✓				✓						
[34]	GA		✓				✓						
[35]	-		✓	✓			✓					✓	
Proposal		✓	✓		✓	✓	✓	✓	✓				

Table 1: Summary of the characteristics in state-of-the-art models and our model proposal.

Our proposal model a problem that optimizes three objective functions: maximum route time, maximum waiting time, and total user satisfaction. It also considers restrictions on time windows, simultaneous visits, and unique visits per vehicle. So far, it is already more complex than the models described in the most relevant articles (Table 1), but the definition of the problem is also different in two additional ways:

- All routes include all POIs only once, which means that each POI is visited K times (once per route). Commonly, VRPs distribute visits to POIs across different routes, i.e., each POI is visited only once.

- The simultaneous visits described in the literature are homogeneous, i.e., all POIs receive the same number of simultaneous visits, while we propose that this value can be configured independently for each POI.

3 Case studies

This optimization problem fits into three case studies: two correspond to the tourism sector and one to warehouse operations. The first case study focuses on Guadalupe Valley: Mexico's main wine-producing region. It is located within the municipality of Ensenada, Baja California, and is famous for its vineyards and growing gastronomic offerings. Over the years, it has positioned itself as one of the most important tourist destinations in the state, offering experiences such as tastings, winery tours, and large harvest festivals. According to reports from the Baja California State Tourism Secretariat (SECTURE), the well-known Guadalupe Valley Wine Route received more than 600,000 visitors in 2017, with a distribution of origin as shown in *Table 2* [31].

Origin	Distribution
Baja California	44.88%
California (USA)	27.20%
Others	27.92%

Table 2: Distribution of visitors to Guadalupe Valley in 2017 [31].

Currently, the rural area of Guadalupe Valley belongs to the Participatory Local Ecological Zoning Program of the Municipality of Ensenada (POELPE), which is an initiative organized by SEMARNAT that seeks to protect and defend the territory in order to preserve the environment and make sustainable use of natural resources, meaning that road modernization is not a priority. For the tourism sector, this complicates access for both local and foreign visitors, since public transportation is scarce and limited, leaving tourists with only two options: travelling in their own vehicle or through a travel agency. According to the records on the official SECTURE website, there are currently 42 active tourism service providers in the state listed in the National Tourism Registry (RNT), not to mention informal or foreign providers. In addition, ProVino, a civil association in the state, reports on its official website that there are at least 84 member wineries that can be considered as places to visit in the area, so assigning routes for travel agencies is not a trivial problem to solve, especially during the harvest season.

The goal of the tour operators is to take a tourist group to the most renowned vineyards in the region. All tour operators have the same objective, so they have to visit all or most important vineyards. In addition, each vineyard has its own characteristics. The most important ones are the service hours (the operational hours of the establishment), the number of tour guides (how many tourist groups can be served simultaneously), and the tour duration (the time a group spends on a tour of the facilities). Therefore, tour operators must organize themselves so that a vineyard does not become overcrowded, making some groups wait, and causing annoyance among tourists.

Another case study is in Mexico City, known for being one of the cities with many museums. According to the National Council for Culture and Arts (CONACULTA) of

Mexico [https://www.cultura.gob.mx], the city currently has 190 museums. Each of these museums has a certain number of tour guides available, who take a certain amount of time to give a tour to a group of people. If at any given time there are more groups than guides, the last groups will have to wait to be served. If each group aims to visit, say, the 10 most important museums in the city, then transportation routes can be designed for each vehicle so that everyone can visit the most important museums while minimizing the waiting time for groups to be attended to by a tour guide and, in turn, minimizing the average time of the routes. This case study can be approached as an optimization problem, where each vehicle transporting a group is considered part of a fleet, and the number of tour guides at each museum represents the number of simultaneous visits allowed at the nodes, so that a penalty time can be considered for groups waiting to be served. Thus, the model proposed in this work is adapted to represent and solve this museum problem.

We could also consider a set of warehouses that can only receive a certain number of suppliers at a time. This operational limitation may be due to the number of available bays in a warehouse or the available staff to unload from the supplier. Each supplier takes a certain amount of time to unload all the goods they must deliver at that warehouse, which can be considered as service time, and if a supplier arrives at a busy warehouse, then this supplier will have to wait from the moment they arrive until the warehouse is free. Thus, the problem boils down to suppliers needing to supply all warehouses in the shortest possible time under these assumptions.

4 Formal definition of the problem

It is not trivial to adapt Euchi's model [9] to this problem, but it is very useful as a starting point. The most challenging adjustment is to change the restriction on synchronized visits that need to be configured independently at each node, so that each point can receive as many simultaneous visits as necessary. Everything else is less difficult.

Let us assume that there are n POIs (nodes) that must be visited by K tourist groups, which travel in K vehicles. Each vehicle seeks to visit all the nodes, but each POI can only receive visits within a certain time frame and only a limited number of visits at a time, so any vehicle that arrives at a busy node must wait until it becomes available. We'll refer to this condition as simultaneous visits. Each time a node is visited, a service time or waiting time (as appropriate) is added. Finally, vehicles can travel down a street (edge) several times without visiting a node. We named this problem the SRP (Service Routing Problem) and propose the following mathematical model to solve it:

Let $G(V, E)$ *be a connected graph representing the underlying topology of a route network, where* $V = \{0,1,2, \dots, n\}$ *is the set of intersections in the network (vertices), with* $P = \{p_1, p_2, \dots, p_Z\}$ *as the POIs to visit,* 0 *as the point where every vehicle starts and finishes their routes, and* $E = \{(i, j) | i, j \in V, i \neq j\}$ *denotes the set of connections (edges) between in* G.

4.1 Solution

A solution to this problem is a set of k routes, defined by r^k for each $k \in \{1, \dots, K\}$, where r^k is a sequence of POIs to visit (in order), and $l(r^k)$ is the length of r^k.

4.2 Decision variables

- $X_{ij}^k \in \{0,1\}$: indicates with 1 if the shortest path from i to j is used and takes 0 otherwise.
- $Y_i^k \in \{0,1\}$: total of times that the vehicle k visits the POI i (given by the route r^k).
- $A_i^k \geq 0$: time of arrival of the vehicle k to the POI i.

4.3 Connections

For each of the edges $(i,j) \in E$, a set of non-negative values is associated with it, $e(i,j) = \{d_{ij}, v_{ij}^{max}\}$:

- d_{ij}: distance between intersections i ,j.
- v_{ij}^{max}: maximum speed allowed on the connection (i,j).

4.4 Points of interest

Each POI $p \in P$ is defined by $p = \{tw_p^{open}, tw_p^{close}, st_p, s_p^{max}\}$, where:

- tw_p^{open}: start time of the operating hours (time window).
- tw_p^{close}: end time of the operating hours (time window).
- $st_p \leq tw_p^{close} - tw_p^{open}$: duration of services at the node p.
- s_p^{max}: maximum number of services that can be handled at the same time at p.

4.5 Travel time

Vehicles can travel at different speeds, so the time it takes each vehicle to travel along a connection in the graph is given by $t_{ij}^k = d_{ij}/\min\{v^k, v_{ij}^{max}\} \geq 0$. From this:

- $T^k = \sum_{i,j \in V} t_{ij}^k \cdot X_{ij}^k$ is the travel time of the vehicle k on the route r^k.
- $T^{\max} = \max_{k \leq K}\{T^k\}$ is the maximum travel time of the K vehicles.

4.6 Satisfaction

It is considered that each group has a preference for visiting POIs on their routes. Let be $score_p^k \in \{0,1\}$ a parameter that indicates if the group associated with the vehicle k wants to visit the pth POI. Then:

- $S^k = \sum_{p \in P}(Y_p^k \cdot score_p^k)$ is the satisfaction of the vehicle k.

- $S = \sum_{k \in K} S^k$ is the total satisfaction.

4.7 Waiting time

If a vehicle k arrives at a POI $p \in P$ and is not available to provide a service, then the vehicle must comply with a waiting time $w_p^k \geq 0$. It can happen because either the vehicle k arrives at the point p before tw_p^{open}, or the POI p has reached the maximum number of services s_p^{max}:

- $W^k = \sum_{p \in P} w_p^k$ is the total waiting time of the vehicle k.
- $\mathrm{W}^{\max} = \max_{\mathrm{k} \leq \mathrm{K}} \{W^k\}$ is the maximum waiting time for the K vehicles.

4.8 Objective functions

A solution to the SRP optimizes the following set of functions:

$$\{T^{max},\ W^{max},\ S\}$$

Where:

- T^{max} is the maximum route time.
- W^{max} is the maximum waiting time.
- S is the total satisfaction.

4.9 Restrictions

- Flow coherence: $\sum_{i \in V \setminus \{p\}} X_{ip}^k = \sum_{j \in V \setminus \{p\}} X_{pj}^k,\ p \in P$
- Unique visit: $Y_p^k = \sum_{j \in V \setminus \{p\}} X_{pj}^k,\ p \in P$
- Depot restriction: $r^k[1] = 0 = r^k[l(r^k)]$
- Arrival sequence: $A_i^k + w_i^k + st_i + t_{ij} \leq A_j^k + M \cdot \left(1 - X_{ij}^k\right),\ \forall i, j \in P$
- Opening coherence: $tw_j^{open} \cdot \sum_{i \in V} X_{ij}^k \leq A_j^k + w_j^k,\ \forall j \in P$
- Closing coherence: $A_j^k + w_j^k + st_j \leq tw_j^{close} + M \cdot \left(1 - \sum_{i \in V} X_{ij}^k\right),\ \forall j \in P$
- Waiting time restriction: $w_p^k \geq tw_p^{open} - A_p^k \geq 0$
- Route time limit: $T^k \leq T^{lim}$

Where
- T^{lim}: limit of time a route can last due to the work schedule.
- M: auxiliary parameter large enough to maintain linearity in the constraints.

4.10 Simultaneous visits

An additional decision variable, $start_p^k \geq \max\{A_p^k, tw_p^{open}\}$ is considered, where $start_p^k$ indicates the moment when the vehicle k starts a service in the POI p:

$$start_p^k \geq \min\{t' | t' > t, OCC(p, t') < s_p^{max}\}$$

Where $OCC(p, t')$ is a function that returns the total of vehicles visiting the POI p in the window $[t', t' + st_p]$.

5 Methodology

Since SVRs are NP-hard problems, finding an optimal solution within the solution space often generates a considerably high computational cost, both in terms of resources and time. To reduce computational costs, strategies and techniques are often used to find approximate solutions. These include:

- **Approximation algorithms:** Primal-Dual, Logarithmic Bound Insertion, or *Minimum Weight Matching*.
- **Exact algorithms:** Mixed-Integer Linear Programming (*MILP*), Dynamic Programming (*DP*), or *Branch-and-Cut*.
- **Heuristics:** *Cheapest* Insertion, *Parallel Savings Heuristic*, or *Nearest Neighbor*.
- **Metaheuristics:** Genetic Algorithms (*GA*), Ant Colony Optimization (*ACO*) or *Tab*u *Search*.

5.1 Metaheuristics

One of the main characteristics of an SRP is that it is implemented on a graph, where the nodes and edges usually represent POIs to visit and streets to travel on, respectively. This means that the graph will become more complex as the urban area to be studied expands. This also affects the solution space of an SRP, exponentially increasing its size. Therefore, metaheuristics based on trajectories, evolutionary processes, or bio-inspired processes are commonly implemented to obtain an approximate solution, such as simulated annealing [29], ant colony optimization [4], genetic algorithms [10], [32], or parallel algorithms [21].

Genetic algorithms (*GA*) particularly simulate Darwin's principle of natural selection [12] by taking a set of solutions to combine and/or modify them, subjecting the set to an evaluation to find candidate solutions that will create a new generation. These types of algorithms excel when seeking to make decisions based on more than one criterion, as they work on the entire solution space.

There is a set of evolutionary algorithms with a multi-objective approach (*MOEAs*), which can be used to find transport routes in SRPs, simultaneously optimizing aspects such as distance, number of vehicles, environmental pollution, energy use, among others. These objectives are often in conflict [7]. MOEAs are considerably useful because they generate a Pareto Front that offers a diverse population for making better decisions; it is possible to achieve solutions from the front in a single execution, and they do not depend on the shape and continuity of the Pareto front [26].

In general, MOEAs have seven basic steps, which are initialization, fitness, selection, offspring, environmental selection, update, and stopping criterion.

The difficulty in designing an evolutionary algorithm lies in representing the SRP problem in the coding, since all its information, including constraints and penalties, must be contained within the coded solutions to be manipulated through the algorithm. In addition, it must be ensured that the entire population is a feasible solution at all times, since the transformation of the population through the processes of descent and mutation could generate elements that are not solutions to the problem if the functions and operators related to these processes are not correctly defined.

5.2 Candidate algorithms

The MOCell (Multiobjective Cellular Genetic Algorithm) algorithm is an evolutionary algorithm based on cellular structures designed to address multiobjective optimization problems. It was introduced by Antonio Nebro [23], and the central idea is to organize the population into a two-dimensional grid structure, where each individual interacts mainly with its neighbors, rather than with the entire global population.

This introduces two key properties: diversity preservation, which, by restricting crossovers and replacements to local neighborhoods, prevents rapid convergence to a single region of the Pareto Front, maintaining broad exploration; and progressive convergence, in the sense that good solutions gradually spread across the grid, achieving a balance between exploration (searching for new regions) and exploitation (refining already promising solutions). MOCell combines several features that make it stand out, which are Parent Selection in local neighborhoods, Controlled Replacement Mechanisms, where a child can replace a neighbor based on dominance and file quality, and an Elitist External File, which stores and manages non-dominated solutions found throughout the execution, ensuring the preservation of high-quality solutions.

Nebro mentions that MOCell is promising because it has balanced control of diversity and convergence, is scalable, has broad exploration capabilities, and proves to be competitive [23]. These qualities are because the algorithm has a cellular structure that prevents premature loss of diversity, a decentralized organization (allowing a natural flow of information and flexibility), and does not depend on a single global population, which allows it to cover more regions of the Pareto Front. MOCell achieved quality results and better dispersion of solutions, making it attractive in applications where it is not enough with a fast convergence, but it is also necessary to cover the front uniformly.

In general, this has not been widely studied. However, Osaba et al. used MOCell along with five other evolutionary algorithms to solve a goods delivery problem, where they consider three objective functions: maximizing the income from a successful delivery, minimizing the distance traveled, and minimizing risk situations, which depend on certain properties of the roads [24]. In this work, three different scenarios with the same characteristics are solved, where the multi-objective algorithms NSGA-II, MOCell, SMPSO, MOEA/D, NSGA-III, and MOMBI2 were implemented, with MOCell showing the best performance under the IQR hypervolume metrics and the median of the solutions for the scenarios. Now, one of the disadvantages of this problem is that we can see in *Table 1* that Osaba's work [24] is not closely related to the SRP.

Peña et al. [27] also use MOCell, which is quite relevant to our research. Here, a bi-objective problem is solved: minimizing operating costs and maximizing the quality of service perceived by users. This study is one of the closest to the proposal (*Table 1*), where they use hypervolume, dispersion, ε-Indicator, and Set-Coverage indicators to compare the performance of MOCell with NSGA-II in solving this problem. From this, MOCell generates solutions where the operating cost is 26% lower than those generated by NSGA-II, in addition to demonstrating a better ability to find different sets of vehicles capable of covering the proposed routes with a lower operating cost.

NSGA-II is one of the most influential and widely used evolutionary algorithms in multi-objective optimization due to its elitism strategy with ranking and crowding distance, good balance between quality and diversity, and its recognized efficiency in multi-objective benchmarks. It is used with problems with one or two objectives, while NSGA-III (a variant of NSGA-II) is used in problems with three or more objectives. Therefore, using it as a standard reference to evaluate the performance of a little-implemented algorithm such as MOCell has two purposes: to demonstrate that MOCell could produce solutions at least as good as a well-established algorithm, and to highlight differences in approach, as NSGA-III operates on a global population with centralized selection and replacement operators, while MOCell shows that a decentralized cellular structure can also be highly competitive.

However, any evolutionary metaheuristic generates an approximate Pareto Front, and the fact that we are searching for solutions for a new variant can be problematic, since without the exact solution as a reference, we will not know how good the solutions generated by the algorithms are. It is necessary to compare the results of these metaheuristics with those of an exact algorithm. On the other hand, given that large-scale multi-objective VRP problems are intractable for exact algorithms, the proposed experimental design is divided into two steps that use the concept of subset validation. The objective is to evaluate the quality of the approximations generated by the metaheuristics under study (NSGA-III and MOCell) against the known Pareto Optimal Front (P*) and extrapolate their performance to more complex instances.

5.3 Two-step evaluation strategy

Step I focus on calibration and obtaining the reference solution. A set of small SRP instances (characterized by a small number of customers, N, and the three-objective space) will be selected to ensure the viability of the exact method. The MOA* (Multi-Objective A*) algorithm will be implemented and executed on these instances to determine the Pareto Optimal Set and its corresponding P*. This P* constitutes the ground truth against which the metaheuristics will be contrasted. Subsequently, NSGA-III and MOCell will be executed M times (e.g., M=30 repetitions) on the same small instances and applying a variance study with ANOVA to calibrate the parameters or variables of the algorithms. The latter is done to guarantee, to a certain extent, good approximations. The quality of the approximate Pareto Fronts (PF) generated will be quantified using convergence and diversity metrics, such as Inverted Generational Distance (IGD) and Hypervolume (HV), thus allowing a direct evaluation of the proximity of each metaheuristic to the known optimum.

Step II addresses the problem of scalability. Medium to large instances of the SRP will be selected, where obtaining the P* using MOA* is computationally unfeasible. Both metaheuristics will be executed using the parameter configurations calibrated in step I, maintaining a criterion of fairness in computation time or number of function evaluations. The analysis in this step will be relative in nature, focusing on the dispersion and comparative dominance between the Pareto Fronts obtained by NSGA-III and MOCell. The same metrics (HV and IGD) will be used to determine which metaheuristic shows superior or more consistent performance in the efficient exploration of the solution space for large-scale problems. This will allow us to validate the hypothesis about the robustness and potential of MOCell compared to the NSGA-III standard in the context of the new SRP.

5.4 Evaluation metrics

To ensure a comprehensive and unbiased evaluation of the quality of the approximate Pareto Fronts generated by NSGA-III and MOCell, we consider a set of metrics that effectively combine the two main objectives in Multi-Objective Optimization: Convergence and Diversity.

The Hypervolume (HV) metric is recognized as the most comprehensive quality metric widely used in the scientific community. Its main strength lies in its ability to simultaneously measure convergence toward F* and diversity across it. A high HV indicates better overall performance of the algorithm, as the front is closer to the optimum and, at the same time, more extended. It is the ideal metric for Step I (direct comparison with the exact) and for Step II (relative comparison), as it correlates directly with the Pareto domain.

Inverted Generational Distance (IGD) complements HV by providing a dual perspective that is crucial when the Optimal Pareto Front is known: it measures the average distance of each point on the Optimal Pareto Front to the nearest point on the approximate Pareto Front. The IGD should be as small as possible, as it is an indicator that penalizes Pareto Fronts that: a) Converge poorly and b) Show little diversity. By basing the measurement on the coverage of the Optimal Pareto Front, the IGD provides a very accurate assessment of how much of the real has been captured by the algorithm. Thus, the IGD is also fundamental in Step I (where the Optimal Pareto Front is known) to quantify the dispersion and convergence of metaheuristics with respect to the theoretical optimum. For Step II, an adapted or relative version of the Generational Distance (GD) metric will be used to compare the relative convergence of one with respect to the other.

By using these two metrics in Step I, the experiment ensures that the validation of metaheuristics is based both on the overall extent of the front and on its point accuracy against the set of Pareto optimal solutions.

6 Conclusions and Future Work

This work has presented the formulation of a highly restricted variant of the Multi-objective SRP, motivated by specific operational needs in the tourism and logistics sectors. The inclusion of three conflicting optimization objectives and realistic parameters, such as concurrent service constraints and user preferences respecting POIs, converts it to the frontier of applied research.

The main contribution of the paper is a complex, realistic SRP model with the justification of a two-step methodology to address this challenge. The complexity of the proposed model and the inherent intractability of large-scale instances make it unfeasible to rely solely on exact search algorithms such as MOA*. Therefore:

- **Need for Validation:** Solving small instances with MOA* is an indispensable methodological step. This step ensures that the model and implementation of metaheuristics can be validated directly against the known optimal Pareto frontier. Without this verification, any results obtained by metaheuristics in large instances would lack a solid quality reference.
- **Need for Metaheuristics:** The future deployment of metaheuristics such as NSGA-III and MOCell is essential to achieve the scalability required by the case studies (Guadalupe Valley, museums, warehouses).

Consequently, this work establishes a methodological basis required to transform a complex practical problem into a research problem with scientifically reproducible results. The justification for this paper lies not only in obtaining final results but in demonstrating the viability, necessity, and methodological soundness of the approach that produces knowledge about the comparative performance of state-of-the-art metaheuristics on a highly complex many-parameter SRP.

As for the future work, we will introduce the following parameters: vehicle capacity (amount of packages, users or any kind of good to transport), vehicle operating cost, and users individual preferences, so the problem will not limit to generate routes and schedules of the POIs, but also to organize users with the vehicles, making groups with similar preferences and aiming to minimize the fleet cost.

7 References

1. Bac, U., & Erdem, M. (2021). Optimization of electric vehicle recharge schedule and routing problem with time windows and partial recharge: A comparative study for an urban logistics fleet. *Sustainable Cities and Society*, *70*, 102883.
2. Baizal, Z., Qatrunnada, A., & Septiana, G. (2025). Hybrid Genetic and Simulated Annealing Algorithm for Vehicle Routing Problem in Recommendation of Tour Route in Days. *International Journal of Intelligent Engineering and Systems*, *18*(3), 257-272.
3. Beheshti, A. K., Hejazi, S. R., & Alinaghian, M. (2015). The vehicle routing problem with multiple prioritized time windows: A case study. *Computers & Industrial Engineering*, *90*, 402-413.
4. Bell, J. E. and McMullen, P. R. (2004). Ant colony optimization techniques for the vehicle routing problem. Advanced Engineering Informatics, 18(1):41–48.

5. Clarke, G. and Wright, J. W. (2012). Scheduling of Vehicles from a Central Depot to a Number of Delivery Points. In Klaus, P. and Müller, S., editors, The Roots of Logistics, pages 229–244. Springer Berlin Heidelberg, Berlin, Heidelberg.
6. Danisa, A., Cosío-León, M. A., Martínez-Vargas, A., Valenzuela-Alcaraz, V. M., & Meza-López, J. H. (2024). T-adaptive an Online Tuning Technique Coupled to MOEA/D Algorithm: A Comparative Analysis with Offline Parameter Tuning Techniques. En A. J. Kulkarni & A. H. Gandomi (Eds.), *Handbook of Formal Optimization* (pp. 719-753). Springer Nature Singapore.
7. Eiben, A. and Smith, J. (2015). Introduction to Evolutionary Computing. Natural Computing Series. Springer Berlin Heidelberg, Berlin, Heidelberg.
8. Elatar, S., Abouelmehdi, K., and Riffi, M. E. (2023). The vehicle routing problem in the last decade: variants, taxonomy and metaheuristics. Procedia Computer Science, 220:398–404.
9. Euchi, J., Zidi, S., and Laouamer, L. (2020). A Hybrid Approach to Solve the Vehicle Routing Problem with Time Windows and Synchronized Visits In-Home Health Care. Arabian Journal for Science and Engineering, 45(12):10637–10652.
10. Fagundez, G., Massobrio, R., and Nesmachnow, S. (2014). Online taxi sharing optimization using evolutionary algorithms. In 2014 XL Latin American Computing Conference (CLEI), pages 1–12, Montevideo, Uruguay. IEEE.
11. Kandakoglu, A., Sauré, A., Michalowski, W., Aquino, M., Graham, J., and McCormick, B. (2020). A decision support system for home dialysis visit scheduling and nurse routing. Decision Support Systems, 130:113224.
12. Kanoh, H. (2007). Dynamic route planning for car navigation systems using virus genetic algorithms. International Journal of Knowledge-based and Intelligent Engineering Systems, 11(1):65–78.
13. Lai, C.-M., Chiu, C.-C., & Chen, T.-L. (2025). A nondominated sorting simplified swarm optimization with local search mechanisms for multi-objective vehicle routing problems with time windows. *Applied Soft Computing*, *174*, 112989.
14. Legrand, C., Cattaruzza, D., Jourdan, L., & Kessaci, M. (2025). Improving neighborhood exploration into MOEA/D framework to solve a bi-objective routing problem. *International Transactions in Operational Research*, *32*(1), 117-143.
15. Lenstra, J. K. and Kan, A. H. G. R. (1981). Complexity of vehicle routing and scheduling problems. Networks, 11(2):221–227.
16. Li, J., Han, Y., Duan, P., Han, Y., Niu, B., Li, C., Zheng, Z., & Liu, Y. (2020). Meta-heuristic algorithm for solving vehicle routing problems with time windows and synchronized visit constraints in prefabricated systems. *Journal of Cleaner Production*, *250*, 119464.
17. Li, X., Gao, C., Wang, J., Tang, H., Ma, T., & Yuan, F. (2025). Research on Multi-Objective Green Vehicle Routing Problem with Time Windows Based on the Improved Non-Dominated Sorting Genetic Algorithm III. *Symmetry*, *17*(5), 734.
18. Luo, H., Dridi, M., & Grunder, O. (2021). An ACO-based heuristic approach for a route and speed optimization problem in home health care with synchronized visits and carbon emissions. *Soft Computing*, *25*(23), 14673-14696.
19. Masmoudi, M. A., Hosny, M., & Koç, Ç. (2022). The fleet size and mix vehicle routing problem with synchronized visits. *Transportation Letters*, *14*(4), 427-445.
20. Masmoudi, M., Jarboui, B., & Borchani, R. (2023). Efficient metaheuristics for the home (health)-care routing and scheduling problem with time windows and synchronized visits. *Optimization Letters*, *17*(9), 2135-2167.
21. Massobrio, R., Fagúndez, G., and Nesmachnow, S. (2014). A parallel micro evolutionary algorithm for taxi sharing optimization. Publisher: Unpublished.

22. Nasution, S. M., Septiawan, R. R., & Azmi, F. (2023). Tourism Itinerary Recommendation using Vehicle Routing Problem Time Windows Based on Shortest Path Method and Analytics Hierarchy Process. Engineering.
23. Nebro, A. J., Durillo, J. J., Luna, F., Dorronsoro, B., & Alba, E. (2009). MOCell: A cellular genetic algorithm for multiobjective optimization. International Journal of Intelligent Systems, 24(7), 726-746.
24. Osaba, E., Del Ser, J., Nebro, A. J., Lana, I., Bilbao, M. N., & Sanchez-Medina, J. J. (2018). Multi-Objective Optimization of Bike Routes for Last-Mile Package Delivery with Drop-Offs. *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 865-870.
25. Osaba, E., Martinez, A. D., Lobo, J. L., Lana, I., & Ser, J. D. (2020). On the Transferability of Knowledge among Vehicle Routing Problems by using Cellular Evolutionary Multitasking. *2020 IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC)*, 1-8.
26. Pena, D., Tchernykh, A., Radchenko, G., Nesmachnow, S., Ley-Flores, J., and Nazariega, R. (2017). Multiobjective Optimization of Greenhouse Gas Emissions Enhancing the Quality of Service for Urban Public Transport Timetabling. In 2017 IVth International Conference on Engineering and Telecommunication (EnT), pages 114–118, Moscow. IEEE.
27. Peña, D., Tchernykh, A., Nesmachnow, S., Massobrio, R., Feoktistov, A., Bychkov, I., Radchenko, G., Drozdov, A. Y., and Garichev, S. N. (2019). Operating cost and quality of service optimization for multi-vehicle-type timetabling for urban bus systems. Journal of Parallel and Distributed Computing, 133:272–285.
28. Polnik, M., Riccardi, A., & Akartunalı, K. (2021). A multistage optimisation algorithm for the large vehicle routing problem with time windows and synchronised visits. *Journal of the Operational Research Society*, *72*(11), 2396-2411.
29. Possel, B., Wismans, L. J. J., Van Berkum, E. C., and Bliemer, M. C. J. (2018). The multi-objective network design problem using minimizing externalities as objectives: comparison of a genetic algorithm and simulated annealing framework. Transportation, 45(2):545–572.
30. Pulido-Gaytan, L. B., Tchernykh, A., Nesmachnow, S., Cristóbal-Salas, A., Avetisyan, A., Barrera, H. E. C., & Hernandez, C. J. B. (2019). Multi-objective Optimization of Vehicle Routing with Environmental Penalty. En M. Torres & J. Klapp (Eds.), *Supercomputing* (Vol. 1151, pp. 147-162). Springer International Publishing.
31. Secretaría de Turismo del Gobierno del Estado de Baja California. (2018). Ensenada. Indicadores 2017. Recuperado de: https://securebc.org/secture/wp-content/uploads/downloads/Ensenada_2018.pdf
32. Sharma, S. and Mathew, T. V. (2011). Multiobjective Network Design for Emission and Travel-Time Trade-off for a Sustainable Large Urban Transportation Network. Environment and Planning B: Planning and Design, 38(3):520–538.
33. Varas, M., Baesler, F., Basso, F., Contreras, J. P., Pezoa, R., Rojas-Goldsack, M. F., & Ronco, R. (2024). A home hospitalization assignment and routing problem with multiple time windows, mandatory returns and perishable biological samples: A Chilean case study. *Computers & Industrial Engineering*, *189*, 109951.
34. Zhou, Y., Kong, L., & Wang, H. (2025). A multiobjective edge-based learning algorithm for the vehicle routing problem with time windows. *Information Sciences*, *715*, 122223.
35. Zuhanda, M. K., Hartono, Hasibuan, S. A. R. S., Muthmainnah, N., & Girsang, M. B. (2025). An evolutionary analytics model for enhancing food distribution through two-tier routing. *Decision Analytics Journal*, *16*, 100621.

Smart energy, energy efficiency, and sustainability

Systematic review of methods for flow optimization in photovoltaic solar pumping systems under conditions of irradiance variability

Manuel González [1][0000-0001-6063-314X] , Félix Santos2[0000-0002-4715-3499], Deyslen Mariano-Hernández[3][0000-0002-4255-3450] , Martin Melendez [3][0000-0003-2218-3445], Armando J Taveras-Cruz[3,4][0000-0001-6538-4728], Miguel Aybar-Mejía [3][0000-0002-4715-3499] Fabio Andrade Rengifo [0000-0002-8859-7336] and Marcel Castro-Sitiriche[4][0000-0002-7650-4112]

[1] Engineering Area, Instituto Tecnológico de Santo Domingo. Santo Domingo, Dominican Republic

[2]Smart tech and sustainability (STECHS INTERFAZ SRL). Santo Domingo, Dominican Republic

[3] Engineering Area, Instituto Tecnológico de Santo Domingo. Santo Domingo, Dominican Republic

[4] Electrical & Computer Engineering Department, University of Puerto Rico, Mayaguez Campus

Miguel.aybar@intec.edu.do

Abstract. Solar Photovoltaic Pumping Systems (SBSFV) represent a sustainable solution for agricultural irrigation in isolated rural areas, where access to the electricity grid is limited. However, its performance is strongly influenced by the variability of solar irradiance, which directly affects the pumped flow. This article aims to evaluate the scientific methods proposed between 2020 and 2025 for optimizing the flow in isolated SBSFVs with hydraulic storage under conditions of irradiance variability. A systematic review was conducted in ScienceDirect and Google Scholar, identifying 38,782 articles. After applying inclusion and exclusion criteria, 72 studies were selected that addressed engine control strategies, energy management with artificial intelligence, system design, efficient operation, and hydraulic evaluation. Data analysis was extracted on rural validation, performance, and applicability. The results show that hybrid control, combining incremental conductance with variable pitch control and direct torque control integrated with fuzzy logic (VSS INC-FDTC), increased the flow rate by 20% and reduced the response time by 46%. Dual-axis solar tracking increased energy capture by more than 90%. Sizing based on Water Supply Shortage Probability (WSP) enabled reducing the system's size by up to 50% without compromising flow rate. PID (Proportional-Integral-Derivative) controller tuning reduced abrupt stops from 40% to 7.3% and improved cloud resistance from 65% to 97.9%. The overall hydraulic efficiency reached 34%, limited by the pump. While the advanced engine controls offer high precision, their structural complexity limits their use in rural areas. Passive strategies such as WSP-based design and power curve alignment are more effective and economical. It is recommended to combine adequate design with straightforward controls and evaluate its economic and climatic viability in different regions.

Keywords: Solar photovoltaic pumping, flow optimization, irradiance variability, motor control, system sizing, energy efficiency

1 Introduction

Agricultural water consumption accounts for approximately 70% of global freshwater use, making it essential to food security and human development [1]. Energy from photovoltaic technology has emerged as one of the alternative sources to meet these needs worldwide [2][3][4][5]. In Latin America and the Caribbean, the penetration of

photovoltaic systems in isolated areas accounted for 1.5% of total generation in 2019, primarily used to electrify homes, health institutions, and schools in isolated or hard-to-reach communities, as well as for water supply and pumping [6][7][8].

In rural and isolated developing countries, limited access to the National Electric Grid forces them to rely on pumps powered by diesel-fueled internal combustion engines. This costly and polluting solution is vulnerable to fuel shortages [9][10] [11] [12] [13][14]. Faced with this challenge, SBSFVs emerge as a sustainable alternative, particularly in countries such as the Dominican Republic, which has a high average solar radiation (5.5 kWh/m²/day) [15] [16] [17].

An SBSFV is composed of a photovoltaic generator, a power conditioner, a motor pump (induction motor and centrifugal pump), a piping system, and an energy storage system [18][19]. The latter can be electric (using batteries) or hydraulic (utilizing an elevated tank). In rural contexts, hydraulic storage is preferable due to its lower cost, greater robustness, and lower maintenance requirements compared to batteries, which are expensive and require specialized maintenance[12] [20]. However, these systems operate only during the day, so it is necessary to maximize the pumped flow under variable irradiance (e.g., due to cloud cover or momentary shading) to ensure tank filling and continuous supply, especially during drought [21] [22].

The main variables involved in this process are:

Flow rate Q, in liters per minute (L/min)
Solar irradiance, in kWh/m²/day,
Electrical power delivered by the photovoltaic generator, in watts (W), which depends linearly on the irradiance,
Speed of the motor pump, in revolutions per minute (rpm),
Power supply frequency of the motor f, in hertz (Hz) [23]

The flow rate (Q) depends linearly on the speed of the motor, according to the laws of affinity. In turn, the motor speed is determined by the supply frequency, while the available power of the photovoltaic generator is a function of solar irradiance [24] [25]. Hydraulic power, according to the same laws, varies with the cube of the speed ($v^{\wedge 3}$). Therefore, by reducing the motor speed through frequency variation, the power required by the pumping system decreases exponentially. This relationship makes optimizing flow through dynamic frequency adjustment—using a variable frequency drive (VFD)—a key strategy for improving system efficiency and operational resilience [26].

Despite its potential, a gap exists in the scientific literature regarding flow optimization in SBSFV, specifically regarding hydraulic storage without electrical storage components. Most studies focus on battery-powered systems, with limited research on configurations suitable for low-income rural communities[27] [28]. This systematic review aims to identify and analyze proposed scientific methods for optimizing flow and improving performance in SBSFVs that use induction motors, centrifugal pumps, and VFDs under transient irradiance variations.

This work is organized into the following sections: 2- Methodology, 3 - Literature review, 4- Discussion and Limitations, 5- 5Conclusions and Future Directions.

2 Methodology

This systematic review was based on the search and analysis of scientific literature published between 2020 and 2025 in the ScienceDirect and Google Scholar databases. The Boolean keywords and combinations used included: "SBSFV", "photovoltaic water

pumping", "induction motor", "centrifugal pump", "variable frequency drive (VFD)", "irradiance variation", "flow optimization", and "energy efficiency".

The inclusion criteria focused on performance optimization, use of VFDs, alternating current (AC) motors, centrifugal pumps, and energy efficiency evaluation. Exclusion criteria included: battery-powered systems, direct current (DC) motors, non-agricultural applications, and duplicates.

Initially, 38.782 articles were identified using the SBSFV keyword. After applying the inclusion and exclusion criteria, which were filtered by time lapse and journal topic, 72 articles were selected for analysis. For each, the following were extracted: control methods, type of validation (simulation, prototype, or case study), and results on flow optimization, system efficiency, and response to irradiance variations. Comparative analysis identified effective strategies to optimize SBSFV performance under variable irradiance conditions, with a focus on systems incorporating hydraulic storage. Figure 1 illustrates the methodology employed in this review.

Identification of papers: After removing duplicates, **38,782** papers were identified	**Screening**: Based on the research question, the number of papers was reduced to **7,806**	**Eligibility assessment**; from these, **264** articles were initially selected for further evaluation	**Full-text assessment**: after reviewing the complete papers, **172** remained	**Final inclusion**: A total of **72** articles were selected for the systematic review

Figure 1. Systematic Literature Review Methodology - Flowchart

3 Literature review

3.1 Induction Motor Control Strategies

Direct Torque Control (DTC) is widely used in three-phase induction motors due to its robustness and dynamic response. However, in SBSFV, torque and flow ripples are observed in response to abrupt irradiance changes, thereby increasing mechanical wear. The 12-sector DTC reduces these ripples (28% in torque and 0.2% in flow), improving accuracy [29]. Integration with fuzzy logic controllers (FLC) fixes the switching frequency and reduces total harmonic distortion (THD) [30][31]. Hybrid strategies, such as VSS INC-FDTC, outperform the Variable Speed Disturbance and Observation (VSS P&O) method by reducing response time by 46% and increasing pumped flow [31]. Meanwhile, the Kalman filter (KF-MPPT) combined with 12-sector DTCs improves reference tracking and current quality [32]. The use of Modified Space Vector Pulse Width Modulation (MSVPWM) optimizes inverter switching, requiring fewer sensors, achieving high efficiency, and minimizing harmonic current [33]. The Modified Particle Swarm Optimization (MPSO) algorithm achieves an accuracy of up to 99.81% in power extraction, although its scalability is limited [34]. Optimal slip control using VFD enables high efficiency under variable irradiance [31], and its combination with a Mandani-type FLC diffuse control reduces the number of serial modules by up to 30% [35]. The High Gain Active Switched Capacitor quasi-Z-Source Inverter (HG-ASCqZSI) achieves a doubled voltage gain, favoring low-power operations [36]. The MEL-ISTSM-LQR (Combination of ISTSM - Integral Super Torque Sliding Surface,

LQR - Linear Related Quadratic Regulator, MEL - Minimum Electrical Loss) hybrid control minimizes electrical losses (MEL), rejects disturbances (ISTSM), and dynamically optimizes error and control energy (LQR) [37], although its computational complexity and cost limit its applicability in rural areas.

Table 1 provides a comparative summary of advanced induction motor control strategies applied in SBSFV systems, highlighting their advantages, limitations, validation methods, and potential applicability in rural contexts.

Table 1. Advanced Induction Motor Control Strategies in SBSFV

Control Category	Representative methods	Key advantage
1º Scalar Control (V/f)	Q V/f + PI [38]; V/f + P&O [39]; V/f + PSO-INC [2]; V/f + VSS-INC [40]; HG-ASCqZSI [36], V/f + HGRI [41]	Simple, low cost, high rural applicability; +20% flow rate vs linear V/f [31]
2º Vector Control / FOC	FOC + IPSO [42]; FOC + P&O + PWM-VSI [43]; Sensorless vector control [44]	High dynamic performance, independent torque-speed control
3º Direct Torque Control (DTC)	DTC 12-sectors [45]; DTC + FLC [46]; DTC-SVM + FLC [46]; FDTC + FSC [47]	Reduces torque/flux ripples, high accuracy, lower stator current
4º Fuzzy Logic (FLC)	Mandani FLC in VFD [35]; FLC + VSS-INC [31]; FLC + VSS-P&O [48]; FLC to DTC [49], FLC +PI [49]	Improves efficiency, reduces losses and current; easy integration into VFDs [20]
5º Advanced MPPT + Hybrid Control	KF-MPPT + DTC [32]; VSS-INC + FLC [29]; NM-MPPT + ISMC [50]; MIWO-P&O + sensor less vector control [51]+PCC- P&O [52]; PTC [53], SMC-P&O [54]	High overall efficiency, rapid convergence, and higher flow rate
6º Robust/Adaptive Control	MEL-ISTSM-LQR [37]; SMC + ZSI [50]; Adaptive Backstepping [55]; Sliding-mode DC regulator [56]	High robustness, tolerance to parameter variations, sensorless in some cases
7º Sensor Reduction Techniques	MSVPWM [50]; Sensorless DTC [74]; SIM [64]; ZSI without DC-DC stage [50]	Lower cost, less hardware, suitable for remote areas

Acronym: MPPT (Maximum Power Point Tracking)., P&O (observation disturbance), VSS (variable step size technique), INC (incremental conductance method), Q V/f + PI (quadratic V/f-based induction motor controller+ proportional integrative controller); DTC (Direct Torque Control), 12-sector DTC, THD (total harmonic distortion), KF-MPPT (Kalman filter), MSVPWM (Modified Space Vector Pulse Width Modulation), MPSO (Modified Particle Swarm Optimization), VFD (variable frequency drive), FLC (Fuzzy logic control), Mandani type FLC fuzzy control, HG-ASCqZSI (High Gain Active Switched Capacitor quasi-Z-Source Inverter), MEL-

ISTSM-LQR (Combination of ISTSM - Integral Super Torque Sliding Surface, LQR - Linear Related Quadratic Regulator, MEL - Minimum Electrical Loss), Q V/f quadratic V/f (control method) + PI (proportional–integral controller), ANFIS (adaptive neuro-fuzzy inference system)., PSO-INC (Particle Swarm Optimization-Incremental Conductance), PSO- PO (Particle Swarm Optimization- and parrow optimization), DTC-SVM (spatial vector modulation in combination with motor control), VSM- FOC (vectorial sliding mode based -field-oriented control) with IPSO (particle swarm optimization), FDTC (Fuzzy direct torque control), FSC (fuzzy speed controller), PWM-VSI (Pulse-Width Modulated Voltage Source Inverter), ISMC (integral sliding mode controller), NM (Nelder-Mead optimization), Adaptive Backstepping controller, MIWO (modified invasive weed optimization algorithm), PCC (Predict current controller), MIWO (modified invasive weed optimization) algorithm), PTC (Predict torque controller), SMC- P&O (Sliding Mode Control and Perturb and Observe algorithms).

3.2 Energy management and prediction

Energy management in solar photovoltaic pumping systems (SBSFV) is evolving towards proactive strategies based on prediction, displacing the traditional reactive approach. Power prediction using artificial intelligence (AI) allows system operation to be optimized through advanced planning, reducing unnecessary start-ups and maximizing pumping hours [57]. Neural networks (ANNs) and adaptive neuro-fuzzy inference systems (ANFIS) stand out for their accuracy in predicting irradiance and temperature. In addition, hybrid models, such as those based on technological and economic coordination using genetic algorithms (GA), allow simultaneous optimization of technical efficiency, profitability, and the system's service life, thereby minimizing unstable operations that can damage the pump [58]. On the other hand, the teaching-learning-based optimization (TLBO) algorithm regulates inverter voltage and frequency to maximize solar power extraction under various weather conditions, keeping motor losses low [19]. Table 2 summarizes the primary energy management and forecasting strategies employed in SBSFV systems, detailing their methodological approaches, enabling technologies, key benefits, and validation methods.

Table 2. Energy Management Strategies and Forecasting

Reference	Approach	Key Technology	Main Benefit	Validation
[24][33][59]	PV Power Prediction	ANFIS	Proactive energy management	Simulation
[60]	Technological and economic coordination, based on GA	Optimization model	Improves efficiency, profitability, and service life	Simulation
[19]	Optimization with TLBO and ANN	Hybrid algorithm	Optimizes inverter performance	Simulation
[45]	ANN-DTC + ANN-MPPT	Neural networks	Increases efficiency and performance	Simulation, dSPACE

As highlighted in Table 2, AI-based prediction and hybrid optimization models significantly enhance system efficiency and reliability; however, most proposals remain at the simulation stage, underscoring the need for experimental validation under operating conditions.

3.3 System sizing and design

Sophisticated controls cannot compensate for inadequate design, so accurate sizing is critical to SBSFV efficiency. Gualteros and Rousse [27] propose a methodology based on the Probability of Water Scarcity (PWS), integrating irradiance simulations, efficiency criteria, component selection, and economic analysis, enabling the system size to be reduced by up to 50% without compromising flow. Alternatively, passive strategies, such as aligning the PV array's power curve with the motor pump's load characteristics [61] or using dual-axis solar tracking [62], optimize performance without the need for active control algorithms. This approach allows operating close to the maximum power point, even without an MPPT, making it ideal for low-cost systems. Additionally, dynamic array reconfiguration via switches, as in the Total Cross Tied (TCT) scheme [55][63], enhances output power under partial shading conditions. These design strategies anticipate the system's behavior, complement active controls, and suggest that their integration could represent the optimal solution for SBSFV in rural contexts. Table 3 presents the main system design and sizing strategies proposed for SBSFV applications, highlighting their methodological approaches, key impacts, evaluation metrics, and validation stage.

Table 3. System design and sizing strategies

Ref.	Strategies	Key Impact	Evaluation Metrics	Validation
[16]	WSP-based sizing	More realistic and economical design	WSP	Simulation
[25]	Solar Tracking (Dual Axis)	+90% Energy Captured	Energy production	Prototype
[25]	VFD with Laminated Conductor	+2–3% VFD efficiency	DC bus losses	Prototype
[26]	PV array configuration (series vs. parallel)	In series at high irradiance hours, the flow rate is reduced better than in parallel.	Gauge height, irradiation, configuration	Simulation
[64]	Dynamic photovoltaic array (DR) reconfiguration	Operates closer to the maximum power point, with higher pumped flow	Flow rate, power, and irradiance matching algorithm	Simulation
[65]	Model based on linear programming to optimize programming and operation	Optimal combination of irrigation sectors that must operate at each time interval	Irradiance, watering, and discharge time	Simulation

As shown in Table 3, design and sizing approaches—ranging from probabilistic methodologies to array reconfiguration and solar tracking—enable significant efficiency gains and cost reductions; however, many strategies remain limited to simulation, with only a few validated through prototypes, underscoring the need for broader experimental assessment in rural settings.

3.4 Operation at points of maximum efficiency

Operation at the point of best efficiency (BEP) ensures maximum flow with minimum energy consumption. However, BEP is defined at nominal frequency and voltage, and under varying irradiance conditions, the sweet spot changes.

To address this, [66] propose the Solar Maximum Efficiency Point (SBEP), a dynamic operating point that maximizes daily energy efficiency by accounting for the irradiance profile, temperature, and affinity laws. This approach improves energy efficiency by 6.7% compared to traditional BEP, without the need for costly experimental testing.

3.5 Evaluation of hydraulic performance: metrics for decision-making

Evaluating hydraulic performance requires built-in metrics to guide design and operation. In [16], the Probability of Water Scarcity (WSP) is introduced, along with a software tool for sizing systems, revealing that only ~10% of the Energy generated by the modules is effectively utilized. In [27], they propose a systematic adjustment of Proportional-Integral-Derivative (PID) controllers using AMIGO (Approximate M-Constrained Integral Gain Optimization) rules, reducing abrupt stops from 40% to 7.3% and from 39.8% to 1.3%, and increasing transient cloud cover resistance from 65% to 97.9% and from 79% to 99.8%. In [14] experimental tests were carried out reporting a flow rate of 61 m³/day, with a yield ratio (PR): 17–19 %) at low pressure (4.1 bar), versus 46 m³/day (PR: 14–15 %) at optimal pressure (4.7 bar), and only 8 m³/day (PR: 3.26 %) on days with low irradiance (2.9 peak sun hours (PSH)). The overall hydraulic efficiency was ~34%, limited by the pump. However, the results confirm the technical feasibility; further research is required to extrapolate the findings to other climatic zones, assess the seasonal impact, and conduct a cost-benefit analysis of oversizing the photovoltaic array. Table 4 summarizes the main metrics and methodologies used to evaluate hydraulic performance in SBSFV systems, detailing the evaluation conditions and the key findings reported in recent studies.

Table 4. Hydraulic Performance Evaluation Metrics

Reference	Main Metric(s)	Evaluation Conditions	Key Finding
[67]	WSP	Annual hourly simulation	Design with WSP: reservoir size, 50% lower nom power
[23]	Flow, pressure, hydraulic efficiency	Days with high and low irradiance	Reducing pressure increases operating hours and pumped volume
[27]	Number of abrupt stops, Cloud resistance ratio, Performance ratio (PR)	Data from 2017 to 2021 for two irrigation systems	Abrupt stops are significantly reduced with systematic method adjustment by 33%
[44]	Performance ratio (PR)	6 months for the irrigation system	Average annual PR of 74.62%., The efficiency of the system is 70%
[68]	Distribution of the generated power	Working frequencies (30–50 Hz), and	There is no power distribution ratio other than 0 and

	maximizes the flow rate from the set of two pumps in parallel	at six pumping heads (18–48 m).	0.5 that maximizes flow rate, except the one necessary to limit the power allocated to one of the pumps to Pmax.
[39]	Discharge (Q), hydraulic power (P_H), pump efficiency ($\acute{\eta}_p$), and overall efficiency ($\acute{\eta}_o$),	Solar radiation (200, 350, 500, 650, and 800 W/m$^{2)}$, Operating head (15, 20, 25, 30, and 35 m)	The best conditions for a solar PV water pump (16.8 m3/h discharge, 1.85 kW hydraulic power, 81.15% pump efficiency, and 16.12% overall efficiency)

As shown in Table 4, hydraulic performance evaluation provides valuable insights into flow rates, pressure optimization, and operational reliability; however, most results remain case-specific, emphasizing the need for broader validation across different climatic zones and cost-benefit scenarios.

4 Discussion

The objective was to identify and analyze scientific methods (2020–2025) for optimizing flow and performance in solar photovoltaic pumping systems with hydraulic storage under irradiance variability. The analysis of 72 articles revealed five key strategies: induction motor control, energy management and solar resource prediction, system sizing and design, operation at maximum efficiency points, and hydraulic performance evaluation.
The findings show that flow optimization in SBSFV requires a comprehensive approach combining design, control, and evaluation—not just isolated technologies. The mention above marks a shift from traditional MPPT-dominated methods [69]. Recent work confirms that MPPT is essential for maximizing PV power, but, alone, it does not address pumping dynamics, irradiance variability, or operational stability in real conditions.

Induction motor control for solar pumps centers on Direct Torque Control (DTC), valued for speed and efficiency but plagued by torque/flow ripples that degrade stability and component life [70]. The 12-sector DTC reduces torque ripple from 56% to 28% and flow ripple from 2.6% to 0.2%, enhancing hydraulic precision.

Intelligent controls include Fuzzy Logic (FLC), which lowers THD and adapts to uncertainty without precise models [19], and MEL-ISTSM-LQR, which optimizes dynamics and disturbance rejection — yet too complex/costly for rural use [23]. MSVPWM cuts sensor count and boosts transient efficiency under cloud cover [22]. For MPPT, VSS INC-FDTC slashes response time by 46% and increases water output by 20% vs. VSS P&O [31]; MPSO achieves 99.81% power accuracy but lacks scalability in remote areas [34]. Optimization must balance robustness, simplicity, and cost per context.

AI-driven energy management predicts irradiance using ANNs/ANFIS to avoid pump cycling, extend lifespan, and ensure supply continuity [26]. Genetic Algorithms (GA) optimize technical/economic/lifespan trade-offs [60]; TLBO improves investor returns and reduces motor losses [30]. However, AI complexity often outweighs benefits in small-scale/rural systems — best reserved for medium/large installations [26].

System sizing advances include the Water Supply Shortage Probability (WSP) method, which reduces system size by 50% without compromising daily delivery, yielding significant cost savings [23]. PV-pump curve matching enables natural MPPT-free operation, ideal for low-maintenance rural setups. Dual-axis tracking boosts energy capture by 90% vs. fixed arrays [25]; Dynamic Reconfiguration (DRA) mitigates shading losses [64] — all improving efficiency without added control complexity [60].

Efficiency operation shifts from a static Best Efficiency Point (BEP) to a dynamic Solar BEP (SBEP), integrating irradiance/temperature/affinity laws to achieve 6.7% overall efficiency — critical under climate variability [29].

Hydraulic performance is validated using metrics such as the Performance Ratio (PR). Tuned PID controllers + AMIGO rules cut abrupt stops from 40% to 7.3% and boost cloud resilience from 65% to 97.9% [57]. Typical overall hydraulic efficiency is ~34%, limited by pump (max 37%) and inverter (91–95%) [23] — significant room for improvement.

Life Cycle & Social Impact: Hybrid LCA shows SPVWPS emit 0.80 g/mm³ over 25 years (55% during manufacturing), 55% lower than diesel (1.46 g/mm³) [71]. Economically, Indian farmers saved 40–50,000 INR/year, with 15–30% higher yields; unit water cost is 0.17 €/kWh — 56% cheaper than diesel [89]. Socially, 70% of farmers were "delighted", 84% had zero operating costs, and local jobs were created — proving transformative for rural development [72].

4.1 Limitations

This systematic review has several limitations:

- Most studies come from specific regions, so results may not apply to other climates or environments.
- Many optimization strategies are only tested in simulations or small labs, with few large-scale, real-world trials in rural areas.
- Most research focuses on technical improvements, but rarely includes cost, economic, or life-cycle analysis.
- Advanced methods are often too complex or expensive for rural use, and there are few direct comparisons between advanced and straightforward solutions in real conditions.

5 Conclusions and Future Directions

Recent advances in solar photovoltaic pumping systems with hydraulic storage (PVWPS) show that better flow and efficiency depend on how motor control, energy management, system design, and hydraulic performance interact. From 2020 to 2025, progress included more reliable motor controls, AI for prediction, low-cost passive designs such as WSP-based sizing, and improved performance metrics.

However, most studies remain limited to simulations or labs, with few real-world comparisons. Seasonal changes, scalability, and economic factors are often overlooked, reducing practical relevance. Many lack cost-benefit or life-cycle analysis.

Promising future directions include combining simple active controls (like fuzzy logic in drives) with passive designs, and exploring photovoltaic array oversizing as a cost-effective alternative. For broader impact, future research should validate solutions across diverse climates and incorporate economic and social sustainability into design choices.

Acknowledgement

The authors are grateful for the support provided by the Thematic Network 723RT0150, "Network for the large-scale integration of renewable energies in electricity systems (RIBIERSE-CYTED)". Fondocyt Grant funded this research No. (FONDOCYT-2023-1-3C1-0547) in the Dominican Republic. This work was supported by the Fondo Nacional de Innovación y Desarrollo Científico–Tecnológico (grant number 2023-1-3C1-0547).

References

1. Banco Mundia: Agua: Panorama general, https://www.bancomundial.org/es/topic/water/overview, last accessed 2025/11/04.
2. Fernández, E.F., Villar-Fernández, A., Montes-Romero, J., Ruiz-Torres, L., Rodrigo, P.M., Manzaneda, A.J., Almonacid, F.: Global energy assessment of the potential of photovoltaics for greenhouse farming. Applied Energy. 309, 118474 (2022). https://doi.org/10.1016/j.apenergy.2021.118474.
3. Solomon, Y., Rao, P.N., Tadesse, T.: A Review on Solar Photovoltaic Powered Water Pumping System for off-Grid Rural Areas for Domestic use and Irrigation Purpose. International Journal of Engineering Research. 10,.
4. Sharma, S., Tomar, A., Bhagat, V.: PV-Based Water Pumping System—A Comprehensive Review. In: Pandey, V.C., Pandey, P.M., and Garg, S.K. (eds.) Advances in Electromechanical Technologies. pp. 389–398. Springer, Singapore (2021). https://doi.org/10.1007/978-981-15-5463-6_34.
5. Vannurappa, D., Reddy, G.P.R., Deepak, K., Hazarathaiah, Y., Prasad, S.S., Malla, S.G.: Small Signal Analysis of Photovoltaic based Water-Pumping System Driven by Induction Motor. In: 2023 7th International Conference on Computing Methodologies and Communication (ICCMC). pp. 15–20 (2023). https://doi.org/10.1109/ICCMC56507.2023.10083929.
6. Miravet-Sánchez, B.L., García-Rivero, A.E., Yuli-Posadas, R.A., Inostroza-Ruiz, L.A., Fernández-Guzmán, V., Chávez-Juanito, Y.A., Rutti-Marin, J.M., Apesteguia-Infantes, J.A.: Solar photovoltaic technology in isolated rural communities in Latin America and the Caribbean. Energy Reports. 8, 1238–1248 (2022). https://doi.org/10.1016/j.egyr.2021.12.052.
7. Sustainability Evaluation of Rural Electrification in Cuba: From Fossil Fuels to Modular Photovoltaic Systems: Case Studies from Sancti Spiritus Province, https://www.mdpi.com/1996-1073/14/9/2480, last accessed 2025/11/04.
8. Fernandez-Fuentes, M.H., Eras-Almeida, A.A., Egido-Aguilera, M.A.: Characterization of Technological Innovations in Photovoltaic Rural Electrification, Based on the Experiences of Bolivia, Peru, and Argentina: Third Generation Solar Home Systems. Sustainability. 13, 3032 (2021). https://doi.org/10.3390/su13063032.
9. Castillo Tapia, R.S., Guerrero Liquet, G.C., Faxas Gúzman, J.G., García Cascales, M.S., Molina García, A.: ASSESSMENT AND PROPOSAL FOR IMPLEMENTATION OF ENERGY COMMUNITIES ON CARIBBEAN ISLANDS: DOMINICAN REPUBLIC APLICATION. In: Proceedings from the International Congress on Project Management and Engineering. pp. 1402–1416. Asociacion Espanola de Direccion e Ingenieria de Proyectos (AEIPRO) (2024).

10.Senthil Kumar, S., Bibin, C., Akash, K., Aravindan, K., Kishore, M., Magesh, G.: Solar powered water pumping systems for irrigation: A comprehensive review on developments and prospects towards a green energy approach. Materials Today: Proceedings. 33, 303–307 (2020). https://doi.org/10.1016/j.matpr.2020.04.092.
11.Pali, B.S., Vadhera, S.: Uninterrupted sustainable power generation at constant voltage using solar photovoltaic with pumped storage. Sustainable Energy Technologies and Assessments. 42, 100890 (2020). https://doi.org/10.1016/j.seta.2020.100890.
12.Ibrahim, E.H., Aslan, S.R.: Solar photovoltaic water pumping system approach for electricity generation and irrigation: Review. Diagnostyka. 24, 1–7 (2023). https://doi.org/10.29354/diag/165849.
13.Belaroussi, O., Terki, A., Dahnoun, I., Lechelah, A.: Optimal DC Voltage Control of a Photovoltaic Water Pumping System for Induction Motor Applications | IIETA. 155–166 (2024). https://doi.org/10.18280/jesa.570116.
14.Valenzuela, B., Muñoz, J., Jaramillo, M., Barrera-Singaña, C., Pavón, W.: Determination of The Maximum Power Point in Solar Photovoltaic Pumping Systems using P&O Perturbation Observation Algorithm. In: 2023 IEEE IAS Global Conference on Renewable Energy and Hydrogen Technologies (GlobConHT). pp. 1–5 (2023). https://doi.org/10.1109/GlobConHT56829.2023.10087762.
15.Congreso de la República Dominicana, Comisión Nacional de Energía: Ley No. 57-07 sobre Incentivo a1 Desarrollo de Fuentes Renovables de Energia y de sus Regimenes Especiales. Comisión Nacional de Energía (2007).
16.Naranjo, L., Willaarts, B.A.: Guía metodológica: diseño de acciones con enfoque del Nexo entre agua, energía y alimentación para países de América Latina y el Caribe. Comisión Económica para América Latina y el Caribe (CEPAL). (2020).
17.Ali, F., Yadav, G., Kumar, J., Yadav, V.: Applications of Photovoltaic Water Pumping System in Developing Countries. Recent Patents on Engineering. (2024). https://doi.org/10.2174/0118722121336917241029062947.
18.Muralidhar, K., Rajasekar, N.: A review of various components of solar water-pumping system: Configuration, characteristics, and performance. International Transactions on Electrical Energy Systems. 31, e13002 (2021). https://doi.org/10.1002/2050-7038.13002.
19.Gevorkov, L., Domínguez-García, J.L., Romero, L.T.: Review on Solar Photovoltaic-Powered Pumping Systems. Energies. 16, 94 (2023). https://doi.org/10.3390/en16010094.
20.Guney, M.S., Tepe, Y.: Classification and assessment of energy storage systems. Renewable and Sustainable Energy Reviews. 75, 1187–1197 (2017). https://doi.org/10.1016/j.rser.2016.11.102.
21.Valverde, J.I.: Bombeo de agua por medio de energía fotovoltaica en zonas rurales, https://alicia.concytec.gob.pe/vufind/Record/UDEP_854ac2ec694b939f898e9ce1155fc6b7, (2006).
22.Ge, L., Liu, J., Wang, Z., Rafiq, M.U.: Optimal Configuration of Off-Grid Hybrid Renewable Energy System with PV/Wind/Hydrogen/Cooling. CSEE Journal of Power and Energy Systems. 11, 2058–2070 (2025). https://doi.org/10.17775/CSEEJPES.2021.07430.
23.Gutierrez, J., Merino, G., Lara, D., Salazar, L.: Hydraulic assessment of a photovoltaic system driving a conventional AC surface electric pump. Sustainable Energy Technologies and Assessments. 45, 101060 (2021). https://doi.org/10.1016/j.seta.2021.101060.
24.Shepovalova, O.V., Belenov, A.T., Chirkov, S.V.: Review of photovoltaic water pumping system research. Energy Reports. 6, 306–324 (2020). https://doi.org/10.1016/j.egyr.2020.08.053.
25.Zahedibialvaei, A., Trojovský, P., Hesari-Shermeh, M., Matoušová, I., Trojovská, E., Hubálovský, Š.: An enhanced turbulent flow of water-based optimization for optimal power flow of power system integrated wind turbine and solar photovoltaic generators. Scientific Reports. 13, 14635 (2023). https://doi.org/10.1038/s41598-023-41749-3.
26.Gan, X., Pei, J., Pavesi, G., Yuan, S., Wang, W.: Application of intelligent methods in energy efficiency enhancement of pump system: A review. Energy Reports. 8, 11592–11606 (2022). https://doi.org/10.1016/j.egyr.2022.09.016.

27.Gualteros, S., Rousse, D.R.: Solar water pumping systems: A tool to assist in sizing and optimization. Solar Energy. 225, 382–398 (2021). https://doi.org/10.1016/j.solener.2021.06.053.
28.Muni, D.T.V., Lalitha, S.V.N.L.: Implementation of Control Strategies for Optimum Utilization of Solar Photovoltaic Systems with Energy Storage Systems. Nanomaterials eJournal. (2020).
29.Errouha, M., Motahhir, S., Combe, Q.: Twelve sectors DTC strategy of IM for PV water pumping system. Materials Today: Proceedings. 51, 2081–2090 (2022). https://doi.org/10.1016/j.matpr.2021.12.214.
30.Errouha, M., Motahhir, S., Combe, Q., Derouich, A.: Intelligent control of induction motor for photovoltaic water pumping system. SN Applied Sciences. 3, 777 (2021). https://doi.org/10.1007/s42452-021-04757-4.
31.Errouha, M., Derouich, A., El Ouanjli, N., Motahhir, S.: High-Performance Standalone Photovoltaic Water Pumping System Using Induction Motor. International Journal of Photoenergy. 2020, 3872529 (2020). https://doi.org/10.1155/2020/3872529.
32.Saady, I., Karim, M., Bossoufi, B., Motahhir, S., Adouairi, M. said, Majout, B., Lamnadi, M., Masud, M., Al-Amri, J.F.: Optimization for a Photovoltaic Pumping System Using Indirect Field Oriented Control of Induction Motor. Electronics. 10, 3076 (2021). https://doi.org/10.3390/electronics10243076.
33.Kalla, U.K., Bhati, N., Chariya, K., Qureshi, I.: Design and Analysis of Solar PV Fed IMD Water - Pumping System. In: 2021 International Conference on Sustainable Energy and Future Electric Transportation (SEFET). pp. 1–6 (2021). https://doi.org/10.1109/SeFet48154.2021.9375744.
34.Rachaputi, B.P., Rathinadurai Louis, J., Sridharan, M.: Solar photovoltaic array fed single-stage sensorless vector control of induction motor drive for water pumping applications. International Journal of Energy and Environmental Engineering. 14, 643–660 (2023). https://doi.org/10.1007/s40095-022-00540-2.
35.dos Santos, W.S., Torres, P.F., Brito, A.U., Galhardo, M.A.B., Macêdo, W.N.: A novel fuzzy controller for photovoltaic pumping systems driven by general-purpose frequency converters. Sustainable Energy Technologies and Assessments. 40, 100758 (2020). https://doi.org/10.1016/j.seta.2020.100758.
36.Gayen, P.K., Das, S.: Wide power control of high-gain and compact quasi Z-source inverter in remote and sustainable solar water pumping system. Sustainable Energy Technologies and Assessments. 74, 104173 (2025). https://doi.org/10.1016/j.seta.2025.104173.
37.Nassiri, Samir., Labbadi, Moussa., Cherkaoui, Mohamed.: Optimal Integral Super-Twisting Sliding-Mode Control for high efficiency of Pumping Systems. IFAC-PapersOnLine. 55, 234–239 (2022). https://doi.org/10.1016/j.ifacol.2022.07.317.
38.Enhanced Quadratic V/f-Based Induction Motor Control of Solar Water Pumping System, https://www.mdpi.com/1996-1073/14/1/104, last accessed 2025/11/04.
39.Swetha, K.T., Reddy, B.V., Jain, R.K.: A Direct Search Nelder Mead MPPT based Induction Motor Drive for Solar PV Water Pumping Systems. In: 2022 22nd National Power Systems Conference (NPSC). pp. 578–583 (2022). https://doi.org/10.1109/NPSC57038.2022.10069049.
40.Shetty, A., Suryanarayana, K., Prabhu, L.V.: Single-Stage Stand-Alone Induction Motor Driven Solar Water Pumping System with Minimal Sensors. In: Chokkadi, S. and Bandyopadhyay, R. (eds.) Smart Sensors Measurement and Instrumentation. pp. 337–351. Springer Nature, Singapore (2023). https://doi.org/10.1007/978-981-19-6913-3_22.
41.Tripathi, P.R., Thakura, P., Laxmi, V., Keshri, R.K.: Stand-alone PV water pumping system based on high-gain resonant inverter fed induction motor serving two-head for permanent water supply. International Journal of Circuit Theory and Applications. 49, 2494–2514 (2021). https://doi.org/10.1002/cta.2983.
42.Abdullateef, S.M., Majeed, Z., Nasir, B.A.: Modeling and Simulation of a PV-Induction Motor Water Pumping System for Off-Grid Irrigation. NTU Journal of Renewable Energy. 9, 48–62 (2025). https://doi.org/10.56286/yefkt888.

43.Singh, B., Sharma, U., Kumar, S.: Standalone Photovoltaic Water Pumping System Using Induction Motor Drive With Reduced Sensors. IEEE Transactions on Industry Applications. 54, 3645–3655 (2018). https://doi.org/10.1109/TIA.2018.2825285.
44.Revathi, V.M., Subramani, K., Krishnakumar, S., Alawadi, A.H., Senthil Kumar, R., Malla, S.G.: A modified invasive weed optimization for MPPT of PV based water pumping system driven by induction motor. Eng. Res. Express. 6, 035311 (2024). https://doi.org/10.1088/2631-8695/ad5cd3.
45.Improving photovoltaic water pumping system performance with ANN-based direct torque control using real-time simulation | Scientific Reports, https://www.nature.com/articles/s41598-025-88330-8, last accessed 2025/11/04.
46.Anbuchandran, S., Stephen, D.S., Babu, M.A., Bhuvanesh, A.: MIWO based MPPT of PV system for induction motor driven water pumping system. e-Prime - Advances in Electrical Engineering, Electronics and Energy. 10, 100835 (2024). https://doi.org/10.1016/j.prime.2024.100835.
47.Sharma, R., Naval, P., Sharma, B., Kulhar, K.S.: Novel control of standalone "Photovoltaic (PV) based Induction Motor Driven Water Pumping System" without battery. E3S Web of Conf. 540, 05003 (2024). https://doi.org/10.1051/e3sconf/202454005003.
48.Errouha, M., Combe, Q., Motahhir, S., Askar, S.S., Abouhawwash, M.: Design and processor in the loop implementation of an improved control for IM driven solar PV fed water pumping system. Sci Rep. 12, 4688 (2022). https://doi.org/10.1038/s41598-022-08252-7.
49.Hamdi, H., Regaya, C.B., Zaafouri, A.: A sliding-neural network control of induction-motor-pump supplied by photovoltaic generator. Protection and Control of Modern Power Systems. 5, 1 (2019). https://doi.org/10.1186/s41601-019-0145-1.
50.Dubey, M., Sharma, S.K., Saxena, R.: Solar Energy based ZSI fed Induction Motor Drive for Water Pumping. In: 2020 IEEE International Conference on Computing, Power and Communication Technologies (GUCON). pp. 135–139 (2020). https://doi.org/10.1109/GUCON48875.2020.9231163.
51.Malla, S.G., Malla, P., Karthik, M., Kumar, D.S., Awad, H.: Modified Invasive Weed Optimization for the Control of Photovoltaic Powered Induction Motor Drives in Water Pumping Systems. Iran J Sci Technol Trans Electr Eng. 47, 925–938 (2023). https://doi.org/10.1007/s40998-023-00589-7.
52.Khan, K., Shukla, S., Singh, B.: Improved Performance Design Realization of a Fractional Kilowatt Induction Motor With Predictive Current Control for Water Pumping. IEEE Transactions on Industry Applications. 56, 4575–4587 (2020). https://doi.org/10.1109/TIA.2020.2968014.
53.Single-Stage Standalone Photovoltaic Water Pumping System Using Predictive Torque Control (PTC) of Induction Machine: Electric Power Components and Systems: Vol 0, No 0 - Get Access, https://www.tandfonline.com/doi/full/10.1080/15325008.2023.2278624, last accessed 2025/11/04.
54.Advanced Control Scheme Optimization for Stand-Alone Photovoltaic Water Pumping Systems, https://www.mdpi.com/2079-3197/12/11/224, last accessed 2025/11/04.
55.Madark, M., Ba-razzouk, A., Abdelmounim, E., Malah, M.E.: Adaptive backstepping control of induction motor powered by photovoltaic generator. International Journal of Electrical and Computer Engineering (IJECE). 11, 2842–2855 (2021). https://doi.org/10.11591/ijece.v11i4.pp2842-2855.
56.Rai, R., Shukla, S., Singh, B.: Sensorless Field Oriented SMCC Based Integral Sliding Mode for Solar PV Based Induction Motor Drive for Water Pumping. IEEE Transactions on Industry Applications. 56, 5056–5064 (2020). https://doi.org/10.1109/TIA.2020.2997901.
57.Ben Ammar, R., Ben Ammar, M., Oualha, A.: Photovoltaic power forecast using empirical models and artificial intelligence approaches for water pumping systems. Renewable Energy. 153, 1016–1028 (2020). https://doi.org/10.1016/j.renene.2020.02.065.
58.Elnozahy, A., Abdel-Salam, M., Abo-Elyousr, F.K.: Optimal techno-economic energy coordination of solar PV water pumping irrigation systems. Energy. 288, 129817 (2024). https://doi.org/10.1016/j.energy.2023.129817.

59.Gadiraju, H.K.V., Barry, V.R., Jain, R.K.: Improved Performance of PV Water Pumping System Using Dynamic Reconfiguration Algorithm Under Partial Shading Conditions. CPSS Transactions on Power Electronics and Applications. 7, 206–215 (2022). https://doi.org/10.24295/CPSSTPEA.2022.00019.
60.Zafar, S., Khalid, M.: Dynamic Performance Assessment Through Simulated Model of Photovoltaic Water Pumping System for Off-Grid Communities. In: 2023 12th International Conference on Renewable Energy Research and Applications (ICRERA). pp. 259–263 (2023). https://doi.org/10.1109/ICRERA59003.2023.10269428.
61.Tiwari, A.K., Kalamkar, V.R., Pande, R.R., Sharma, S.K., Sontake, V.C., Jha, A.: Effect of head and PV array configurations on solar water pumping system. Materials Today: Proceedings. 46, 5475–5481 (2021). https://doi.org/10.1016/j.matpr.2020.09.200.
62.Shaik, R., Beemkumar, N., Adharsha, H., Venkadeshwaran, K., Dhass, A.D.: Efficiency enhancement in a PV operated solar pump by effective design of VFD and tracking system. Materials Today: Proceedings. 33, 454–462 (2020). https://doi.org/10.1016/j.matpr.2020.05.035.
63.Zavala, V., López-Luque, R., Reca, J., Martínez, J., Lao, M.T.: Optimal management of a multisector standalone direct pumping photovoltaic irrigation system. Applied Energy. 260, 114261 (2020). https://doi.org/10.1016/j.apenergy.2019.114261.
64.Errouha, M., Derouich, A., Nahid-Mobarakeh, B., Motahhir, S., El Ghzizal, A.: Improvement control of photovoltaic based water pumping system without energy storage. Solar Energy. 190, 319–328 (2019). https://doi.org/10.1016/j.solener.2019.08.024.
65.Eltawil, M.A., Alhashem, H.A., Alghannam, A.O.: Design of a solar PV powered variable frequency drive for a bubbler irrigation system in palm trees fields. Process Safety and Environmental Protection. 152, 140–153 (2021). https://doi.org/10.1016/j.psep.2021.05.038.
66.Sena dos Santos, W., Ferreira Torres, P., Ubaiara Brito, A., Arrifano Manito, A.R., Figueiredo Pinto Filho, G., Leão Monteiro, W., Negrão Macêdo, W.: A novel method to determine the optimal operating point for centrifugal pumps applied in photovoltaic pumping systems. Solar Energy. 221, 46–59 (2021). https://doi.org/10.1016/j.solener.2021.04.005.
67.Rachaputi, B.P., Rathinadurai Louis, J., Sridharan, M.: Novel sliding mode control of single-stage induction motor drive for solar water pumping applications. Electrical Engineering. 105, 3019–3032 (2023). https://doi.org/10.1007/s00202-023-01851-1.
68.Gasque, M., González-Altozano, P., Gutiérrez-Colomer, R.P., García-Marí, E.: Comparative evaluation of two photovoltaic multi-pump parallel system configurations for optimal distribution of the generated power. Sustainable Energy Technologies and Assessments. 48, 101634 (2021). https://doi.org/10.1016/j.seta.2021.101634.
69.Narendra, A., Venkataramana Naik, N., Panda, A.K., Tiwary, N.: A Comprehensive Review of PV Driven Electrical Motors. Solar Energy. 195, 278–303 (2020). https://doi.org/10.1016/j.solener.2019.09.078.
70.Mollik Babu, R., Alam, M.S., Islam, A., Kamrul Islam, Md.: Toward Sustainable and Clean Energy Futures: A Techno-Economic Review of Solar PV Systems, Challenges, and Opportunities. IEEE Access. 13, 169720–169757 (2025). https://doi.org/10.1109/ACCESS.2025.3614771.
71.Cao, L., Wang, S., Zhu, J., Wang, X., Hu, W.: Life cycle assessment of solar PV water pumping system for carbon emissions reduction at a location: a case study. Energy Sources, Part A: Recovery, Utilization, and Environmental Effects. 46, 4704–4717 (2024). https://doi.org/10.1080/15567036.2024.2329811.
72.Habib, S., Liu, H., Tamoor, M., Zaka, M.A., Jia, Y., Hussien, A.G., Zawbaa, H.M., Kamel, S.: Technical modelling of solar photovoltaic water pumping system and evaluation of system performance and their socio-economic impact. Heliyon. 9, e16105 (2023). https://doi.org/10.1016/j.heliyon.2023.e16105.

Analysis of the Impact of the Penetration of Variable Renewable Energies on the Inertia of the National Interconnected Electric System (SENI) of the Dominican Republic and Mitigation Proposals

Miguel Aybar-Mejía [1,2*][0000-0002-4715-3499], Onésimo Ramirez[2,3][0009-0008-5225-0817], Rony Montero [2][0009-0002-3800-7286], Deyslen Mariano-Hernández[1][0000-0002-4255-3450] , Roberto Eduardo Quintal Palomo [4][0000-0001-8265-3977], Armando J Taveras-Cruz[1][[0000-0001-6538-4728] and Máximo Alberto Domínguez Garabitos[5][0000-0001-6948-7337]

[1] Área de Ingeniería, Instituto Tecnológico de Santo Domingo. Santo Domingo, República Dominicana

[2] Escuela de Ingeniería, Universidad Acción Pro Educación y Cultura (UNAPEC), Santo Domingo, República Dominicana.

[3] Organismo Coordinador del Sistema Eléctrico Interconectado (OC-SENI). Santo Domingo, República Dominicana.

[4]Autonomous University of Yucatan, UADY, Mérida, México

[5] Área de Ciencias Básicas, Instituto Tecnológico de Santo Domingo. Santo Domingo, República Dominicana

* miguel.aybar@intec.edu.do

Abstract. The growing integration of Variable Renewable Energies (VRE) in the National Interconnected Electric System (SENI) of the Dominican Republic, driven by Law 57-07, has transformed its energy matrix towards sustainability, going from 33 MW in 2011 to 1122 MW in 2023. However, this transition poses critical technical challenges, primarily due to the reduction of system inertia, as power converters connect VRE (wind and solar) and do not provide, mechanical inertia, unlike conventional plants. This loss of inertia increases the Rate of Change of Frequency (RoCoF) and the risk of instability, especially in the 2025-2026 horizon, considering ongoing projects that will improve the ERV share from 12 to 18%. The primary objective of this research is to assess the impact and propose mitigation measures. The methodology used included the tuning of real events in the SENI and dynamic studies using the Power Factory DIgSILENT software to quantify the impact of VRE and determine a critical level of inertia. The results demonstrate that inertia below 17 GW·s causes frequency deviations below regulatory limits, thereby increasing the risk of system collapse. In conclusion, the value was determined that integrating 60 MW and 120 MW Battery Energy Storage Systems (BESS) significantly improves stability, reducing the performance of the Automatic Load Shedding Scheme (ALSS) and the dumping of renewable energy.

Keywords: DIgSILENT, Synchronous Inertial Response, Power systems, Battery Energy Storage Systems.

Introduction

The Dominican Republic is no stranger to this trend, having undergone a notable transformation of its National Interconnected Electric System (SENI) since the enactment of Law 57-07 in May 2007. The growing evolution of installed capacity in VRE, from 33 MW in 2011 to 1122 MW in 2023, has supported the country's decision to achieve the climate action goal committed to in the United Nations 2030 Agenda and the Paris Agreement, reinforcing the need to integrate clean energy, allowing ERV participation to increase by up to 30% by 2030 [1].

Despite the undeniable benefits of VRE integration, this transition is not without technical challenges. One of the most significant impacts is the penetration of variable renewable energies on the inertia and stability of the electricity system. Conventional thermal power plants, which have historically been the backbone of electrical systems, provide inertia to the system thanks to the rotating mass of their synchronous generators. In contrast, renewable generation plants, such as wind and solar farms, connect their power through power electronics converters, meaning they do not contribute to system inertia in the same way. This inherent characteristic of VRE increases the risk of frequency fluctuations and compromises SENI's operational stability.

Loss of inertia is a critical issue that affects the system's dynamic stability and transient response to contingencies. Inertia is defined as the ability of the system to withstand rapid changes in frequency due to the kinetic energy stored in the rotating masses of conventional generators [2].

For any period, the synchronous inertial response (SIR) of the generators is calculated as expressed by the following ratio [3]:

$$E_{kin}^{sys} = \sum_{i=1}^{n} H_i \, S_{r,i} \tag{1}$$

Where:

H_i is the inertia constant of generator i., $S_{r,i}$ is the nominal apparent power of the generator expressed in (VA). E_{kin}^{sys} is the total rotational kinetic energy of the system expressed in (Ws) and n is the number of generators online each hour.

In different power systems, inertia values vary according to technology and scale. For example, in ERCOT, nuclear plants with capacities between 1,100 and 1,300 MVA present inertia constants of 3.8–4.34 seconds, contributing around 4,180–5,642 MW·s [4]. Charcoal units of 500–900 MVA provide 1,450–4,050 MW·s with inertia between 2.9–4.5 seconds, while combined-cycle plants (200–700 MVA) range from 200 to 3,780 MW·s with H values of 1–5.4 seconds. Simple combustion turbines with capacities of 50–200 MVA can contribute between 50 and 2,500 MW·s, depending on inertia values of 1–12.5 seconds. Hydroelectric synchronous units (50–400 MVA) typically show inertia constants of about 2 seconds, contributing 100–800 MW·s. In contrast, inverter-based resources such as solar, wind, and batteries provide no physical inertia. At the system level, CAISO (WECC) estimates a total base of ~80,000 MVA, with inertia ranging from 2.5 to 6.5 seconds, equivalent to 200,000–520,000 MW·s [5] [6]. Similarly, EirGrid and SONI report a complete system inertia of about 23,000–25,000

MW·s, considering ~6,000 MVA in Ireland and ~4,000 MVA in Northern Ireland, with an equivalent Hsys of 4.0 seconds [7]. The SEM Committee sets a minimum procurement requirement of 10,000 MW·s [8], while the UK's National Grid ESO reports synchronous generators in the range of 100–1,000 MVA with inertia between 3–9 seconds, contributing 300–9,000 MW·s [9] . At the European level, ENTSOE reports system-dependent values that vary by country [10], while NREL (USA) documents conventional thermal plants of 300–900 MVA, with inertia constants between 2–7 s, corresponding to 600–6,300 MW·s [11].

Low inertia increases the Rate of Change of Frequency (RoCoF) after disturbances, reducing the time available for frequency regulation and increasing the risk of massive disconnections..

The Dominican Republic shares these structural characteristics and faces a projected rise in renewable penetration for 2025–2028, making stability measures indispensable. According to OC-SENI's long-term plan, two scenarios are outlined: one with high integration of Battery Energy Storage Systems (BESS), and another with greater reliance on thermal generation (coal, oil, and natural gas, the latter considered a transition fuel). In both cases, thermal plants retain a central role, but storage emerges as a key element for managing variability and maintaining security of supply.

In this context, the objective of this research is to evaluate the impact of VRE penetration on SENI's inertia and to propose mitigation measures to ensure safe and stable operation amid growing renewable participation. The article is organized as follows: Section 2 reviews the theoretical framework; Section 3 describes the Dominican electricity market; Section 4 details the methodology; Section 5 presents results; and Section 6 concludes with recommendations and limitations.

2. Methodology

This research evaluates the impact of variable renewable energy (VRE) on SENI's inertia and proposes mitigation measures. The approach combines the tuning and validation of real events, the quantification of VRE impact through dynamic studies, and the determination of the system's critical inertia threshold. The methodology is structured into two main components: (i) an energy perspective, focused on optimizing generation dispatch by minimizing operating costs based on declared production costs, generation availability, and demand projections, like the model proposed in [1] but excluding its other cost components. The objective function is thus formulated as shown in Equation 1, and (ii) an electrical perspective, which validates system performance in terms of voltage and frequency stability.

$$\min(\$)\ OF = \sum_{g,p} VPC_g \cdot e_{g,p} \ + +VOLL \cdot \sum_{b,p} nsp_{b,p} \quad (1)$$

Where:

OF represents the objective function, VCP the variable production cost, *e* the estimated energy, VOLL the value of lost load defined by the regulatory authority, *nsp* the non-served power, *g* the thermal generator, *p* the period, and *b* the bus at the connection point.

2.1. Tuning the Simulated Case with Real Events.

The first step of the methodology involves tuning and validating the dynamic database used in the SENI simulations. This process aims to ensure that the models accurately reflect the system's behavior under critical contingencies, such as generator trips or demand fluctuations. The tuning was performed by adjusting key parameters in the dynamic model, including the time constants of the frequency and voltage control systems, adjustments to the Automatic Load Shedding Scheme (ALSS), and parameters of generators and their regulation systems.

The accuracy of the simulations was evaluated using the Mean Squared Error (MSE), a widely used metric for quantifying the difference between simulated and observed values. The expression calculates the EMC:

$$EMC = \sqrt{\frac{1}{N}\sum(Y_{REAL} - Y_{SIMULADO})^2} \quad (2)$$

Where YREAL represents the real values recorded, Y represents the values generated by the model, and N represents the total number of observations. This metric penalizes significant errors more severely due to the squaring of differences, making it especially useful when it needs to reflect substantial deviations more sensitively.

According to Hodson [13], the use of RMSE (Root Mean Square Error) is particularly suitable in contexts where errors are normally distributed and where it is required to capture both the magnitude and the dispersion of differences between real and simulated values. In his review, the author argues that RMSE, unlike MAE (Mean Absolute Error), provides a more comprehensive view of model performance in the presence of extreme errors and is preferable when seeking measures sensitive to these cases.

2.2. Determination of the Critical Inertia Level of the SENI

To determine the critical level of inertia of the SENI, the challenges associated with low inertia in other electrical systems were considered, recognizing that there is no universal criterion for this determination. Each system defines this level based on its characteristics, the size of the most critical contingencies, and the capabilities of its frequency control and regulation schemes. Different regions have reported crucial challenges associated with low inertia. In Ireland, a minimum of 23 GW·s is required for a maximum demand of 6 GW, since RoCoF ≥ 0.5 Hz/s can cause massive disconnections of unsynchronized generation, leading the operator to guarantee a minimum inertia level [14] [15] [16]. In the United Kingdom, where the minimum inertia is around 135 GW·s, values of RoCoF above 0.125 Hz/s activate protections that risk generation disconnection, making it essential to maintain inertia to avoid system collapses [17] [18]. In the Nordic countries, dominated by hydroelectric generation, the slow response

makes UFLS ineffective when RoCoF exceeds 3 Hz/s, thereby increasing instability risks under high renewable penetration [19]. In Québec (Canada), with a demand of 38 GW, hydro generation dominates but has limited primary response, requiring complementary technologies such as synthetic inertia and fast controls [20]. Finally, South Australia operates with only 2 GW·s of inertia, facing RoCoF between 1–3 Hz/s after contingencies; traditional UFLS schemes fail under these conditions, highlighting the need for storage, fast control, and synthetic inertia to prevent blackouts [21] [22].

2.3 Fundamental Parameters of Inertia and Frequency Regulation in Electrical Systems with High Penetration of Renewable Energies.

The impact of inertia on frequency stability is crucial. Inertia determines the initial rate of frequency drop (RoCoF) after a sudden loss of generation. Low inertia increases RoCoF, which can delay the response of frequency response mechanisms, leading to unintentional disconnection of low-frequency loads and compromising reliability. Synchronously connected generators and motors store kinetic energy that they automatically release in the event of a loss of generation, slowing down and causing a decrease in frequency, a phenomenon known as inertial response. This response is critical for reliability in the initial moments of a disturbance, as it determines the initial RoCoF.
The amount of inertia in a system depends on the number and size of synchronously connected generators and driving loads. The calculation of the synchronous inertial response (SIR) is performed by:

$$E_{kin}^{sys} = \sum_{i=1}^{n} H_i \, S_{r,i} \tag{3}$$

Where:

H_i is the inertia constant of generator I, $S_{r,i}$ is the nominal apparent power of the generator expressed in (VA). E_{kin}^{sys} is the total rotational kinetic energy of the system expressed in (Ws) and n is the number of generators online each hour.
It has been observed that technologies such as hydroelectric, combustion turbines, and steam and coal plants have lower inertia constants than nuclear and combined cycle plants. However, non-synchronous inverter-based resources, such as wind and solar generation, do not contribute to synchronous inertia.

3. Case Study: The Dominican Electricity Market

The growing penetration of variable renewable energy (VRE) in the Dominican Republic's SENI poses a challenge to system stability due to the progressive reduction of inertia. Unlike synchronous generators, which stabilize the grid through the kinetic energy of their rotating masses, VRE connected via power converters provide no inertia, increasing the system's sensitivity to frequency disturbances. The Dominican Republic's National Interconnected System (SENI), with an installed capacity of 5,985.34 MW (2024), integrates various technologies, with the main ones by capacity being Internal Combustion Engine (23.7%), Combined Cycle (18.9%), and Steam Turbine (18.8%). Its transmission system comprises lines totaling more than 5,800 km,

operating at 345 kV, 230 kV, 138 kV, and 69 kV; it has 529 connection points.For the 2025–2026 horizon, higher VRE shares could raise destabilization risks if no mitigation measures are adopted. Reduced inertia is reflected in higher RoCoF values and deeper frequency nadirs during contingencies, which may prematurely trigger the Automatic Load Shedding Scheme (ALSS).

To address this, Dominican regulation (RLGE No. 125-01) establishes frequency control requirements, supported by several mechanisms: Primary and Secondary Frequency Regulation, Revolving Reserves, Tertiary Frequency Control (though not formally regulated), Synthetic Inertia, and Energy Storage. Additionally, SENI must be able to withstand the simultaneous loss of the two largest units—Punta Catalina (2 × 376 MW, 13.25% of installed capacity)—requiring a minimum system inertia of 17 GW·s dispatch restrictions, a trend more pronounced during peak demand, when the share of low-inertia renewable generation is highest.

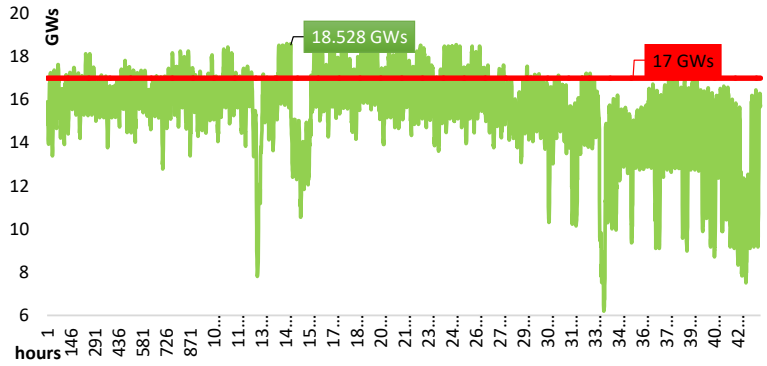


Figure 1. Behavior evolution of inertia without restriction in the office

Figure 2 shows the behavior of the evolution of inertia with the restriction on dispatch

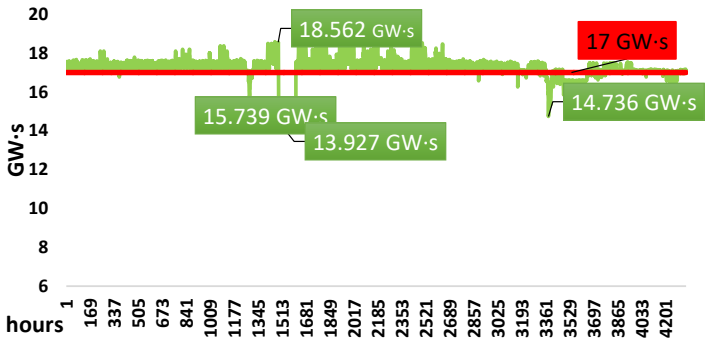


Figure 2. Behavior evolution of inertia with restriction in dispatch

In the analysis model used, the inertial contribution of each unit is calculated proportionally to its dispatched energy, using historical generation factors for wind, solar, and hydroelectric technologies, and considering the nominal capacity for thermal power plants. Demand is projected based on hourly profiles constructed from historical data. To validate the model, two real events that affected the stability of SENI were tuned:

- Firing of the AES Andrés Central (January 12, 2023):
  - Generation: 141.46 MW, Demand: 2448.37 MW, Activation: third step of the Automatic Load Shedding Scheme (ALSS), Nadir: 59.093 Hz, RoCoF: –0.1948 Hz/s, Unballasted load: 221.54 MW, simulation Correlation: 99.98%
- Firing of the Punta Catalina 1 unit (March 6, 2023):
  - Generation: 330.13 MW, Demand: 2654.16 MW, Activation: Fifth step of the Automatic Load Shedding Scheme (ALSS), Nadir: 58.864 Hz, RoCoF: –0.392 Hz/s, Unballasted load: 431.76 MW, Simulation Correlation: 99.98%

To evaluate the accuracy of the model used in the frequency simulations, the mean square error (RMSE) was used as a validation metric. The RMSE quantifies the difference between simulated and observed or reference values, offering a robust measure of the model's fit. In Figure ***3*** (a) and 3(b), comparisons are presented between the actual behavior and the simulations, demonstrating the overperformance of the Automatic Load Shedding Scheme (ALSS). In both cases, the over-frequency levels exceeded the expected values, reaching approximately 60.2 Hz in the first event (AES) and more than 61.3 Hz in the second (Punta Catalina). However, in terms of frequency drop, a high concordance was obtained between the real and simulated events, with a correlation of 99.98%. These results suggest that the tuning of the base case is adequate.

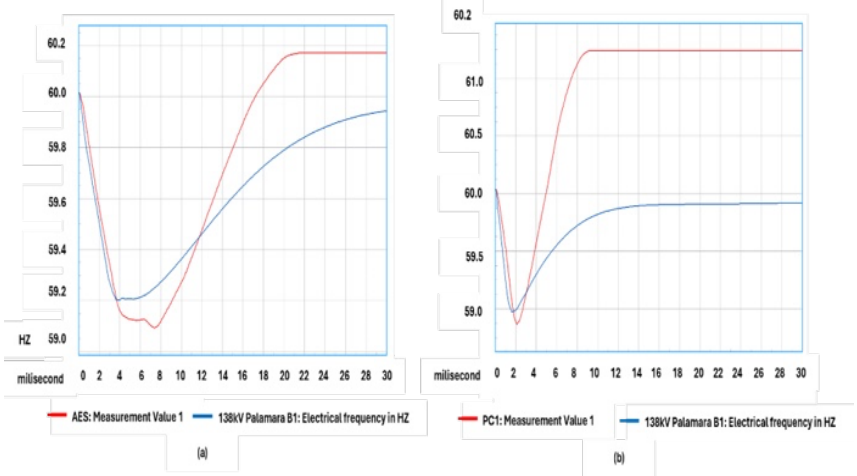


Figure 3. Real vs simulated event comparison. (a) AES ANDRÉS (b) Punta Catalina

These results demonstrate the importance of adequate frequency regulation mechanisms and minimum levels of inertia to preserve SENI stability in the face of high-impact contingencies.

5. Simulations and Results

The simulations enabled the determination of the critical level of inertia for the SENI, the evaluation of the impact of inertia restriction on system operation, and the assessment of the effectiveness of the BESS as a mitigation measure.

5.1. SENI Critical Inertia Level

In the "SENI Automatic Load Shedding Scheme (ALSS) Update", the maximum level of inertia identified was approximately 18.78 GW·s, with an inertia constant (H) of 5.42 seconds. Simulating the simultaneous loss of both units at Punta Catalina (the plant with the highest installed capacity), a RoCoF of -1.778 Hz/s at 100 ms and a minimum nadir of 58.92 Hz were obtained. For the loss of a single unit from Punta Catalina, the average RoCoF was approximately -0.711 Hz/s, and for the simultaneous loss of both, it averaged -1.44 Hz/s. In Figure 4, the behavior of the frequency is presented when there is a simultaneous loss of both units of the Punta Catalina Thermal Power Plant. In Figure 5, the behavior of the frequency with the loss of a single unit of the Punta Catalina Thermal Power Plant is observed.

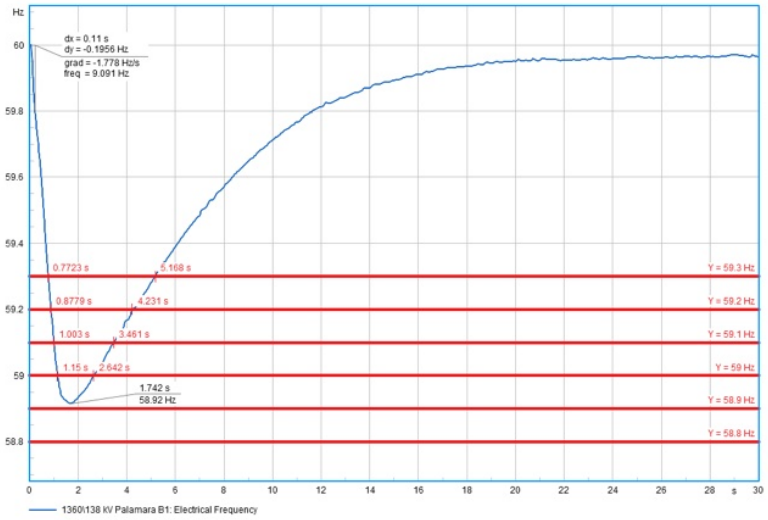


Figure 4. Simultaneous loss simulation of both units, Punta Catalina Thermal Power Plant

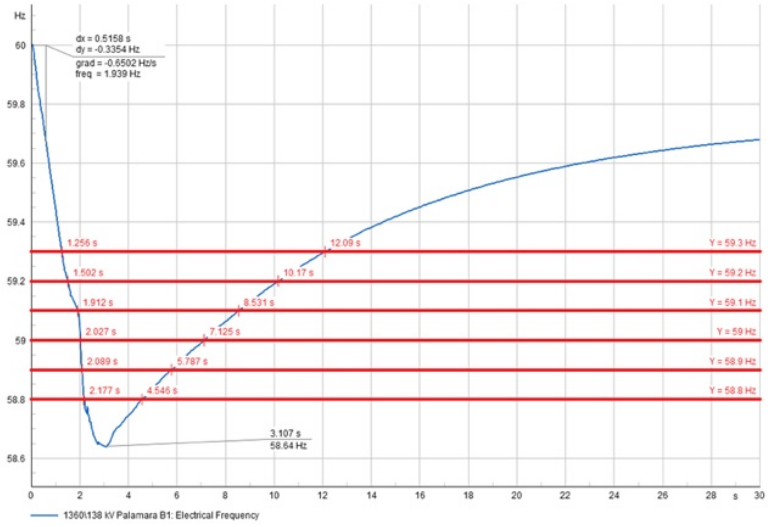


Figure 5. Simulation of a single unit lost at the Punta Catalina Thermal Power Plant

Simulations indicate that with inertia below 17 GW·s, frequency exceeds regulatory limits and recovery after contingencies is inadequate. Figures 6 and 7 show the loss of one or both Punta Catalina units under these conditions, confirming the system's vulnerability below the minimum inertia threshold.

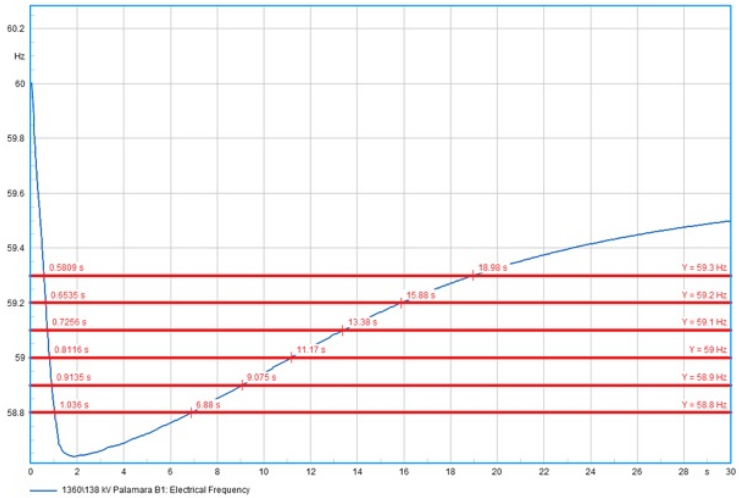


Figure 6. Simultaneous loss of simulation of both units of the Punta Catalina Thermal Power Plant with an inertia level <17 GW·*s*

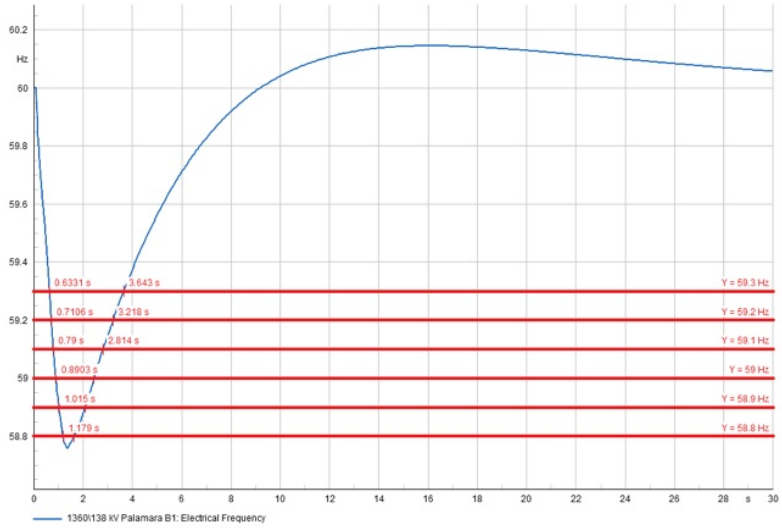


Figure 7. Simulation of one of the units of the Punta Catalina Thermal Power Plant with an inertia level of <17 GW·s

Implementing a minimum inertia restriction of 17 GW·s in the SENI planning model has significant effects on the operation of the electricity system. First, there is evidence of an increase in operating and marginal costs, as dispatch must prioritize synchronized thermal units capable of providing inertia, even if they are less efficient or more expensive than non-conventional renewable sources. This condition limits the firm's economic flexibility and increases the overall system cost.

Secondly, there is a greater discharge of VRE, especially from wind and solar sources. Because inertial stability concerns prevent renewable energy from being dispatched in favor of thermal units, available clean energy is wasted, leading to financial losses and an adverse environmental impact that restricts the use of sustainable sources.

Lastly, a significant change in the system's inertia behavior is noted. In unrestricted situations, inertia gradually diminishes over the semester, especially when demand is high and renewable energy sources are predominant. With the restriction active, however, inertia remains consistently above the 17 GW·s threshold, contributing to greater resilience to shocks but at the cost of economic and environmental commitments.

5.3. Battery Energy Storage Systems

Dynamic simulations were performed to evaluate the contribution of BESS to frequency stability. Three scenarios were compared: without BESS, with a 60 MW BESS, and with a 120 MW BESS, each with 710 MW output events. The battery system was modeled to be activated by a RoCoF threshold of -1 Hz/s, allowing it to be connected before the Automatic Load Shedding Scheme (ALSS) Step 0 actuation.

Figure 8, Figure 9 , and Figure 10 visually illustrate how the frequency response of SENI improves with the implementation of BESS in different projected months and years, demonstrating its effectiveness in mitigating the impacts of low inertia. Additionally, the BESS reduces renewable energy discharge to near-zero levels, except during specific periods, thereby enhancing both economic efficiency and environmental benefits. In Figure 8, the frequency response of the SENI is shown, in the event of an output of 710 MW, with and without a battery system – January 2025

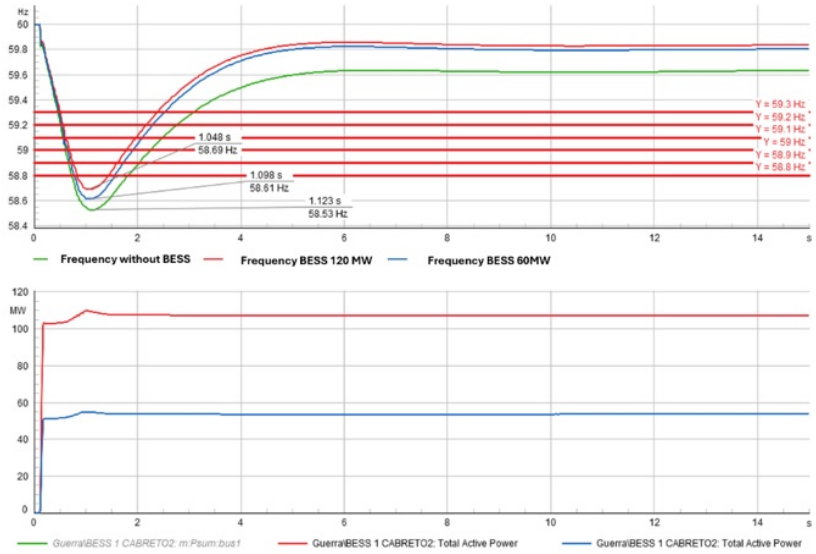


Figure 8. Frequency response of the SENI, in the face of a 710 MW event, with and without Battery System – JAN 2025

In Figure 9, the frequency response of the SENI is shown, in the event of an output of 710 MW, with and without the Battery System – July 2025

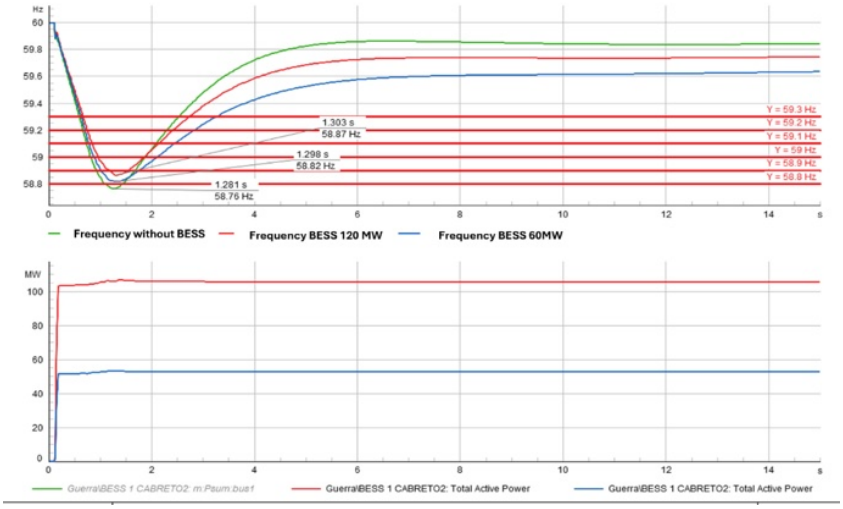


Figure 9. SENI Frequency Response, given a 710 MW output event, with and without Battery System – July 2025

Figure 10 shows the frequency of the SENI, in the event of an event of 710 MW, with and without a Batteries – JUN 2026

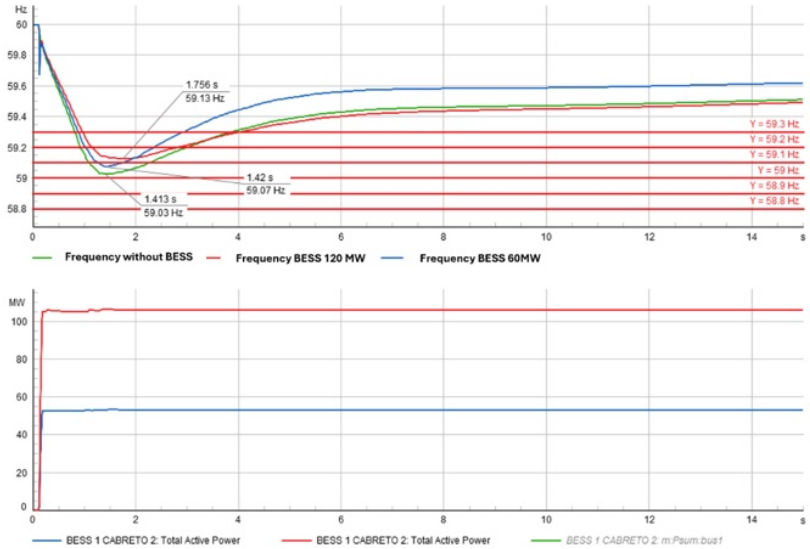


Figure 10. Frequency response of the SENI, in the event of 710 MW, JUN 2026

Simulations of 710 MW generation loss events for January 2025, July 2025, and June 2026 confirm that Battery Energy Storage Systems (BESS) improve frequency

stability under different operating conditions. In January 2025, with a demand of 2646 MW and inertia of 7.82 s, the frequency nadir dropped to 58.53 Hz without BESS, but improved to 58.61 Hz and 58.69 Hz with 60 MW and 120 MW BESS, respectively. In July 2025, despite higher demand (3082 MW) and slightly lower inertia (7.59 s), the nadir improved from 58.76 Hz without BESS to 58.82 Hz and 58.87 Hz with storage. Finally, in June 2026, under the highest demand (4135 MW) and inertia of 8.76 s, the nadir increased from 59.03 Hz without storage to 59.07 Hz and 59.13 Hz with BESS. These results show that the impact of storage is greater under low-inertia conditions, while in higher-inertia scenarios, the effect, although positive, is less pronounced. In all cases, BESS also slightly reduced nadir occurrence time, providing a faster system response to disturbances.

6. Conclusions, Future Work, and Limitations

This study shows that while the increasing integration of variable renewable energy (VRE) into the Dominican Republic's National Interconnected Electric System (SENI) supports sustainability, it also creates significant challenges for operational stability due to reduced system inertia. Maintaining minimum inertia levels above 17 GW·s requires dispatching synchronized thermal units, which are more costly, increasing overall operating costs, and limiting the use of low-cost renewable energy. In high renewable penetration scenarios, this can lead to substantial renewable energy curtailment, reducing both economic efficiency and environmental benefits.

Battery Energy Storage Systems (BESS) offer a viable, cost-effective solution. Simulations with 120 MW of BESS show clear improvements in frequency stability, reduce the need for advanced load-shedding schemes, and minimize renewable energy curtailment. Therefore, it is recommended to prioritize BESS implementation at this capacity, strengthen system stability studies as renewable penetration grows, and update operational codes to address these challenges. Continuous technical training, advanced monitoring and simulation tools, improved renewable generation forecasting, and exploring other operational flexibility resources are also essential.

The study has some limitations. It focuses solely on the Dominican SENI, so the findings are specific to the country's technical, regulatory, and renewable resource context and may not apply elsewhere. Only non-conventional renewable sources (wind and solar PV) were considered, excluding other potentially significant technologies. System inertia estimates relied on dynamic studies with specialized tools, whose accuracy depends on data quality. Finally, no full economic-financial analysis was conducted—only marginal and operational costs were considered—leaving room for future research on economic, regulatory, and investment aspects.

Acknowledgements:

To the Organismo Coordinador del Sistema Eléctrico Interconectado (OC) for the collaboration in the simulations of the proposed events. The authors acknowledge the support provided by the Thematic Network 723RT0150 "Red para la integración a gran escala de energías renovables en sistemas eléctricos (RIBIERSE-CYTED)", financed by the call for Thematic Networks of the CYTED (Ibero-American Program of Science and Technology for Development) for 2022.

References

1. M. A. Dominguez-Garabitos, V. S. Ocana-Guevara, F. Santos-Garcia, A. Arango-Manrique, and M. Aybar-Mejia, "Demand-Shifting Strategies to Optimize the Perfor-mance of the Wholesale Electricity Market: A Dominican Republic Study Case," IEEE Access, vol. 10, pp. 44362–44383, 2022, doi: 10.1109/ACCESS.2022.3169696.
2. P. Tielens and D. Van Hertem, "The relevance of inertia in power systems," Renewa-ble and Sustainable Energy Reviews, vol. 55, pp. 999–1009, Mar. 2016, doi: 10.1016/j.rser.2015.11.016.
3. J. Zhang and H. Xu, "Online Identification of Power System Equivalent Inertia Con-stant," IEEE Transactions on Industrial Electronics, vol. 64, no. 10, pp. 8098–8107, Oct. 2017, doi: 10.1109/TIE.2017.2698414.
4. N. Mago and ERCOT, "Inertia: Basic Concepts and Impacts on the ERCOT Grid," TEXAS, 2018. Accessed: Jul. 22, 2025. [Online]. Available: https://www.er-cot.com/files/docs/2018/04/04/Inertia_Basic_Concepts_Impacts_On_ERCOT_v0.pdf
5. C. Colbert, CAISO, and M&IP, "Frequency Response Phase 2 Issue Paper," Dec. 2016. Accessed: Jul. 22, 2025. [Online]. Available: https://www.caiso.com/Documents/IssuePa-per_FrequencyResponsePhase2.pdf
6. N. W. Miller, M. Shao, and S. Venkataraman, "California ISO (CAISO) Frequency Re-sponse Study Final Draft," Nov. 2011. Accessed: Jul. 22, 2025. [Online]. Available: https://www.caiso.com/Documents/Report-FrequencyResponseStudy.pdf
7. S. Tweed and Eirgrid, "Inertia Management on the Power Systems of Ireland and Northern Ireland," Mar. 2024. Accessed: Jul. 22, 2025. [Online]. Available: https://globalpst.org/wp-content/uploads/EIRGRID-G-PST-Inertia-Management-on-the-Power-Systems-of-Ireland-and-Northern-Ireland.pdf
8. EirGrid as the Transmission System Operator (TSO) for Ireland and SONI Ltd as the Trans-mission System Operator (TSO) for Northern Ireland, "Recommendation on pro-curement of Low Carbon Inertia Services (LCIS) Requirements and Procurement Ap-proach," Nov. 2022. Accessed: Jul. 22, 2025. [Online]. Available: https://consult.soni.ltd.uk/consulta-tion/consultation-low-carbon-inertia-service-lcis-competitive-procurement-0
9. A. Johnson and National Grid, "Grid Code Frequency Response Working Group Sys-tem Inertia Overview Background to System Inertia," 2010. Accessed: Jul. 22, 2025. [Online]. Available: https://www.nationalgrid.com/sites/default/files/documents/16890-Meet-ing%208%20-%20Inertia%20presentation.pdf
10. European Network of Transmission System Operators for Electricity (ENTSON-E), "In-ertia and Rate of Change of Frequency (RoCoF)," Dec. 2020. Accessed: Jul. 22, 2025. [Online]. Available: https://eepublicdownloads.entsoe.eu/clean-docu-ments/SOC%20documents/Inertia%20and%20RoCoF_v17_clean.pdf

11. P. Denholm, T. Mai, R. W. Kenyon, B. Kroposki, and M. O'malley, "Inertia and the Power Grid: A Guide Without the Spin," 2020, Accessed: Jul. 22, 2025. [Online]. Available: www.nrel.gov/publications.
12. T. O. Hodson, "Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not," Geosci Model Dev, vol. 15, no. 14, pp. 5481–5487, Jul. 2022, doi: 10.5194/gmd-15-5481-2022.
13. European Network of Transmission System Operators for Electricity (ENTSOE) and System Operations Committee, "2021 Load-Frequency Control Annual Report," Sep. 2022. Accessed: Jul. 25, 2025. [Online]. Available: https://eepublicdownloads.entsoe.eu/clean-documents/SOC%20documents/LFC/2021_ALFC_Report_FOR_APPROVAL.pdf
14. System Operator for Northern Ireland (SONI), "Winter Outlook 2024/25," 2024. Accessed: Jul. 25, 2025. [Online]. Available: https://cms.soni.ltd.uk/sites/default/files/media/documents/SONI-Winter-Outlook-2024-25.pdf
15. EIRGRID and SONI, "DS3 System Services Protocol-Regulated Arrangements DS3 System Services Implementation Project," Oct. 2024. Accessed: Jul. 25, 2025. [Online]. Available: https://cms.soni.ltd.uk/sites/default/files/media/documents/DS3-SS-Protocol-v4.1.pdf
16. National Grid ESO, "Past Frequency Events. A system operability framework doc-ument," 2019. Accessed: Jul. 25, 2025. [Online]. Available: https://www.neso.energy/document/156761/download
17. National Grid ESO, "A System Operability Framework document Operating a Low Inertia System," 2020. Accessed: Jul. 25, 2025. [Online]. Available: https://www.neso.energy/document/164586/download
18. European Network of Transmission System Operators for Electricity (ENTSOE), "Overview of Frequency Control in the Nordic Power System," 2024. Accessed: Jul. 25, 2025. [Online]. Available: https://eepublicdownloads.blob.core.windows.net/public-cdn-container/clean-documents/SOC%20documents/Nordic/2024/Overview_of_Frequency_Control_in_the_Nordic_Power_System.pdf
19. M. Asmine, C. Langlois, and N. Aubut, "Inertial response from wind power plants during a frequency disturbance on the Hydro-Quebec system – event analysis and vali-dation," IET Renewable Power Generation, vol. 12, no. 5, pp. 515–522, Apr. 2018, doi: 10.1049/iet-rpg.2017.0392.
20. Australian Energy Market Operator, "AEMO | Renewable Integration Study (RIS)." Accessed: Jul. 25, 2025. [Online]. Available: https://aemo.com.au/energy-systems/major-publications/renewable-integration-study-ris
21. Australian Energy Market Operator (AEMO), "Renewable Integration Study: Stage 1 report," Apr. 2020. Accessed: Jul. 25, 2025. [Online]. Available: https://aemo.com.au/-/media/files/major-publications/ris/2020/renewable-integration-study-stage-1.pdf
22. Superintendence of Electricity (SIE), "General Electricity Law No. 125-01 and its Implementing Regulations," Santo Domingo, 2012. Accessed: Jul. 29, 2025. [Online]. Available: https://ayuntamientomoca.gob.do/wp-content/uploads/2018/05/LeyGeneraldeElecctricidadNo.125-01.pdf

Agrivoltaic evaluation of horticultural crops for Balearic tourism: Comparative analysis of lettuce, tomato and zucchini

Alicia Nadal-Salas [1[0009-0001-1461-3727]], Marta Pons-Nieto [1[0000-0002-5702-1481]], Esperanza Dalmau [2[0000-0002-0078-0911]], Antoni Capellà-Ferretjans [2[0009-0002-0970-7945]] and Susana Simal [2[0000-0002-4883-3533]]

[1] Clúster de Transición Ecológica de las Illes Balears (TEIB), Palma, España
[2] Universitat de les Illes Balears (UIB), Palma, España
anadal@clusterteib.com
Susana.simal@uib.es

Abstract. Tourism, the main economic driver of the Balearic Islands, generates structural dependence on external food supply and intensifies energy demand. The potential of agrivoltaic systems to simultaneously enhance food security and energy sustainability in this insular context is examined through a structured review of the agrivoltaics literature and a mathematical simulation framework. The analysis focuses on three representative horticultural crops (lettuce, tomato, and zucchini), using consumption data from collaborating hotel facilities as the reference context.

The results show differentiated crop responses to partial shading: lettuce is the most suitable candidate for agrivoltaic integration, tomato maintains conditional viability subject to shading thresholds, and zucchini exhibits significant limitations. In parallel, the energy assessment indicates that agrivoltaics can contribute substantially to local electricity supply while maintaining agricultural production.

Overall, agrivoltaics emerges as a strategic approach to optimize land use, strengthen food supply resilience, and accelerate the decarbonization of the island's energy system, offering a pathway toward a more sustainable and resilient tourism model.

Keywords: Agrivoltaics, Tourism agriculture, Shade tolerance, Crop production.

1 Introduction

The Balearic Islands represent only 1% of Spanish territory [1], yet they receive more than 18 million tourists each year, making tourism their main economic driver [2]. This massive influx, concentrated in the high season, exerts extraordinary pressure on food

and energy supply systems, with visible impacts on land use, the local economy, and the environment.

One of the most recent studies on food self-sufficiency in the Balearic Islands places this value at around 11.6% [3], implying that nearly 88% of food consumed must be imported. Seasonality exacerbates this dependence, as peaks in tourist demand coincide with periods of lower local agricultural production, increasing the need for external transport and, consequently, the carbon footprint of the food system.

In the energy field, the situation is similar. In 2024, only 14.6% of the electricity consumed in the islands came from renewable sources [4]. Demand is concentrated in summer, driven by the use of air conditioning and tourism-related activities, which requires maintaining oversized infrastructure, increases supply costs, and raises greenhouse gas emissions.

In addition, the area of usable agricultural land has been steadily decreasing, having fallen by 16% over the past decade [5]. In an island territory with limited space and subject to urban pressures, this loss further undermines the agricultural sector's capacity to respond to the food demand generated by tourism.

In this context, agrivoltaics, the integration of agricultural and photovoltaic production within the same area, emerges as a promising strategy to generate clean energy without displacing agricultural activity, while improving water efficiency and protecting crops from thermal stress under high solar radiation conditions. Its potential is particularly relevant in island territories such as the Balearics, where solar radiation is abundant and dependence on external resources is critical.

Based on these premises, this research aims to examine the potential of agrivoltaics in the Balearic Islands as a tool to improve land-use efficiency, enable renewable energy generation, and help meet the growing food demand driven by tourism in Mediterranean regions. To this end, the study analyses the performance of three representative horticultural crops (lettuce, tomato and zucchini) through an integrated simulation model that estimates the synergies between photovoltaic generation and agricultural yield.

2 Conceptual framework and state of the art

2.1 Agrivoltaics: principles, typologies and agronomic factors

According to the definition provided in *DIN SPEC 91434:2021-05*, the concept of agrivoltaics refers to the shared use of the same area of land for agricultural and photovoltaic production. One of its core principles is the primacy of agricultural activity: power generation must adapt to crop cycles and minimize interference with agricultural operations.

Agrivoltaic systems are generally implemented through two fundamental strategies: row (or stripe) distribution, where panels are sandwiched between crops, or *overhead systems*, where panels are superimposed on crops. They can also be classified according to the arrangement of the solar panels: (1) horizontal, including fixed panels with a specific tilt angle (the configuration considered in this study) or tracking panels with one axis (north-south orientation) or two axes (north-south and east-west orientation). (2) Vertical, characterized by the use of bifacial panels, which exceed the two previous aspects but result in lower electricity generation.

The agricultural integration of PV modules is closely dependent on the physiological characteristics of the crops, as the modules act as an additional source of shade, altering growth cycles. Several studies have pointed out that certain legumes (such as chickpeas [6]) and some fruit trees (such as the olive tree [7]) show low tolerance to shade, with negative consequences for soil productivity and resource use efficiency. Conversely, shading can reduce soil temperature, reduce water evaporation, and decrease crop water requirements, providing benefits for coping with extreme local weather conditions. This is especially relevant in the Balearic Islands, where warm summers are characterized by minimum temperatures exceeding 27 ºC in Mallorca and 24 ºC in the rest of the islands during July and August, while the maximum temperatures in both months are close to or above 40 ºC [[8], [9]]. In addition to these conditions, adverse weather phenomena such as strong winds are added, especially in Menorca due to its flat topography.

2.2 Agrivoltaic viability by crop type

Based on the effects of shading described above, the viability of agrivoltaics in the Balearic context is also analysed by considering socio-economic factors, particularly the relevance of crops in both the tourist diet and local horticultural production [10]. Accordingly, the study focuses on three representative horticultural species – lettuce, tomato and zucchini – which exhibit differentiated physiological responses to shading (high, moderate and low, respectively), allowing a comparative analysis of their performance under agrivoltaic systems.

- **Lettuce.** This crop is a fast-growing crop with good tolerance to shading, capable of maintaining stable yields even under reduced radiation conditions. These characteristics, together with its short growth cycle, enhance its nutritional and regional significance, particularly during the peak season when the demand for fresh produce increases. Field trials support this adaptive capacity: under varying levels of shading generated by photovoltaic panels, the crop maintains acceptable yields and improves its efficiency in the use of diffuse radiation [11].

- **Tomato.** This crop shows moderate suitability for agrivoltaic systems, with performance strongly influenced by the design of the photovoltaic structure (panel height, shading degree, and spatial arrangement). In this regard, it has been observed that moderate levels of shading can cause a significant decrease in yield and fruit quality when the reduction in radiation is excessive [12]. However, the crop may benefit from a more stable microclimate, characterized by lower extreme temperatures and more efficient water use [10].

- **Zucchini.** A crop that is only viable in agrivoltaic systems with low levels of shading or with structures that allow adequate light penetration [13]. Studies show that a reduction in incident radiation affects physiological parameters such as colour and sugar content [14], and that yield under full sunlight is significantly higher than under photovoltaic modules, even when these exhibit varying degrees of transparency [15].

Overall, the selection of these three crops provides a representative framework for assessing the agronomic and energetic viability of agrivoltaic systems in the Balearic context, highlighting both integration opportunities and key technical constraints for territorial planning.

3 Methodology

The methodological design of the study was structured in three interrelated phases that combine territorial analysis, energy modelling and agronomic assessment. The purpose was to examine in an integrated way the potential of agrivoltaic systems to strengthen the food and energy security of the Balearic Islands, taking as a reference the three selected horticultural crops.

Based on this premise, the methodology articulates an approach that combines the collection and analysis of geospatial and statistical data on local agricultural production, the estimation of food demand derived from tourism and the simulation of agrivoltaic scenarios using a model developed specifically for this study. This approach makes it possible to quantify the synergies and commitments between agricultural production and photovoltaic generation, as well as to assess the feasibility of their implementation in the Balearic territorial context.

3.1 Local agricultural production of the three crops

The analysis of local horticultural production constituted the first phase of the study and aimed to determine the cultivated area and the average annual yield of the selected crops in the Balearic Islands. This information served as a baseline for assessing the capacity of local agriculture to meet the food demand associated with tourism.

To this end, geospatial metadata corresponding to the 2023 and 2024 agricultural seasons from the Geographic Information System of Agricultural Plots (SIGPAC), managed by the Ministry of Agriculture, Fisheries and Food (MAPA, 2025), were collected and processed. This database provides detailed information on plot boundaries, their unique identification, agricultural land uses, and complementary cartography (national orthophotos and reference maps).

Specific scripts were developed in MATLAB (version R2025a, MathWorks) in order to automate the download, classification and analysis of the data, allowing to accurately quantify the area assigned to each crop on each island of the archipelago and estimate its total annual production.

The yield values per hectare were obtained from the statistical reports of the Institute for Agri-Food and Fisheries Research and Training of the Balearic Islands (IRFAP) for the years 2021, 2022 and 2023, which provide official and consolidated data on agricultural, livestock and fisheries production in the region.

Overall, this phase made it possible to establish a quantitatively and territorially disaggregated reference of Balearic horticultural production, an essential basis for the subsequent analysis of food self-sufficiency and agrivoltaic integration.

3.2 Estimated tourism demand by crop

The objective of this phase was to quantify the food demand generated by tourism in the Balearic Islands during the year 2024 for the three products. The estimation was made based on the product between the average consumption per overnight stay (kg) and the total number of overnight stays registered in the same period, thus allowing to directly relate tourist activity with the consumption of fresh food of local origin.

The adoption of this approach responds to the limited availability of public and experimental data at the island level, both from national and regional sources. The lack of official disaggregated records on tourist food consumption or on local supply flows makes it difficult to calibrate specific empirical models for the Balearic context. Therefore, an indirect estimate was chosen, based on representative empirical information obtained from a real case study and complemented with official tourism statistics, which allows generating a robust and coherent analysis framework with the objectives of the work.

Firstly, to establish the average consumption per overnight stay of the three products, two sources of information were used: (1) actual consumption data in 2024 on the kilograms of these products used for catering by customers provided by Garden Hotels during its active period, and (2) the estimated number of overnight stays registered in this hotel chain during the same period.

The choice of Garden Hotels is justified by its representative profile of the predominant holiday accommodation model in the archipelago, which allows its consumption patterns to be considered as a valid approximation to the average behaviour of the Balearic tourism sector.

To calculate the number of overnight stays in 2024 in Garden Hotels (p), the following formula was applied:

$$p\,(2024) = g \cdot t \cdot c \tag{1}$$

where g represents the general occupancy rate of tourist accommodations in the Balearic Islands for the designated period, obtained from the 2024 annual report published by AETIB [2]; *t* corresponds to the operating period of Garden Hotels, set at 200 days, which reflects not only the days of effective activity of the hotel complex, but also the average duration of opening of tourist establishments in the archipelago (~7-8 months); and *c* is the number of tourist accommodations or available beds in this hotel chain, obtained from the 2023 Sustainability Report published by the chain in 2024.

Once the estimated number of overnight stays in the Garden Hotels chain for 2024 was obtained, the consumption data provided by the company were used to calculate the average amount of each product consumed per overnight stay. This value was then extrapolated to the entire Balearic archipelago using the total number of overnight stays reported by the AETIB in its 2024 annual tourism report [2], allowing the estimation of total annual consumption attributable to tourism.

3.3 Agrivoltaic simulator

In order to analyze agricultural and energy production under different photovoltaic configurations in an integrated way, a proprietary agrivoltaic simulator called ATERRA v.01 was developed, implemented in MATLAB (R2025a, MathWorks). This model was designed to estimate the spatial and temporal distribution of solar irradiance in agrivoltaic systems, incorporating geometric, climatic and agronomic parameters specific to the Balearic context.

The simulator allows you to define key variables such as the geometry of the photovoltaic array, the simulation period, the geographical location, the resolution of the terrain and the surface albedo, through an interactive graphical interface. The terrain is discretized into a two-dimensional mesh, and the solar position and shading patterns are calculated at 15-minute intervals using a lightning intersection algorithm adapted to elevated PV structures.

For incident radiation estimation, ATERRA uses a PVGIS input mode, in which hourly values of direct normal (DNI), diffuse horizontal (DHI), and global horizontal (GHI) radiation are extracted from the PVGIS typical weather year (TMY) dataset and interpolated with the 15-min temporal resolution. The simulation uses the isotropic sky

model of Liu and Jordan (1960) [16] to calculate the diffuse and reflected fraction on inclined surfaces, integrating the effects of shadow, transparency of the modules and visibility of the sky at each point of the terrain. Cumulative irradiance values are aggregated by panel and by terrain cell at monthly intervals.

Crop yields were estimated based on relative solar radiation (RSR) received at ground level. Each crop was assigned to a predefined shade tolerance group based on the results of the literature. The percentage of shading at each point of the terrain was calculated from the maximum accumulated irradiance in the mesh. This allowed the derivation of the distribution of RSR across the terrain. Crop yield (%) was then estimated using the empirical equations proposed by Laub et al. (2022) [17] specific to each tolerance group.

Together, the ATERRA simulator provides a flexible tool to link photovoltaic geometry with agricultural productivity, allowing for quantifying the interactions between energy generation and crop yield. Although it does not replace experimental validation, it offers a solid and reproducible quantitative basis to guide the design of future agrivoltaic trials and pilot projects in Mediterranean territories.

4 Results

4.1 Tourist demand and local horticultural production

The data provided by Garden Hotels showed that during the active period of 2024, total consumption amounted to 11,457.10 kg for lettuce, 59,596.47 kg for tomato, and 8,471.27 kg for zucchini. Applying Equation (1), a value of p = 571,773.6 was estimated in 2024 for Garden Hotels. Based on the report published by AETIB [2], it was assumed that 120 million overnight stays were recorded in the Balearic territory in 2024.

The combination of these results made it possible to estimate a total annual demand in the Balearic Islands of 2,404.5 t of lettuce, 12,507.7 t of tomato and 1,777.9 t of zucchini, evidencing the predominance of tomatoes over the rest of the vegetables analysed. Figure 1 shows the comparison between local production and the estimated demand of tourists for food consumption of the three crops in the Balearic Islands.

Calculations reveal local supply deficits of 57.0 % for tomatoes and 60.9 % for lettuce, while zucchini shows a minor mismatch of 18.7%, remaining close to equilibrium. These findings are consistent with previous studies that have pointed to the structural dependence of the Balearic food system on external sources of supply [3], and reinforcing the idea that tourism acts as a pressure factor that accentuates this dependence.

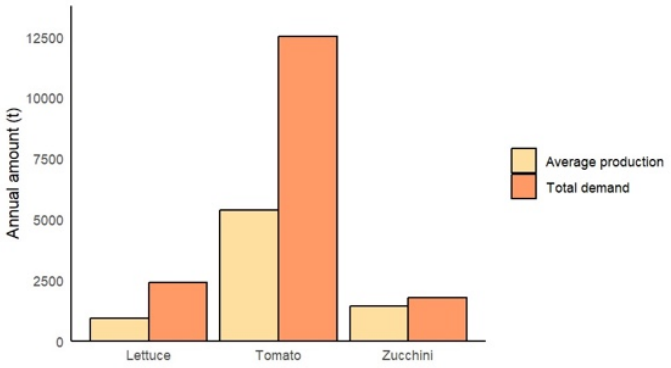


Figure 1. Local production and estimated tourists demand for the food consumption of lettuce, tomato and zucchini in the Balearic Islands in 2024.

4.2 Simulation of a case study

To analyse the simultaneous integration of agricultural production and power generation, a representative simulation was developed with the ATERRA model, based on the parameters summarized in Table 1, which includes the characteristics of the site, the type of crop and the specifications of the photovoltaic system [18]. The photovoltaic configuration was taken from a project planned on a small plot in Palma, Mallorca (Spain) by the company ESTEL, a specialist in photovoltaic installations. ESTEL used the commercial software PV*SOL to obtain detailed energy-generation parameters, which served as reference data for the present simulation. Figure 2 shows the representation of the simulated agrivoltaic plant.

Table 1. Input parameters related to the photovoltaic installation used in the simulation.

Input	Value	Input	Value
Available area (ha)	0.273	**Number of columns**	57
Grid (10-250)	50	**Number of rows**	5
Latitude	39.56101	**Row spacing (m)**	5
Longitude	2.70880	**Panel transparency (%)**	0
Albedo (0-1)	0.3	**Panel efficiency (%)**	23.5
Azimut (º)	189	**Panel height (m)**	4.764*
Crop type:	lettuce, tomato & zucchini	**Panel width (m)**	1.134
Base yield (kg/ha)	24,370; 23,330; 23,280	**Panel elevation (m)**	4
Hotel type (stars)**	4	**Panel tilt – beta (º)**	30
Hotel food demand (kg/y)	6,589 ; 5,472 ; 4,279	**Hotel energy demand (kWh/y)**	1,914,500
Start day	01-Jan-2024	**End day**	31-Dec-2024

* To represent the configuration provided by ESTEL, which includes two strings of modules arranged both in columns and rows, the panel height parameter was doubled in the simulation to more accurately approximate the actual layout.

** The simulator allows the establishment's energy demand (kWh/y) to be incorporated and, based on reference data, to estimate the relative efficiency according to the hotel category, providing a comparative indicator of the potential impact of the agrivoltaic system.

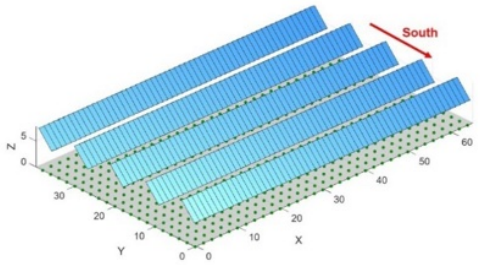


Figure 2. Depiction of the simulated agrivoltaic plant in ATERRA simulator.

Irradiance distribution at terrain level. The cumulative spatial distribution of the irradiance on the ground (Figure 3) was estimated considering the shading effects generated by the design of the photovoltaic panel. A marked heterogeneity is observed in the incident irradiance, with higher values in the areas between rows and significant reductions below the projection of the photovoltaic structures. The simulated values ranged from a maximum of 3,384 kWh/m² to a minimum of 1,028 kWh/m², reflecting the contrast between fully exposed areas and those subjected to higher levels of shading.

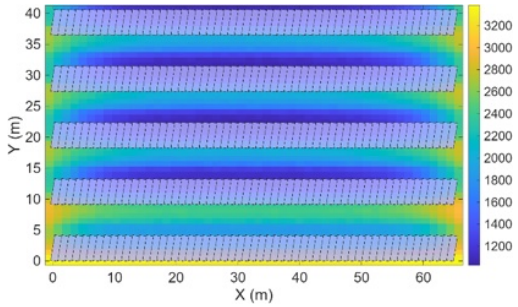


Figure 3. Accumulated irradiance (kWh/m2) distribution at terrain level using the ATERRA agrivoltaic simulator, the PVGIS TMY data and input data from Table 1.

Energy production from PV panels. The monthly generation of electricity as a function of the irradiance distribution, the geometry of the panel and the efficiency of the module shows a clear seasonality, with maximums during the summer months (June-August) and minimums in winter (December-February), in accordance with the solar resource available in Mediterranean latitudes (Figure 4).

Quantitatively, ATERRA produced slightly higher values than PV*SOL, with deviations generally less than 10%, more pronounced in periods of maximum irradiance. These differences can be attributed to the treatment of diffuse radiation and the parameterization of shading losses in each model. However, the satisfactory agreement between the two tools supports the robustness of the simulator as a suitable option for the design and preliminary evaluation of agrivoltaic systems, while providing greater flexibility for the integration of crop-related variables into energy analyses.

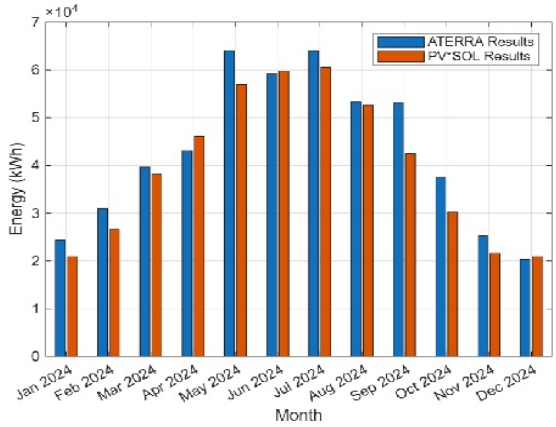


Figure 4. Monthly energy production (kWh) estimated by the ATERRA simulator and by the commercial software PV*SOL.

Crop-specific yield responses under shading conditions. To evaluate the influence of crop type on agrivoltaic yield, simulations were performed using identical system parameters varying the species grown (lettuce, tomato and zucchini).

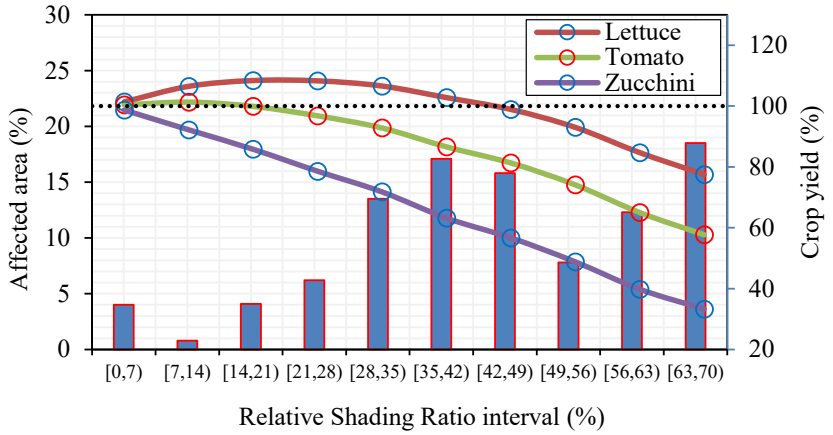


Figure 5. Simulated results (ATERRA simulator) of crop-specific yield responses.

Figure 5 shows the relationship between the *relative shading ratio (RSR)* and the *corresponding crop yield* response for lettuce, tomato, and zucchini, along with the *proportion of field area* that falls within each RSR interval. The blue bars (left y-axis) indicate the percentage of the cultivated area affected by each shading interval, expressed as the proportion of the total area of the field. These values illustrate how often each shading condition occurs within the agrivoltaic design. Solid lines with open symbols (right y-axis) represent the relative crop yield (% of baseline yield in full sunlight) as a function of RSR. Each curve corresponds to a crop: lettuce (red), tomato (green) and zucchini (blue). The dotted horizontal line marks the 100% yield (baseline in unshaded conditions).

Lettuce exhibited the most favourable yield, with yield increases of up to 108% in moderate shade conditions (RSR 14-28%). Beyond the RSR values of 35%, yields progressively declined, although even in the most restrictive range (RSR 63-70%) they remained around 77% of the reference value.

The tomato showed an intermediate response. Under low shading (RSR < 21%) yields remained close to baseline, but notable reductions appeared from RSR 35-42%, where they fell to around 87%. Under intense shading (RSR 63-70%), yields fell further, reaching about 58% of the benchmark.

Zucchini was the most sensitive crop, with rapid yield losses as shade increased. While in low shading intervals (RSR < 14%) yields were still above 90%, in moderate

conditions (RSR 35-42%) they fell to around 63%, and in the shaded scenario (RSR 63-70%) they fell dramatically to just 33% of the baseline.

4.3 Comparative analysis of agrivoltaic scenarios

The calculations were made considering the basal yields of each crop and the percentage variations as a function of shading intensity (Figure 5), which allowed estimating the total production of lettuce (6,589 kg), tomato (5,472 kg) and zucchini (4,279 kg) in the study area (0.273 ha).

In terms of energy, the estimated monthly production with the simulator resulted in an annual generation of 514,863 kWh. This value was compared with reference data on the average annual electricity consumption of hotels of similar category, estimated at 1,914,500 kWh/year. Based on this comparison, the simulated agrivoltaic system would cover approximately 27% of the establishment's energy demand.

At the same time, the simulated agricultural production was compared with the annual consumption data provided by the same hotel. The results indicate that the modelled system would cover approximately 57 % of the demand for lettuce, 9 % of the demand for tomatoes and 50 % of the demand for zucchini in the study area, thus determining the degree of self-sufficiency achievable in each case.

Finally, the results were expressed in relation to the available area, reaching values of 24,186 kg/ha for lettuce, 20,085 kg/ha for tomatoes and 15,706 kg/ha for zucchini, together with 1,887,000 kWh/ha for total energy production. These indicators not only allow for standardized comparisons between different scenarios but also provide an objective basis for assessing land-use efficiency under a dual production approach.

5 Discussion

This study offers an exploratory approach to the potential of agrivoltaics as a transition tool towards more sustainable agri-food and energy systems in island contexts. The methodology applied is based on empirical modelling and mathematical simulation, which allows the integration of information from heterogeneous sources and the evaluation of representative scenarios even in the absence of local experimental data. This approach responds to a structural reality of the Balearic Islands, where the scarcity of reliable and high-resolution data on agricultural productivity and solar radiation at crop scale limits the validation of physical models.

Although the model does not incorporate specific agronomic variables that directly affect plant growth such as irrigation, fertilization, type of substrate or pest incidence, its basis in empirical relationships derived from bibliographic meta-analyses allows a reasonable representation of the average response of crops to shading. This approach

provides a comparative analysis tool that identifies trends and feasibility thresholds, constituting a solid and methodologically coherent proof of concept.

The results obtained are consistent with the international literature: lettuce shows a positive behaviour under moderate shade, tomato maintains acceptable yields in intermediate conditions and zucchini shows a marked sensitivity to direct radiation reduction. These coincidences reinforce the validity of the model and its applicability to the Balearic context, suggesting that agrivoltaic systems can be effectively integrated into crops with greater tolerance to shade.

The added value of this work lies in its integrative approach. The combination of data on tourism consumption, agricultural production and energy yield allows the development of a reproducible and scalable methodology to analyse food self-sufficiency in island tourist destinations. In this way, the study not only provides quantitative results, but also offers an operational framework to guide public policies, plan pilot projects and design future research aimed at the experimental calibration of the model.

Overall, the agrivoltaic simulation developed demonstrates that, even under simplified assumptions, it is possible to obtain a coherent and useful view of the interaction between agricultural production and energy generation. This approach represents a solid starting point for the design of agrivoltaic strategies adapted to island territories with space, data and resource constraints, such as the Balearic Islands.

6 Conclusions

Agrivoltaic systems can make a significant contribution to the sustainable supply of food and energy in the Balearic Islands, strengthening territorial resilience and reducing dependence on external resources. The results highlight lettuce as the species with the greatest capacity for integration, while tomatoes and zucchini have a more limited viability.

To realize this potential, agrivoltaic deployment must be adapted to local agronomic and socioeconomic conditions and accompanied by public policies and coordinated investment. Likewise, the developed methodology and the ATERRA simulator can serve as preliminary tools to support the decision, guiding crop selection and energy planning.

Future research and pilot projects could help expand the range of crops considered, scale systems to real operational dimensions, and integrate agrivoltaics into regional planning strategies that recognize the multifunctionality of the land. In this sense, technology can serve as a strategic tool to move towards a more sustainable and resilient tourism model in the Balearic Islands.

Acknowlegments. The authors would like to acknowledge the financial support of the *Agència d'Estratègia Turística de les Illes Balears (AETIB)* through the *Govern de les Illes Balears*, in the framework of the *Plan de Recuperación, Transformación y Resiliencia* funded by the

European Union (NextGenerationEU). Sincere appreciation is also extended to *Garden Hotels* and *Estel* for their collaboration and for making available key data that enabled the development of this study.

7 References

1. Instituto Geográfico Nacional (IGN) Información de datos geográficos. https://www.ign.es/web/ane-datos-geograficos/-/datos-geograficos/datosPoblacion?tipoBusqueda=CCAA. Accessed 21 Aug 2025
2. Agència d'Estratègia Turística Illes Balears (AETIB) El turismo en datos 2024. https://www.caib.es/sites/estadistiquesdelturisme/es/nota_coyuntura_turistica-36521/. Accessed 7 Nov 2025
3. Agroassessor, Govern de les Illes Balears (2020) Un estudio revela el consumo agroalimentario en las Baleares. https://www.caib.es/pidip2front/jsp/es/ficha-convocatoria/strongspan-stylecolor1c697dcovid-19-spanun-estudio-revela-que-ennbsplas-baleares-se-consumioacute-un-20-maacutes-de-producto-local-durante-los-meses-de-confinamientonbspstrong. Accessed 22 Aug 2025
4. Red Eléctrica (2024) El sistema eléctrico en las Islas Baleares 2024
5. INEbase / Agricultura y medio ambiente /Agricultura /Encuesta sobre la estructura de las explotaciones agrícolas / Resultados. https://www.ine.es/dyngs/INEbase/es/operacion.htm?c=Estadistica_C&cid=1254736176854&menu=resultados&idp=1254735727106#_tabs-1254736194989. Accessed 10 Sep 2025
6. Verghis TI, McKenzie BA, Hill GD (1999) Effect of light and soil moisture on yield, yield components, and abortion of reproductive structures of chickpea (Cicerarietinum), in Canterbury, New Zealand. N Z J Crop Hortic Sci 27:153–161. https://doi.org/10.1080/01140671.1999.9514091;REQUESTEDJOURNAL:JOURNAL:TNZC20;WGROUP:STRING:PUBLICATION
7. Mouhib E, Fernández-Solas Á, Pérez-Higueras PJ, et al (2024) Enhancing land use: Integrating bifacial PV and olive trees in agrivoltaic systems. Appl Energy 359:122660. https://doi.org/10.1016/J.APENERGY.2024.122660
8. Ministerio para la Transición Ecológica y el Reto Demográfico (2024) Avance climatológico mensual - Julio 2024 Illes Balears
9. Ministerio para la Transición Ecológica y el Reto Demográfico (2024) Avance climático mensual - Agosto 2024
10. Ministerio de Agricultura P y A Consumo alimentario extradoméstico. https://www.mapa.gob.es/es/alimentacion/temas/consumo-tendencias/panel-de-consumo-alimentario/hosteleria-y-restauracion
11. Marrou H, Wery J, Dufour L, Dupraz C (2013) Productivity and radiation use efficiency of lettuces grown in the partial shade of photovoltaic panels. European Journal of Agronomy 44:54–66. https://doi.org/10.1016/J.EJA.2012.08.003
12. Abdel-Mawgoud AMR, El-Abd SO, Singer SM, et al (1996) Effect of shade on the growth and yield of tomato plants. Acta Hortic 434:313–319. https://doi.org/10.17660/ACTAHORTIC.1996.434.38

13. Del Ángel-Hernández M, Zermeño-Gonzalez A, Melendres-Alvarez AI, et al (2017) Characteristics of a tunnel cover effect on radiation, chlorophyll and zucchini yield. Rev Mex De Cienc Agric 8:1127–1142. https://doi.org/10.29312/REMEXCA.V8I5.113
14. Consumer Study of Agrivoltaics Food Products Including Tomato, Basil, Potato, Bean, and Squash - ProQuest. https://www.proquest.com/openview/4af147150e6f97786d424b0b29f6908f/1?pq-origsite=gscholar&cbl=18750&diss=y. Accessed 4 Sep 2025
15. Hickey T, Uchanski M, Bousselot J (2024) Vegetable crop growth under photovoltaic (PV) modules of varying transparencies. Heliyon 10:e36058. https://doi.org/10.1016/j.heliyon.2024.e36058
16. Liu BYH, Jordan RC (1960) The interrelationship and characteristic distribution of direct, diffuse and total solar radiation. Solar Energy 4:1–19. https://doi.org/10.1016/0038-092X(60)90062-1
17. Laub M, Pataczek L, Feuerbacher A, et al (2022) Contrasting yield responses at varying levels of shade suggest different suitability of crops for dual land-use systems: a meta-analysis. Agron Sustain Dev 42:1–13. https://doi.org/10.1007/S13593-022-00783-7/FIGURES/5
18. LONGI Hi-MO X10 Explorer LR7-72HVH 635M

Estimation of Heat Transfer Coefficients In A Dual Falling Film Condenser

Ortíz-Huerta A.[0000-0002-3523-718] and Morales L.I.[0000-0001-7500-6202]

Facultad de Ciencias Químicas e Ingeniería, Universidad Autónoma del Estado de Morelos, Av. Universidad No. 1001. Col Chamilpa, Cuernavaca, Morelos, México
huerta_aoh@uaem.mx, laura.morales@uaem.mx

Abstract. In this study, an empirical correlation for the Nusselt number was developed using the least-squares method implemented in MATLAB to determine the Nusselt number and calculate the local heat transfer coefficient of a dual falling-film condenser. The correlation was established based on the experimental results of 22 tests performed on a dual helical condenser, with water serving as the working fluid. The proposed equation adopts the general form of Nusselt number correlations as a function of the dimensionless Reynolds and Prandtl numbers. The Reynolds number was defined for a helical geometry, employing the critical Reynolds definition. The maximum deviation between the predicted and experimental values obtained from the correlation was 0.28%, with a correlation coefficient of $R^2 > 0.999$. The correlation that minimizes prediction error and can be used reliably in the design and optimization of such condensers. This correlation can be applied in the design of equipment with identical geometry and for various refrigeration applications.

Keywords: Dual Condenser, Falling Film, Correlation, Nusselt Number, Heat Transfer Coefficients.

1 Introduction

Energy is a fundamental element for human development, as it drives industrial, agricultural, technological, and social processes that sustain global economic dynamics. Over the past decades, increasing energy demand has been closely linked to improvements in quality of life and economic growth. However, the intensive use of conventional energy sources, particularly fossil fuels, has caused a significant environmental imbalance. Numerous studies indicate that, while energy consumption promotes progress, it also contributes substantially to environmental degradation, highlighting the urgent need to transition toward more sustainable energy models [1]. In this context, industrial processes often involve significant heat losses, which reduce overall efficiency and increase operational costs and environmental impact. Implementing heat recovery systems is crucial for capturing and reusing this energy, leading to substantial energy savings, reduced fuel consumption, and lower greenhouse gas emissions. By transferring recovered energy to preheat incoming streams or drive secondary processes, facilities can improve overall thermal efficiency. Recent studies have demonstrated that the integration of effective heat recovery technologies, such as heat exchangers and falling-film condensers, can achieve energy savings of up to 30–40% in

certain industrial sectors, emphasizing both their economic and environmental benefits [2].

The body of literature addressing dual systems—devices that integrate two unit operations within a single shell—remains relatively scarce. Reported examples include an absorber–evaporator with vertical tubes operating under falling film conditions [3], another absorber–evaporator with helicoidal tubes [4], an absorption machine with a duplex system arranged as horizontal tubes in a shell [5], and an absorption heat transformer with duplex generator–condenser and absorber–evaporator systems operating on helical tubes under falling film conditions [6,7]. Delgado-Gónzaga et al [8] evaluate the behavior of a novel dual component: generator-condenser applied to a single-stage absorption heat transformer. The results indicate a good coupling between the GE-CO and the AHT, with maximum values of COP and desalinated water flow of 0.49 and 3.91×10−4 kg/s. The maximum specific desalination rate (SDR) of the AHT-WP was 1.10×10−4 kg/s/kW.

Falling-film helical coil condensers (FFHCCs) have emerged as a particularly effective solution for industrial heat recovery applications, including refrigeration, desalination, and chemical processing, due to their compact design and superior heat transfer efficiency [9]. The helical geometry induces secondary flows within the liquid film, enhancing turbulence and promoting uniform heat distribution along the coil surface, which improves condensation rates and maximizes thermal energy recovery from exhaust streams [10,11]. The performance of these condensers is strongly influenced by the flow dynamics of the liquid film and the thermal properties of the working fluid [12]. By optimizing the Nusselt number through empirical correlations, engineers can design FFHCCs tailored to specific process conditions, reducing energy consumption, lowering operational costs, and supporting sustainable industrial practices [13]. Essalhi et al. [14] designed a helical condenser with a thermal load of 10.5 kW, employing a water–lithium bromide solution as the working fluid. Empirical correlations were utilized to determine both local and overall heat transfer coefficients. The study reported an external convective coefficient of 8458 W/m²·°C, an internal convective coefficient of 2303 W/m²·°C, and an overall heat transfer coefficient of 1043 W/m²·°C, corresponding to a heat transfer area of 3.082 m². The proposed helical condenser configuration demonstrates significant potential for domestic applications, owing to its compact design and minimal structural weight. Liu et al [15] conducted a study to improve the performance of an air conditioning system with an independent dual evaporative condenser. The results showed that the COP increased by up to 13.1%, and that the ambient temperature affected the COP value and the heat transfer coefficient was 232 to 409 W/m^{2}K. Nannan and Shuhong [16] developed a numerical model in MatLab for a helical condenser submerged in a tank, integrated with a heat pump. The model defined the flow conditions at the interface of the two-dimensional helicoid, and heat transfer was simulated using CFD. The results demonstrated good agreement with experimental data. The study reported that the heat transfer coefficient increased by 19.06% relative to the helical diameter, while the coefficient of performance (COP) improved by 3.97%. Hernández-Magallanes et al [17] realized an experimental study with a 5 kW dual helical condenser, obtaining local heat transfer coefficients with values between 1288 and 1615 W/m^2°C and global heat transfer coefficients between 441

and 614 W/m^2°C, which are similar to the values reported in the literature for equipment with helical geometry.

The bibliographic review indicated above, shows the relevance and limitation about helical coil condensers and the scarcity of papers about nested helical coil condensers adapted to absorption heat transformer. In this work, a correlation for the Nusselt number was developed for a dual helical falling-film condenser. The equipment was experimentally evaluated, and 22 data points were obtained for both the Nusselt number and the local film heat transfer coefficient. The correlation was derived using the least-squares method, achieving good agreement and reproducibility with the experimental data. This study provides a correlation that can be applied to the design of equipment with helical geometry and falling-film condensation.

2. Materials and Methods

2.1 Dual Condenser Device and Experimental Test

The dual condenser consists of two shells joined by a screwed flange, the evaporator is located in the lower part while the condenser is located in the upper part to facilitate the rise of steam (see Fig. 1). The shells have a diameter of 0.32 m and a height of 0.33 m each. On the outside, the shell has 3 borosilicate peepholes to observe the internal process. The condenser has a condensate hood which allows the passage of steam and prevents the return of condensate to the evaporator. In the internal part of the condenser there is a bank of concentric (nested) helical coil smooth tubes formed by 7 caliber 18 stainless steel 304, pipes of 0.0127 m (0.5 in), on which condensation occurs. Cold water from a commercial cooler (chiller) flows from the inside of the tubes, while steam from the evaporator condenses on the outside. Fig. 1 shows the dual condenser device, as well as its internal and external component.

A total of 22 experiments were conducted to characterize the dual-condenser system, varying source temperatures (65, 75, and 80 °C) and volumetric flow rates (10–12-14-16-18-20 L/min) in Table 1 shows experimental data. The experimental setup consisted of a perforated steel frame (1.02 × 1.4 × 0.51 m) equipped with calibrated instrumentation: four flow meters, nine type-J thermocouples, a pressure transducer, a hand-held vacuometer, and two positive displacement pumps (Table 2). Data were acquired using a Keysight system with a 20-channel multiplexer and BenchLink Data Logger 3 software at 10 s intervals. The test rig and process diagram are shown in Figure 2. Armaflex insulation was applied to minimize thermal losses.

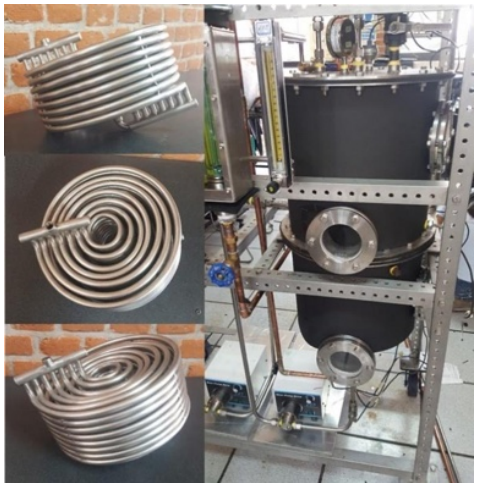


Fig. 1. Dual Condenser Device.

Table 1. Experimental conditions of the experimental dual condenser.

Test Number	T_{source} (°C)	Hot Water (L/min)	Absolute Pressure (kPa)	$T_{e,int}$ (°C)	$T_{s,int}$ (°C)	$T_{e,ext}$ (°C)	$T_{s,ext}$ (°C)
1	65	16	12.2	58.7	52.6	18.9	31.3
2	65	18	12.7	59.1	53.0	19.7	32.7
3	65	20	13.5	59.8	53.0	21.2	34.3
4	65	14	11.7	58.2	51.6	18.2	29.5
5	75	12	23.7	68.8	57.7	17.4	28.7
6	75	16	25.2	69.8	60.7	19.5	31.9
7	75	10	23.9	69.2	58.4	16.4	29.6
8	75	18	26.1	70.7	60.0	17.4	30.5
9	75	14	24.9	69.6	59.8	18.9	30.4
10	75	20	25.9	70.6	60.2	17.7	30.1
11	80	10	26.9	71.2	61.2	20.9	37.6
12	80	12	27.5	71.1	61.5	20.4	37.3
13	80	14	28.2	71.0	61.1	19.0	35.9
14	80	16	28.5	71.5	61.2	19.4	37.4
15	80	18	26.8	72.0	62.2	19.6	39.1
16	80	20	26.9	72.3	62.3	20.1	39.8
17	80	10	26.8	71.1	61.2	20.4	37.3
18	80	12	26.9	71.2	61.5	20.9	37.6
19	80	14	26.9	71.0	61.2	18.9	35.8
20	80	16	27.5	71.4	61.2	19.3	37.3
21	80	18	28.2	72.0	62.2	19.6	39.0
22	80	20	28.5	72.2	62.2	20.0	39.8

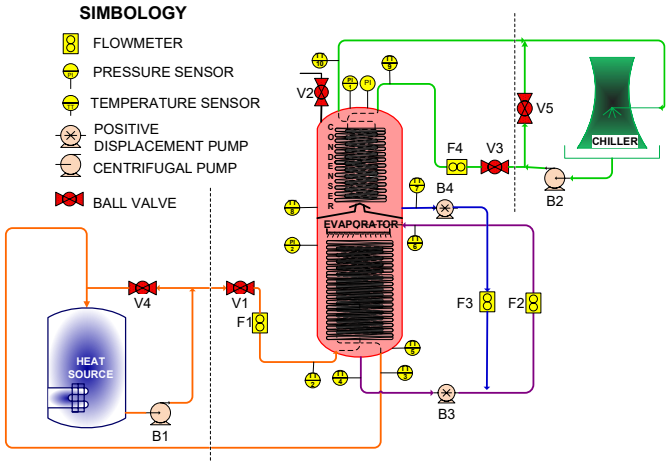


Fig. 2. Instrumentation Diagram.

Table 2. Measurement equipment characteristics.

VARIABLE	SENSOR	FEATURES
Temperature	West Brand Type J Thermocouple	Temperature range 0-120 ˚C Accuracy ±0.01˚C
Pressure	Cole-Parmer Brand Pressure Transducer	Temperature range -20 a 125 ˚C Pressure Range -14.7 a 15 psi. Accuracy ±0.25%
Mass flow	Cole-Parmer Brand Analog Flowmeters of 1200 mL/min, 6 gal/min, 12 gal/min	Temperature range -26 a 121 ˚C Accuracy ±0.5%

In Table 3 is specified the percentage of uncertainty for all variables.

Table 3. Uncertainty for variables.

Measured variables.	Uncertainty (%)
Hot water (L/min)	0.4
Cool water (L/min)	1.2
Absolute Pressure (kPa)	0.1
Local heat coefficient h_i (W/m^2 °C)	10.2
Local heat coefficient h_o (W/m^2 °C)	11.7
Heat capacity (W)	3.8
ΔT_{lm} (° C)	0.005
Overall heat coefficient (W/m^{2}K)	11.7

2.2 Ordinary Least Squares Linear Regression Methodology

It is a statistical method used to model the relationship between a dependent variable and one or more independent variables. The main goal is to find the coefficients of the linear equation that best fit the observed data. For a simple regression with one independent variable X, the relationship is expressed as $Y = aX + b$, where $a$is the slope and $b$is the intercept. In multiple regression, several predictor variables are included, and the equation is generalized as $Y = a_1X_1 + a_2X_2 + \ldots + a_nX_n + b$[18][19].

The ordinary least squares (OLS) method aims to minimize the sum of the squared differences between the observed values and the values predicted by the model. Mathematically, if we have a set of data (X_i, Y_i), the error is defined as $e_i = Y_i - \hat{Y}_i$, and OLS minimizes $\sum_i e_i^2$. This becomes an optimization problem that can be solved analytically using linear algebra with the normal equation: $(X^TX)\beta = X^TY$, where βcontains the coefficients to be estimated.

In practice, OLS is widely used because it is simple, efficient, and provides unbiased and consistent estimates of the coefficients under certain conditions, such as linearity, independence of errors, homoscedasticity, and normality of errors [20].

One of the main advantages of OLS is its simplicity and ease of implementation. Under the right assumptions, it provides unbiased and efficient estimates of the model coefficients. Moreover, the estimated coefficients are easy to interpret, which helps understand the relationship between variables. OLS is also computationally efficient, making it suitable for large datasets.

Despite its advantages, OLS has some limitations. It is sensitive to outliers, which can significantly distort the results. In addition, the method assumes that errors are independent and have constant variance (homoscedasticity); if these assumptions are violated, the estimates may be biased or inefficient [18][21].

2.3 Calculation of statistical errors

To verify the similarity between the simulator results and the data from the previously mentioned articles, different deviation indicators were used to evaluate the goodness of fit between both sets of data [22]. The deviation indicators are defined as follows:

- Deviation (% Error): This indicator shows how far an approximate value is from an exact value.

$$\text{Deviation (\%)} = \frac{|\mathrm{X_{sim}} - \mathrm{X}|}{\mathrm{X}} * 100 \qquad (1)$$

- Root Mean Square Error (RMSE): This measures the amount of error between two data sets, providing an indication of how close the observed data points are to the estimated values. Lower RMSE values indicate a better fit, and an

RMSE of 0 represents a perfect match between the observed and estimated series.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{N}(\mathrm{X_{sim,i}} - \mathrm{X_i})^2}{\mathrm{N}}} \quad (2)$$

- Mean Bias Error (MBE): This is used to validate the model results against experimental data. It represents the degree of agreement between a prediction and an observation, indicating whether a model tends to overestimate or underestimate the observed values.

$$\mathrm{MBE} = \frac{\sum_{i=1}^{N}(\mathrm{X_{sim,i}} - \mathrm{X_i})}{\mathrm{N}} \quad (3)$$

- Coefficient of Determination (R^2): This indicates the goodness of fit of the model. Its values range from 0 to 1, where 0 means the proposed model does not improve prediction compared to the mean model, and 1 represents a perfect prediction.

$$\mathrm{R^2} = 1 - \left(\frac{\sum_{i=1}^{N}(\mathrm{X_{sim,i}} - \mathrm{X_i})^2}{\sum_{i=1}^{N}(\mathrm{X_i} - \mathrm{X_{ave}})^2}\right) \quad (4)$$

Where N represents the amount of data, $\mathrm{X_{sim,i}}$ and $\mathrm{X_i}$ are each of the COP values obtained in the simulator and reported in the articles respectively, while $\mathrm{X_{ave}}$ indicates the average of the data referenced in the studies.

2.4 Nusselt Number Correlation

Experimental evidence shows that the Nusselt number is strongly influenced by the Reynolds and Prandtl numbers. This adimensional numbers was calculated by equations 4, 5 and 6.

Prandtl number: $Pr = \frac{\mu Cp}{k}$ (5)

Reynolds' number: $Re = \frac{\rho v D}{\mu} = \frac{4\dot{m}}{\pi D_{int}\,\mu}$ (6)

Due to the helical geometry of the condenser, a critical Reynold's number must be calculated [23]

Critical Reynolds' Number: $Re_{crit} = 2300\left[1 + 8.6\left(\frac{D_i}{D_{hel}}\right)^{0.45}\right]$ (7)

The former represents the intensity of mixing caused by the fluid flow, while the latter describes the relationship between the fluid's viscosity and thermal conductivity [24]. Consequently, the obtained values can be fitted using a correlation of the form:

$$\mathrm{Nu} = \mathrm{f}(Re, Pr) \tag{8}$$

The general form of the correlation function relating the Nusselt number (Nu) to the Reynolds (Re) and Prandtl (Pr) numbers is usually expressed as a power-type empirical relationship:

$$Nu = C \cdot Re^m \cdot Pr^n \tag{9}$$

where:

C= empirical constant (determined experimentally),
m= exponent reflecting the influence of flow (Reynolds),
n= exponent reflecting the influence of fluid properties (Prandtl).

This type of correlation is obtained by fitting the experimental data using linear regression in logarithmic scale:

$$\ln(Nu) = \ln(C) + m\ln(Re) + n\ln(Pr) \tag{10}$$

In this way, C, m and n can be determined directly from the experimental results.
The reason for using the form $Nu = C \cdot Re^m \cdot Pr^n$ to correlate the Nusselt number with Reynolds and Prandtl is based on the physics of convective heat transfer and experimental experience:

Dependence on turbulence and flow:
The Reynolds number (Re) represents the ratio of inertial to viscous forces in the fluid. Its influence on heat transfer is reflected in how the flow mixes the fluid near the wall. Therefore, Nugenerally increases with Re, and the exponent m quantifies this dependence.

Influence of fluid properties:
The Prandtl number (Pr) relates momentum diffusion (viscosity) to heat diffusion (thermal conductivity). Hence, the exponent ncaptures how the combination of viscosity and conductivity affects the thermal boundary layer and, consequently, the Nusselt number.

The thermophysical properties of water used in the calculations were obtained from the NIST (National Institute of Standards and Technology) database.

Empirical form and flexibility:
The power-law relationship is a well-established empirical form that allows easy fitting of experimental data under different flow conditions, geometries, and fluids. When transformed logarithmically, it becomes a linear equation in $\ln(Nu)$ vs $\ln(Re)$ and $\ln(Pr)$, facilitating regression techniques to determine C, m and n.

Universal applicability:
Many classical correlations in textbooks [24] and experimental studies follow this form because it correctly captures the underlying physics without requiring complex analytical solutions of the Navier-Stokes and energy equations.

In summary, this form combines physical basis, ease of experimental fitting, and flexibility for different flows and fluids, making it the most practical and widely used approach in heat transfer.

By solving equation (9) using ordinary least squares linear regression, the following correlation was obtained:

$$Nu = 20.6382\ Re^{-0.0184} Pr^{02682} \tag{11}$$

The equation is valid for the following ranges of Re and Pr.

$$2.5 \leq Pr \leq 3.5\ ; \quad 113 \leq Re \leq 283$$

This correlation describes the behavior of the Nusselt number as a function of the Reynolds and Prandtl numbers. The constant $C = 20.6382$ sets the baseline value of Nu for the experimental conditions. The very small negative Reynolds exponent ($m = -0.0184$) indicates that Nu slightly decreases as Re increases, suggesting that in this experimental regime or geometry, the effect of flow on convection is minimal. On the other hand, the Prandtl exponent ($n = 0.2682$) shows that Nu moderately increases with Pr, reflecting that the fluid's relative viscosity and thermal conductivity have a more significant impact on heat transfer. Overall, this power-law correlation indicates that, in this case, fluid properties are more important than flow velocity in determining convective heat transfer.

3. Results

The Table 4 presents the statistical error results obtained for the proposed correlation. It can be observed that the obtained errors are within acceptable limits.

Table 4. Table of statistics errors values.

Variable	RMSE	MBE	R^2
Nu	0.006	0.00092	0.9999998
h_o	0.41	0.00092	0.9999998

When analyzing the results obtained in Table 5 from equation (1), it can be seen that there is agreement between the experimental and theoretical values for both the Nusselt number (Nu) and the heat transfer coefficient (h_o). The calculated relative errors remain within a very low range, from 0.01% to 0.28%, which indicates that the differences are practically negligible.

The highest error value appeared in test 3 with 0.28%, while in most measurements the errors were below 0.15%, and in several tests the error was as small as 0.01%. These results show that the theoretical model reproduces the experimental data with errors lower than 0.28%, confirming its validity for calculating the heat transfer phenomenon using the correlation proposed in this work. This provides confidence in both the methodology and the quality of the calculations carried out.

The errors (%) of h_o shown in Table 5 indicate that the differences between the experimental and theoretical values are very small, ranging from 0.01% to 0.28%. The highest error appears in the third measurement, with a value of 0.28%, while in test 4 there is also a relatively high error of 0.26%. This discrepancy can also be seen in Fig. 3 and Fig. 4. In contrast, the lowest errors, only 0.01%, are found in tests 6, 15, 17, and 21, which practically reflects a coincidence between the results. In the first five trials, the errors tend to be higher, ranging between 0.14% and 0.28%, which could be due to small initial experimental variations. However, from test 6 onwards, the errors decrease and remain within a much narrower range, mostly between 0.01% and 0.15%. For example, in test 10 the error was only 0.06%, and in test 12 it was 0.02%, showing high accuracy. Even in cases where the error approaches 0.15%, such as in test 14 (0.15%) and test 20 (0.14%), the differences remain minimal. Overall, the errors never exceed 0.3%, confirming that the theoretical values obtained from the correlation represent the experimental results. These data indicate that the model used is reliable and adequately describes the heat transfer phenomenon, with remarkable agreement between theory and practice.

Table 5. Experimental and Theoretical Percent Errors for Nu and h_o

Test Num-ber	$Nu_{experimental}$ **(-)**	$Nu_{,theoritical}$ **(theoretical) (-)**	**Error (%)**	$h_{o,experimental}$ **(W/m² K)**	$h_{o,theoretical}$ **(W/m² K)**	**Error (%)**
1	25.4	25.4	0.03	1615.0	1615.5	0.03
2	25.3	25.3	0.14	1613.4	1611.1	0.14
3	25.3	25.2	**0.28**	1612.1	1607.5	**0.28**
4	25.4	25.5	0.26	1617.4	1621.6	0.26
5	24.4	24.4	0.22	1574.1	1570.6	0.22
6	24.1	24.1	0.01	1559.2	1559.4	0.01
7	24.4	24.3	0.11	1570.7	1568.9	0.11
8	24.1	24.1	0.02	1558.2	1558.0	0.02
9	24.2	24.2	0.09	1563.6	1562.3	0.09
10	24.1	24.1	0.06	1558.1	1557.2	0.06
11	24.0	24.0	0.08	1551.5	1552.7	0.08
12	24.0	24.0	0.02	1550.6	1550.9	0.02
13	24.0	24.0	0.11	1552.7	1551.0	0.11
14	24.0	23.9	0.15	1550.4	1548.1	0.15
15	23.8	23.8	0.01	1543.6	1543.7	0.01
16	23.8	23.8	0.05	1542.4	1543.2	0.05
17	24.0	24.0	0.01	1551.6	1551.4	0.01
18	24.0	24.0	0.11	1550.4	1552.1	0.11
19	24.0	24.0	0.10	1552.2	1550.7	0.10
20	24.0	23.9	0.14	1550.5	1548.3	0.14
21	23.8	23.8	0.01	1543.9	1543.8	0.01
22	23.8	23.8	0.05	1542.7	1543.5	0.05

Regarding the variation of the Nusselt number (Nu) as a function of the number of tests carried out (Fig. 3), Nu is represented on the vertical axis, while the test numbers, from 1 to 23, appear on the horizontal axis. The experimental values are shown by the black line, while the theoretical values are indicated by blue squares. In the first tests, Nu values range between 25.2 and 25.5, showing slight oscillations. However, from the fifth test onwards, a noticeable decrease is observed toward values close to 24. As the number of tests increases, the values tend to stabilize within a range between 23.8 and 24.1, showing consistency between the theoretical model and the experimental results. This downward trend suggests system stability due to the heat transfer conditions. The closeness between the curves confirms the validity of the theoretical model to describe the experimental behavior.

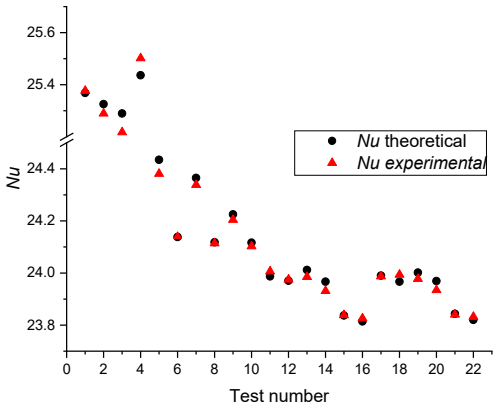


Fig. 3. Nu experimental y theoretical

Fig. 4 shows the variation of the convective heat transfer coefficient, h_o, as a function of the number of tests performed. On the vertical axis, the values of h_o are represented in W/m²·K, while the horizontal axis shows the number of tests, ranging from 0 to 25. The experimental results are represented by a solid black line, whereas the theoretical values are shown as blue squares. Initially, the h_o values are located around 1610 to 1620, but from the fifth test onward, a marked downward trend is observed. As the tests progress, both the experimental and theoretical values converge toward a more stable region near 1540–1560 W/m²·K. Although slight oscillations exist in the experimental data, the theoretical points closely follow the same behavior. This agreement suggests that the theoretical model is representative of the experimental observations. However, the slight initial dispersion indicates possible variations in the experimental conditions, which are inherent to the behavior of the experiment.

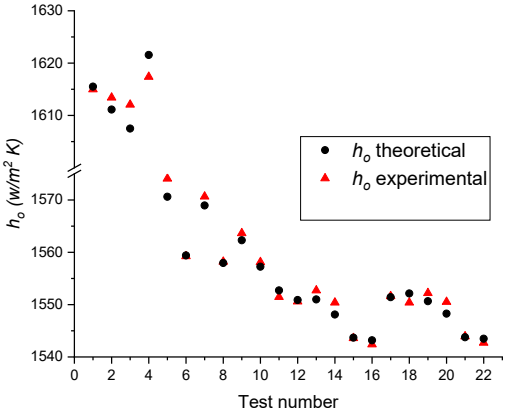


Fig. 4. h_o experimental and theoretical.

Fig. 5 and 6 show the errors obtained, illustrating the range of variation (positive and negative) associated with each experimental and theoretical data point. These figures allow the dispersion of the results to be observed and the accuracy of the model to be evaluated in comparison with the measured values. However, due to the density of the data points and the scale used, Figure 4 does not clearly show the trend of the predicted values.

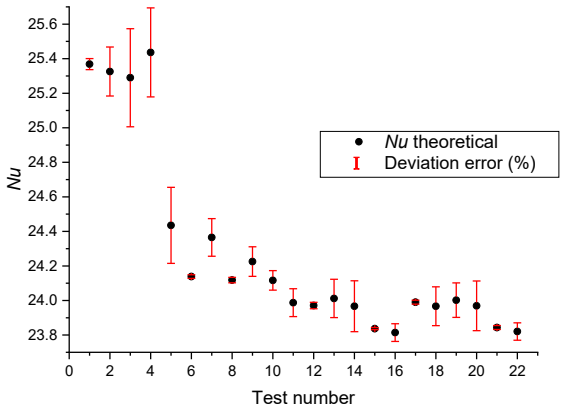


Fig. 5. Error of experimental and theoretical Nu.

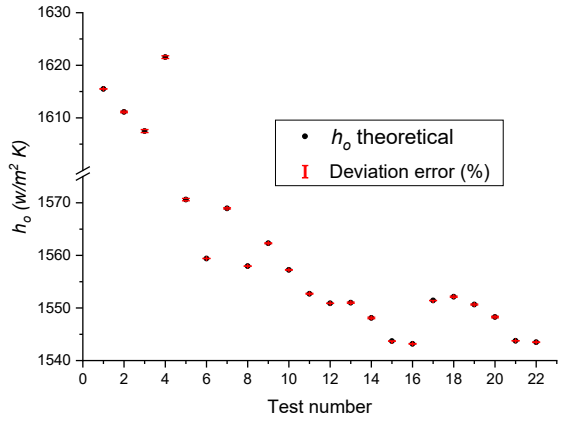


Fig. 6. Error of experimental and theoretical h_o

4. Conclusions.

A dual falling-film helical condenser was experimentally evaluated through 22 tests conducted at three temperature levels and six flow rates. The experimental falling-film heat transfer coefficients ($h_{o,exp}$) were determined, and a correlation for the Nusselt number (Nu) as a function of the Reynolds and Prandtl numbers was developed using the least-squares method. The correlation predicted Nu values with errors ranging from 0.01% to 0.28% relative to the experimental data, with most deviations concentrated around 0.15%. The maximum error of 0.28% was observed in Test 3. The results in Table 3 show that the correlation demonstrates highly reliable performance: the RMSE indicates very low average errors for both Nu (0.006) and h_o (0.41), confirming high accuracy; the MBE of 0.00092 for both variables indicates that there is no significant bias, ruling out tendencies of overestimation or underestimation; and finally, the R^2 value of 0.9999998 reflects an almost perfect fit, explaining practically all the variability of the data. These results demonstrate that the proposed correlation reproduces the experimental data with high accuracy, providing a reliable tool for the design of equipment with similar helical geometries.

5. References

1. He, J., Zhuang, T., & Xie, X. Energy consumption, economic development and environmental improvement in China. Energy & Environment, 25(8), 1345–1361 (2014). https://doi.org/10.1260/0958-305X.25.8.1345
2. Nguyen, T. N., Tran, T. T., & Le, H. H.: Machine learning-based Nusselt number prediction for falling-film evaporators. Energy Reports 10, 118–126 (2024). https://doi.org/10.1016/j.egyr.2023.11.105
3. Tesern I., Lin H, Chou S. Heat and mass transfer in vertical twin-parallel-tube absorber-evaporator unit. Journal of the Chinese Institute of Engineers. 24 (2001) 581-593.
4. Yoon J, Phan TT., Moon C.G., Lee H., Jeong. S.K. Heat and mass transfer characteristics of a horizontal tube falling film absorber with small diameter tubes. Heat and Mass Transfer. 44 (2006) 437-444. https://doi.org/10.1007/s00231-007-0261-8
5. Pospisil J., Balas M., Baxa M., Fortelny Z. Working characteristics of small-scale absorption unit two-cylinder design. Wseas Transaction on Heat and Mass Transfer. 3 (2009) 77-86.
6. Morales L.I., Conde-Gutiérrez R.A., Hernández J.A., Huicochea A., Juárez-Romero D., Siqueiros J., Optimization of an absorption heat transformer with two-duplex components using inverse neural network and solved by genetic algorithm Applied Thermal Engineering. 85 (2015) 322-333.DOI: 10.1016/j.applthermaleng.2015.04.018
7. Demesa N., Hernández J.A., Siqueiros J., Huicochea A., Heat transfer coefficients for helical components inside an absorption heat transformer. International Journal of Heat and Mass Transfer. 120 (2018) 342-349. https://doi.org/10.1016/j.ijheatmasstransfer.2017.12.038
8. Delgado-Gonzaga J, Juárez-Romero D., Saravanan R., Márquez-Nolasco A., Huicochea A., Morales L.I., Performance analysis of a dual component generator-condenser of an absorption heat transformer for water desalination. Thermal Science and Engineering Progress 20 (2020). https://doi.org/10.1016/j.tsep.2020.100651
9. Canela-Sánchez, I. J., et al. (2023). Modeling and simulation of a helical falling film evaporator. *Energy*, 263, 125433. https://doi.org/10.1016/j.energy.2022.125433

10. Kwon, K., & Jeong, S. (2001). Effect of vapour flow on the falling-film heat transfer of a helical coil. *International Journal of Refrigeration*, 24(6), 575–583. https://doi.org/10.1016/S0140-7007(00)00061-4
11. Zhao, C. Y. (2024). Correlations of falling film hydrodynamics and heat transfer. *Renewable and Sustainable Energy Reviews*, 197, 113884. https://doi.org/10.1016/j.rser.2024.113884
12. Fang, J., et al. (2019). Establishment of the falling film evaporation model and correlation of the overall heat transfer coefficient. *Applied Thermal Engineering*, 156, 1–10. https://doi.org/10.1016/j.applthermaleng.2019.04.028
13. Fernández-Seara, J., et al. (2014). Experimental analysis of a helical coil heat exchanger. *Applied Thermal Engineering*, 63, 498–506. https://doi.org/10.1016/j.applthermaleng.2013.11.055
14. Essalhi, A., et al.: Design of a helical coil condenser of a small-capacity Water/LiBr absorption cooling machine. Applied Thermal Engineering 123, 1223–1231 (2017). https://doi.org/10.1016/j.applthermaleng.2017.05.049
15. Liu, H., Zhou, Q., Liu, Y., et al.: Experimental study on cooling performance of air conditioning system with dual independent evaporative condenser. International Journal of Refrigeration 55(C), 85–92 (2015). https://doi.org/10.1016/j.ijrefrig.2015.03.012
16. Nannan D, Shuhong L. Simulation and performance analysis on condenser coil in household heat pump water heater. Sustainable Cities and Society. 36 (2018) 176-184. 10.1016/J.SCS.2017.10.020
17. Hernández-Magallanes, J. A., Siqueiros, J., Sarmiento-Bustos, E., Juaréz-Romero, D., & Morales, L. I.: Experimental study of a dual condenser (EVACON) with concentric helicoids in use with an absorption heat transformer. Case Studies in Thermal Engineering 30, 101794 (2022). https://doi.org/10.1016/j.csite.2022.101794
18. Gujarati, D. N., & Porter, D. C.: *Basic Econometrics*, 5th ed., McGraw-Hill (2009).
19. Wooldridge, J. M.: *Introductory Econometrics: A Modern Approach*, 6th ed., Cengage Learning (2016).
20. Novales, A.: *Análisis de Regresión*. Universidad Complutense de Madrid (2003). https://www.ucm.es/data/cont/docs/518-2013-11-13-Analisis%20de%20Regresion.pdf
21. Vadillo, F.: *Introducción a los problemas de mínimos cuadrados*. Universidad de La Laguna (2022). https://www.ehu.eus/~mepvaarf/ficheros/minimos_cuadrados.pdf
22. Lugo, S., Morales, L. I., Best, R., Gómez, V. H., & García-Valladares, O.: Numerical simulation and experimental validation of an outdoor-swimming pool solar heating system in warm climates. *Solar Energy* 189, 45–56 (2019). https://doi.org/10.1016/j.solener.2019.05.023
23. Fernández-Seara J, Piñeiro-Pontevedra C, Dopazo JA. On the performance of a vertical helical coil heat exchanger. Numerical model and experimental validation. Applied Thermal Engineering. 62 (2014) 680-689. DOI:10.1016/j.applthermaleng.2013.09.054
24. Bergman, T. L., Lavine, A. S., Incropera, F. P., & DeWitt, D. P.: *Fundamentals of Heat and Mass Transfer*, 8th ed., Wiley (2020). ISBN 978-1119722489

Urban decarbonization through PV installations on building facades: the real-life case of a university building in southern Spain

Carlos Gilabert-Torres[1,2[0009-0002-5647-1648]], Giuseppe M. Tina[3[0000-0003-3896-1838]], G. Jiménez-Castillo[1,2[0000-0002-5370-7283]], F.J. Muñoz-Rodriguez[2,3[0000-0002-0833-0528]] and Catalina Rus-Casas[1,2[0000-0002-6982-4054]]

[1] University of Jaén Las Lagunillas Campus, A3 Building, 23071 Jaén, Spain *gilabert@ujaen.es (C. G.T) fjmunoz@ujaen.es (F.J.M.R) crus@ujaen.es (C.R.C)*
[2] Centre for Advanced Studies in Energy and Environment CEACTEMA, Universidad de Jaén, 23071 Jaén, Spain
[3] Department of Electrical, Electronics and Informatics Engineering, University of Catania, Catania, Italy g*iuseppe.tina@unict.it (G.M.T)*
Department of Electrical Engineering, University of Jaen, Jaén, Spain
gjimenez@ujaen.es (G.J.C)

Abstract. The growing interest in renewable energies has driven the implementation of solar photovoltaic energy, especially in urban environments, where Building-Integrated Photovoltaics (BIPV) and Building-Applied Photovoltaics (BAPV) systems are key solutions. Large-scale integration, however, presents challenges for the grid, such as the "duck curve". This study presents an energy and environmental analysis of a BAPV installation on a university building facade in southern Spain, which includes six subsystems with unconventional orientations. Two new parameters are also introduced to evaluate the homogeneity of the PV generation profile.

Generation and consumption data for the year 2024 show that self-consumed photovoltaic energy represents 65% of the annual total. During solar hours, self-sufficiency reached 69%, indicating good coupling between generation and demand. Compared to a conventional rooftop PV system, the facade system showed a 7.4% higher energy-to-peak power ratio (EPPR), demonstrating its ability to generate energy during early and late hours. However, its mean-to-peak power ratio (MPPR) was 6% lower, suggesting that while multi-oriented PV systems can flatten the generation curve, their design must be optimized to achieve this goal. The installation resulted in a total emissions reduction of 3.2 tCO2eq/year. Furthermore, data indicates that the hourly emission factor of grid energy decreases by 44% to 64% during midday hours compared to its peak value.

These findings confirm that multi-oriented facade PV systems are a viable strategy for urban decarbonization and for aligning generation with both consumption and the emission profile of the grid.

Keywords: Photovoltaics; Zero Energy Building; Building-Integrated Photovoltaics; Building-Applied Photovoltaics.

1 Introduction

In recent years, the installed capacity of photovoltaic (PV) energy has grown considerably, with a projected increase from 666 GW in 2024 to 935 GW by 2030 [1]. This growth is driven by a drop in costs, with utility-scale installation costs falling from 1044 USD/kW in 2020 to 691 USD/kW in 2024 [2], as well as a greater environmental awareness and the need to reduce greenhouse gas (GHG) emissions [3, 4]. The trend of integrating PV into buildings is particularly strong, with Germany projecting over 200 GWp by 2050, a pattern also seen in countries like Japan and Switzerland [5].

Despite this growth, the large-scale integration of PV presents challenges for the grid. The "duck curve" effect, for example, highlights the mismatch between peak solar generation at midday and electricity demand patterns [6], which can lead to grid congestion and instability [7]. To mitigate these issues, it is crucial to adopt strategies that improve the alignment of generation and consumption, such as load-shifting, to avoid excessive PV energy injection into the grid during midday [8, 9].

Building-integrated photovoltaics (BIPV) and building-applied photovoltaics (BAPV) are key to urban decarbonization and efficient space use [10]. BAPV systems are integrated into the building's available areas (e.g., facades, rooftops), whereas BIPVs become an integral component of the structure. Beyond mere energy generation, BIPV and BAPV systems facilitate urban decarbonization by replacing or integrating into conventional building components like facades [11]. This supports the goal of Zero Energy Buildings (ZEBs) and the urban energy transition. These systems also provide aesthetic and multifunctional solutions that balance energy production, savings, visual comfort, and architectural integration [12, 13].

However, integrating PV modules into the envelope of a building often means using tilt angles and orientations that are not optimal for maximum annual energy production [14, 15]. While this can be a constraint, non-optimal orientations can offer the advantage of better coupling the generation curve with the consumption curve [16]. For instance, east-west orientations have proven to be both efficient and viable, reducing module operating temperatures and optimizing the use of the available surface area [17, 18]. Accordingly, methodologies have been developed to optimize the design of BIPV-BAPV systems, with the objective of addressing effects such as shading and the incidence angle modifier [19].

In addition to increasing photovoltaic self-consumption, BIPV/BAPV systems with multiple orientations have the potential to achieve a flatter generation profile than conventional PV systems. By distributing energy production over a wider period, the midday generation peak is reduced, minimizing its impact on the electrical grid and system stability [20]. The assessment of the amount of photovoltaic energy utilized by a building can be facilitated by the implementation of indices, such as the self-consumption rate. However, extant literature exhibits a gap in terms of quantifying how flat the daily photovoltaic generation profile is.

Consequently, this study proposes two novel metrics to assess the shape of the PV generation profile: the energy-to-peak power ratio (EPPR) and the mean-to-peak power

ratio (MPPR). The parameters in question are incorporated into the analysis of the energy and environmental performance of a BAPV installation on a university building in southern Spain, which features six subsystems with non-optimal orientations.

2 Materials and methods

2.1 Case study

This study analyzes a four-story building at the University of Jaén (37.7875, -3.7736°), which serves as a research center and student laboratory. Its design enabled the installation of a facade-mounted photovoltaic system with six different orientations The 3D model and PV system of the building are shown in **Fig. 1**.

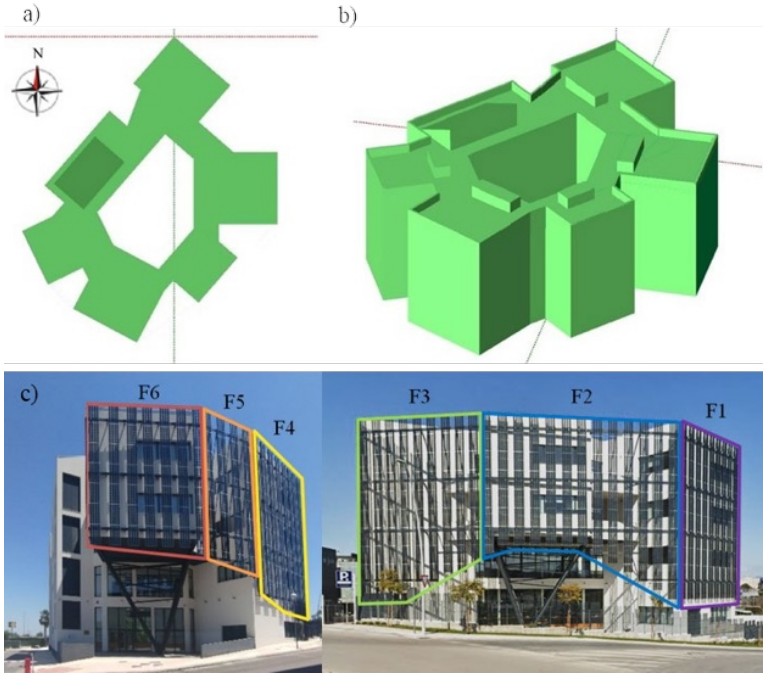


Fig. 1. (a) Plan view of the building, (b) perspective view of the building and (c) pictures of the actual PV installation.

The PV generator, anchored to the façade of the building, uses two types of modules: Onyx GL.01 ("Module I") with a 166 Wp peak generation, and Onyx GL.02 ("Module II") with a 74 Wp peak generation under standard conditions [21, 22]. The photovoltaic system consists of four distinct subsystems (SS), each connected to its own dedicated inverter. **Table 1** show the technical characteristics of the photovoltaic system under

study. Module orientations range from east to southwest, with 64% of the installed capacity facing east-southeast. This system occupies a footprint of 32 m², a detail that is relevant for evaluating its spatial efficiency.

Table 1. PV system characteristics.

System	Facade	Strings	Orientation (°)	Modules I	Modules II	Capacity (kWp)[1]
SS1	F1	3	-90	90	0	14,9
SS2	F2	4	-45	100	4	16,9
SS3	F3	3	-33,5	81	2	13,6
	F5	1	45	31	2	5,3
SS4	F4	3	28	75	2	12,6
	F6	2	64	48	0	8
Total				425	10	71,3

The system includes four Schneider CL-33 inverters, each with three MPPTs and a nominal AC output of 33 kW [23]. PV generation data was monitored using the software of the inverter manufacturer, which provides the AC output power of each inverter [24]. Consumption was monitored with an ABB M4M30 network analyzer [25], which has a 0.5s class accuracy [26]. **Fig. 2** shows the average seasonal generation and consumption profiles with a resolution of 10 min.

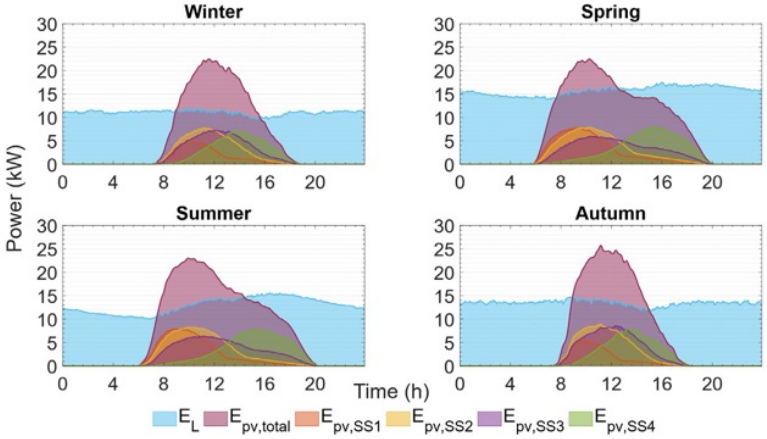


Fig. 2. Average building PV generation and energy consumption per season.

Building energy consumption stays relatively constant throughout the year, with average powers of 11.1 kW in winter, 15.7 kW in spring, 12.9 kW in summer, and 13.4 kW

[1] Data under standard conditions defined in Standard IEC 61724-1 [22].

in fall. In contrast, generation varies significantly by season. During periods of lower solar irradiance (fall and winter), the generation profile is bell-shaped with a daily average of 147 kWh. During periods of higher solar irradiance (spring and summer), the profile is flatter, with a daily average of 188 kWh. This seasonal variation in generation is attributed to the increased contribution of subsystems 1 (F1) and 4 (F4-F6), which increase their output by 100% and 28% respectively in spring and summer. These increases are significantly higher than those of subsystems 2 (F2) and 3 (F3-F5), which increase by only 17% and 8% respectively.

2.2 Performance analysis

Energy and environmental parameters were analyzed. Regarding energy parameters, there is a gap in the literature concerning the characterization of the daily photovoltaic generation profile. For this reason, this study defines two new metrics: the energy-to-peak power ratio (EPPR) and the mean-to-peak power ratio (MPPR), which are defined in Equation 1 and Equation 2, respectively.

$$EPPR_d = \frac{\sum_{k\in d} P^k_{pv,total}\cdot\Delta t}{\max\left(P^{k\in d}_{pv,total}\right)} \quad (1)$$

$$MPPR_d = \frac{\sum_{k\in d} P^k_{pv,total}/T_{SH,d}}{\max\left(P^{k\in d}_{pv,total}\right)} \quad (2)$$

Where P_{pv} is the output PV power, k is the time step number ($k = 1,2,\dots,N$), Δt is the time step, $T_{SH,d}$ is the solar hours duration and d is the day number ($d = 1,2,\dots,365$). EPPR correlates the daily generated PV energy with the recorded daily maximum photovoltaic power value. In contrast, MPPR relates the daily mean and maximum photovoltaic power values. A high value for these parameters indicates a flatter generation curve, with the key distinction that EPPR is dependent on the number of solar hours, whereas the MPPR is not. The self-consumed solar photovoltaic energy was calculated using Equation 3, which groups the subsystems ($ss = 1,2,\dots,4$) and time steps.

$$E_{PV,SC} = \sum_{ss=1}^{4}\sum_{k=1}^{N} \min(P^k_{pv,ss}, P^k_L)\cdot\Delta t \quad (3)$$

Where P_{pv} is the output PV power, P_L is the power consumed by the building and Δt is the temporal resolution of the data. Two significant parameters derived from the amount of self-consumed energy are the self-consumption ratio (SCR) and self-sufficiency ratio (SSR). The self-consumption ratio is defined as the amount of photovoltaic energy consumed by the building, while the self-sufficiency ratio is the fraction of the building total energy demand met by the photovoltaic system. SCR and SSR were calculated using Equations 4 and 5, respectively.

$$SCR = \frac{E_{PV,SC}}{\sum_{ss=1}^{4}\sum_{k=1}^{N} \min(P^k_{pv,ss},P^k_L)\cdot\Delta t} \quad (4)$$

$$SSR = \frac{E_{PV,SC}}{\sum_{k=1}^{N} P^k_L\cdot\Delta t} \quad (5)$$

$$SSR_{SH} = \frac{E_{PV,SC}}{\sum_{k=1}^{N} P_{L_{SH}}^{k} \cdot \Delta t} \quad (6)$$

Additionally, the self-consumption ratio in solar hours was calculated using Equation 6 to characterize more precisely the potential for solar self-consumption [27]. Furthermore, the final yield, one of the most common energy metrics defined in the IEC 61724-1 standard, was analyzed. This index relates the photovoltaic generation to the installed power of the system (P_0), according to Equation 7.

$$Y_f = \sum_{ss=1}^{4} \sum_{k=1}^{N} P_{pv,ss}^{k} \cdot \Delta t / P_0 \quad (7)$$

$$Y_{f,s} = Y_f / S_{fp} \quad (8)$$

To complement this analysis, the final yield per unit of occupied surface (footprint area) of the system, calculated with Equation 8, was also included [28]. This facilitates a more precise evaluation of the efficiency of space utilization, a critical factor in urban environments.

2.3 Environmental analysis

A key benefit of renewable energy generation is the reduction of greenhouse gas (GHG) emissions. To estimate the environmental impact of a photovoltaic system, the total emissions reduction (TER) is calculated using Equation 9. The hourly CO_2 emission factor (HEF) represents the tons of CO_2 emitted per MWh consumed from the grid. This factor is calculated following the method described in a previous study [28], as shown in Equation 10.

$$TER = \sum_{k=1}^{N} E_{PV,SC,k} \cdot HEF_h \quad (9)$$

$$HEF_h = \frac{GHG_d}{E_{gen,nr,d}} \cdot \frac{E_{nrg,h}}{E_{tg,h}} \quad (10)$$

Where GHG_d is the daily CO_2 emissions associated with generation, $E_{gen,nr,d}$ is the daily non-renewable energy generated, $E_{nrg,h}$ is the hourly non-renewable generation, and $E_{tg,h}$ is the total hourly generation of the electrical system. To assess the efficiency of the photovoltaic system in urban decarbonization, the TER has been normalized based on the PV system capacity and the area it occupies, as shown in Equation 11.

$$TER_{P0,s} = TER / (S_{fp} \cdot P_0) \quad (11)$$

This parameter is a useful metric for comparing the environmental efficiency of various BIPV and BAPV configurations, particularly for optimizing space in urban environments.

3 Results and discussion

This analysis is based on the generation and consumption data of the building in the year 2024. The study is structured in three parts: first, we evaluate the performance of each photovoltaic subsystem; second, we analyze the self-consumed energy and its contribution to the total consumption of the building; and finally, we evaluate the environmental impact of the installation. **Fig. 3** shows the final yield of the photovoltaic system, aggregated by subsystem and month.

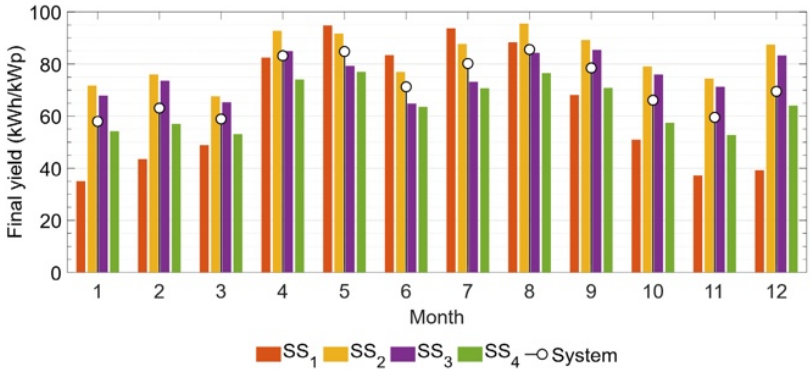


Fig. 3. Final yield per month and subsystem.

From September to April, the average final yield of subsystems 2 (80 kWh/kWp) and 3 (76 kWh/kWp) was higher than that of subsystems 1 (51 kWh/kWp) and 4 (60 kWh/kWp). For the remainder of the year (May to August), the average final yield of subsystem 1 (90 kWh/kWp) and 2 (88 kWh/kWp) exceeded that of subsystems 3 (75 kWh/kWp) and 4 (72 kWh/kWp). These results are attributed to the southeast orientation of subsystem 2, which allows it to receive more solar radiation than the east-facing subsystem 1. Additionally, its lower ambient temperature, compared to the west-facing subsystems (3 and 4), reduces temperature-related losses in the modules. It is important to note that the final yield of subsystem 1 only exceeds that of subsystem 2 from May to July due to earlier sunrise times.

The annual final yield of the installation is 858 kWh/kWp, which is slightly higher than values reported in other studies [18, 29], a difference attributed to the more favorable climate of Spain compared to Germany. The value in question ranges between 50 and 60% of the maximum final yield for an ideal rooftop PV system in Spain (800 to 975 kWh/kWp). This conclusion is consistent with findings derived from earlier studies [30]. Considering the spatial footprint of the photovoltaic system, the final yield per unit of installed power and square meter is 26.8 kWh/kWp·m². This value falls within the range of 21 to 50 kWh/kWp·m² reported in previous studies for optimized BAPV systems [28]. **Fig. 4** shows the photovoltaic generation of the system, characterized by its use.

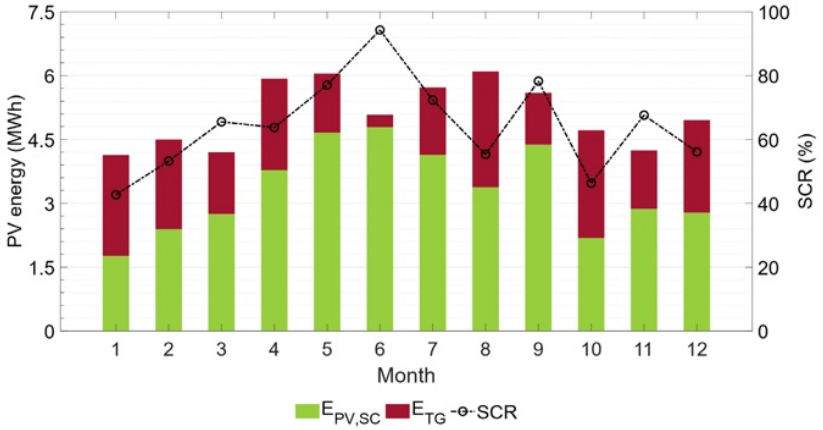


Fig. 4. PV energy self-consumed, fed to grid and SCR per month.

Total generation showed a clear seasonality, being higher in spring and summer (34 MWh) than in autumn and winter (27 MWh), a difference attributed to greater solar radiation and longer daylight hours. Self-consumption was lower in autumn (54%) and winter (56%) compared to spring (78%) and summer (68%), suggesting a weaker coupling between photovoltaic generation and demand during months with less sunlight. The monthly self-consumption index exhibited fluctuations between 43% and 94%, with an annual average of 65%. In configurations optimized to minimize exchanges with the grid, the self-consumption ratio can reach between 50% and 70%, depending on the load profile and available surface area [31]. Consequently, the findings of this study are consistent with those of previous studies. As shown in **Fig. 2**, the flatter generation profile during spring and summer better aligns with the constant energy consumption of the building. A preceding examination of PV installations at the University of Jaén [32] revealed an aggregate self-consumption rate of 97% on the university campus. This discrepancy can be attributed to the consumption of the PV surplus from individual buildings by other on-campus loads, which complements the PV generation consumption.

Fig. 5 presents the daily EPPR and MPPR values to provide a more detailed characterization of the generation profile. In this case, two PV configurations are compared: the façade system presented in this work, and the conventional rooftop system described in [28], which was designed for the same building and has optimal orientation and tilt angles.

The data reveal that the energy-to-peak power ratio for both PV systems increases during the spring and summer. The façade PV system performs notably better (7.3 h and 7.6 h) than the conventional system (6.4 h and 6.7 h). For the remainder of the year, both systems achieve a similar EPPR. This is attributed to enhanced ability of the façade system to harness irradiance during the early and late hours of the day, which boosts its energy generation compared to the conventional system during the spring and summer months, when daylight hours are longer.

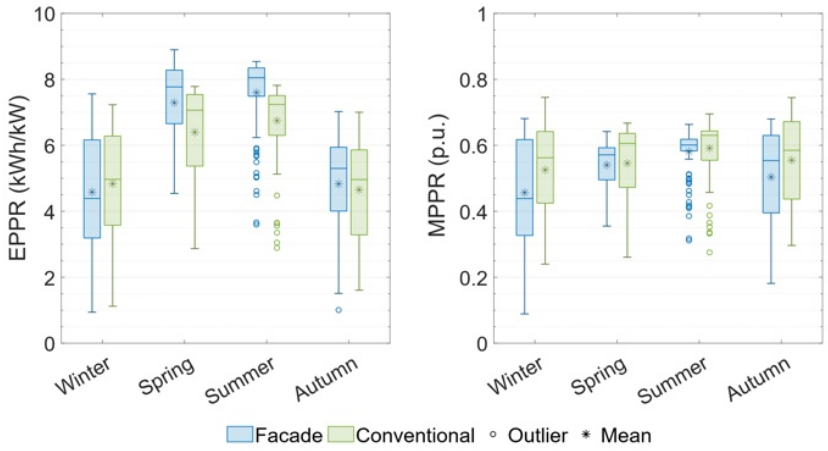


Fig. 5. EPPR (a) and MPPR (b) aggregated by season and PV layout.

An analysis of the seasonal mean-to-peak power ratio for the façade PV system reveals a similar seasonal trend, with higher values in spring (0.54) and summer (0.58) compared to winter (0.46) and autumn (0.5). This confirms that the generation profile of the façade system is flatter during spring and summer, as illustrated in **Fig. 2**. A comparison of the MPPR for both systems shows no significant differences in spring and summer, but notable disparities are observed in winter (Façade: 0.46, Conventional: 0.53) and autumn (Façade: 0.5, Conventional: 0.55). These data indicate that the generation profile of the conventional system is flatter than that of the façade system during these seasons. This occurs because the façade PV system generates very low power during the early and late hours of the day, which consequently decreases its average value. In contrast, the power generated by the conventional system increases and decreases more rapidly during the same periods, resulting in a higher and more consistent average power output. These findings demonstrate that façade systems have the potential to flatten the generation profile, but their design must be optimized to achieve this.

Fig. 6 illustrates the energy consumed by the building by source and solar availability. The monthly self-sufficiency ratio from the PV installation ranged from 23% to 46%, with an annual average of 34%. This value is proximate to the 30% achieved by other buildings equipped with BAPV [3 1]. SSR was higher in spring (38%) and summer (42%) than in autumn (29%) and winter (26%), due to greater PV generation. Similar BAPV systems integrated into university buildings have achieved a campus-wide SSR of 35% [33].

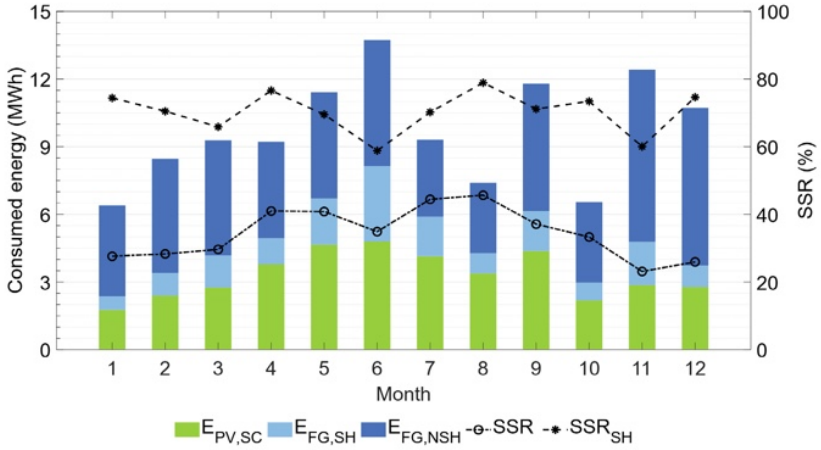


Fig. 6. Energy consumed from PV system and grid and SSR aggregated by solar hours and months.

An average of 51% of the total energy consumption of the building occurred during non-solar hours, an amount that cannot be met without energy storage. The self-sufficiency index during solar hours showed similar values across seasons, with 69% in autumn, 67% in spring, 73% in summer, and 68% in winter. Achieving a 69% self-sufficiency during solar hours with a multi-oriented facade highlights a strong coupling between generation and demand and provides a model for future BAPV system designs in urban settings.

Fig. 7 illustrates the relationship between the hourly emission factor and the self-consumed PV energy profile, aggregated by hour and season. The reduction in emissions from grid consumption is determined by multiplying these two curves. The total reduction in CO2 emissions per season was 535 kgCO2eq in winter, 930 kgCO2eq in spring, 890 kgCO2eq in summer, and 802 kgCO2eq in autumn. This indicates a yearly reduction of 3.2 tCO2eq in total annual emissions, which corresponds to a 24% decrease in GHG emissions, similar to previous works [34]. The annual emission reduction, considering the footprint and installed capacity of the PV system, is 1.4 kgCO2eq/kWp·m². This value falls within the range of 1.3 to 3.1 kgCO2eq/kWp·m² reported for optimized BAPV installations [28]. The figure demonstrates that the contribution of conventional PV systems to the Spanish distribution network reduces the hourly emission factor by 44% to 64% during peak daytime hours compared to its maximum value.

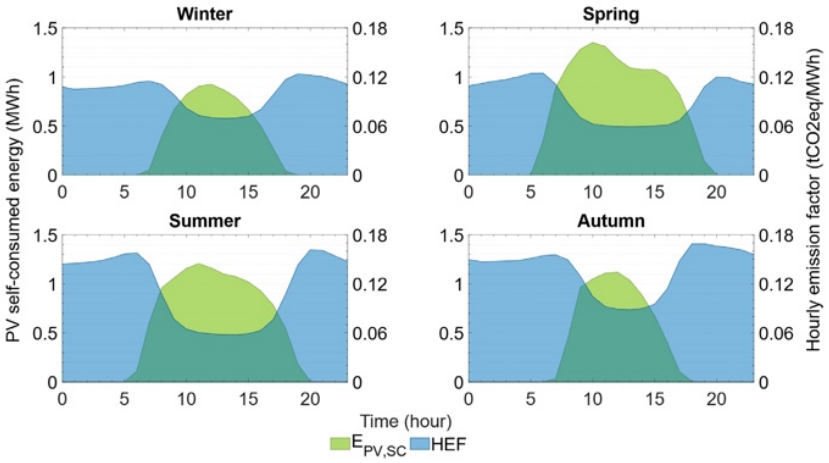


Fig. 7. PV self-consumed energy and average emission factor per hour and season.

Therefore, generators with unconventional orientations, whose generation peak does not coincide with noon, have significant potential to reduce GHG emissions associated with electricity grid consumption.

4 Conclusions

This study analyzed the energy performance and environmental impact of a BAPV installation on a university building facade, confirming its viability for urban decarbonization. The multi-orientation configuration distributed energy generation throughout the day, mitigating the impact of generation peaks on the grid and optimizing self-consumption. The installation achieved an annual self-consumption rate of 65% and a self-sufficiency of 69% during solar hours. These findings demonstrate that facade PV systems, even with unconventional orientations, can meet a significant portion of the energy demand of the building. This self-consumption not only reduces the energy bill of the building, but also contributes directly to the reduction of GHG emissions, with a total annual reduction of 3.2 t CO2eq.

The ability of these installations to generate energy during hours of high grid emission intensity is particularly relevant. While conventional PV systems reduce the grid emission factor by 44% to 64% during peak radiation hours, east-west oriented installations offer additional potential by mitigating emissions outside of this peak. This highlights the need to optimize PV generation to align with both the consumption of the building and the emission profile of the grid, thereby improving the overall environmental contribution of the building.

The analysis of the new EPPR and MPPR metrics provided a detailed characterization of the daily generation profile. The results demonstrate that the façade system has a flatter generation profile in spring and summer, with EPPR values up to 7.6 h and 7.3

h, respectively, compared to the 6.7 h and 6.4 h obtained by the conventional system. In contrast, in winter and autumn, the conventional system exhibits a flatter profile, as indicated by its higher MPPR values of 0.53 and 0.55, respectively, compared to the façade system (0.46 and 0.5). These findings confirm the ability of façade PV systems to generate energy during early and late hours. However, its mean-to-peak power ratio was 6% lower, suggesting that while multi-oriented PV systems can flatten the generation curve, their design must be optimized to achieve this goal.

The integration of multi-oriented PV systems on building envelopes is an effective strategy for urban decarbonization and efficient space utilization. This case study serves as a tangible model that can guide future building designs in dense urban environments and establish a precedent for future research exploring the impact of different configurations on energy and environmental performance.

Acknowledgments. The authors would like to thank the program that finances projects aimed at ecological and digital transition (Grant No. TED2021-131137B-I00: "Contribution to the Ecological Transition of the Industrial Sector through Photovoltaic Self-consumption") and the research program "Investigo" (2022-c23-i01-p03.s0020-0000590), which was financed by the European Union with Next Generation EU funds. The authors also acknowledge the support provided by the Thematic Network 723RT0150 "Red para la integración a gran escala de energías renovables en sistemas eléctricos (RIBIERSE-CYTED)" financed by the call for Thematic Networks of the CYTED (Ibero-American Program of Science and Technology for Development) for 2022.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. IEA (2024): Renewables 2024. Paris (2024).
2. IRENA: Renewable power generation costs in 2024. Abu Dhabi (2025).
3. UN: Transforming our world: the 2030 Agenda for Sustainable Development. United Nations: New York, NY, USA. (2015).
4. Ministerio para la Transición Ecológica y el Reto Demográfico: Plan Nacional Integrado de Energía y Clima (PNIEC 2023-2030), https://www.miteco.gob.es/es/energia/estrategia-normativa/pniec-23-30.html, last accessed 2025/07/10.
5. Kuhn, T.E., Erban, C., Heinrich, M., Eisenlohr, J., Ensslen, F., Neuhaus, D.H.: Review of technological design options for building integrated photovoltaics (BIPV). Energy Build. 231, 110381 (2021). https://doi.org/10.1016/J.ENBUILD.2020.110381.
6. CAISO: What the duck curve tells us about managing a green grid. , Folsom, CA, USA (2016).

7. Smith, O., Cattell, O., Farcot, E., O'Dea, R.D., Hopcraft, K.I.: The effect of renewable energy incorporation on power grid stability and resilience. Sci Adv. 8, 6734 (2022). https://doi.org/10.1126/SCIADV.ABJ6734.
8. Zhou, H., Lu, L., Wei, M., Shen, L., Liu, Y.: Robust Scheduling of a Hybrid Hydro/Photovoltaic/Pumped-Storage System for Multiple Grids Peak-Shaving and Congestion Management. IEEE Access. 12, 22230–22242 (2024). https://doi.org/10.1109/ACCESS.2023.3346670.
9. Calero, I., Cañizares, C.A., Bhattacharya, K., Baldick, R.: Duck-Curve Mitigation in Power Grids with High Penetration of PV Generation. IEEE Trans Smart Grid. 13, 314–329 (2022). https://doi.org/10.1109/TSG.2021.3122398.
10. Wang, W.T., Yang, H., Xiang, C.Y.: Green roofs and facades with integrated photovoltaic system for zero energy eco-friendly building – A review. Sustainable Energy Technologies and Assessments. 60, 103426 (2023). https://doi.org/10.1016/J.SETA.2023.103426.
11. Shukla, A.K., Sudhakar, K., Baredar, P.: Recent advancement in BIPV product technologies: A review. Energy Build. 140, 188–195 (2017). https://doi.org/10.1016/J.ENBUILD.2017.02.015.
12. Basher, M.K., Nur-E-Alam, M., Rahman, M.M., Alameh, K., Hinckley, S.: Aesthetically Appealing Building Integrated Photovoltaic Systems for Net-Zero Energy Buildings. Current Status, Challenges, and Future Developments—A Review. Buildings 2023, Vol. 13, Page 863. 13, 863 (2023). https://doi.org/10.3390/BUILDINGS13040863.
13. Xu, K., Song, S., Xiang, C.: Parametric design for combined solar facades for high-rise residential buildings. Sustainable Energy Technologies and Assessments. 74, 104167 (2025). https://doi.org/10.1016/J.SETA.2024.104167.
14. Goncalves, A., Rativa, D., Gomez-Malagon, L.A.: Model-Based Assessment of the Incident Angle Modifier on the Annual Angular Losses and Gain of PV Modules in Tracking Systems. IEEE J Photovolt. 14, 185–193 (2024). https://doi.org/10.1109/JPHOTOV.2023.3323802.
15. Shahrestani, M., Yao, R., Essah, E., Shao, L., Oliveira, A.C., Hepbasli, A., Biyik, E., Caño, T. del, Rico, E., Lechón, J.L.: Experimental and numerical studies to assess the energy performance of naturally ventilated PV façade systems. Solar Energy. 147, 37–51 (2017). https://doi.org/10.1016/J.SOLENER.2017.02.034.
16. Martín-Chivelet, N., Kapsis, K., Wilson, H.R., Delisle, V., Yang, R., Olivieri, L., Polo, J., Eisenlohr, J., Roy, B., Maturi, L., Otnes, G., Dallapiccola, M., Upalakshi Wijeratne, W.M.P.: Building-Integrated Photovoltaic (BIPV) products and systems: A review of energy-related behavior. Energy Build. 262, 111998 (2022). https://doi.org/10.1016/J.ENBUILD.2022.111998.
17. Carrera, L.A.I., Molina-Santana, E., Alvarez-Alvarado, J.M., Garcia-Martinez, J.R., Rodriguez-Resendiz, J.: Energy Efficiency Analysis of East-West Oriented Photovoltaic Systems for Buildings: A Technical-Economic-Environmental Approach. IEEE Access. 11, 137660–137679 (2023). https://doi.org/10.1109/ACCESS.2023.3340145.

18. Albinius, N., Rau, B., Riedel, M., Ulbrich, C.: A Comprehensive Case Study of a Full-Size BIPV Facade. Energies 2025, Vol. 18, Page 1293. 18, 1293 (2025). https://doi.org/10.3390/EN18051293.
19. Ren, H., Sun, Y., Norman Tse, C.F., Fan, C.: Optimal packing and planning for large-scale distributed rooftop photovoltaic systems under complex shading effects and rooftop availabilities. Energy. 274, 127280 (2023). https://doi.org/10.1016/J.ENERGY.2023.127280.
20. Biyik, E., Araz, M., Hepbasli, A., Shahrestani, M., Yao, R., Shao, L., Essah, E., Oliveira, A.C., del Caño, T., Rico, E., Lechón, J.L., Andrade, L., Mendes, A., Atlı, Y.B.: A key review of building integrated photovoltaic (BIPV) systems. Engineering Science and Technology, an International Journal. 20, 833–858 (2017). https://doi.org/10.1016/J.JESTCH.2017.01.009.
21. Onyx Solar: Building Integrated Photovoltaics Solutions, https://onyxsolar.com/, last accessed 2025/09/11.
22. International Electrotechnical Commission: IEC 61724-1; Photovoltaic System Performance-Part 1: Monitoring. Geneva, Switzerland (2021).
23. Schneider Electric Solar Europe: CL 33/50 String Inverte, https://solar.se.com/eu/es/product/inversor-string-cl33-50/, last accessed 2025/09/11.
24. Schneider Electric Solar Europe: Insight Energy Management Platform, https://solar.se.com/eu/en/product/insight/, last accessed 2025/09/11.
25. ABB: M4M 30 MODBUS, https://new.abb.com/products/2CSG274761R4051/m4m-30-modbus, last accessed 2025/09/11.
26. International Electrotechnical Commission: IEC 62053-22:2020; Electricity metering equipment - Particular requirements - Part 22: Static meters for AC active energy (classes 0,1S, 0,2S and 0,5S). (2020).
27. José Muñoz-Rodríguez, F., Snytko, A., de la Casa Hernández, J., Rus-Casas, C., Jiménez-Castillo, G.: Rooftop photovoltaic systems. New parameters for the performance analysis from monitored data based on IEC 61724. Energy Build. 295, 113280 (2023). https://doi.org/10.1016/J.ENBUILD.2023.113280.
28. Gilabert-Torres, C., Hontoria-García, L., Fernández-Carrasco, J.I., Mellit, A., Rus-Casas, C.: Integration of PV Systems in Urban Environments: Complementary Metrics to Analyze Their Performance. Energies 2025, Vol. 18, Page 4411. 18, 4411 (2025). https://doi.org/10.3390/EN18164411.
29. Schram, W.L., Shirazi, E.: PV on façades: A financial, technical and environmental assessment. Energy Build. 328, 115010 (2025). https://doi.org/10.1016/J.ENBUILD.2024.115010.
30. Virtuani, A., Borja Block, A., Wyrsch, N., Ballif, C.: The carbon intensity of integrated photovoltaics. Joule. 7, 2511–2536 (2023). https://doi.org/10.1016/J.JOULE.2023.09.010.
31. Thebault, M., Gaillard, L.: Optimization of the integration of photovoltaic systems on buildings for self-consumption – Case study in France. City and Environment Interactions. 10, 100057 (2021). https://doi.org/10.1016/J.CACINT.2021.100057.

32. Muñoz-Rodríguez, F.J., Gómez, P., Fernández-Carrasco, J.I., Tina, G.M., Jiménez-Castillo, G.: A new approach to analyse from monitored data the performance, matching capability and grid usage of large Rooftop Photovoltaic systems. Case of study: Photovoltaic system of 1.05 MW installed at the campus of University of Jaén. Renew Energy. 239, 121947 (2025). https://doi.org/10.1016/J.RENENE.2024.121947.
33. Gao, L., Wang, S., Mao, M., Liu, C., Li, T.: Study on the Energy Consumption Characteristics and the Self-Sufficiency Rate of Rooftop Photovoltaic of University Campus Buildings. Energies 2024, Vol. 17, Page 3535. 17, 3535 (2024). https://doi.org/10.3390/EN17143535.
34. Amini Toosi, H., Lavagna, M., Leonforte, F., Del Pero, C., Aste, N.: Building decarbonization: Assessing the potential of building-integrated photovoltaics and thermal energy storage systems. Energy Reports. 8, 574–581 (2022). https://doi.org/10.1016/J.EGYR.2022.10.322.

Regeneration of absorbent solution in a desorption column for post-combustion CO_2 capture at laboratory scale.

K.A. Altamirano-Tlali[1[0000-0001-9064-3541]] Laura Castro[1[0000-0003-2507-0067]], Miriam Navarrete Procopio[2[0000-0002-0247-9647]] Víctor Zezatti[1[0000-0001-6470-6772]] and J.C. Garcia[1[0000-0002-1650-2932]].

[1] Centro de Investigación en Ingeniería y Ciencias Aplicadas (CIICAp) Universidad Autónoma del Estado de Morelos (UAEM)
[2] Facultad de Ciencias Químicas e Ingeniería (FCQeI) Universidad Autónoma del Estado de Morelos (UAEM)
miriam.navarrete@uaem.mx

Abstract. Global warming is caused by an excess of greenhouse gases. CO_2 is one of the gases that directly contribute to this problem. In Mexico, the main source of CO_2 emissions is industry, which generates electricity and heat by burning fossil fuels (oil, natural gas, and coal). As a result, CO_2 capture technologies have been developed to mitigate this issue. Among these technologies is post-combustion carbon capture, which involves an absorption-desorption process using an amine absorbent solution to capture CO_2. One disadvantage of this technology is the energy consumption required to regenerate the absorbent solution. The desorption process is carried out in packed columns to regenerate the CO_2-rich solution from the absorption process. This work presents experimental tests on a laboratory-scale post-combustion CO_2 capture system for the regeneration of the monoethanolamine (MEA) absorbent solution at different concentrations by weight (10, 15, 20, 25, and 30%). The results presented demonstrate the regeneration of the absorbent solution at these different concentrations; their validation is achieved through refractive index, saturation, and reverse engineering tests.

Keywords: Desorption process, CO_2 capture, absorbent solution, post-combustion CO_2 capture.

1 Introduction

Environmental pollution is a persistent and growing problem that affects the entire planet and the beings that inhabit it [1,2]. Among the many types of pollution, one of the main ones is air contamination, which is mainly caused by anthropogenic activities that generate greenhouse gases, including Carbon Dioxide (CO_2) [3]. These emissions are the primary contributor to climate change, as they are a greenhouse gas that traps heat in the atmosphere and contributes to global warming [4]. The increase in atmospheric CO_2 concentration is due to the burning of fossil fuels [5]. Carbon dioxide (CO_2) emissions from thermoelectric power plants are one of the most significant sources of greenhouse gases worldwide, as they use coal, natural gas, and oil to generate electricity [6]. In Mexico, the Federal Electricity Commission (CFE) has pointed out that thermoelectric power plants account for a significant fraction of national CO_2 emissions, especially those that use fuel oil and coal as their primary energy source [7]. Given this situation, it is essential to advance in the development and implementation of mitigation and adaptation strategies that reduce CO_2 emissions. One of the proposals to mitigate this problem is CO_2 capture technologies, also known as CCS (Carbon Capture and Storage), whose main objective is to prevent CO_2 generated from industrial processes or energy generation from being released into the atmosphere. These technologies play an important role in reducing emissions and mitigating climate change. [8]. CO_2 capture technologies are grouped into three categories [9]: Pre-combustion (prior to combustion), oxy-combustion (oxygen combustion) and post-combustion (after combustion). Post-combustion capture technology uses chemical absorption with aqueous amine solutions. It is currently the most developed and feasible option for removing CO_2 from the gases emitted by fossil fuel-fired thermoelectric plants. [10]. This technique is particularly useful for treating gas streams with low partial pressures of CO_2, as is often the case in the waste gases from conventional power plants [11]. One of the main advantages of this post-combustion CO_2 capture method is its ability to be integrated into existing power plants without requiring significant modifications. CO_2 is removed from the combustion gas using an amine-based solvent, which is regenerated at high temperatures and reused continuously. However, its main disadvantage lies in the high energy consumption required to regenerate the solvent [12]. Recent studies of the desorption process focus on minimizing the energy required to regenerate the absorbent solution by using additives to the MEA solution or other absorbent solutions in the post-combustion process [13-16].In this work, tests are carried out on the desorption of saturated CO_2 solutions at different weight concentrations of monoethanolamine (MEA) in a laboratory-scale desorption process.

2 Methodology

2.1 System description

The desorption process is a unit operation that separates a gaseous solute dissolved in a liquid solution [17]. This process is carried out in a packed column operating with countercurrent gas and liquid flows. Saturated steam is used as the stripping agent and is subsequently condensed. Figure 1 shows the configuration of the desorption system.

The system consists of two main streams. In Stream 1, the solution to be desorbed is pumped (P-101) from the storage tank (T.K.-105). In Stream 2, steam is supplied through a steam generator (C.V.-101). The steam enters through the bottom of the packed column (P.T.-102). At the outlet (Stream 3), the steam is condensed using a heat exchanger (C.N.-101) and stored in a tank (T.K.-106). The liquid from the desorption process exits in stream 4 and is collected in another storage tank (T.K.-104).

Temperature is monitored by four sensors installed at the inlet and outlet of both the steam and the solution streams.

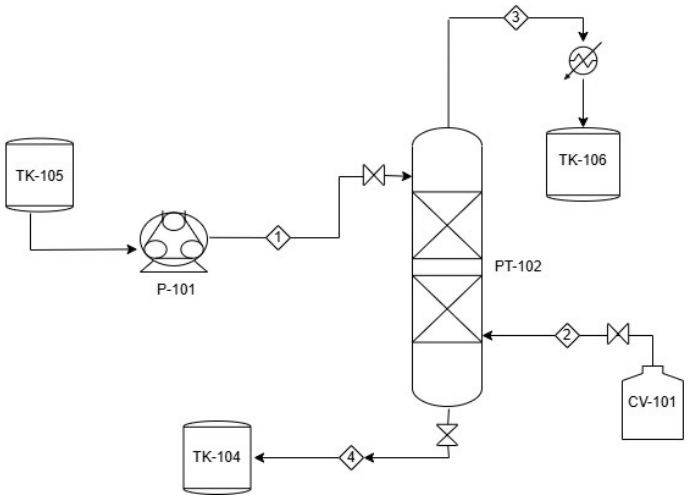


Fig. 1. Diagram of the laboratory-scale desorption process.

2.2 Experimental setup

In the desorption process, there are three types of variables: controlled, uncontrolled, and output. Controlled variables are those that clarify their behavior when observed with different values. Uncontrolled variables, on the other hand, depend on external factors, such as climate and steam inlet conditions. Output variables are the result of the desorption process.

For the tests, changes were applied to the input variables to evaluate the sensitivity of the desorption system. In the context of the desorption process, a total of 15 variables are considered, classified into three groups. Figure 2 illustrates the variables of the desorption process. The first group includes controlled variables, such as steam opening, temperature, and cooling fluid flow. The second group consists of uncontrolled variables, such as steam generator pressure, steam inlet temperature, and humidity. The third group includes output variables, such as solution and steam temperature, as well as condensate temperature and the volume of condensate and solution.

The experimental design enables the identification of variables that exhibit the highest sensitivity within the desorption process, thereby facilitating the establishment of optimal operating conditions. The volume of condensate is a particularly significant metric for evaluating the regeneration efficiency of the absorbent solution. The volume of condensate is a particularly significant metric for evaluating the regeneration efficiency of the absorbent solution.

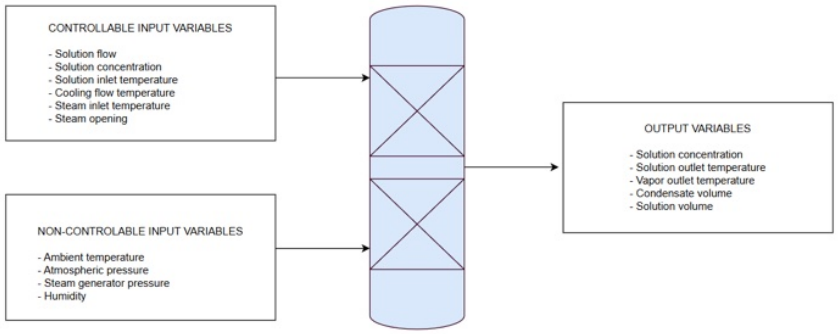


Fig. 2. Experimental Setup

2.3 Experimental methodology of the absorption process

The experimental procedure involves preparing solutions in weight percentage (w/w) of monoethanolamine (MEA) and water (H_2O) at varying concentrations and operating an absorption tower to evaluate the CO_2 capture process. Initially, the tower valves are checked to ensure they are closed, and the compressor is turned on to establish the air flow. The peristaltic pump is turned on to fill the tower at a controlled flow rate. Once the operating level is reached, the initial temperature (t = 0) is recorded and the CO_2 supply is opened. The air and CO_2 flows are adjusted using solenoid valves and measured with flow meters, considering different concentrations of MEA and CO_2 according to the test conditions.

The temperature is monitored at 4-second intervals using a data acquisition system. The instruments used for the collection of experimental data of the post-combustion process on a laboratory scale have been previously calibrated. [18]. The results (temperature and concentration of absorbed CO_2) are recorded until a stable state is reached. At the end of the experiment, the CO_2 flow is shut off, maintaining only the air flow, the outlet valve is opened, and the absorbent solution is transferred to a storage tank.

2.4 Experimental methodology of the desorption process

Once the absorbent solution (MEA + H_2O + CO_2) is saturated, it undergoes a desorption process to regenerate it. For this purpose, a saturated steam generator is used to feed the desorption column. The system must be thermally stabilized for 12 minutes, after which the CO_2-rich solution enters the top of the column via a peristaltic pump, coming into contact with the rising steam. In this process, the steam carries away the released CO_2, which then passes through a heat exchanger, while the CO_2-poor solution is collected at the bottom.

The temperature is monitored at 3-second intervals using a data acquisition system. Finally, the steam generator and valves are turned off, and the experimental results for temperatures, condensates (CO_2-poor solution), and refractive index are analyzed to determine the new concentration of the regenerated solution.

Tables 1 and 2 show the operating conditions of the controlled and non-controlled variables, respectively. The flows for operating the desorption column are established from the absorption process to ensure saturation of the absorbent solution [19]. The concentrations of the absorbent solution are less than 30% by weight (MEA-H_2O) to avoid corrosion problems [20,21]. The solution to be desorbed enters at room temperature, and the cooling fluid to condense the steam at the outlet of the desorption column was controlled at room temperature. For the desorption process, saturated steam was used.

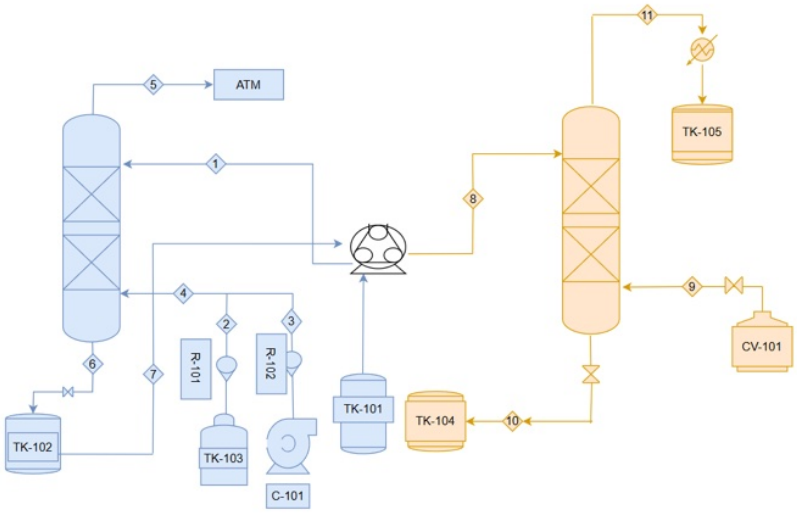


Fig. 3. Laboratory-scale CO_2 capture process.

Table 1. Controllable Operating Conditions

Variables	Value	Units
Solution Flow	65, 41, 36, 31 and 26	mL/min
Solution Concentration	10, 15, 20, 25, and 30	% w/w
Solution inlet temperature	20	°C
Cooling Flow temperature	20	°C
Steam Inlet temperature	95	°C
Steam flow	31.6	mL/min

Table 2. Non-controllable operating conditions

Variables	Value	Units
Ambient temperature	18-20	°C
Atmospheric pressure	0.8	atm
Steam generator pressure	4.4	atm
Humidity	40-70	%

3 Results and discussion

In this section, the regeneration of the absorbent solution in a packed desorption column from the post-combustion CO_2 capture process will be demonstrated on a laboratory scale. The methods used are refractive index measurements and absorption tests.

3.1 Refractive index analysis.

Table 3 shows the results of the refractive index measurements from the samples of solutions to be desorbed (saturated with CO_2) and the samples after the desorption process (free of CO_2). The results were compared with the data reported in the literature. The last column of the results table shows the new weight composition of the regenerated solution. With the new concentration of the regenerated solution, it can be seen that the solution is diluted due to the steam condensed during the desorption process from the steam generator. The volume condensed in the desorbed solution tank (TK-104) was 280 mL.

Table 3. Refractive index of the desorbed solution

Initial Concentration from the absorbent solution %	CO_2 saturated IR [22]	CO_2 free IR [22]	CO_2 saturated IR (exp)	CO_2 free IR (exp)	New concentration %
10	1.3550	1.3465	1.3552	1.3541	8.72
15	1.3660	1.3520	1.3593	1.3492	12.54
20	1.3770	1.3590	1.3682	1.3592	16.29
25	1.3870	1.3645	1.3850	1.3641	24.63
30	1.3985	1.3720	1.3965	1.3621	23.40

Table 4 shows the new concentrations obtained from the desorption process. The regeneration concentrations of the absorbent solutions are lower than the initial concentrations of the absorbent solutions prepared for the post-combustion CO_2 capture process. The concentrations obtained by refractive index readings are compared with the analytically calculated concentrations. The analytical calculations considered the 280 mL of steam condensed in the regenerated solution. These results provide the new concentrations of the regenerated solutions.

Table 4. Concentration comparison

Initial Concentration %	Regenerated concentration (I.R) %	Regenerated concentration (analytic) %
10	8.72	7.86
15	12.54	11.79
20	16.29	16.92
25	24.63	19.37
30	23.40	23.57

3.2 Saturation Tests Analysis

To validate that absorbent solution was regenerated at the new concentrations, tests were carried out in the absorption column. The tests consisted of absorbing CO_2 with the new regenerated solutions, expecting saturation behavior. Saturation of the MEA solution occurs when the maximum load is reached according to the stoichiometric reaction (0.5 mol-CO_2/mol-MEA). As this is an exothermic reaction, an increase in temperature in the MEA solution indicates that CO_2 is being absorbed. A decrease in temperature indicates that no absorption occurred because there are no more moles of MEA to absorb the moles of CO_2. Saturation time is when the maximum temperature increase of the absorbent solution is reached. The saturation of the regenerated solutions is shown in Figure 4. The regenerated solutions are capable of absorbing CO_2 again until they become saturated once more.

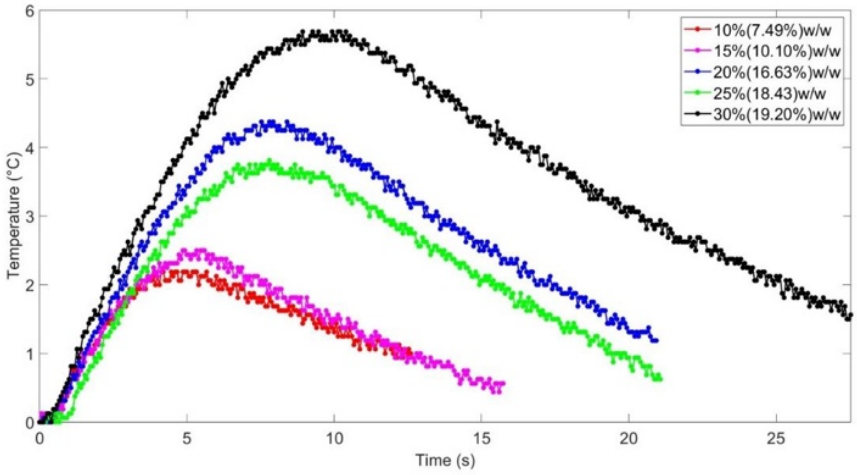


Fig. 4. Regenerated solutions saturation

Theoretical saturation time.

Saturation of the solution occurs when the maximum load of 0.5 (molCO_2/molMEA) is reached according to the stoichiometric reaction:

$$2\,MEA + CO_2 \leftrightarrow MEACOO^- + MEA^+ \text{ [23]}$$

For the analytical estimation of the saturation time of each MEA solution, the equation where MEA (kg) is the total amount of MEA, P (kPa) is the operating pressure, T (K) is the gas stream temperature, and v CO_2 (L/min). The mathematical expression for this results from a combination of the stoichiometry of the reaction, the continuity equation, and the consideration of ideal gases.

$$t_{sat} = \frac{67.17 mol_{MEA} T_{gas}}{P v_{CO_2}} \quad (1)$$

Table 5 shows the results of the concentration obtained from the regenerated solution based on saturation times; the desorbed solutions are diluted in the process, which indicates a new concentration. The saturation times were obtained experimentally, and the concentrations from those times using the stoichiometry of the reaction (0.5 mol-CO_2/mol-MEA) and the molar flow of CO_2 used for saturation (1 L/min). It can be observed that the concentrations calculated with the saturation times are close to those calculated from the condensed volume and those obtained with the refractive index. (Table 4).

Table 5. New concentration of the regenerated solution by saturation time

Initial concentration %	Experimental saturation time (min)	Regenerated concentration (%)
10	4.06	7.49
15	5.47	10.10
20	9.01	16.63
25	10.00	18.43
30	10.40	19.20

4 Conclusion

The regeneration of the absorbent solution was demonstrated in a laboratory-scale desorption column integrated into the post-combustion CO_2 capture process. The concentrations of the regenerated solution were obtained using refractive index and analytical calculations based on the volumes condensed in the desorption process and the saturation times when reusing the regenerated solution in the absorption process. The regenerated solutions decreased their initial concentrations from the CO_2 capture process due to the volumes condensed in the desorption process using saturated steam from the steam generator. These results reveal the new concentrations of the regenerated solutions for the post-combustion CO_2 capture process, which can then be fed back into the process. Characterizing the free and saturated CO_2 absorbent solution is important to verify that the separation process is occurring and, subsequently, to optimize the technology to minimize the energy required to desorb the absorbent solution.

Disclosure of Interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Our World in Data. (s. f.). Our World in Data. Global Change Data Lab. Recuperado de https://ourworldindata.org/
2. Apesteguia, C., et al.: Greenhouse gases emissions and global climate change: Examining the influence of CO_2, CH_4, and N_2O. *Science of The Total Environment* (2024).
3. International Energy Agency (IEA): Global Energy–Related CO_2 emissions increased by 0.8% in 2024, reaching 37.8 Gt CO_2. IEA, Paris (2025).
4. INECC-SEMARNAT. Primer Informe Bienal de Actualización ante la Convención Marco de las Naciones Unidas sobre el Cambio Climático, México, 2015; INECC: Mexico City, Mexico, 2023
5. Intergovernmental Panel on Climate Change (IPCC): Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge (2021).
6. Martínez, F.G.: Cambio climático, calentamiento global y efecto invernadero, ¿Cuál es cuál? Ciencia (2021).
7. Comisión Federal de Electricidad (CFE): Informe Anual 2023. CFE, México (2023), https://www.cfe.mx, last accessed 2025/09/12.
8. Instituto Nacional de Ecología y Cambio Climático (INECC): https://datos.gob.mx/busca/organization/about/inecc, last accessed 2025/09/12
9. Clemente Jul, María del Carmen and Rodrigo Naharro, Julio (2009). Comparación de las tecnologías de captura y almacenamiento de CO2 en la generacion de energía.. "Anales de la Real Academia de Doctores de España", v. 13 (n. 2); pp. 19-29. ISSN 1138-2414.

10. Wang, Y., Zhao, L., Otto, A., Robinius, M., Stolten, D.: A review of post-combustion CO_2 capture technologies from coal-fired power plants. Energy Procedia 114, 650–665 (2017). https://doi.org/10.1016/j.egypro.2017.03.1209
11. Mores, P., Scenna, N., Mussati, S.: CO_2 capture using monoethanolamine (MEA) aqueous solution: Modeling and optimization of the solvent regeneration and CO_2 desorption process. Energy 45, 1042–1058 (2012). https://doi.org/10.1016/j.energy.2012.06.038
12. Faber, R., Köpcke, M., Biede, O., Knudsen, J.N., Andersen, J.: Open-loop step responses for the MEA post-combustion capture process: Experimental results from the Esbjerg pilot plant. Energy Procedia 4(1), 1558–1565 (2011). https://doi.org/10.1016/j.egypro.2011.02.025
13. Elmarghni, S., Ansarpour, M., & Borhani, T. N. (2025). Effect of Different Amine Solutions on Performance of Post-Combustion CO2 Capture. *Processes*, *13*(8), 2521.
14. Han, J., Qiu, Z., Chen, Y., Gui, X., & Chen, X. (2024). A Novel Phase Change Absorbent for CO2 Capture with Low Viscosity and Effective Absorption–Desorption Properties. *Energy Technology*, *12*(9), 2400114.
15. Zhang, W., Deng, Z., Wang, Q., Li, J., & Qiu, X. (2021). Experimental research on chemical desorption based on CO2-rich absorption solutions. *International Journal of Greenhouse Gas Control*, *109*, 103356.
16. Bai, Y., Wang, K., Wang, L., Zeng, S., Bai, L., & Zhang, X. (2023). Experimental investigation on post-combustion CO2 capture for [Bpy][NO3] and MEA aqueous blends with lower regeneration energy. *Separation and Purification Technology*, *327*, 124846.
17. Mores, P., Scenna, N., & Mussati, S. (2012). CO2 capture using monoethanolamine (MEA) aqueous solution: Modeling and optimization of the solvent regeneration and CO2 desorption process. Energy, 45(1), 1042-1058. https://doi.org/10.1016/j.energy.2012.06.038
18. León, M. F. S., Castañeda, E. O., Flores, Z., & Becerro, Á. T. Instrumentación y control de un proceso de absorción a escala laboratorio. 1 Avances, 17.
19. Navarrete Procopio, M., Urquiza, G., & Castro, L. (2023). Analysis of Absorber Packed Height for Power Plants with Post-Combustion CO2 Capture. Sustainability, 15(12), 9536. https://doi.org/10.3390/su15129536
20. He, S., Qiu, Y., Sun, Y., Zhang, Z., Cheng, J., Gao, C., & Zhao, Z. (2020). Corrosion behavior of AISI 1020 steel in MEA and [Bmim]BF4 mixed solution containing saturated CO2. International Journal of Greenhouse Gas Control, 94, 102931. https://doi.org/10.1016/j.ijggc.2019.102931
21. Kittel, J., Idem, R., Gelowitz, D., Tontiwachwuthikul, P., Parrain, G., & Bonneau, A. (2009). Corrosion in MEA units for CO2 capture: Pilot plant studies. Energy Procedia, 1(1), 791-797. https://doi.org/10.1016/j.egypro.2009.01.105
22. Procopio, M. N., Urquiza, G., Castro, L., & Zezatti, V. (2023). Saturation of the MEA solution with CO2: Absorption prototype and experimental technique. Results in Engineering, 19, 101286. https://doi.org/10.1016/j.rineng.2023.101286
23. Fair, J.R.; Seibert, A.F.; Behrens, M.; Saraber, P.P.; Olujic, Z. Structured packing performance experimental evaluation of two predictive models. Ind. Eng. Chem. Res. 2000, 39, 1788–1796.

Optimizing Solar Panel Allocation in Smart-City Buildings Using Genetic Algorithms

Ricardo Alcaraz Fraga
Jorge Ponciano Escamilla Ambrosio

September 2025

Abstract

As climate change consequences become more serious than ever, renewable energies arise as an alternative to satisfy electric energy necessities in a sustainable way. One of these alternatives is photovoltaic energy via solar panels where concepts such as distributed generation appear. Distributed generation is a really important concept for designing smart cities solar panel's distribution as a good allocation can account for most of the city's electric energy necessities. Therefore, this papers presents a genetic algorithm approach to solve solar panel allocation in smart city buildings. Simulation results show that the proposed method achieves significant reductions in energy losses and demonstrates scalability from small test cases to larger urban scenarios.

1 Introduction

The consequences of climate change are becoming increasingly severe. According to a recent report, 48% of the Earth's surface has experienced extreme drought, and "the health threats posed by climate change have already reached record levels," as denounced in the new climate crisis report prepared by 122 international experts from 57 major institutions (including the WHO) and published in the medical journal The Lancet[1].

In light of these consequences and concerns, it is essential to seek alternatives for energy production, focusing on renewable sources such as photovoltaic energy. As the United Nations has highlighted[2], fossil fuels account for 80% of global energy production, and their use exacerbates the aforementioned consequences since they are responsible for more than 75% of total global greenhouse gas emissions.

According to two UN reports[3], in 2024 renewable energies accounted for 74% of global electricity growth, with 92.5% of newly added capacity coming from green energy. Furthermore, solar energy is 41% cheaper than the least expensive fossil fuel, while wind power is 53% cheaper. Investments in renewables reached two trillion dollars.

Based on this data, a future urban scenario can be envisioned in which the use of solar panels is a common and standardized practice, ensuring interoperability across devices (i.e., solar panels from different providers are mutually compatible). This would make it feasible to relocate panels from one building to another and, therefore, to pose the question: How should solar panels be distributed among buildings to best meet electricity needs through photovoltaic energy?

2 Related Work

Genetic Algorithms (GAs) have been extensively used for energy management in microgrids and buildings to minimize operating cost, grid purchases, and emissions under operational constraints. In grid-connected PV–battery microgrids, Nouar et al. formulate a GA-based day-ahead scheduling EMS and report cost reductions against a heuristic baseline, directly aligning with objectives that penalize energy bought from the grid.[4] At the building scale, Gong integrates rule-based control with a GA inside a Building Energy Management System and shows lower operating expenditures versus SCADA/BAS–PSO/LSTM-CNN baselines, which is relevant when the primary objective is to reduce net imports while maintaining feasibility.[5]

A closely related line is optimal component sizing and energy management of hybrid microgrids using multiobjective or hybrid GAs. Jarso et al. employ a hybrid GA/NSGA-II to size a PV–wind–diesel–battery system, optimizing annualized cost and reliability; the workflow generalizes to grid-connected contexts by substituting reliability metrics with net exchange or cost-of-purchase terms.[17] Beyond sizing, GA-based frameworks have also been combined with simulation/learning stacks (e.g., EnergyPlus/JEPlus and ML) to jointly optimize energy use and comfort; such architectures motivate extending a single-objective "minimize grid purchases" loss into multiobjective formulations.[17]

At distribution-network level, GA/NSGA-II has been applied to PV siting and sizing to reduce losses and improve voltage profiles—an analogue to minimizing purchased energy insofar as technical losses and voltage quality shape the required imports. Alizadeh et al. solve multiobjective placement/sizing on IEEE test feeders with NSGA-II and report improved loss and voltage indices.[13] From the territorial/planning perspective, Nagkoulis et al. couple GA with GIS to select PV sites that minimize social/visual disturbance (SDIS) under land-use and techno-economic constraints, illustrating how spatial/regulatory objectives can be integrated into GA-based energy planning.[8]

Regarding genetic operators, real-coded GAs commonly use parent-centric (e.g., SBX/PCX) and mean-centric (e.g., BLX-α/UNDX/SPX) crossovers. Deb and Jain show that self-adapting between parent- and mean-centric behaviors (SA-SBX) improves performance across problem phases.[9] If our implementation uses a proportion-blend crossover, it is natural to relate it to BLX-α (controlled mixing around parents) and SBX, discussing exploration (out-of-range sampling), diversity preservation, and convergence in energy-cost landscapes.[16,

3 Theoretical Framework

3.1 Renewable Energy and Smart Cities

The urgency of climate change has accelerated the adoption of renewable energy sources as a cornerstone of smart-city planning. Among these, photovoltaic (PV) systems stand out due to their modularity, declining costs, and ease of deployment in urban environments.[11] Smart cities integrate distributed PV generation to reduce greenhouse gas emissions, lower dependence on fossil-based centralized plants, and enhance resilience.

3.2 Distributed Generation and Energy Management

Distributed generation refers to the production of electricity close to the point of consumption. In urban contexts, panel allocation directly determines the share of demand covered locally versus purchased from the grid.[12] The total grid exchange can be modeled as:

$$G = \sum_{i=1}^{N} \max\{0, D_i - P_i^{\mathrm{PV}}\},$$

where D_i is demand and P_i^{PV} is PV generation in building i. Minimizing G directly reduces grid purchases and operational costs.

3.3 Optimization Problem Formulation

The allocation problem can be expressed as a constrained non-linear optimization, where decision variables are the PV capacities x_i assigned to each building. Constraints typically include rooftop area, investment budgets, and technical regulations.[13] This combinatorial nature makes deterministic approaches (e.g., LP, MILP) less effective in realistic settings.

3.4 Genetic Algorithms for Energy Optimization

Genetic Algorithms (GAs) are population-based metaheuristics inspired by natural evolution, well-suited for non-linear and discrete problems.[14] Each individual encodes an allocation vector $\mathbf{x}$, and operators evolve the population:

- **Selection:** Tournament selection favors fitter individuals (lower grid purchase).[15]
- **Crossover:** Operators such as Blend Crossover (BLX-α) or Simulated Binary Crossover (SBX) generate offspring by mixing parental traits.[16]
- **Mutation:** Introduces diversity to escape local optima.

3.5 Advantages of Metaheuristics

Compared with exact optimization, GAs and other metaheuristics provide robustness in handling non-convex search spaces, hybrid decision variables, and multi-objective tradeoffs.[17] They have been successfully applied to microgrid scheduling, PV siting, and hybrid renewable sizing, supporting their suitability for the smart-city PV allocation problem.

4 Methodology

4.1 Problem Definition

The problem addressed in this work is the optimal allocation of photovoltaic (PV) panels across smart-city buildings in order to minimize energy purchased from the grid. The city is represented as a set of N buildings, each with a specific demand profile D_i and rooftop availability. The objective function is formulated as the total purchased energy:

$$f(\mathbf{x}) = \sum_{i=1}^{N} \max\{0, D_i - P_i^{\mathrm{PV}}(x_i)\},$$

where $\mathbf{x} = (x_1, x_2, \ldots, x_N)$ denotes the decision vector representing the number of panels (or PV capacity) allocated to each building. Constraints include maximum rooftop area and non-negativity of the decision variables.

4.2 Genetic Algorithm Framework

Given the non-linear and combinatorial nature of the allocation problem, a Genetic Algorithm (GA) was employed to search for near-optimal solutions. The GA implementation follows the standard evolutionary cycle: initialization, fitness evaluation, selection, crossover, mutation, and termination.

4.2.1 Encoding

Each candidate solution (chromosome) is represented as an integer vector $\mathbf{x}$, where gene x_i encodes the number of panels assigned to building i. This representation directly matches the decision variables of the optimization problem.

4.2.2 Fitness Function

The fitness function is defined as the negative of the objective function, so that lower grid purchases correspond to higher fitness:

$$\mathrm{Fitness}(\mathbf{x}) = -f(\mathbf{x}).$$

This ensures that individuals yielding less purchased energy are favored during selection.

4.2.3 Population Initialization

The initial population of size M is generated randomly within feasible bounds. Each gene x_i is initialized to a value consistent with rooftop capacity and project constraints, ensuring all individuals are valid solutions.

4.2.4 Selection Mechanism

Tournament selection was used to choose parents for reproduction. In each tournament, a subset of individuals is sampled randomly from the population, and the one with the highest fitness is selected. This balances exploitation of good solutions with exploration of diverse candidates.

4.2.5 Crossover Operator

A Proportion Blend Crossover (PBC) operator was implemented to generate offspring. Given two parents, the operator produces a child by interpolating gene values proportionally between the parental genes. This operator allows the offspring to inherit characteristics of both parents while exploring new allocation patterns.

4.2.6 Mutation Operator

To maintain diversity, mutation was applied with a small probability p_m to each gene. Mutation perturbs x_i by adding or subtracting a small random integer within valid bounds. This prevents premature convergence and helps explore unvisited regions of the solution space.

4.2.7 Termination Criteria

The algorithm runs for a fixed number of generations G or until the best fitness does not improve over a given number of consecutive generations. The best solution found is reported as the approximate optimum.

4.3 Implementation

The GA was implemented in Python within a Jupyter notebook. The codebase includes modular functions for energy modeling, population management, selection, crossover, and mutation, ensuring reproducibility and transparency. Computational experiments were executed with consistent random seeds for comparability across runs.

4.4 Validation Approach

To validate the methodology, multiple runs of the GA were performed with varying population sizes and generation counts. Results were compared to heuristic baselines (random allocation and uniform distribution) to confirm that the GA consistently achieves lower grid purchases.

5 Results

5.1 Energy Production and Consumption Profiles

The first step in the analysis consisted of simulating photovoltaic (PV) generation for each building using the `simulate_production` function. Figure 1 shows the temporal production profiles obtained for the set of buildings, highlighting differences in magnitude due to panel allocation.

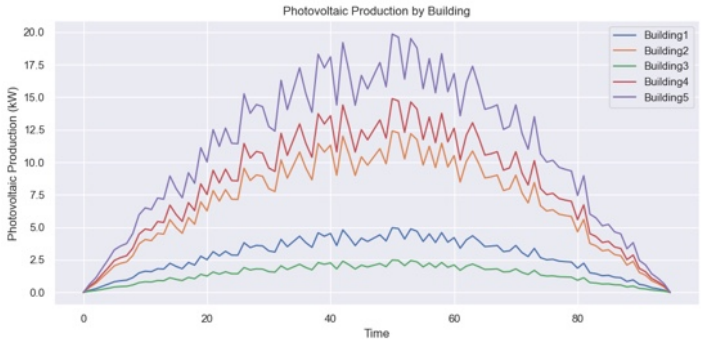


Figure 1: Simulated photovoltaic production per building.

Consumption was modeled with stochastic variability using the `simulate_consumption` function. Figure 2 presents the resulting load profiles, which serve as the demand baseline against which production is compared.

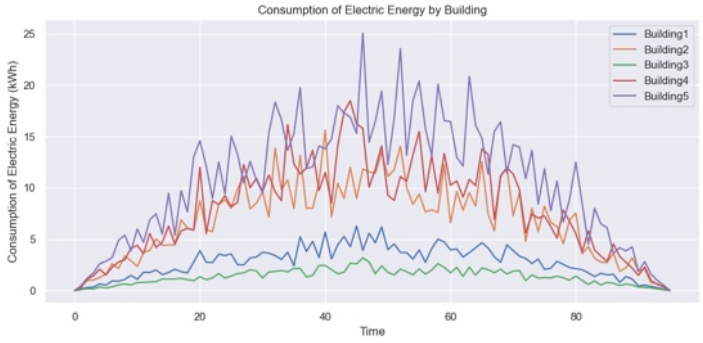


Figure 2: Simulated consumption per building with stochastic variation.

5.2 Net Energy Balance

By aligning production and consumption, the net exchange with the grid was computed. For each building, intervals where demand exceeded PV generation

resulted in energy purchased from the grid, while surpluses were considered fully self-consumed. Figure 3 illustrates this balance for a representative building, highlighting both PV utilization and purchased energy.

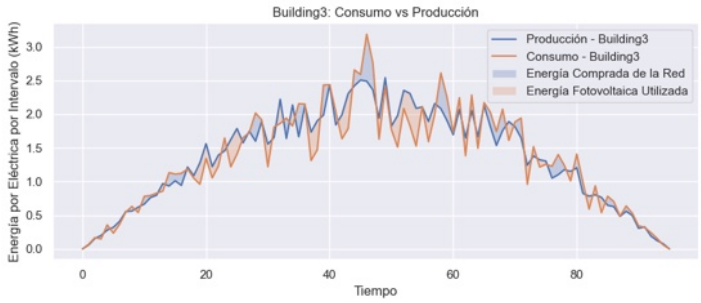


Figure 3: Consumption vs. production for Building 3. The shaded areas represent purchased energy (red) and self-consumed PV (green).

The aggregate purchased energy across all buildings was quantified using the `get_loss` function, which sums the positive demand–production differences. This value constitutes the loss metric minimized by the optimization procedure.

5.3 Genetic Algorithm Optimization

The complete GA was applied to optimize the allocation of PV panels across buildings. The population was initialized with feasible allocations using a multinomial distribution. Tournament selection, proportion blend crossover, and mutation were iteratively applied. The fitness of each individual was evaluated by the total purchased energy.

Although stochastic by nature, the GA consistently reduced the loss metric across generations. Figure 4 depicts the convergence trend for one representative run, showing a decrease in purchased energy as the algorithm progressed.

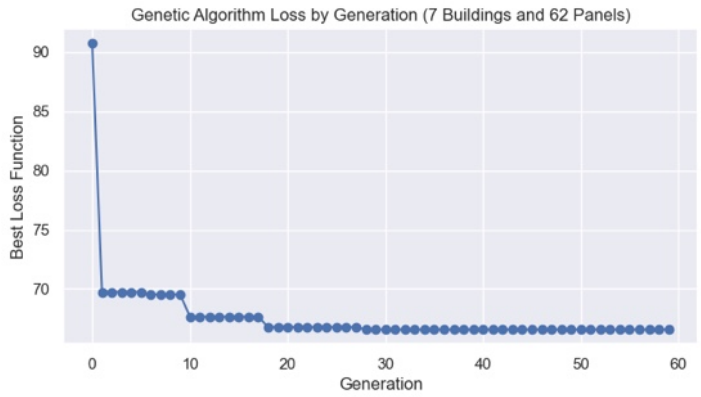


Figure 4: Convergence of the GA: reduction of purchased energy across generations.

5.4 Final Allocation

Figure 5 presents the optimal PV panel distribution across the seven buildings obtained by the GA. The allocation concentrates more panels on Buildings 4 and 7, reflecting their relative contribution to reducing grid purchases given their demand and feasible rooftop capacity.

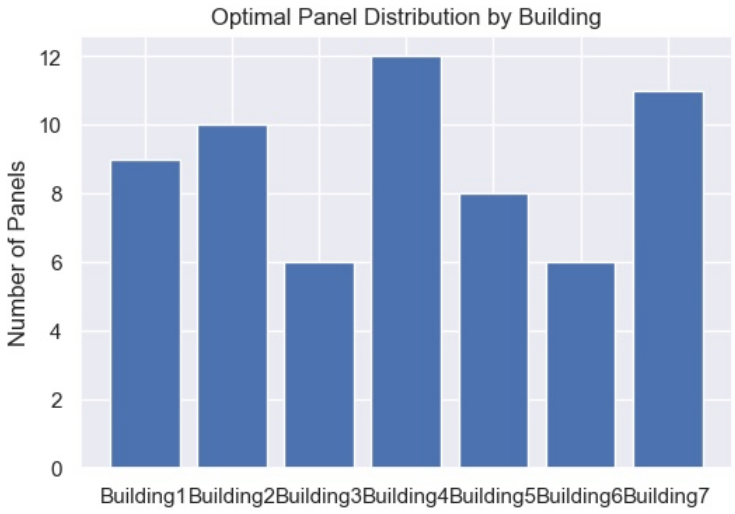


Figure 5: Optimal panel distribution by building obtained with the GA.

Table 1 reports the same allocation in tabular form for reproducibility and

reference.

Table 1: Optimized PV panel allocation per building

Building	**Allocated PV Panels**
Building1	9
Building2	10
Building3	6
Building4	12
Building5	8
Building6	6
Building7	11
Total	**62**

6 Conclusions

This paper presented a genetic-algorithm (GA) approach to optimally allocate photovoltaic (PV) panels across smart-city buildings with the explicit objective of minimizing energy purchased from the grid. The problem was formulated directly on the operational metric of interest—the positive part of demand minus PV generation—and solved with an integer, real-coded GA that employs tournament selection, a proportion-blend crossover, mutation with feasibility repair, and a loss function equal to the aggregate grid imports. The methodology was implemented in Python within a reproducible notebook, including explicit simulation of production and consumption profiles and a transparent evaluation of the objective at each generation.

Conceptually, the contribution is twofold. First, the optimization is anchored to a physically and economically meaningful quantity (purchased energy), avoiding surrogate objectives that may dilute operational relevance. Second, the combination of proportion-blend crossover with tournament selection proved effective at exploring diverse allocation patterns while preserving feasibility, leading to stable solutions that align PV capacity with demand heterogeneity and rooftop constraints across buildings.

From a planning perspective, the resulting allocations concentrate panels where marginal reductions in purchased energy are highest, reflecting both load profiles and rooftop limits. This behavior is consistent with prior evidence that metaheuristics are well suited to non-convex, constrained city-scale energy problems, while keeping the modeling and implementation lightweight enough for rapid iteration and scenario testing.

Limitations. The present study focuses on PV-only allocation without storage, tariff structures, curtailment modeling, or network constraints (e.g., voltage limits, feeder capacities). PV surplus is assumed self-consumed within each building (no peer-to-peer exchanges), and the demand/production profiles are simulated rather than measured. These choices simplify inference but limit operational fidelity.

Future work. Natural extensions include: (i) incorporating storage and time-of-use tariffs to capture arbitrage and peak-shaving effects; (ii) adding distribution-network constraints or co-optimization of siting/sizing to internalize technical limits; (iii) multiobjective formulations that trade off purchased energy with cost, emissions, or comfort; (iv) calibration with measured data and uncertainty propagation; and (v) benchmarking the GA against alternative heuristics (e.g., NSGA-II, PSO) and exact methods where tractable. These extensions would increase realism and help position the proposed approach within a broader decision-support toolkit for urban renewable integration.

References

[1] Cadena SER, "La mitad de la Tierra sufre ya sequía extrema por el cambio climático," Oct. 30, 2024. [Online]. Available: https://cadenaser.com/nacional/2024/10/30/la-mitad-de-la-tierra-sufre-ya-sequia-extrema-por-el-cambio-climatico-cadena-ser/

[2] United Nations, "Renewable energy – powering a safer future," [Online]. Available: https://www.un.org/es/climatechange/raising-ambition/renewable-energy

[3] Associated Press, "Climate change report: Solar, wind power surged as fossil fuels faltered," Oct. 30, 2024. [Online]. Available: https://apnews.com/article/climate-change-solar-wind-power-fossil-fuels-6aca4846e594ea8405f91edda39a03ad

[4] Y. Nouar, A. Boukadoum, and O. Boudebbouz, "Optimization of Microgrid Energy Management using a Genetic Algorithm," *Engineering, Technology & Applied Science Research*, vol. 15, no. 3, pp. 23742–23747, Jun. 2025. [Online]. Available: https://etasr.com/index.php/ETASR/article/view/10278 DOI: https://doi.org/10.48084/etasr.10278

[5] J. Gong, "Building energy management model integrating rule-based control algorithm and genetic algorithm," *International Journal of Renewable Energy Development*, vol. 14, no. 1, pp. 136–145, Jan. 2025. [Online]. Available: https://ijred.cbiore.id/index.php/ijred/article/view/60628 DOI: https://doi.org/10.61435/ijred.2025.60628

[6] A. K. Jarso, G. Jin, and J. Ahn, "Hybrid Genetic Algorithm-Based Optimal Sizing of a PV–Wind–Diesel–Battery Microgrid: A Case Study for the ICT Center, Ethiopia," *Mathematics*, vol. 13, no. 6, Art. 985, 2025. [Online]. Available: https://www.mdpi.com/2227-7390/13/6/985 DOI: https://doi.org/10.3390/math13060985

[7] S. Alizadeh, M. Mahdavian, and E. Ganji, "Optimal placement and sizing of photovoltaic power plants in power grid considering multi-objective optimization using evolutionary algorithms," *Journal of Electrical Systems*

and Information Technology, vol. 10, Art. 7, 2023. [Online]. Available: https://jesit.springeropen.com/articles/10.1186/s43067-023-00073-6 DOI: https://doi.org/10.1186/s43067-023-00073-6

[8] N. Nagkoulis, E. Loukogeorgaki, and M. Ghislanzoni, "Genetic Algorithms-Based Optimum PV Site Selection Minimizing Visual Disturbance," *Sustainability*, vol. 14, no. 19, Art. 12602, Oct. 2022. [Online]. Available: https://www.mdpi.com/2071-1050/14/19/12602 DOI: https://doi.org/10.3390/su141912602

[9] K. Deb and H. Jain, "Self-Adaptive Parent to Mean-Centric Recombination for Real-Parameter Optimization," KanGAL Report No. 2011001, IIT Kanpur, 2011. [Online]. Available: https://www.egr.msu.edu/~kdeb/papers/k2011001.pdf

[10] K. Deb and R. B. Agrawal, "Simulated Binary Crossover for Continuous Search Space," *Complex Systems*, vol. 9, no. 2, pp. 115–148, 1995. [Online]. Available: https://content.wolfram.com/sites/13/2018/02/09-2-2.pdf

[11] International Energy Agency, "World Energy Outlook 2023," IEA, Paris, 2023. [Online]. Available: https://www.iea.org/reports/world-energy-outlook-2023

[12] S. Chowdhury, S. P. Chowdhury, and P. Crossley, *Microgrids and Active Distribution Networks*, 2nd ed. London, UK: Institution of Engineering and Technology, 2011. [Online]. Available: https://digital-library.theiet.org/content/books/po/pbpo052e

[13] S. Alizadeh, M. Mahdavian, and E. Ganji, "Optimal placement and sizing of photovoltaic power plants in power grid considering multi-objective optimization using evolutionary algorithms," *Journal of Electrical Systems and Information Technology*, vol. 10, Art. 7, 2023. [Online]. Available: https://jesit.springeropen.com/articles/10.1186/s43067-023-00073-6 DOI: https://doi.org/10.1186/s43067-023-00073-6

[14] D. E. Goldberg, *Genetic Algorithms in Search, Optimization, and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

[15] J. H. Holland, *Adaptation in Natural and Artificial Systems*. Ann Arbor, MI: University of Michigan Press, 1975.

[16] K. Deb and R. B. Agrawal, "Simulated Binary Crossover for Continuous Search Space," *Complex Systems*, vol. 9, no. 2, pp. 115–148, 1995. [Online]. Available: https://content.wolfram.com/sites/13/2018/02/09-2-2.pdf

[17] A. K. Jarso, G. Jin, and J. Ahn, "Hybrid Genetic Algorithm-Based Optimal Sizing of a PV–Wind–Diesel–Battery Microgrid: A Case Study for the ICT Center, Ethiopia," *Mathematics*, vol. 13, no. 6, Art. 985, 2025. [Online]. Available: https://www.mdpi.com/2227-7390/13/6/985 DOI: https://doi.org/10.3390/math13060985

Quantifying Environmental Benefits of Energy Refurbishment: From Fossil Fuel to Biomass Heating in University Buildings - A Case Study

Diana Alexandra Murcia-Velasco [1], Adriana Correa-Guimarães[1], Luis Manuel Navas-Gracia[1]

[1]TADRUS Research Group, Department of Agricultural and Forestry Engineering, Universidad de Valladolid, 34004 Palencia, Spain
dianaalexandra.murcia@uva.es

Abstract. This study presents an environmental assessment of the energy refurbishment of a university educational building using real operational data and an attributional life cycle assessment (LCA) approach. Prior to the refurbishment (2007–2014), the building relied predominantly on heating gasoil, with annual consumption averaging 428,500 kWh, while electricity use averaged 230,000 kWh/year. The integration of biomass in 2015 partially displaced fuel oil, and post-refurbishment (2016–2021), biomass became the primary heating source with an average annual consumption of 287,146 kWh, while fuel oil use was eliminated and electricity demand decreased. Energy carriers were converted to useful energy accounting for boiler efficiency, distribution losses, and heat exchanger performance. CO_2-equivalent emissions were calculated using an attributional approach, considering fuel oil emissions of 0.267 kg CO_2/kWh and biomass as carbon-neutral. The transition from fuel oil to biomass resulted in substantial reductions in fossil CO_2 emissions, approaching zero by 2017, and overall reductions in primary energy consumption. Comparative analysis with pre-refurbishment conditions and existing benchmarks indicates that the refurbishment achieved near-complete fossil fuel substitution and performance levels approaching nearly zero-energy building (NZEB) standards. These findings demonstrate that combining renewable heating sources with operational efficiency measures can significantly reduce the environmental impact of educational buildings, providing actionable insights for campus-scale decarbonization strategies.

Keywords: Life cycle assessment, Energy retrofit, Biomass heating, Educational buildings.

1 Introduction

Decarbonizing the building sector is imperative if Europe is to meet its climate neutrality goals, since buildings account for roughly 40 % of final energy consumption and about one-third of CO_2 emissions in the EU [1]. Retrofitting existing stock is more challenging but essential, because new construction is insufficient to offset historic emissions and because operational lifetimes of buildings extend multiple decades.

In this context, universities represent particularly suitable testbeds: compact campuses, stable occupancy patterns, and institutional commitment allow controlled interventions and systematic monitoring of performance [2]. Moreover, many university buildings in Spain and Europe still rely on individual fossil fuel boilers for heating and hot water, even as centralized low-carbon heating options become viable.

District Heating (DH) systems, especially those using renewable biomass, are increasingly proposed as sustainable alternatives to fossil-fueled boilers. Several studies across Europe show that DH can reduce climate change impacts compared to standalone systems, albeit with sensitivity to assumptions on losses, fuel supply chains, and emissions accounting [3]. A recent review on DH and LCA emphasizes a wide range of published emission factors for DH systems (from 11 to 470 g CO_2e/kWh_th) depending on technology, boundary choices and method [4]. This high variability underscores the need for case-specific LCA modeling rather than reliance on benchmark averages.

Despite the potential, few published studies have robustly quantified the life cycle and operational emissions savings from converting existing campus buildings from fossil boilers to biomass-based DH under actual monitored conditions. To fill that gap, this paper presents an LCA-based comparative analysis for the Aulario building at the Campus Esgueva, University of Valladolid (Spain), which in 2015 transitioned from gasoil boilers to a biomass-powered DH network. We evaluate the environmental performance per unit of useful heat delivered, explore methodological choices in emissions accounting (e.g. biogenic carbon treatment), and assess the sensitivity of results to key parameters.

The specific contributions of this study are fourfold. First, it applies life cycle assessment in a transparent manner, grounded on actual fuel consumption records and the detailed configuration of the local district heating network. Second, it contrasts different approaches to greenhouse gas accounting, including the treatment of biogenic carbon, attributional modeling, and dynamic perspectives. Third, it performs a sensitivity and uncertainty evaluation of the most influential parameters in the model. Finally, it extracts insights that can inform the design and replication of campus-scale retrofit strategies in temperate European contexts.

2 Material and methods

The methodological framework of this study follows the ISO 14040 and 14044 14044 [5], [6] standards for life cycle assessment (LCA). A comparative analysis was conducted to evaluate the environmental performance of two alternative heating configurations for the Esgueva Campus Classroom building at the University of Valladolid: (i) the pre-retrofit system based on on-site gasoil boilers, and (ii) the post-retrofit system connected to a biomass-fueled district heating (BioDH) network.

2.1 Goal and scope

The main goal is to quantify and compare the environmental impacts of the heating retrofit, with emphasis on greenhouse gas (GHG) emissions and energy efficiency improvements. The scope includes direct fuel consumption, district heating supply, and auxiliary electricity demand for pumps and control systems. The environmental assessment was conducted using an attributional LCA approach, considering biomass as carbon-neutral based on established biogenic carbon accounting principles [7], [8].

2.2 System boundaries

The system boundaries cover the heating supply chain from fuel extraction to useful heat delivery inside the building. For the pre-retrofit case, the system includes crude oil extraction, refining gasoil, transport, on-site storage, boiler operation, and auxiliary electricity. For the post-retrofit case, the system includes biomass harvesting and chipping, fuel transport, central CHP plant operation, distribution network, heat exchanger, and auxiliary electricity [9], [10]. Infrastructure (boilers, pipes, exchangers) was excluded based on expected low contribution per functional unit [11].

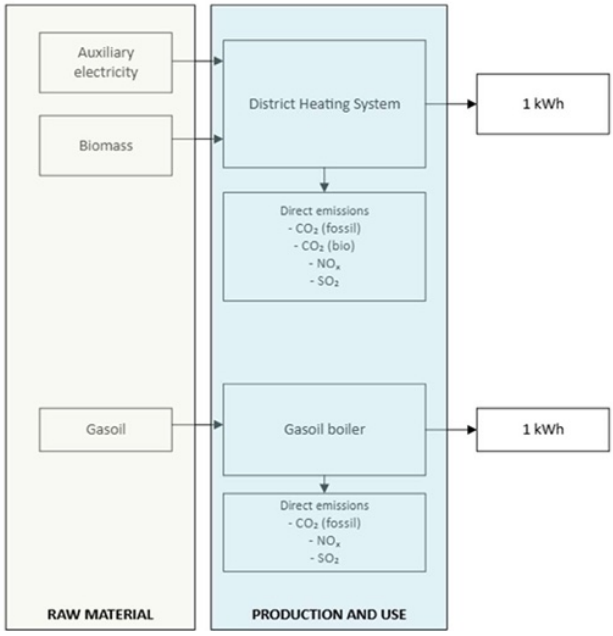


Fig. 1. System boundary diagram for heating system LCA comparison: (a) pre-retrofit gasoil boilers, (b) post-retrofit biomass district heating

2.3 Functional unit

The functional unit (FU) was defined as 1 kWh of useful heat delivered to the building over one year of operation. This unit enables direct comparison of system performance and alignment with energy service provided, avoiding bias introduced by fuel properties or boiler efficiencies [12].

2.4 Life Cycle Inventory (LCI)

The LCI was developed using measured fuel consumption data from the building's energy audits (2007 - 2015) and project documentation for the retrofit. Heating gasoil consumption was 46,451 L/year (≈ 490,336 kWh LHV) with boiler efficiency of 0.70, resulting in 343,235 kWh useful heat. For the BioDH system, the delivered heat was modeled considering a central biomass boiler efficiency of 0.85, distribution network losses of 10%, and heat exchanger efficiency of 0.99. Auxiliary electricity was modeled at 0.00714 kWh per kWh of heat delivered, with the Spanish grid mix emission factor of 0.25 kgCO_2-eq/kWh. All background processes were sourced from ecoinvent database, applying the attributional cut-off model [13].

Table 1. Life cycle inventory (LCI) for 1 kWh useful heat delivered to the Esgueva Campus Classroom building

Flow / Parameter	Unit per UF	Gasoil boilers	Biomass DH	Additional data
Fuel input	kWh (fuel)	1.43	1.18	η_boiler = 0.70 η_central = 0.85 η_exch = 0.99 loss_DH = 0.10
Heating Gasoil consumption	L	0.143	–	PCI_gasoil = 10 kWh/L
Biomass consumption	Kg (dry)	–	0.34	PCI_woodchips = 3.5 kWh/kg
Auxiliary electricity	kWh (elec)	–	0.0071	District heating pumps
Direct emissions (combustion)	kg CO_2 (fossil)	0.381	0.002	EF_gasoil = 2.67 kg CO_2/L; EF_biomass fossil share = 0.007 kg/MJ
Biogenic CO_2	kg CO_2 (bio)	–	1.10	-
NO_x emissions	g	0.35	0.28	-
SO_2 emissions	g	0.42	0.06	-
Upstream processes				
Fuel supply chain (refining, transport)	MJ	0.52	–	Database

Biomass supply chain (harvesting, chipping, transport)	MJ	–	0.27	Database
Electricity upstream emissions (E_aux)	kg CO_2eq	–	0.0018	EF_el Spain = 0.25 kg CO_2/kWh
Total, primary energy demand	kWh_p	1.57	1.23	Calculated
Total, GHG emissions	kg CO_2eq	0.42	0.004 (excl. bio CO_2)	Calculated

2.5 Life cycle impact assessment (LCIA)

GHG emissions were quantified using ReCiPe 2016 Midpoint (H) characterization factors for climate change [14].

3 Results and discussion

The building energy consumption data reveals a significant transition from fossil-based to renewable energy sources post-refurbishment. Prior to 2015, the building's energy demand was met predominantly through heating gasoil with annual consumption ranging between 303,000 and 505,000 kWh, averaging approximately 428,500 kWh/year, alongside an average electricity consumption of 230,000 kWh/year. In 2015, biomass was introduced into the energy mix, partially displacing fuel oil; consumption figures for that year were 101,000 kWh of fuel oil, 127,270 kWh of biomass, and 234,327 kWh of electricity. Post-refurbishment (2016–2021), biomass became the primary heating source, with average annual consumption of 287,146 kWh, while gasoil usage was eliminated after 2017. Electricity demand exhibited a decreasing trend, reflecting improvements in system efficiency and operational practices. To standardize energy use and facilitate life cycle assessment (LCA), all energy carriers were converted to useful energy using system-specific efficiencies: boiler efficiency (η_boiler = 0.70), distribution losses (10%), and heat exchanger efficiency (0.99). CO_2-equivalent emissions were calculated using an attributional approach, assigning 0.267 kg CO_2/kWh to fuel oil, considering biomass as biogenic carbon-neutral, and applying the Spanish electricity mix for electrical consumption. The transition from fuel oil to biomass led to a substantial reduction in fossil CO_2 emissions, effectively eliminating them post-2017, and the concurrent decrease in electricity use further reduced the building's carbon footprint. Comparative analyses with pre-refurbishment conditions indicate that the building achieved near-complete substitution of fossil fuels, significant reductions in primary energy consumption, and lower associated CO_2 emissions, aligning with findings from other university campus district heating systems employing biomass, which report 30–50% reductions in fossil energy use [3], [15]. Moreover, the post-refurbishment performance approaches nearly zero-energy building (NZEB) benchmarks for similar educational buildings in Southern Europe [16]. While these results demonstrate the effectiveness of biomass integration and energy efficiency measures, the attributional approach assumes carbon neutrality of biomass and does not account for temporal dynamics of carbon release and sequestration, which could be addressed in

future studies using dynamic carbon accounting in accordance with ISO 14067 [8]. The refurbishment illustrates that targeted energy interventions can substantially reduce the environmental impact of existing educational buildings and provide insights for the design of low-carbon campus energy systems.

4 Conclusions

The refurbishment of the building demonstrates the potential of targeted energy interventions to reduce the environmental impact of existing educational buildings. The integration of biomass as the primary heating source, combined with improvements in electricity efficiency, effectively eliminated fossil fuel consumption post-2017 and contributed to a substantial reduction in the building's carbon footprint under the attributional LCA framework. Standardization of energy carriers to useful energy enables consistent comparison and highlights the benefits of system efficiency improvements, including boiler and heat distribution performance. Comparative assessment with pre-refurbishment conditions and existing benchmarks of university campus district heating systems indicates that the building achieved near-complete substitution of fossil fuels and reductions in primary energy use, approaching performance levels of nearly zero-energy buildings (NZEB) in Southern Europe. The attributional approach applied in this study treats biomass as carbon neutral. Future research could explore dynamic carbon accounting approaches to capture temporal dynamics of biogenic carbon release and sequestration, providing additional insights into the climate impact timing of biomass-based heating systems. Overall, the study illustrates that combining renewable heating sources with operational efficiency measures provides a robust pathway for decarbonizing educational buildings, offering insights applicable to broader campus-scale energy planning and NZEB retrofitting strategies.

Acknowledgments. Diana Alexandra Murcia Velasco was financed on the grounds of University of Valladolid predoctoral contracts and co-financed by Banco Santander.

Disclosure of Interests. The authors declare no conflict of interest.

References

[1] A. Fnais *et al.*, "The application of life cycle assessment in buildings: challenges, and directions for future research," *Int J Life Cycle Assess*, vol. 27, no. 5, pp. 627–654, May 2022, doi: 10.1007/s11367-022-02058-5.

[2] A. Arias, I. Leon, X. Oregi, and C. Marieta, "Environmental Assessment of University Campuses: The Case of the University of Navarra in Pamplona (Spain)," *Sustainability*, vol. 13, no. 15, p. 8588, Aug. 2021, doi: 10.3390/su13158588.

[3] J. M. Rey-Hernández, F. J. Rey-Martínez, C. Yousif, and D. Krawczyk, "Assessing the performance of a renewable District Heating System to achieve nearly zero-energy status in renovated university campuses: A case study for

Spain," *Energy Convers Manag*, vol. 292, p. 117439, Sep. 2023, doi: 10.1016/j.enconman.2023.117439.

[4] K. Autelitano, J. Famiglietti, M. Aprile, and M. Motta, "Towards Life Cycle Assessment for the Environmental Evaluation of District Heating and Cooling: A Critical Review," *Standards*, vol. 4, no. 3, pp. 102–132, Jul. 2024, doi: 10.3390/standards4030007.

[5] International Organization for Standardization (ISO), "Environmental management—Life cycle assessment—Principles and framework (ISO 14040:2006)," Geneva, Switzerland, 2006.

[6] International Organization for Standardization (ISO), "Environmental management—Life cycle assessment—Requirements and guidelines (ISO 14044:2006)," Geneva, Switzerland, 2006.

[7] E. Hoxha *et al.*, "Biogenic carbon in buildings: a critical overview of LCA methods," *Buildings and Cities*, vol. 1, no. 1, pp. 504–524, Aug. 2020, doi: 10.5334/bc.46.

[8] F. Cherubini, N. D. Bird, A. Cowie, G. Jungmeier, B. Schlamadinger, and S. Woess-Gallasch, "Energy- and greenhouse gas-based LCA of biofuel and bioenergy systems: Key issues, ranges and recommendations," *Resour Conserv Recycl*, vol. 53, no. 8, pp. 434–447, Jun. 2009, doi: 10.1016/j.resconrec.2009.03.013.

[9] G. Guest, R. M. Bright, F. Cherubini, O. Michelsen, and A. H. Strømman, "Life Cycle Assessment of Biomass-based Combined Heat and Power Plants," *J Ind Ecol*, vol. 15, no. 6, pp. 908–921, Dec. 2011, doi: 10.1111/j.1530-9290.2011.00375.x.

[10] C. Jeandaux, J.-B. Videau, and A. Prieur-Vernat, "Life Cycle Assessment of District Heating Systems in Europe: Case Study and Recommendations," *Sustainability*, vol. 13, no. 20, p. 11256, Oct. 2021, doi: 10.3390/su132011256.

[11] K. Autelitano, J. Famiglietti, M. Aprile, and M. Motta, "Towards Life Cycle Assessment for the Environmental Evaluation of District Heating and Cooling: A Critical Review," *Standards*, vol. 4, no. 3, pp. 102–132, Jul. 2024, doi: 10.3390/standards4030007.

[12] S. Lo Piano and L. Benini, "A critical perspective on uncertainty appraisal and sensitivity analysis in life cycle assessment," *J Ind Ecol*, vol. 26, no. 3, pp. 763–781, Jun. 2022, doi: 10.1111/jiec.13237.

[13] G. Wernet, C. Bauer, B. Steubing, J. Reinhard, E. Moreno-Ruiz, and B. Weidema, "The ecoinvent database version 3 (part I): overview and methodology," *Int J Life Cycle Assess*, vol. 21, no. 9, pp. 1218–1230, Sep. 2016, doi: 10.1007/s11367-016-1087-8.

[14] M. A. J. Huijbregts *et al.*, "ReCiPe2016: a harmonised life cycle impact assessment method at midpoint and endpoint level," *Int J Life Cycle Assess*, vol. 22, no. 2, pp. 138–147, Feb. 2017, doi: 10.1007/s11367-016-1246-y.

[15] A. Zajacs, A. Borodinecs, and N. Vatin, "Environmental Impact of District Heating System Retrofitting," *Atmosphere (Basel)*, vol. 12, no. 9, p. 1110, Aug. 2021, doi: 10.3390/atmos12091110.

[16] “Nearly-zero energy and zero-emission buildings - European Commission.” Accessed: Aug. 27, 2025. [Online]. Available: https://energy.ec.europa.eu/topics/energy-efficiency/energy-performance-buildings/nearly-zero-energy-and-zero-emission-buildings_en?prefLang=es

Operational Lessons from a Commercial Solar Photovoltaic and Battery System in Jamaica: Curtailment, Frequency, and Demand Charge Management

Toni Ann Hendricks[1], Jayaka D. Campbell[1], Dudley Williams[1], A. Anthony Chen[1], Kevin Mills[2], Randy Koon Koon[3], Aston Stephens[4]

[1] Department of Physics, University of the West Indies, Mona Campus, Jamaica
[2] Managing Director of Renewable Energy Design & Installation Services LTD, Jamaica
[3] School of Mechanical Engineering, Coventry University, U.K
[4] Managing Director of Empowered Caribbean Communities, Jamaica

Abstract. This study evaluates the performance of a 500 kWp photovoltaic array and 1 MWh battery energy storage system (BESS) installed under a commercial tariff structure in Jamaica, highlighting technical, economic, and policy challenges of renewable energy integration in small island contexts. In the absence of net billing or net metering, excess PV generation produced unrealized revenue, with the BESS mitigating but not fully eliminating curtailment. Curtailment remained a major operational constraint, as the battery could not absorb all surplus generation under the zero-export restriction; modelling indicated that site demand could have been reduced by an additional 13% with larger storage capacity. Operational data further revealed that the BESS helped mitigate frequency violations during morning ramp-up and midday peaks, though advanced control strategies are still required for complete stability. A key benefit was observed in demand shaving: under PV-only operation, sudden afternoon drops in solar output caused demand spikes up to 137 kVA, incurring charges of ~US$ 1,900, whereas PV-BESS operation limited imports to <65 kVA, reducing charges to ~US $900. On-site weather data also showed lower peak sun hours than solar databases, with implications for system design, financial projections, and payback periods. These findings emphasize the importance of supportive policy, adequate storage sizing, and site-specific resource assessment for advancing PV-BESS deployment in the Caribbean.
Keywords: Solar photovoltaic (PV) systems, Battery Energy Storge Systems (BESS), Renewable energy adoption

1 Introduction

1.1 The Caribbean's energy transition efforts

The Caribbean, like many other regions, has pledged to decarbonize its energy systems through the adoption of renewable energy (RE) and low carbon technologies. The region has set an ambitious RE target of 47% of electricity generation from renewable

sources by 2027, with some countries aiming for 100% by 2030 [1]. For small island developing states (SIDS), this transition is particularly urgent given their disproportionately high energy costs, driven by heavy reliance on imported fossil fuels [2]. The challenge is compounded by their acute climate sensitivity, as aging grid infrastructure is frequently disrupted by tropical cyclones and other extreme weather events [3].

RE technologies such as wind and solar are often considered more climate-resilient than fossil fuel plants due to their adaptability, low dependence on fuel supply chains, and decentralized nature [4], [5]. While not immune to storm damage, their modularity and distributed deployment can enable faster recovery compared to centralized fossil-based systems. Coupled with their increasing cost competitiveness [6], RE technologies offer a strategic pathway for Caribbean SIDS to bolster energy independence, energy security, and improve resilience to environmental shocks.

Despite these advantages, the region's transition has lagged, constrained by structural and technical barriers. Key obstacles include high upfront investment costs, fragmented policy frameworks, and small-scale centralized grids that were not originally designed to accommodate bidirectional energy flows [7]. Moreover, the islands differ significantly in population, terrain, and economic conditions, including electricity pricing structures, necessitating tailored, country-specific approaches to the design and deployment of these RE technologies.

1.2 Renewable energy integration challenges

RE capacity in the Caribbean has expanded markedly in recent years, driven largely by solar photovoltaic (PV) deployment, which grew from under 1 GW in 2015 to approximately 4 GW by 2023 [8]. Several utility-scale PV projects have been commissioned or are underway in countries such as the Dominican Republic, Jamaica, St. Kitts, and Trinidad. While PV systems serve as an effective platform for distributed energy generation by enabling localized and clean electricity production, they also present a range of technical and operational challenges. PV systems are often prone to cloud induced transients that can have negative implications for a power grid, causing rapid and unpredictable fluctuations in voltage and frequency levels [2], [9], [10]. Such variability poses a serious challenge to existing grids, where reliance on slow-response spinning reserves or generators often leads to voltage fluctuations and frequency violations, highlighting their inadequacy in addressing solar intermittency [2].

Tariff structure considerations. PV systems offer attractive benefits, including reduced electricity costs and greater energy independence for both residential and commercial consumers. Although high upfront capital requirements can deter investment, their adoption becomes particularly compelling in countries with high electricity tariffs. For residential customers, they are billed solely on cumulative energy consumed over time, measured in kilowatt-hours (kWh). In contrast, commercial and industrial customers face additional demand charges based on the maximum apparent power drawn measured in kilovolt-amperes (kVA). This demand charge reflects the

impact of inductive equipment on grid capacity and frequently represents a substantial portion of the monthly bill.

Tariff structures vary significantly across the region which can have direct implications on the economic viability of RE systems. For example, in Jamaica, commercial customers with a minimum demand of 25 kVA up to 2 MVA are billed a demand charge of US$ 13.88 per kVA and an energy charge of US$ 0.032 per kWh as specified in the Office of Utilities Regulation tariff schedule[11]. A lower rate is observed in Trinidad where commercial customers with demand between 50 kVA and 350 kVA are charged US$7.50 per kVA and US$ 0.029 per kWh [12]. Additionally, in some islands like Grenada, commercial customers are not differentiated by minimum kVA usage, instead a flat commercial tariff of US$ 0.162 per kWh is applied [13]. These differences highlight that the economic performance of RE systems depends not only on energy production but also on their ability to manage demand charges.

Curtailment. A further challenge with PV systems arises from seasonal and diurnal energy production, which can result in surplus generation during periods of high irradiance. Solar resource availability fluctuates throughout the year, directly influencing system sizing to meet dynamic building loads. PV systems are often sized according to user preferences; for example, in off-grid applications, systems are frequently oversized to ensure that reduced irradiance in winter still provides sufficient power to meet demand. This approach, however, leads to overproduction in summer. If load does not increase during these high-production periods, or if batteries are not available to store excess energy, the surplus generation is curtailed by the system.

The U.S. National Renewable Energy Laboratory defines curtailment as the reduction in power output from a source that could otherwise produce given the available resources [14]. For RE systems, curtailment reduces revenue, especially in markets without robust feed-in tariffs or net-metering. Regions with high renewable penetration have adopted flexible mechanisms to address this challenge. In California, excess solar generation is managed through dynamic pricing, demand response, and large-scale storage, while Germany relies on guaranteed feed-in tariffs and grid compensation schemes to ensure producers are not penalized [14],[15]. These approaches highlight the importance of supportive regulatory frameworks in reducing the economic burden of curtailment. By contrast, Jamaica's net billing arrangement with the island's sole utility, the Jamaica Public Service (JPS), applies only to systems under 100 kW[17]. This restriction prevents larger or community-scale projects from exporting surplus energy, requiring careful planning to minimize curtailment. These dynamics underscore the need for both adequate storage and enabling policy frameworks in the Caribbean to maximize the economic and operational value of PV systems.

1.3 Energy storage solution

For small island grids with growing PV penetration, Battery Energy Storage Systems (BESS) is presented as a promising solution to some of the challenges associated with

intermittent PV generation. BESS offers services such as fast frequency response, voltage and reactive power support, power reserves, peak shaving, grid forming, and synthetic inertia [10], [18]. BESS not only provides these technical services but also reduces reliance on costly diesel reserves, thereby supporting both grid stability and economic resilience. Globally, BESS has become integral in realizing the full potential of PV systems, with costs declining by 93% since its average of US$ 2,571 in 2010 to US$ 192 per kWh in 2024 [19]. Yet in the Caribbean, limited economies of scale, high implementation costs, and logistical barriers keep prices well above global benchmarks[20].

Beyond cost, there is also a lack of detailed case studies on PV-BESS performance at commercial and utility scale in the region [21]. Similar challenges in RE integration have been documented across other tropical and islanded grids. Prakash et al.[9] review the role of BESS in distribution networks, emphasizing their importance for ancillary services such as frequency response and voltage regulation. Their findings highlight that reliable deployment depends on accurate characterization of local operating conditions, a theme relevant to islanded networks. Ochoa-Correa et al.[21] examine pathways to 100% renewable energy in island systems, identifying irradiance variability, curtailment, and storage costs as key challenges.

Building on these insights, PV-BESS systems in the Caribbean cannot be deployed as one-size-fits-all solutions derived from global standards. Their assessment must reflect the region's diverse microclimates, country specific load profiles, tariff structures, and grid conditions, and individual use cases.

1.4 Impact of localized meteorological data on PV-BESS system sizing

Accurate system sizing for PV-BESS systems depends fundamentally on the quality of the meteorological data used, primarily solar irradiance, temperature, and wind speed. The reliability of this data is shaped by the source, duration, and resolution of the measurements [22]. Standardized global datasets provide valuable long-term meteorological averages, their coarse spatial resolution often masks the fine scale variability that shapes energy yields SIDS. In such contexts, localized weather data can provide a more accurate basis for PV-BESS system sizing, supporting robust performance in topographically diverse islands where microclimates strongly affect irradiance and efficiency [23].

Studies have shown that reliance on non-site specific or low-resolution data can misinform system design choices, leading to inappropriate tilt angles or inverter sizing [22], [24], [25]. This further emphasizes the need for context specific studies that incorporate high resolution meteorological data, which are critical for ensuring optimal sizing and cost-effectiveness, as both directly shape performance and investment outcomes [20], [22], [24], [26]. Regional research has further demonstrated the value of incorporating localized meteorological measurements into PV system sizing, underscoring their role in improving design accuracy and performance predictions [2], [27], [28]. In line with these findings, the present study integrates on-site weather data with opensource datasets to evaluate their influence on PV-BESS sizing for a Jamaican

small and medium-sized enterprises (SME), thereby extending regional evidence to this context.

Building on these challenges, this paper aim to contribute to the literature on PV-BESS deployment in Caribbean SIDS by examining the combined effects of tariff structures, localized meteorological conditions, and commercial applications. Through assessment of a Jamaican SME installation, it assesses operational performance, curtailed energy, and economic implications. The findings provide context-specific insights to support more resilient, cost-effective, and reliable RE deployment in the region.

2 Methodology

This study analysed the performance of a roof-mounted, grid-connected self-consumption system consisting of a 500 kWp PV array and a 1 MWh lithium battery storage unit, located in St. Catherine, Jamaica. Operational and meteorological data were collected over a 16-month period (April 2024 - July 2025) using two primary monitoring platforms: the AGETO Renewable Controller (ARC) and a Davis Vantage Pro2 weather station (WS).

2.1 Energy Management system

The ARC functions as the site's energy management system, recording electrical variables at 30-minute intervals, including facility load (kW), PV generation (kW), inverter charge and discharge power (kW), battery state of charge SOC (%), voltage (V), and frequency (Hz). The ARC is connected to the other system components as seen in Fig.1, and data retrieved via an online portal. These variables capture the complete energy balance of the PV-BESS system, supporting analysis of generation, storage behaviour, and demand interactions.

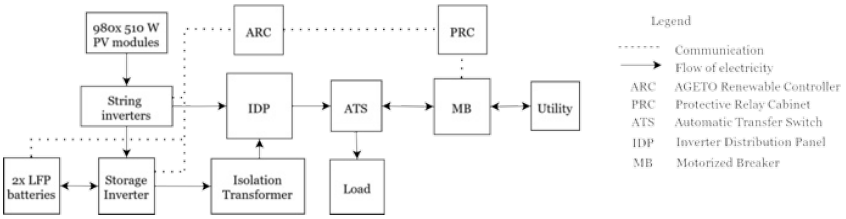


Fig. 1. Diagram of the PV+ System Configuration

In operation, the ARC follows a simple control logic: once the PV is generating, it simultaneously supplies the facility load and charges the battery. During this period, the facility still imports a baseload of roughly 30 - 50 kW from the utility. Once the battery is fully charged, utility import is reduced to 5 kW, and the PV output is used to meet the remaining load. An on-site WS was installed approximately 61 m from the first panel and 100 m from the end of the last row of panels. Data was recorded at 15-

minute intervals. For this study, the primary variable of interest was global horizontal irradiance (GHI), measured with an accuracy of ±5%. GHI was used to estimate expected PV system output, curtailment, and peak sun hours (PSH). To ensure data quality and consistency, days with fewer than 90 data points were excluded from the analysis (Fig. 2.).

For the monthly curtailment and PSH calculations, April 2024 and July 2025 were omitted due to substantial gaps in the WS record. After data screening, the WS was resampled to 30 minutes to match the ARC time interval. Subsequent analyses were performed only in periods where both datasets provided complete records. Capturing short-interval irradiance variability was essential for identifying curtailment events and assessing battery utilization under zero-export restrictions, ensuring that system performance was evaluated against actual site conditions rather than generalized resource profiles.

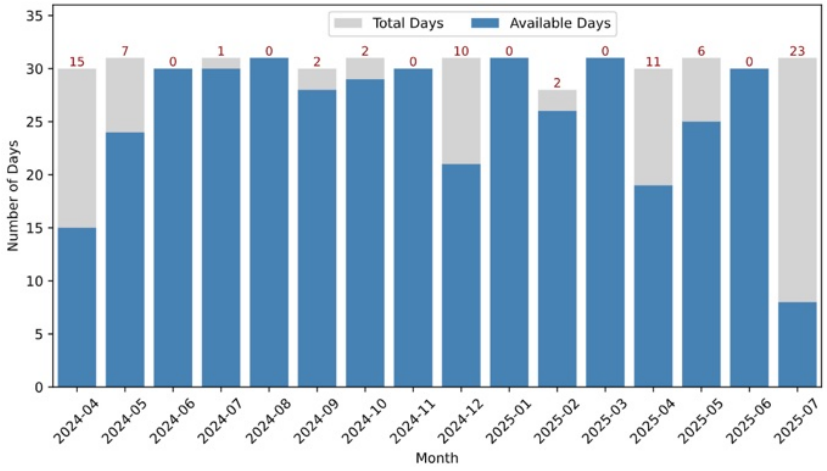


Fig. 2. Total number of available days in the dataset, based on complete 24-hour records (96 points per day at 15-minute intervals). The number of missing days (red) for each month based on the criteria given.

2.2 Curtailment Model

Daily PV curtailment was estimated by first modelling the irradiance incident on the panels using the opensource pvlib Python library [29]. Plane-of-array irradiance was derived from measured global horizontal irradiance through the Erbs decomposition and transposition models, accounting for the installed tilt (10°) and azimuth (180°). Cell temperature was then estimated with the Sandia Array Performance Model (SAPM) using measured ambient temperature and wind speed. Finally, the expected AC power, $EPV(t)$, was calculated as:

$$EPV(t) = \frac{(G_{POA}(t) \cdot A \cdot DF \cdot \eta_{PV} \cdot \eta_{inv})}{1000} \quad (1)$$

Where: $G_{POA}(t)$ is the plane-of-array irradiance in Wm^{-2}, A is the total panel area in m^2. Derating factor, DF was set at 0.85 in this study, consistent with values commonly applied in similar analyses [30], [31], while η_{PV} is the module efficiency, and η_{inv} is the inverter efficiency.

Modelled PV generation estimates were compared with measured data from the ARC monitoring platform, which recorded actual solar output (ASO), facility load, and battery state of charge. A temporal offset was identified after assessing agreement between ASO and EPV, attributed to the 60 - 100 m separation between the array and the WS, where cloud-cover effects were detected at different moments. Shifting WS data by 30 minutes relative to the EPV series improved the correspondence between datasets. Curtailment was defined as periods when EPV exceeded the ASO produced by the system which occurred when the battery was fully charged or offline. Under these conditions, the inverter reduced the PV output to match the facility load, in accordance with the site's zero-export agreement with the utility. Following the temporal adjustment, an uncertainty analysis was conducted for non-curtailed periods. The results showed a mean bias of -1.055 kW, a root mean square error (RMSE) of 38.728 kW, a standard deviation of 38.798 kW, and a coefficient of determination (R^2) of 0.874.

Two additional constraints were applied to estimate the daily potential curtailment, C_{pot}. The first constraint was setting an allowable margin between the load, L(t) and ASO(t) represented by ϵ. This accounted for the conversion and system losses between the parameters. This margin was set at 10kW based on a parametric study conducted. The second constrain was imposed to avoid misclassification during morning ramp-up and evening ramp-down, thus curtailment was only recorded when the threshold condition was satisfied for at least two consecutive timesteps.

$$C_{\text{pot}}(t) = \begin{cases} 1, & if\ \text{EPV}(t) > \text{ASO}(t)\ \Lambda\ |\text{ASO(t)} - L(t)| \leq\ \epsilon \\ 0, & otherwise \end{cases} \quad (2)$$

The magnitude of the curtailment was then quantified as the difference between EPV and ASO during identified curtailment periods. For each timestep, $EPV(t) - ASO(t)$ was recorded whenever $C_{pot}(t) += 1$. These values were summed across the analysis period and multiplied by the timestep duration (30 minutes) to convert from instantaneous kW to cumulative curtailed energy in kWh as demonstrated in Fig. 3. The cost of the estimated curtailed energy was then calculated using the average commercial rate.

2.3 Peak sun hours

Accurate estimation of PSH is critical for PV system design and performance assessment, as it provides a normalized measure of solar resource availability. To examine the importance of localized PSH values, three data sources were compared: (i) on-site weather station measurements, (ii) a multiyear TMY dataset from PVWatts, and (iii) a single year (2021) timeseries dataset from Solargis. A paired sample *t*-test and the Wilcoxon signed rank test were applied to the results to test for agreement. This

comparison was undertaken to evaluate the extent to which localized climate conditions align with or diverge from widely used global datasets. By integrating site-specific irradiance into the analysis, the study ensured that subsequent assessments of PV output, curtailment, and battery utilization were based on actual operating conditions rather than generalized resource assumptions. The statistical tests therefore served not only to quantify agreement among datasets but also to establish the reliability of localized measurements as the basis for performance evaluation.

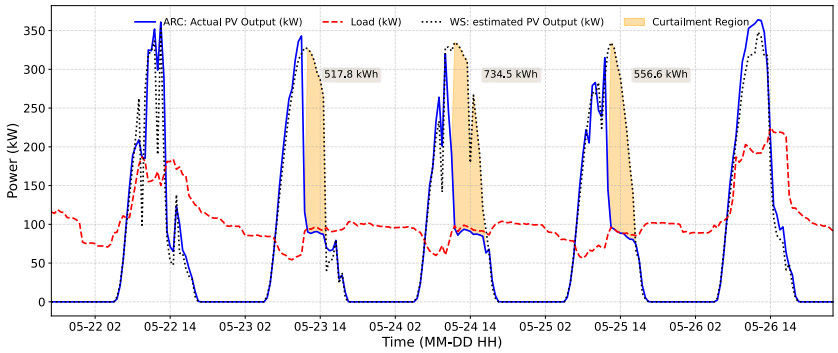


Fig. 3. A time-series comparison of ASO recorded from the ARC, estimated PV output from the on-site WS data, and load over a five-day period (May 22 - 26, 2025). The shaded region represents total curtailed energy in kWh calculated as the difference between ASO and estimated PV output when load is met almost entirely by the PV modules and battery is full.

3 Results and discussion

3.1 Demand Charges

In Jamaica, demand charges based on maximum kVA account for a large share of commercial electricity costs. Unlike energy charges calculated from monthly kWh use, demand charges reflect the highest 15-minute average kVA, which then sets the billing rate for the next five months, or 80% of that value if not exceeded.

To demonstrate the role of BESS in demand shaving, two different system operations were compared: PV only and PV-BESS. Fig. 4 presents data from May 13, 2025 (PV only), when the BESS was unavailable to provide load support. On this day, the utility recorded a maximum demand of 137 kVA sustained for approximately eight minutes, coinciding with a significant drop in PV output between 15:00 and 16:00. In contrast, on May 21, 2025 (PV-BESS), when the BESS was available, it supplied over 140 kVA during a similar PV fluctuation, thereby limiting utility demand to less than 65 kVA. The financial implications of these scenarios are substantial: the peak on May 13 would have resulted in a kVA demand charge of approximately US$ 1,900, compared to US$ 900 on May 21 with BESS support, a 52% reduction if these values had represented the monthly maximum.

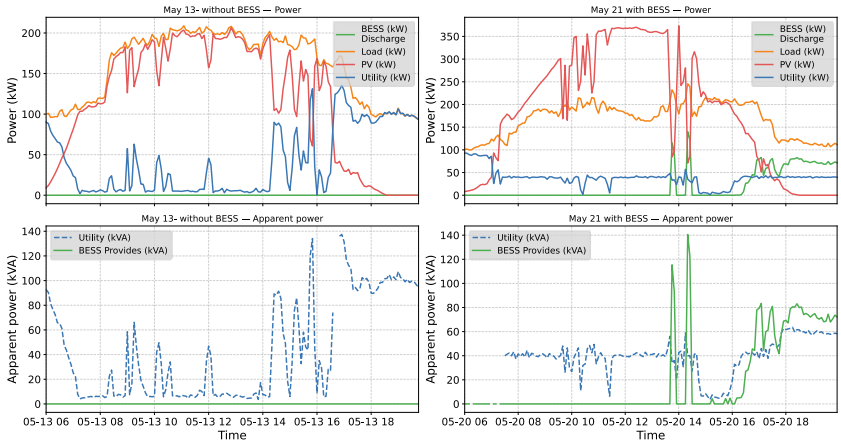


Fig. 4. Comparison of power and apparent power flows in a 500 kWp PV system at a Jamaican SME. Data on May 13 (without BESS) and May 21 (with BESS) illustrate how storage integration stabilizes utility supply.

3.2 Frequency support

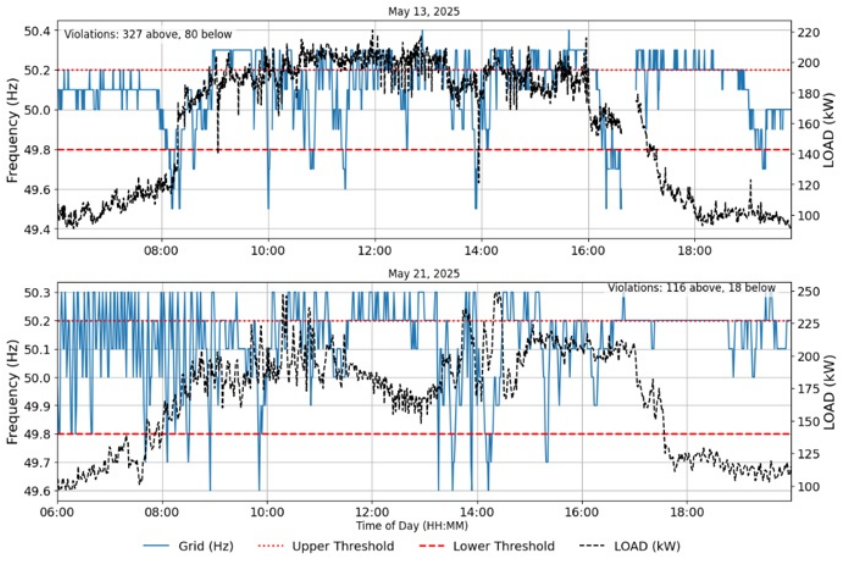


Fig. 5. Grid frequency and load profiles for May 13 and May 21, 2025, with operational frequency thresholds (49.5 - 50.5 Hz) and number of violations noted for each day.

Fig. 5 illustrates the system's frequency response over a single day under two configurations: PV only (May 13) and PV-BESS (May 21). On May 13, the upper frequency threshold was breached 327 times, primarily between 10:00 and 18:00 when system load exceeded 180 kW. In contrast, the PV-BESS system on May 21 experienced 116 breaches, a 60% reduction, despite peak loads reaching 250 kW. Notably, most violations occurred between 06:00 and 10:00, coinciding with business opening hours, solar ramp-up, and battery charging. These overlapping dynamics may have challenged the system's frequency regulation as generation was greater than load demand, highlighting the need for refined control strategies within the ARC framework

3.3 Curtailment

The commercial facility deployed a 500 kWp PV system sized to meet its 250 kW peak load while charging a 1 MWh battery. Under typical operating conditions, the battery reached full SOC by early afternoon, around 13:00. on weekdays and as early as 11 a.m. on weekends. Once the battery reached full capacity, the inverter curtailed PV output to match instantaneous load, thereby ensuring compliance with zero export requirements. Battery discharge was primarily utilized to shave short term demand spikes and to sustain nighttime baseload. This operating strategy, however, also led to energy curtailment, with monthly curtailed energy peaking at 36 MWh (equivalent to an estimated revenue loss of US$ 4,752) and averaging 11.4 MWh (US$ 1,505) over the study period, Fig. 6.

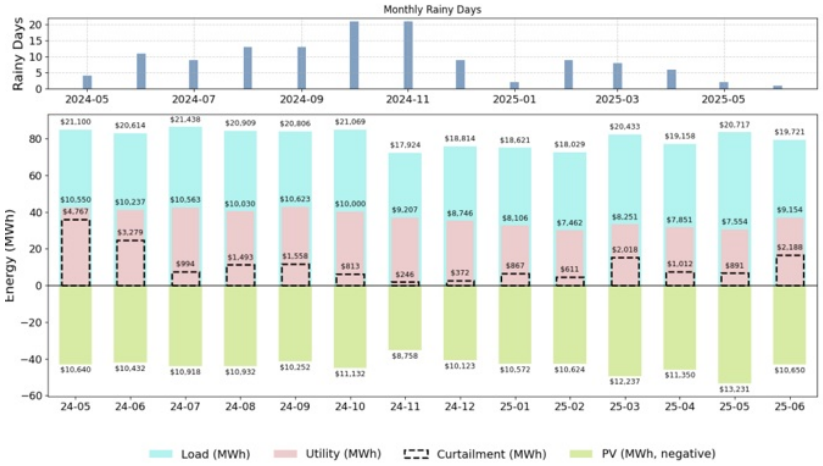


Fig. 6. Monthly rainy days and PV system energy statistics from May 2024 to May 2025. The upper panel shows rainfall frequency, while the lower panel presents monthly energy use (MWh) from the utility, PV generation, and calculated curtailed energy, with associated cost estimates highlighted.

Over the 14-month analysis period, cumulative curtailed energy totalled 160 MWh, corresponding to a foregone revenue of approximately US$ 21,120 at the average net billing rate of US$ 0.132/kWh over the period August 2024 to July 2025. During this period, the facility consumed 1,160 MWh, of which 612 MWh (52%) was supplied by the PV system. If curtailed energy had been utilized, PV could have supplied an additional 13% of total demand. These findings underscore the trade-off between oversizing PV capacity for battery charging and the resulting curtailment once storage is saturated, consistent with observations in similar studies [32]. It should be noted that the BESS was not commissioned until July 2024, which explains the elevated curtailment observed during the first two months of operation. The subsequent reduction in curtailment between August and November 2024 coincides with periods of reduced solar irradiance, as each of these months experienced more than 10 days of rainfall.

3.4 PSH calculation

Open-source tools such as PVWatts and Solargis are widely used in feasibility studies; however, their reliance on generalized datasets may overlook site-specific atmospheric conditions. To evaluate the potential benefits of localized weather data, on-site irradiance measurements were compared with these two commonly applied sources.

Seasonal variation in PSH is illustrated in Fig. 7. PVWatts consistently overestimates PSH relative to on-site measurements, with average percentage differences of approximately 20% across most months. Exceptions occur in May and March, when the difference narrows to below 10%. In contrast, Solargis demonstrates stronger alignment with observed data, particularly between March and August, where deviations remain consistently under 10%. Nonetheless, both datasets fail to capture winter variability, with December discrepancies reaching as high as 40%. Such discrepancies can lead to inaccurate projections of energy yield and suboptimal system sizing, a pattern that has also been noted in previous studies. [22], [24].

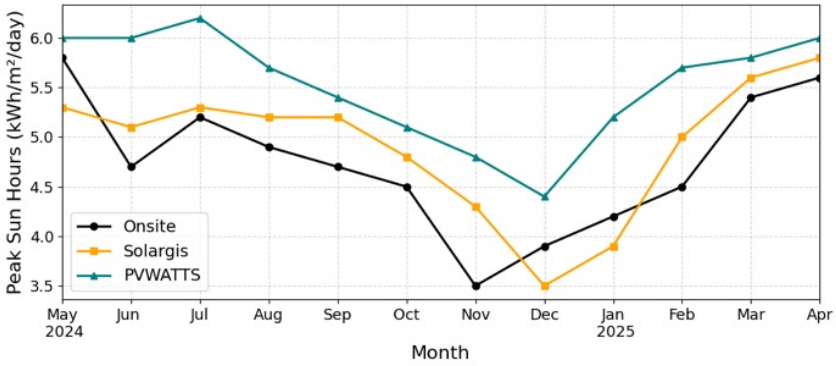


Fig. 7. Seasonal variation in monthly PSH in Jamaica from May 2024 to April 2025, comparing onsite measurements with Solargis and PVWatts estimates.

To quantify these discrepancies, a paired t-test and Wilcoxon signed-rank test were conducted, comparing on-site weather station data with PVWatts and Solargis outputs (Table 1). Both tests returned p-values < 0.05, confirming that the observed differences are statistically significant. The magnitude of the t-statistics further illustrates the divergence: PVWatts PSH values (t = 7.02) deviate more strongly from on-site measurements than Solargis (t = 3.32). The closer alignment of Solargis is likely attributable to its use of sub-hourly irradiance time series, whereas PVWatts relies on typical meteorological year (TMY) datasets, which may not adequately capture site-specific microclimatic effects. These results highlight the risks of relying solely on generalized datasets for PV system design in Caribbean contexts. Incorporating onsite or high-resolution meteorological data is essential to avoid mis-sizing, reduce curtailment risk, and ensure cost effective deployment of PV-BESS systems in small, resource constrained grids.

Table 1. Statistical comparison of monthly peak sun hours from Solargis and PVWatts against onsite measurements in Jamaica, using paired t-tests and Wilcoxon signed-rank tests.

Configuration	**Paired T-test**		**Wilcoxon**
	P-value	T	P-value
Solargis vs on-site	0.0068	3.32	0.0103
PVWatts vs onsite	0.000	7.02	0.0005

4 Conclusion

This study has presented key lessons from the operation of a commercial scale PV-BESS installation in Jamaica, offering insights into both the technical and economic considerations of renewable energy integration in SIDS. The absence of net billing or net metering for commercial clients emerged as a critical barrier, limiting the economic value of PV generation and resulting in unrealized revenue when excess electricity could not be exported. Although the BESS absorbed a substantial share of surplus production, its limited capacity was insufficient to prevent curtailment, particularly on nonwork days. Modelling indicated that electricity purchased from the utility could have been reduced by ~13% if sufficient storage had been available to capture curtailed energy, underscoring the importance of appropriately sizing storage systems.

Operational data further revealed that frequency thresholds were frequently violated during morning ramp up periods, when the energy management system attempted to balance rising building loads with simultaneous battery charging. In the absence of storage, additional breaches occurred during midday peaks, highlighting the role of BESS in mitigating but not fully eliminating grid stability challenges. Importantly, the study also demonstrated the value of BESS in demand shaving within a tariff environment where demand charges outweigh energy charges. Under PV-only operation, sudden afternoon drops in solar output caused demand spikes up to 137 kVA, resulting in a demand charge of approximately US$ 1,900. With the PV-BESS system,

the battery reduced the spike in kVA during a similar event, limiting utility imports to below 65 kVA and reducing the demand charge to about US$ 900. This illustrated the significant economic advantage of storage in managing demand related costs.

The study also highlighted the importance of site-specific solar resource assessment. Onsite WS data showed statistically significant differences from widely used databases, with lower PSH values recorded during winter months and extended rainfall periods between August and November. If system design had relied solely on TMY estimates, actual performance would have been up to 40% lower than projected, which can have serious implications for financial viability and payback periods. This underscores the risks of overreliance on generalized datasets and the necessity of localized measurements for accurate system design and investment planning.

Together, these findings emphasize four overarching lessons. First, supportive policy frameworks are essential to unlock the full economic potential of commercial PV-BESS systems. Second, storage must be adequately sized and paired with robust control strategies to minimize curtailment and frequency instability. Third, site-specific resource data are indispensable for realistic performance and financial projections. Fourth, BESS provides a critical economic function in demand charge reduction, directly improving project viability in tariff structures dominated by demand costs. Addressing these dimensions will be critical for scaling PV-BESS adoption in Jamaica and similar tropical contexts. Ultimately, the integration of PV and storage represents not only a technical solution to curtailment and reliability challenges but also a strategic pathway toward more resilient, sustainable, and economically viable energy futures in the Caribbean.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. G. Leon, “Keynote- Energy Transition Imperatives for the Caribbean: Post COP26 Reflections | Caribbean Development Bank,” Caribbean Development Bank. Accessed: July 02, 2025. [Online]. Available: https://www.caribank.org/newsroom/news-and-events/speeches/keynote-energy-transition-imperatives-caribbean-post-cop26-reflections
2. A. A. Chen, A. J. Stephens, R. Koon Koon, M. Ashtine, and K. Mohammed-Koon Koon, “Pathways to climate change mitigation and stable energy by 100% renewable for a small island: Jamaica as an example,” *Renew. Sustain. Energy Rev.*, vol. 121, pp. 2–7, Apr. 2020, doi: 10.1016/j.rser.2019.109671.
3. C. T. Lewis, “Climate Change and the Caribbean: Challenges and Vulnerabilities in Building Resilience to Tropical Cyclones,” *Climate*, vol. 10, no. 11, p. 178, Nov. 2022, doi: 10.3390/cli10110178.
4. Y. Shang, S. Sang, A. K. Tiwari, S. Khan, and X. Zhao, “Impacts of renewable energy on climate risk: A global perspective for energy transition in a climate adaptation framework,” *Appl. Energy*, vol. 362, p. 122994, May 2024, doi: 10.1016/j.apenergy.2024.122994.
5. S. Han, B. H. Bin Kamaruddin, X. Shi, and J. Zhu, “Achieving energy resilience: Studying renewable and fossil fuel energy generation drivers and COPE-28 pathways of China,” *Energy Strategy Rev.*, vol. 58, pp. 4–7, Mar. 2025, doi: 10.1016/j.esr.2025.101669.

6. IPCC *et al.*, "Climate Change 2023: Synthesis Report. Contribution of Working Groups I, II and III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Team, H. Lee and J. Romero (eds.)]. IPCC, Geneva, Switzerland.," Intergovernmental Panel on Climate Change (IPCC), July 2023. doi: 10.59327/IPCC/AR6-9789291691647.
7. P. Blechinger, K. Richter, and O. Renn, "Barriers and Solutions to the Development of Renewable Energy Technologies in the Caribbean," in *Decentralized Solutions for Developing Economies*, S. Groh, J. Van Der Straeten, B. Edlefsen Lasch, D. Gershenson, W. Leal Filho, and D. M. Kammen, Eds., in Springer Proceedings in Energy. , Cham: Springer International Publishing, 2015, pp. 267–284. doi: 10.1007/978-3-319-15964-5_24.
8. IRENA, "Renewable capacity statistics 2024." International Renewable Energy Agency, 2024.
9. K. Prakash *et al.*, "A review of battery energy storage systems for ancillary services in distribution grids: Current status, challenges and future directions," *Front. Energy Res.*, vol. 10, Sept. 2022, doi: 10.3389/fenrg.2022.971704.
10. E. Nyemah, "Use of Battery Systems for VAR Support in Con Edison's Distribution Network/Substation," 2022.
11. OUR, "JPS Rate Schedule 2023." Jamaica Public Service Company, 2023. [Online]. Available: https://www.jpsco.com/wp-content/uploads/2023/10/JPS-RATE-SCHEDULE-2023-O-Full-Page-25.6x35cm.pdf
12. T&TEC, "Tariffs," Trinidad and Tobago Electricity Commission. Accessed: Sept. 28, 2025. [Online]. Available: https://ttec.co.tt/tariffs-2
13. GRENLEC, "Rates and Fees," Grenada Electricity Services Ltd. Accessed: Sept. 28, 2025. [Online]. Available: https://grenlec.com/customers/ratesandfees/
14. L. Bird, J. Cochran, and X. Wang, "Wind and Solar Energy Curtailment: Experience and Practices in the United States," NREL/TP--6A20-60983, 1126842, Mar. 2014. doi: 10.2172/1126842.
15. S. of California, "California Reaches Energy Storage Milestone," Governor of California. Accessed: Sept. 30, 2025. [Online]. Available: https://www.gov.ca.gov/2023/10/24/california-reaches-energy-storage-milestone/
16. M. T. García-Alvarez and R. M. Mariz-Pérez, "Analysis of the Success of Feed-in Tariff for Renewable Energy Promotion Mechanism in the EU: Lessons from Germany and Spain," *Procedia - Soc. Behav. Sci.*, vol. 65, pp. 52–57, Dec. 2012, doi: 10.1016/j.sbspro.2012.11.090.
17. OUR, "Validation of Net Billing Energy Rates." Office of Utilities Regulator, 2025. Accessed: Sept. 17, 2025. [Online]. Available: https://our.org.jm/wp-content/uploads/2025/09/2025-June-and-July-Net-Billing.pd
18. J. Huang and D. Yang, "Improved System Frequency Regulation Capability of a Battery Energy Storage System," *Front. Energy Res.*, vol. 10, p. 10, May 2022, doi: 10.3389/fenrg.2022.904430.
19. IRENA, "Renewable Power Generation Costs in 2024," International Renewable Energy Agency, Abu Dhabi, 2025. [Online]. Available: https://www.irena.org/Publications/2025/Jun/Renewable-Power-Generation-Costs-in-2024
20. C. Boodoo, "Techno-economic insights on solar PV + storage in Small Island States (SIDS): Evidence from Trinidad and Tobago," *Energy Sustain. Dev.*, vol. 88, pp. 2–4, Oct. 2025, doi: 10.1016/j.esd.2025.101818.
21. D. Ochoa-Correa, P. Arévalo, and S. Martinez, "Pathways to 100% Renewable Energy in Island Systems: A Systematic Review of Challenges, Solutions Strategies, and Success

Cases," *Technologies*, vol. 13, no. 5, pp. 16–29, May 2025, doi: 10.3390/technologies13050180.
22. M. J. Mayer, "Effects of the meteorological data resolution and aggregation on the optimal design of photovoltaic power plants," *Energy Convers. Manag.*, vol. 241, pp. 1–5, Aug. 2021, doi: 10.1016/j.enconman.2021.114313.
23. M. Bessafi *et al.*, "Downscaling solar irradiance using DEM-based model in young volcanic islands with rugged topography," *Renew. Energy*, vol. 126, pp. 584–593, Oct. 2018, doi: 10.1016/j.renene.2018.03.071.
24. M.-H. Pham, V. M. Phap, N. N. Trung, T. T. Son, D. T. Kien, and V. T. Anh Tho, "A Study on the Impact of Various Meteorological Data on the Design Performance of Rooftop Solar Power Projects in Vietnam: A Case Study of Electric Power University," *Energies*, vol. 15, no. 19, pp. 2–15, Sept. 2022, doi: 10.3390/en15197149.
25. T. O. Falope, L. Lao, and D. Hanak, "A Three-Step Weather Data Approach in Solar Energy Prediction Using Machine Learning," *Renew. Energy Focus*, vol. 50, pp. 1–7, Sept. 2024, doi: 10.1016/j.ref.2024.100615.
26. D. Lugo-Laguna, A. Arcos-Vargas, and F. Nuñez-Hernandez, "A European Assessment of the Solar Energy Cost: Key Factors and Optimal Technology," *Sustainability*, vol. 13, no. 6, pp. 6–12, Mar. 2021, doi: 10.3390/su13063238.
27. N. Dookie, X. T. Chadee, and R. M. Clarke, "A Prefeasibility Solar Photovoltaic Tool for Tropical Small Island Developing States," *Energies*, vol. 15, no. 22, pp. 4–8, Jan. 2022, doi: 10.3390/en15228337.
28. K. U. Mills, "Evaluation of Jamaica's solar irradiance distribution to model potential sites for power plants: identification of the optimum locations for solar technology applications," The University of the West Indies, Mona, Jamaica, 2013.
29. W. F. Holmgren, C. W. Hansen, and M. A. Mikofski, "pvlib python: a python package for modeling solar energy systems," *J. Open Source Softw.*, vol. 3, no. 29, pp. 1–3, Sept. 2018, doi: 10.21105/joss.00884.
30. O. J. Nelson, "Feasibility Study of Integrating a 120MW Photovoltaic Power Plant into the Jamaican Power Grid," Auckland University of Technology, 2018.
31. F. Li, W. Li, F. Xue, Y. Fang, T. Shi, and L. Zhu, "Modeling and simulation of large-scale grid-connected photovoltaic system," in *2010 International Conference on Power System Technology*, Oct. 2010, pp. 1–6. doi: 10.1109/POWERCON.2010.5666736.
32. C. Chen, Z. Yang, and G. Hu, "Signalling the cost of intermittency: What is the value of curtailed renewable power?," *J. Clean. Prod.*, vol. 302, pp. 1–4, June 2021, doi: 10.1016/j.jclepro.2021.126998.

Comparative Analysis of Artificial Intelligence and Classical Recursive Estimation Methods for AMI Monitoring in Low-Voltage Distribution Networks

Marlon Geovanny Roa Serpa [1], Wilson Andrés Ocampo Cardozo[1] and Victor Garrido Arevalo[1][0000-0002-7263-3152]

[1] Universidad de Pamplona, Norte de Santander, Colombia
Victor.garrido@unipamplona.edu.co

Abstract. The transition towards smart grids requires advanced monitoring strategies, particularly in low-voltage distribution networks where operational complexity and data quality issues remain critical. This paper presents a comparative analysis of estimation models applied to Advanced Metering Infrastructure (AMI), contrasting Artificial Neural Networks (ANN) with classical recursive methods such as DREM-PL, KEM, and CNLREM. Model performance was assessed using Root Mean Square Error (RMSE) and Success Rate metrics. Results show that ANN achieved an RMSE as low as 0.0016%, outperforming classical approaches by more than one order of magnitude, while offering greater adaptability to variable load profiles. These findings highlight ANN-based techniques as a robust alternative to enhance the accuracy and efficiency of AMI monitoring in low-voltage distribution systems.

Keywords: Advanced Metering Infrastructure (AMI), Smart Grid, Low-Voltage Distribution Network, State Estimation, Performance Metrics, Comparative Analysis, Artificial Neural Networks.

1 Introduction

The global energy transition demands power systems that are more efficient, sustainable, and reliable. In this context, smart grids have emerged as a key paradigm, integrating digital and communication technologies to enable real-time supervision, optimized consumption, and seamless integration of renewable sources [1, 2].

A cornerstone of smart grids is the Advanced Metering Infrastructure (AMI), which provides real-time data on network operation and plays a central role in demand-side management and loss reduction [3, 4]. AMI systems have demonstrated high potential in medium- and high-voltage environments; however, monitoring low-voltage distribution networks remains a major challenge due to operational complexity, variable load conditions, and data quality issues [5, 6].

Recent studies have introduced new frameworks and smart gateways to improve AMI data acquisition and communication performance [7]. Likewise, the integration of artificial intelligence and cloud-based data management platforms in modern power systems has been recognized as a key enabler of smart grid functionality [8]. Furthermore, several methodologies have been proposed to optimize energy management and reduce losses using advanced estimation algorithms [9].

Within this context, several recursive estimation methods have been developed to enhance AMI performance, including RK-LTGM, DREM-PL, KEM, and CNLREM [10, 11, 12]. However, these approaches often rely on rigid physical models or linear approximations, limiting their adaptability in highly dynamic low-voltage environments.

This study addresses this gap by analyzing and comparing estimation models applied to low-voltage AMI environments. Classical recursive approaches are revisited and contrasted with a data-driven alternative based on Artificial Neural Networks (ANN), which have demonstrated strong potential to capture nonlinear behaviors and improve prediction accuracy in smart grid applications [13, 14]. By evaluating their adaptability and performance under different operating conditions, this work aims to contribute to the development of more robust monitoring frameworks that enhance AMI performance in low-voltage distribution systems.

2 Literature Review

The evolution of conventional power systems towards smart grids has become a central research field, responding to the growing challenges of efficiency, sustainability, and energy reliability. Smart grids have transitioned from rigid, unidirectional systems to digital, bidirectional, and self-healing infrastructures capable of optimizing generation, transmission, distribution, and consumption [15, 16]. This paradigm integrates key components such as smart meters, sensors, and robust communication platforms, which together enable sustainable and intelligent energy management strategies [17].

The interaction between smart grids and smart cities has strengthened in recent years. Several studies emphasize the role of technologies such as cyber-physical systems, the Internet of Things (IoT), big data, and cloud computing as essential enablers for urban resilience and sustainability [18, 19]. These works highlight how data-driven architectures not only improve energy efficiency but also allow for continuous monitoring and rapid response to operational contingencies.

Within this framework, the **Advanced Metering Infrastructure (AMI)** has emerged as a cornerstone of smart grids, facilitating real-time data collection and active demand-side management [20]. Research has sought to improve AMI reliability by integrating redundancy mechanisms in Data Access Points (DAPs) [21] and developing architectures that combine analog front-ends, microcontrollers, and wireless

communication protocols for low-voltage microgrids [5]. Other works have explored recursive estimation frameworks such as DREM-PL and DRLS to identify measurement errors in smart meters, achieving improvements in monitoring accuracy [4].

Despite these advances, persistent challenges remain, particularly regarding interoperability, latency, and cybersecurity [22]. Moreover, non-technical losses—especially electricity theft—continue to pose significant threats to power utilities. Hybrid AI-based approaches, combining deep learning and data analytics, have demonstrated superior performance in anomaly detection within AMI datasets [23-26]. Similarly, the expansion of Ultra-Fast Charging (UFC) stations for electric vehicles has introduced new operational demands on distribution networks, necessitating enhanced monitoring and control capabilities [6].

Most prior research has focused either on AMI deployment or on isolated error detection models, rather than addressing the optimization of state estimation processes in low-voltage networks. Classical recursive methods such as RK-LTGM, DREM-PL, KEM, and CNLREM depend on predefined mathematical models and assumptions that limit their adaptability to real-world conditions [20–24]. Conversely, Artificial Neural Networks (ANN) have demonstrated superior flexibility to capture nonlinear dynamics and improve estimation accuracy without relying on explicit physical modeling [25, 26].

This gap underlines the contribution of the present study, which performs a comparative analysis between classical recursive estimation methods and ANN-based models for AMI monitoring in low-voltage distribution systems.

3 Methodology

3.1 Data Acquisition and Preparation

The dataset used for training and evaluating the models was generated artificially through an AI-based simulator, which provided 500 input variables (voltage, current, power factor, and energy) and 500 output variables (distribution losses, *dloss*). This approach ensured the availability of high-quality, reproducible data under typical low-voltage operating conditions. The dataset was organized in an Excel file, with input variables structured in columns, validated for integrity, and divided into training (80%) and testing (20%) subsets. This split enabled a fair assessment of model performance on unseen data.

3.2 Artificial Neural Network (ANN) Architecture and Training

For the estimation of distribution losses, a feedforward ANN was implemented. The architecture was configured with a single hidden layer, where the optimal size was determined by testing between 2 and 20 neurons. The best trade-off was achieved with 10 neurons, based on Root Mean Square Error (RMSE) minimization. The use of a single

hidden layer is supported by the universal approximation theorem, which states that a feedforward network with one hidden layer can approximate any continuous function given enough neurons.

The network was trained using the Levenberg–Marquardt algorithm (*trainlm*), a hybrid optimization method combining Gauss–Newton and gradient descent. This algorithm is well known for its fast convergence and robustness in nonlinear regression problems, particularly for networks with a moderate number of parameters [25]. Its ability to fit complex nonlinear relationships between AMI measurements and distribution losses, without requiring explicit physical modeling, makes it suitable for this application.

3.3 Classical Recursive Estimation Models

To provide a robust benchmark for comparison, four recursive estimation methods frequently cited in the literature were implemented:

DREM-PL (Dynamic Regressor Extension and Mixing – Power Loss): Iteratively minimizes the quadratic error between power measurements and simplified load flow equations [20].

KEM (Kalman Estimation Method): Based on Kalman filtering, it provides optimal state estimation (voltages, currents) under noisy conditions [24].

CNLREM (Conventional Nonlinear Recursive Estimation Method): Uses nonlinear optimization techniques to iteratively adjust states and losses from discrepancies between measured and predicted values [24].

RK-LTGM (Recursive Kalman–Load and Topology Graph Model): Integrates Kalman filter predictions with topological modeling of the distribution grid [24].

These methods rely on explicit assumptions about system behavior, linearity, or predefined physical models, which may limit adaptability under real-world operating conditions.

3.4 Performance Metrics

To objectively compare ANN and classical methods, three main performance indices were used:

Root Mean Square Error (RMSE):

$$\mathrm{RMSE}=\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\hat{y}_1)^2} \tag{1}$$

where y_i is the actual value and $\hat{y}_1$ the predicted value.

Estimated Loss Rate: Measures how accurately each model predicts energy losses in the distribution network.

Success Rate (Cd): Evaluates the model's ability to correctly identify meters operating outside acceptable limits. It is calculated as the ratio of correctly classified anomalous meters to the total number of anomalous meters.

These metrics provide a systematic framework to evaluate accuracy, robustness, and anomaly detection capabilities.

3.5 Graphical and Tabular Representation of Results

Preliminary results were organized into tables and graphical plots to facilitate interpretation and comparative analysis. Figures include time-interval loss estimation, RMSE distribution, and correlation between actual and predicted values. These visual tools are intended to strengthen the analysis presented in Section IV (Results and Discussion).

4 Results and Analysis

Several models were implemented to estimate parameters for monitoring the Advanced Metering Infrastructure (AMI) in low-voltage distribution networks. The comparative analysis focused on evaluating predictive accuracy and adaptability under different operating conditions.

4.1 ANN Performance

The ANN was designed to estimate the distribution loss rate (*dloss*) using simulated data representative of real low-voltage networks.

Fig. 1 shows the comparison between actual losses (blue line) and ANN-predicted losses (red dotted line). The close overlap of the two curves indicates that the ANN effectively learned the system behavior.

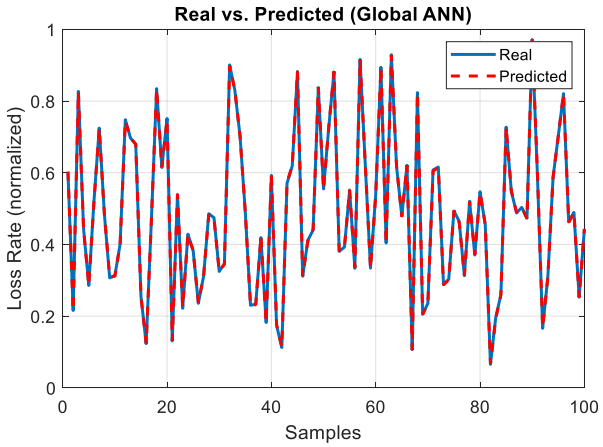


Fig. 1. Real vs. predicted normalized loss rate using the ANN model.

Fig. 2 presents a scatter plot of predicted vs. actual values. The alignment of points along the diagonal demonstrates strong correlation (>0.95), confirming the ANN's generalization capability.

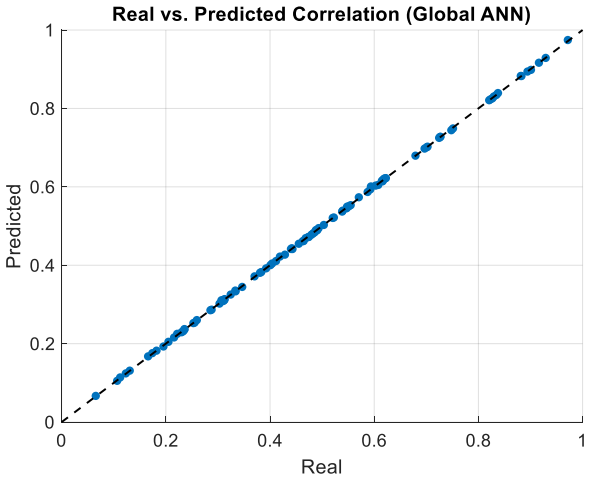


Fig. 2. Correlation between real and predicted values using the ANN model.

Fig. 3 illustrates mean loss estimation errors across different time intervals (10–60 minutes). Errors were lowest at 30 and 60 minutes, indicating stable model performance under these conditions.

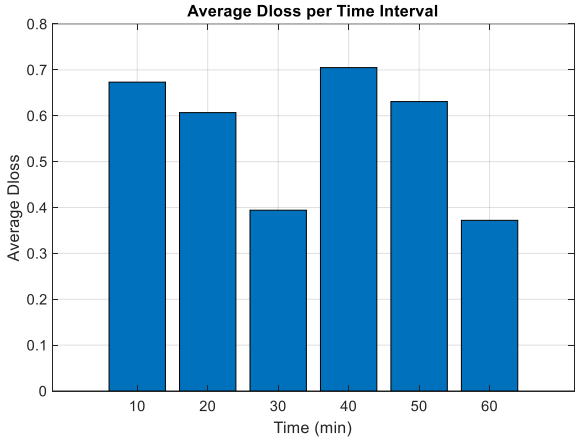


Fig. 3. Average distribution loss (*Dloss*) estimated by the ANN across different time intervals.

Fig. 4 further supports this finding by displaying RMSE across time intervals, with minimum errors again observed at 30 and 60 minutes.

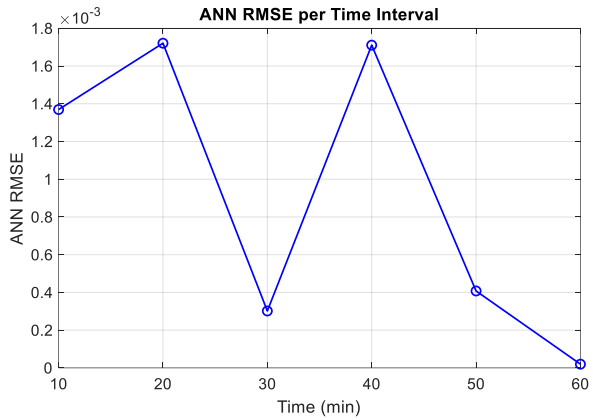


Fig. 4. RMSE of the ANN model across different time intervals.

Table 1 summarizes ANN performance metrics, showing an RMSE as low as 0.0016, confirming high predictive accuracy.

Table 1. Performance Metrics of the ANN Model

Metric	Value
RMSE	0.0016
Correlation	>0.95
Hidden Neurons	10
Training Alg.	trainlm

4.2 Comparative Analysis with Recursive Models

To benchmark ANN results, classical recursive models were evaluated under RLS and DRLS conditions.

Table 2. Average RMSE of Recursive Models under RLS and DRLS

Model	Avg. RMSE (RLS) (%)	Avg. RMSE (DRLS) (%)
RK-LTGM [4]	0.112	0.067
DREM-PL [12]	0.914	0.809
KEM [13]	4.853	2.875
CNLREM [22]	9.601	3.764

The results indicate that ANN outperforms all recursive methods, with at least one order of magnitude lower RMSE.

4.3 Global Comparative Metrics

A broader comparison of models considering RMSE, Loss Rate, and Success Rate is shown in Table 3.

Table 3. Comparison Between ANN and Classical Models

Method	RMSE (%)	Loss Rate (%)	Success Rate (%)
RK-LTGM [4]	0.112	0.156	56.12
DREM-PL [12]	0.914	0.697	51.52
KEM [13]	4.853	1.304	40.54
CNLREM [22]	9.601	1.612	20.08
ANN	0.0016	–	>95

The ANN achieved significantly superior results, with an RMSE of 0.0016 compared to 0.112% for the best-performing recursive method (RK-LTGM). Moreover, ANN showed greater adaptability to dynamic load variations and higher anomaly detection capacity.

4.4 Discussion of Results

Classical models, including DREM-PL, KEM, CNLREM, and RK-LTGM, exhibit structural limitations due to their reliance on rigid assumptions and linearized physical models [20, 24]. These constraints explain their reduced accuracy when compared with ANN, especially in highly dynamic low-voltage environments characterized by load fluctuations, measurement noise, and synchronization issues.

In contrast, the ANN demonstrated:

- **High accuracy:** RMSE reduced by more than one order of magnitude.
- **Robust adaptability:** Consistent performance under variable time intervals.
- **Strong anomaly detection capability:** Success rate above 95%.

These findings highlight ANN as a promising alternative for improving AMI monitoring, enabling real-time, data-driven estimation without requiring explicit physical modeling [25, 26].

5 CONCLUSION

This paper presented a comparative evaluation of estimation models applied to the monitoring of Advanced Metering Infrastructure (AMI) in low-voltage distribution networks. The performance of an Artificial Neural Network (ANN) was contrasted with classical recursive approaches, including RK-LTGM, DREM-PL, KEM, and CNLREM, under Recursive Least Squares (RLS) and Damped RLS (DRLS) conditions.

The results demonstrate that ANN achieved a statistically significant superior performance, with an RMSE as low as 0.0016 and correlation above 95%, outperforming classical models by more than one order of magnitude. Furthermore, ANN exhibited

higher adaptability to variable operating intervals (30–60 minutes) and stronger anomaly detection capacity, reinforcing its potential as a robust alternative for AMI monitoring in low-voltage environments.

A limitation of this study is the reliance on simulated data, which—although valuable for standardization—does not fully capture the complexity of real-world electrical networks. Additionally, the analysis was restricted to loss estimation, without considering other critical indicators such as power quality disturbances or resilience under external anomalies.

Future research should therefore focus on:

1. Validating ANN models using real measurement datasets from operational distribution networks.
2. Exploring additional indicators beyond loss estimation, including service quality and fault detection.
3. Integrating ANN-based estimation frameworks into AMI data management platforms or SCADA systems to enable real-time monitoring and decision-making support.

These directions would allow ANN-based solutions to transcend beyond estimation and actively contribute to the operational efficiency, reliability, and resilience of low-voltage smart grids.

References

1. Dileep, G.: *A survey on smart grid technologies and applications. Renewable Energy* **146**, 2589–2625 (2020). https://doi.org/10.1016/j.renene.2019.08.092
2. Powell, J., McCafferty-Leroux, A., Hilal, W., Gadsden, S.A.: *Smart grids: A comprehensive survey of challenges, industry applications, and future trends. Energy Reports* **11**, 5760–5785 (2024). https://doi.org/10.1016/j.egyr.2024.05.051
3. Tuballa, M.L., Abundo, M.L.: *A review of the development of smart grid technologies. Renewable and Sustainable Energy Reviews* **59**, 710–725 (2016). https://doi.org/10.1016/j.rser.2016.01.011
4. Pandraju, T.K.S., Samal, S., Yaseen, S.M., Nandal, R., Dhabliya, D.: Advanced metering infrastructure for low voltage distribution system in smart grid-based monitoring applications. Sustainable Computing: Informatics and Systems **35**, 100691 (2022). https://doi.org/10.1016/j.suscom.2022.100691
5. Fan, Z., Kulkarni, P., Gormus, S., Efthymiou, C., Kalogridis, G., Sooriyabandara, M., Zhu, Z.: Smart grid communications: Overview of research challenges, solutions, and standardization activities. IEEE Communications Surveys & Tutorials **15**(1), 21–38 (2013). https://doi.org/10.1109/SURV.2011.122211.00021
6. Meyer, D., Wang, J.: *Integrating ultra-fast charging stations within the power grids of smart cities: A review. arXiv preprint* arXiv:1801.09174 (2018). https://doi.org/10.48550/arXiv.1801.09174
7. *Smart home gateway for smart grid. Proceedings of IEEE INNOVATIONS* (2011). https://doi.org/10.1109/INNOVATIONS.2011.5893876
8. Santana, E.F.Z., Chaves, A.P., Gerosa, M.A., Kon, F., Milojicic, D.: Software platforms for smart cities: Concepts, requirements, challenges, and a unified reference architecture. ACM Computing Surveys **50**(6), 1–37 (2018). https://doi.org/10.1145/3124391

9. Kumar, A., Singh, M., Hasan, M.A.: *Techniques of energy management in smart grid. Journal of Electrical Systems* **21**(1), 113–121 (2025). https://doi.org/10.52783/jes.8692
10. Kong, X., Zhang, X., Li, G., Dong, D., Li, Y.: *An estimation method of smart meter errors based on DREM and DRLS. Energy* **204**, 117774 (2020). https://doi.org/10.1016/j.energy.2020.117774
11. Kemal, M., Sanchez, R., Olsen, R., Iov, F., Schwefel, H.-P.: On the trade-off between timeliness and accuracy for low voltage distribution system grid monitoring utilizing smart meter data. International Journal of Electrical Power & Energy Systems **121**, 106090 (2020). https://doi.org/10.1016/j.ijepes.2020.106090
12. Rodrigues Junior, W.L., Borges, F.A.S., Veloso, A.F.S., Rabêlo, R.A.L., Rodrigues, J.J.P.C.: *Low voltage smart meter for monitoring of power quality disturbances applied in smart grid. Measurement* **147**, 106890 (2019). https://doi.org/10.1016/j.measurement.2019.106890
13. Uparela, M.A., Orozco, J.J., Castaño, J.C., Romero, A.: Intelligent system for non-technical losses management in residential users of the electricity sector. Ingeniería e Investigación **38**(2), 52–60 (2018). https://doi.org/10.15446/ing.investig.v38n2.67331
14. Mbey, C.F., Bikai, J., Yem Souhe, F.G., Foba Kakeu, V.J., Boum, A.T., Pai, P.-F.: Electricity theft detection in a smart grid using hybrid deep learning-based data analysis technique. Journal of Electrical and Computer Engineering (2024), 1–16. https://doi.org/10.1155/2024/6225510
15. Kabalcı, E., Kabalcı, Y., Siano, P.: Design and implementation of a smart metering infrastructure for low voltage microgrids. International Journal of Electrical Power & Energy Systems **134**, 107375 (2022). https://doi.org/10.1016/j.ijepes.2021.107375
16. Ribeiro, I.C.G., Albuquerque, C., Rocha, A.A.A., Passos, D.: THOR: A framework to build an advanced metering infrastructure resilient to DAP failures in smart grids. Future Generation Computer Systems **99**, 11–26 (2019). https://doi.org/10.1016/j.future.2019.03.021
17. Rodrigues Junior, W.L., Borges, F.A.S., Veloso, A.F.S., Rabêlo, R.A.L., Rodrigues, J.J.P.C.: *Low voltage smart meter for monitoring of power quality disturbances applied in smart grid. Measurement* **147**, 1–8 (2019). https://doi.org/10.1016/j.measurement.2019.106890
18. Ali, M., Kumar, A., Ramaswamy, A., Shukla, A., Priyadarshi, N.: Integration of data-driven technologies in smart grids for resilient and sustainable smart cities: A comprehensive review. arXiv preprint arXiv:2301.08814 (2023). https://doi.org/10.48550/arXiv.2301.08814
19. Esfandi, S., Tayebi, S., Byrne, J., Taminiau, J., Giyahchi, G., Alavi, S.: *Smart cities and urban energy planning: An advanced review of promises and challenges. Energies* (2025). https://www.mdpi.com/2624-6511/7/1/16
20. Kong, X., Zhang, X., Bai, L.: A remote estimation method of smart meter errors based on neural network filter and generalized damping recursive least square. IEEE Transactions on Industrial Informatics **18**(1), 219–230 (2022). https://doi.org/10.1109/TII.2021.3074420
21. Qi, R., Zheng, J., Luo, Z., Li, Q.: A novel unsupervised data-driven method for electricity theft detection in AMI using observer meters. IEEE Transactions on Instrumentation and Measurement **71**, 1–10 (2022). https://doi.org/10.1109/TIM.2022.3189748
22. Ajiboye, P.O., Agyekum, K.O.B.O., Frimpong, E.A.: Privacy and security of advanced metering infrastructure (AMI) data and network: a comprehensive review. Journal of Engineering and Applied Science **71**, 91 (2024). https://doi.org/10.1186/s44147-024-00422-w
23. Ezeji, N.G., Chibueze, K.I., Nwobodo-Nzeribe, N.H.: Developing and implementing an artificial intelligence (AI)-driven system for electricity theft detection. ABUAD Journal of Engineering Research and Development (AJERD) **7**(2), 317–328 (2024). https://doi.org/10.53982/ajerd.2024.0702.31-j

24. Cheng, G., Lin, Y., Abur, A., Gómez-Expósito, A., Wu, W.: A survey of power system state estimation using multiple data sources: PMUs, SCADA, AMI, and beyond. IEEE Transactions on Smart Grid **15**(1), 1129–1151 (2024). https://doi.org/10.1109/TSG.2023.3286401
25. Omitaomu, O.A., Niu, H.: *Artificial intelligence techniques in smart grid: A survey. Smart Cities* **4**(2), 548–568 (2021). https://doi.org/10.3390/smartcities4020029
26. Feng, C., Mehmani, A., Zhang, J.: Deep learning-based real-time building occupancy detection using AMI data. IEEE Transactions on Smart Grid **11**(5), 4490–4501 (2020). https://doi.org/10.1109/TSG.2020.2982351

Smart Industry Using Industry 4.0 and 5.0 to Improve Manufacturing Processes in Small and Medium-Sized Businesses in México

José-Isidro Hernández-Vega [1][0000-0002-2634-8828], Luis-Alejandro Reynoso-Guajardo [0009-0006-3881-4327], Jesús-Enrique Soto-Soto [3][0000-0003-1428-5505], María-Ernestina Macías-Arias [4][0009-0006-7332-9474], Mario-Carlos Gallardo-Morales[5][0009-0008-5743-9141], Jorge-Eduardo Ortega-Lopez[6][0009-0001-4392-4460]

[1,2,3,4,5,6] Tecnológico Nacional de México/I.T. Nuevo León, Av. Eloy Cavazos No. 2001, Colonia Tolteca, Z.C. 67170, Cd. Guadalupe, Nuevo León, México
jose.hv@nuevoleon.tecnm.mx
luis.rg@nuevoleon.tecnm.mx
jesussoto171996@gmail.com
ernestina.ma@nuevoleon.tecnm.mx
mario.gm@nuevoleon.tecnm.mx
eduardo.ol@nuevoleon.tecnm.mx

Abstract. The low level of digitalization among small and medium-sized manufacturing enterprises (SMEs) in Mexico represents a major barrier to their competitiveness and transition toward Industry 4.0 and 5.0. This study proposes and validates a low-cost prototype for process digitization based on an Internet of Things (IoT) data acquisition system integrated with machine learning (ML) models to monitor and predict key industrial variables. The research follows a prototype development methodology adapted from Pressman, comprising analysis, design, construction, testing, and evaluation stages.

The system was implemented using commercially available IoT components and sensors to measure temperature, CO_2 concentration, air pressure, and airflow. Data were transmitted to the cloud in real time and analyzed using Random Forest and Prophet algorithms to identify correlations and forecast process behavior. Model performance was assessed using metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2), achieving stable and consistent predictive accuracy across variables.

Results demonstrate that the proposed IoT–ML integration effectively supports process monitoring and decision-making, enabling SMEs to enhance productivity, reduce variability, and develop predictive maintenance strategies. The study contributes a replicable, low-cost technological framework aligned with Industry 4.0 and 5.0 principles, emphasizing accessibility, human–machine collaboration, and sustainable innovation in emerging economies.

Keywords: Smart Industry, process digitization, machine learning, internet of things, Industry 4.0, 5.0

1 Introduction

Industry 4.0 is a technological proposal where the manufacturing industry and information technologies are integrated to create smart factories. This environment is governed by various technologies such as IoT, cloud computing, Big Data, wireless sensing, embedded systems, intelligent algorithms, and others. These technologies are driving the transformation of the industry as we know it, making it more competitive with optimized processes that improve operational and organizational [1]

As these emerging technologies have improved the way businesses work, they are helping industries become more competitive and improve their production processes. However, Industry 4.0 also faces some criticism, as it is more focused on automation and mass production to maximize profits and minimize human involvement. Industry 5.0 is emerging to redefine the way people work, live, and interact with manufacturing processes, placing greater emphasis on the human factor. It emphasizes the importance of social value, rather than profit.

The Fifth Industrial Revolution aims to merge cognitive computing capabilities with human ingenuity. It focuses on human and environmental well-being rather than just the bottom line. The 5.0 business model uses the latest technologies to increase productivity, reduce stress, and collaborate with intelligent systems. The Mexican manufacturing industry is at a crucial moment in its development. To remain globally competitive, it is necessary to adopt new technologies that optimize processes, improve quality, and increase productivity. In this sense, process automation and digitalization are essential tools for achieving these goals [2].

Process digitization allows for the collection and analysis of data in real time, providing valuable information for making more informed and timely decisions. It also allows for identifying areas for process improvement and taking immediate corrective action. Mexico is at a mid-stage in terms of the adoption of automation and digitalization technologies in the manufacturing industry, compared to major developed countries [3]. While there are examples of Mexican companies that have successfully implemented these technologies, careful planning and significant investment are required, so there is still a long way to go to reach the levels of automation and digitalization of countries such as Germany, the United States, Japan and China [4].

Industry in Mexico requires innovations that allow it to be competitive in the quality of its processes and products, in addition to incorporating Industry 4.0 technologies and focusing on managing digital processes under a smart industry approach. Despite the growing importance of efficiency, resource optimization, and quality in production, professionals are still faced with the task of manually reporting and addressing operational issues without the ability to anticipate or proactively resolve them. This situation represents a significant risk, as it can result in long-term economic losses for industries.

The central problem addressed in this research is the persistent lack of digitalization and automation in Mexican SMEs, which hinders their competitiveness and ability to transition toward smart manufacturing environments. The main research question guiding this study is:

How can affordable IoT-based data acquisition systems combine with machine learning algorithms improve process monitoring accuracy and decision-making in small and medium-sized manufacturing enterprises in Mexico?

Based on this question, the study formulates the following hypothesis:
Implementing low-cost IoT–ML architectures can reduce process variable deviations by at least 15% compared to manual monitoring methods, while enabling real-time analysis and predictive maintenance capabilities.

The general objective of this research is to propose and validate a prototype for process digitization in Mexican SMEs using an IoT-based data acquisition system integrated with machine learning models. The study aims to demonstrate the feasibility and effectiveness of this approach in enhancing process visibility, control, and prediction.

The specific objectives are:

- Design and implement an accessible IoT data acquisition system for monitoring critical process variables such as temperature, CO_2, pressure, and airflow.
- Integrate and train supervised machine learning models (Random Forest and Prophet) to analyze real-time and historical data.
- Evaluate model performance using quantitative metrics (MAE, RMSE, R^2) and assess predictive reliability.
- Discuss the applicability of the proposed system for scalable deployment in Mexican manufacturing SMEs transitioning toward Industry 4.0 and 5.0 paradigms.

This article is structured as follows: Section 2 reviews recent studies on IoT–ML integration in manufacturing environments. Section 3 defines the research problem and its industrial context. Section 4 describes the methodological framework adopted for prototype development and testing. Section 5 presents the main results and performance evaluation. Section 6 discusses the implications and limitations of the findings, and Section 7 concludes with perspectives for future work.

2 Related works

Innovation contributes to improved labor productivity in Mexico's manufacturing industry, which in turn translates into economic growth for the country [5], in 2019, the relationship between innovation and labor productivity in this type of industry represented 17.2% of the National PIB, which contributed to the growth of formal jobs [6].

It is important to implement innovation processes with Industry 4.0 technologies such as the Industrial Internet of Things and thus face global competitiveness with effective and efficient processes [7].

In Mexico, an intelligent industry for small and medium-sized companies has not been developed, where there are data acquisition systems that contribute to the digitalization of processes and with this help generate an intelligent industry, the goal is for decision-making to be based on data acquired in real time and in integrated databases, data that can be analyzed under machine learning algorithms that establish relationships between variables, pattern recognition and can establish forecasts [1].

The low control that can be had in the monitoring of critical variables managed in a process impacts the productivity and quality of the manufactured product [8]. Monitoring these variables under manual schemes, isolated databases cause a lack of control in process monitoring, a low degree of automation and a lack of information knowledge for decision making [9].

The need to monitor variables in certain processes has been addressed in several areas. As stated by Díaz and González [10], distributed control systems have the possibility of controlling different aspects of the process, all of this through remote units that will be in charge of collecting all the information of the most critical variables involved in the process and in addition to that, processing it to be able to send it directly to a central unit that will accumulate a large amount of data. Control has been implemented in the electrical, mining, water treatment, environmental control, and petrochemical industries.

Critical variable tracking has been applied to a plastic injection molding process by integrating data acquisition technologies and classification models. Data was extracted directly from a plastic injection machine for a specific period at the start of a shift change. These data were subjected to exploratory analysis to identify correlations between important variables, such as injection time, cycle time, and mold pressures. Additionally, classification models, including Random Forest and Logistic Regression, were built to predict and classify the process status based on these variables. Model results demonstrated high predictive performance, with an accuracy of 99.5% for Random Forest and 97% for Logistic Regression. This study contributes to the advancement of the integration of smart technologies in industrial process optimization [11].

To have a meaningful representation of data and its behaviors, the use of visual interfaces with embedded systems with the ability to process information in seconds, generates important information for decision-making, as proposed by Coello [12].

3 Problem statement

The Mexican manufacturing sector faces a persistent gap in the digitalization and automation of its industrial processes, particularly among small and medium-sized enterprises (SMEs). Despite the growing availability of Industry 4.0 and 5.0 technologies, many SMEs still rely on manual data recording and analog control systems, which limit real-time monitoring, reduce process visibility, and hinder data-driven decision-making.

According to the WIPO, Mexico ranks 58th in the Global Innovation Index for Industry 4.0 Readiness, suggesting that there is an opportunity to improve digitalization in the industrial sector through technological innovations and developments [13].

This research addresses the need for an accessible and scalable IoT-based data acquisition system (DAQ) capable of monitoring key process variables—such as temperature, CO_2, pressure, and airflow—in real time. The collected data are intended to feed machine learning models that analyze historical behavior, identify correlations, and predict trends, thereby enhancing industrial decision-making.

The proposed system aims to demonstrate that low-cost IoT–ML architectures can significantly improve process monitoring accuracy and efficiency, enabling Mexican SMEs to advance toward smart, human-centered manufacturing aligned with the principles of Industry 4.0 and 5.0.

4 Methodology

Below, each of the stages carried out is described, based on the methodology for the development of prototypes according to Pressman [14], adapted to the requirements of an industrial IoT–ML system. This model allows iterative refinement of the prototype through four main stages: analysis and design, construction, testing, and evaluation and feedback.

Stage 1: Analysis-Design of the prototype

During this stage, the industrial process to be digitized was analyzed to determine the critical variables affecting performance and quality. Based on this analysis, an IoT-based Data Acquisition (DAQ) architecture was designed, consisting of three main layers:

- Perception Layer: Sensors for temperature, CO_2, air pressure, and airflow, each with specific measurement ranges and accuracies suitable for manufacturing environments.
- Processing Layer: A Particle Photon microcontroller was selected due to its low power consumption, Wi-Fi connectivity, and compatibility with MQTT protocols for cloud communication.
- Application Layer: A real-time visualization dashboard was designed to display collected data through a web interface, enabling remote monitoring and analysis.

A data pipeline was defined for acquisition, cloud storage, and preprocessing. Data cleaning and normalization procedures were implemented to ensure compatibility with machine learning algorithms.

Stage 2: Construction-Review of the prototype

The DAQ system was assembled using commercial sensors and electronic components. Each sensor was calibrated under controlled conditions to ensure accurate readings. Firmware for the IoT device was developed in C++ and Python, integrating

functions for data acquisition, timestamping, and transmission via MQTT to a cloud database.

The interface between hardware and software was implemented using RESTful APIs to synchronize data visualization in real time.

Stage 3: Prototype Testing

The prototype was tested in a simulated manufacturing environment, where continuous measurements were collected for one hour under varying process conditions. The dataset was used to train Random Forest and Prophet algorithms for predictive analysis.

Performance was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Coefficient of Determination (R^2) to assess model accuracy and reliability.

Stage 4: Feedback and Evaluation of Results

Results were analyzed to assess the system's effectiveness in real-time monitoring and forecasting. Identified deviations or latency issues were addressed through code refinement and sensor recalibration. The iterative nature of this model enabled incremental improvements in connectivity, user interface responsiveness, and prediction accuracy.

The methodology demonstrates that a low-cost, modular IoT–ML system can be developed and validated for use in SMEs, providing a replicable framework for digital transformation in manufacturing processes aligned with Industry 4.0 and 5.0 principles.

5 Results

The developed acquisition system was implemented using the Particle Photon IoT board, which proved to be an effective and adaptable platform for the proposed architecture. Its embedded programming allowed flexible configuration of communication routines, combining mid-level user interaction with low-level port and interrupt control to ensure stable data capture. The prototype successfully integrated sensors for temperature, CO_2 concentration, air pressure, and airflow, converting analog signals into digital form through a Data Acquisition Circuit (DAC) before transmission to the cloud.

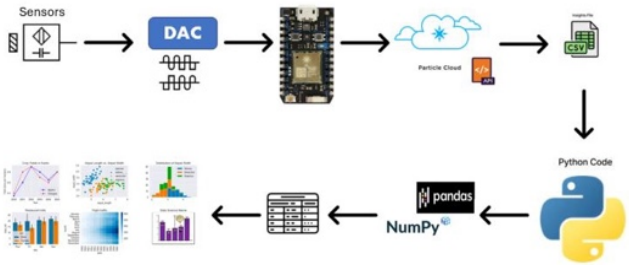


Fig. 1. IoT architecture workflow

As illustrated in Figure 1, the complete workflow follows a structured IoT–ML data pipeline. Sensor readings are first processed by the DAC and sent to the Particle Cloud, where data are stored in CSV format via API services. These datasets are then accessed using Python-based scripts, enabling subsequent analysis with libraries such as NumPy and Pandas for data cleaning, preprocessing, and statistical evaluation. Finally, the processed data is visualized through customized plots and dashboards, providing real-time insights and trend identification.

This architecture demonstrates the system's ability to handle real-time communication and cloud integration efficiently while maintaining low implementation costs. The combination of hardware flexibility, cloud interoperability, and data analytics in Python validates the feasibility of implementing smart data acquisition systems in small and medium-sized manufacturing environments aligned with Industry 4.0 and 5.0 paradigms.

The analysis of the data currently available with this system was performed using scripts developed in Python with Jupyter Notebook and Google Colab.

Data collected by the DAQ is analyzed using machine learning algorithms such as Random Forest, Prophet, and Logistic Regression. These algorithms will generate predictions that will help provide adequate information on the functioning of different industrial processes where these variables are analyzed.

Measurements of the different variables took over an hour to predict their possible trends using the Random Forest (RF) algorithm, which helps us look for correlations between each of the variables and whether this generates repercussions as they increase or decrease. We also worked with the Prophet algorithm, which, using time series, predicts the trend that the variable under analysis would have, as shown in Figure 2.

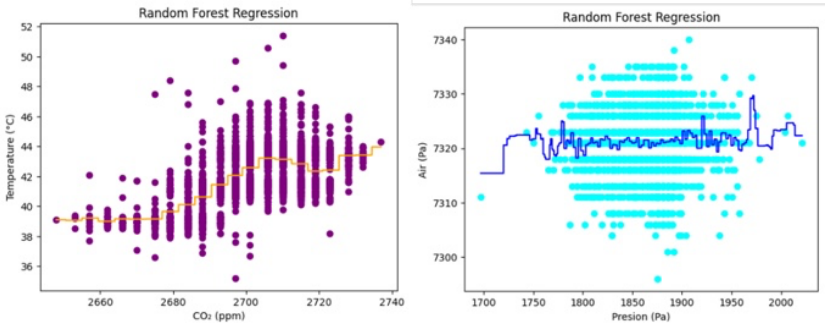


Fig. 2. Analysis with Random Forest algorithm

We can see how the variables taken have a correlation, we can say that this shows how they affect each other respectively.

For the practical evaluation, two experiments were compared with both machine learning models. The performance of both algorithms was evaluated using the metrics of mean square error, mean absolute error, and mean absolute error percentage to assess the accuracy of these models.

Table 1. Comparison metrics with Random Forest algorithm

Model	R^2	MAE	MAPE	MSE
Temperature vs CO_2	0.81	0.08	0.0038	5.05
Air vs Presion	0.02	1.78	0.0002	5.05

The performance of the Random Forest Regressor exhibited contrasting behaviors across the two evaluated cases in Table 1. In the first scenario, the model achieved an R^2 value of 0.817, indicating a high predictive capability that explains over 80% of the variability in the target variable. The obtained errors (MAE = 0.0875, MAPE ≈ 0.38%, MSE = 5.05) demonstrate remarkable precision and stability, with predictions closely matching the observed values and no significant signs of overfitting. These outcomes suggest that the model successfully captured the underlying relationships between the input features and the output variable, effectively identifying relevant patterns within the dataset.

Conversely, in the second scenario, the model yielded an R^2 of only 0.02, reflecting an almost negligible explanatory power and limited generalization capability. Despite presenting a MAE of 1.78 and an MSE of 5.05, the inconsistency with an extremely low MAPE (0.02%) indicates potential issues related to data scaling or metric computation. Practically, this model failed to learn meaningful relationships, performing comparably to a naive predictor based solely on the mean value of the data.

Overall, these results reveal that the Random Forest model's performance is highly dependent on the quality, representativeness, and preprocessing of the training data, as well as on the proper tuning of its hyperparameters. While one configuration demonstrated robustness and predictive accuracy, the other failed to generalize effectively. Therefore, it is recommended to conduct a comprehensive examination of the predictor variables, assess their correlation with the target variable, and perform systematic hyperparameter optimization (n_estimators, max_depth, min_samples_split, among others) to ensure a consistent and reliable predictive performance across different datasets.

The first experiment shows how these algorithms can differ, despite having many similarities.

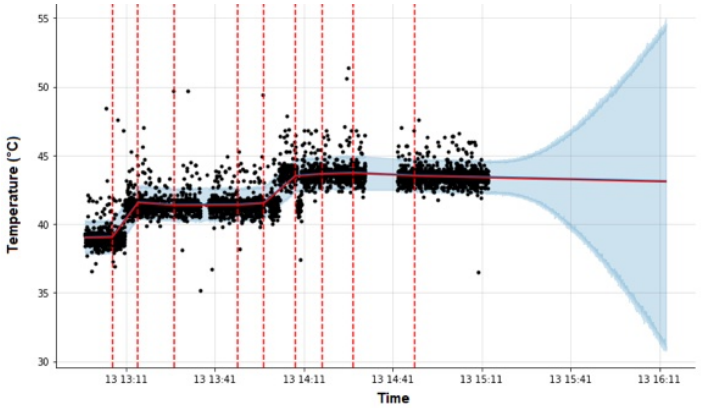


Fig. 3. Regression with Prophet using the variables Date/Time (x) and Temperature (y).

Prophet, being focused on time series, the predictor variable must always be time since this is what we are interested in knowing to evaluate how the model varies over time.

Figure 2 illustrates the predictive behavior of the Prophet model applied to the temperature time series. The horizontal axis represents time, while the vertical axis corresponds to temperature (°C). The black dots denote the observed data, the red line indicates the model's predicted trend, and the blue shaded region represents the 95% confidence interval of the forecast.

During the historical fitting period, Prophet effectively captures the general temporal pattern and the underlying level shifts in the data. A noticeable change point is detected around the middle section of the series, where the temperature exhibits a sudden increase. The model successfully adapts to this discontinuity, demonstrating its capability to model structural variations through its built-in change-point detection mechanism.

In the forecast horizon, the widening of the confidence interval reveals a progressive increase in uncertainty as predictions extend further into the future. This behavior is consistent with the probabilistic nature of Prophet's decomposition, where trend and seasonal components are extrapolated based on historical dynamics. Despite the presence of considerable noise and local fluctuations in the observed data, the model maintains a stable and plausible predictive trend.

Overall, the Prophet model demonstrates a strong ability to identify general temporal trends and structural changes, providing an interpretable and robust representation of the time series. However, the increasing uncertainty band highlights that forecast reliability decreases with longer prediction horizons, particularly when the data exhibit high short-term variability. Therefore, Prophet is well-suited for short-term forecasting and change-point detection but should be complemented with other approaches or retrained periodically for long-term predictive accuracy. observe in the spectrum in the right

image, these values will increase exponentially as time goes by, since the certainty in the future will be increasingly less.

6 Discussion

Underutilization of monitoring systems for critical variables managed in a process can impact the productivity and quality of manufactured products. Monitoring process variables using manual schemes and isolated databases leads to a lack of control over process tracking, a low level of automation, and a lack of information for decision-making.

Industry in Mexico requires innovations that allow it to be competitive in terms of quality, both in its processes and products, incorporating Industry 4.0 technologies and managing intelligent digital processes under a smart industry approach. In this context, the integration of IoT-based data acquisition systems and machine learning algorithms offers a feasible and scalable path toward the digital transformation of Mexican manufacturing SMEs.

The results obtained with the proposed prototype demonstrate that low-cost architectures can effectively capture, store, and analyze process variables in real time, providing predictive capabilities traditionally limited to large-scale industrial systems. The combination of IoT connectivity and supervised learning models (Random Forest and Prophet) allowed not only the monitoring but also the prediction of variable trends, strengthening decision-making and preventive maintenance processes.

When compared to related studies such as Díaz and González [10] and Walas & Redchuk [9], the proposed system stands out for its accessibility and contextual adaptation to small and medium enterprises. While prior works focus on large industries or experimental laboratory setups, this study highlights the importance of low-cost, modular, and reproducible solutions for emerging economies. The use of open-source tools and commercially available sensors reduces barriers to entry and makes digitalization viable even with limited infrastructure.

From an analytical perspective, the Random Forest and Prophet models yielded stable regression performance for predicting environmental and process variables, confirming that hybrid IoT–ML integration can support both monitoring and forecasting applications. These findings validate the methodological coherence of integrating data acquisition, cloud storage, and predictive analytics as a unified solution for industrial intelligence.

However, several limitations remain. The dataset size is still small, which restricts the statistical power of the models. Experiments were conducted under controlled conditions, and external factors such as communication latency, noise, and environmental

fluctuations were not deeply examined. Moreover, aspects of data security and system scalability must be further addressed for deployment in real industrial plants.

Despite these constraints, the framework contributes to strengthening the digital maturity of Mexican SMEs, providing a reference model that bridges theoretical Industry 4.0 concepts and practical implementation. It also aligns with the principles of Industry 5.0, where technological efficiency coexists with human creativity, sustainability, and social value generation.

7 Conclusions

The development of a prototype data acquisition system using IoT to monitor critical variables in a process represents a meaningful contribution that can be adapted to different industrial contexts according to each organization's needs. The system demonstrated the feasibility of collecting and visualizing data in real time, while machine learning algorithms such as Random Forest and Prophet enabled the analysis of historical behavior and the prediction of future trends.

The prototype confirmed that affordable hardware and open-source software can be effectively used to enhance industrial monitoring capabilities, allowing small and medium-sized enterprises to transition toward smart manufacturing. These results suggest that such systems can reduce process variability, improve accuracy in decision-making, and increase process transparency with limited technological resources.

The main contributions of this research can be summarized as follows:

- Technical contribution: Development of a modular and replicable IoT–ML architecture for real-time process monitoring and predictive analysis.
- Scientific contribution: Empirical validation of supervised machine learning models applied to industrial variables, providing insight into predictive performance and methodological viability.
- Industrial contribution: A practical and low-cost framework that supports the digital transformation of Mexican SMEs, strengthening their competitiveness and sustainability.

The iterative prototype development methodology (based on Pressman) proved effective for refining both hardware and software components, enabling continuous improvements in data accuracy and user interface responsiveness. This cyclic refinement allows future scalability for broader datasets, more complex processes, and integration with industrial communication standards (Modbus, OPC-UA).

Nevertheless, future research should focus on extending the dataset, implementing cross-validation and benchmarking against additional algorithms, and incorporating cybersecurity mechanisms to ensure data integrity and secure transmission. Furthermore,

integrating edge computing architectures could minimize latency and enhance system autonomy.

Ultimately, this study contributes to making smart industry solutions more accessible in emerging economies. By combining IoT, machine learning, and human-centered design principles, the proposed approach supports the evolution toward Industry 5.0, where technology enhances not replaces human expertise, fostering sustainability, innovation, and social value within the industrial ecosystem.

References

1. Rozo-García, F. (2020). “Revisión de las tecnologías presentes en la industria 4.0,”. Revista UIS Ingenierías, vol. 19, no. 2, pp. 177–191. https://doi.org/10.18273/revuin.v19n2-2020019.
2. Méndez Vázquez, E. A. (2023). “Desarrollo del modelo de negocios para la creación de una empresa especializada en la comercialización de productos y servicios para entregar soluciones a la industria manufacturera,” Instituto Tecnológico y de Estudios Superiores de Monterrey.
3. López, D. E. Z., Leyva, S. L. & Romero-Rubio, . (2023) S. A. “La necesidad de una educación 4.0 en México para adentrarse en la industria 4.0,” Revista de Investigación en Tecnologías de la Información. vol. 11, pp. 62–75.
4. Navalpotro, F. R. D. & Pastor, J. M. A. (2024) “Tecnología global, IV revolución industrial y mutaciones sociales,” Historia Actual Online, no. 63, pp. 183–198.
5. Jiménez Castañeda, R. y García González, E. (2022). El capital social como determinante del desempeño de la industria manufacturera en México, 2004-2020. Paradigma económico. Revista de economía regional y sectorial, 14(1), 129-158.
6. INEGI, Instituto Nacional de Estadística y Geografía. (2019). “Censos económicos 2019. Micro, pequeña, mediana y gran empresa. Estratificación de los establecimientos”. Censos Económicos 2019 / Instituto Nacional de Estadística y Geografía. México INEGI, c2020.
7. Aguilar L. J., (2021). “Internet de las cosas: Un futuro hiperconectado: 5G, Inteligencia Artificial, Big Data,Cloud, Blockchain, Ciberseguridad”. Alpha Editorial.
8. Corzo, G. D. & Alvarez-Aros, E. L. (2020). “Estrategias de competitividad tecnológica en la conectividad móvil y las comunicaciones de la industria 4.0 en Latinoamérica”. Información tecnológica, vol. 31, pp. 183–192. https://doi.org/10.4067/S0718-07642020000600183.
9. Walas, F., & Redchuk, A. . (2021). “IIoT/IoT and Artificial Intelligence/Machine Learning as a Process Optimization driver under industry 4.0 model”. Journal of Computer Science and Technology, 21(2), e15. https://doi.org/10.24215/16666038.21.e15.
10. Diaz, A., & González, H. (2022). “Diseño de un sistema de comunicación iot para el monitoreo y control de variables de procesos industriales. Revista Colombiana de Tecnologías de Avanzada (RCTA)”, 1(39), 32–36.
11. Hernández-Vega, J. I., Reynoso-Guajardo, L. A., Gallardo-Morales, M. C., Macias-Arias, M. E., Hernández, A., de la Cruz, N., Soto-Soto, J. E., & Hernández-Santos, C. (2024). Plastic Injection Molding Process Analysis: Data Integration and Modeling for Improved Production Efficiency. Applied Sciences, 14(22), 10279. https://doi.org/10.3390/app142210279

12. Coello Galindo, L. R., Cruz-Segura, E., & Santillan-Ramírez, O. (2020). “Sistema de monitoreo y comunicación en tiempo real mediante una interfaz de usuario, para mejorar el proceso de estibación y el registro del producto embotellado en la industria Pascual Tizayuca”. Boletín Científico INVESTIGIUM De La Escuela Superior De Tizayuca, 5(10), 1-6. https://doi.org/10.29057/est.v5i10.3923.
13. Dutta, S., Lanvin, B., Rivera, L., & Wunsch-Vincent, S. (2023). “Global Innovation Index 2023. Innovation in the face of uncertainty. WIPO.
14. Pressman, Roger S. (2010). Ingeniería de Software. Un enfoque práctico. Editorial Mc Graw Hill, 7º ed. Madrid, España.

Observación Territorial Integral: Análisis de Pendiente, Radiación y Confort Térmico como Determinantes de la Caminabilidad en Entornos Urbanos

Cinthya L. Butron-Revilla[1,2][0000-0002-5485-9772], Liliana S. Valverde-Caballero[1][0000-0003-0308-4580], Yolanda Y. Pamo-Diaz[1][0009-0005-9728-7603]

[1] Universidad Nacional de San Agustín de Arequipa, Arequipa, Perú
[2] Universidad Pablo de Olavide, Sevilla, España
lvalverdec@unsa.edu.pe

Resumen. La evaluación de la caminabilidad urbana en ciudades intermedias latinoamericanas enfrenta limitaciones metodológicas al basarse exclusivamente en métricas morfológicas tradicionales que no consideran las condiciones ambientales de contextos urbanos, generando evaluaciones incompletas que sobreestiman la caminabilidad en territorios con topografías accidentadas y climas áridos de altura. Esta problemática resulta evidente en el Área Metropolitana de Arequipa, donde la alta dependencia del transporte motorizado y la falta de datos sistemáticos sobre condiciones ambientales urbanas limitan la implementación de políticas de movilidad sostenible adaptadas al entorno. El objetivo de esta investigación es desarrollar una metodología que incorpore métricas de confort ambiental derivadas de datos de observación territorial, pendientes y temperatura de superficie terrestre (LST), como factores de corrección al índice convencional de caminabilidad, implementando un enfoque híbrido que combina análisis morfológico tradicional con procesamiento de datos satelitales, y plataformas de Big Data para generar métricas ambientales sistemáticas. Los resultados revelan discrepancias entre la caminabilidad teórica y las condiciones ambientales reales en el centro histórico, donde valores óptimos de conectividad y densidad comercial coexisten con limitaciones térmicas asociadas al efecto isla de calor urbano y barreras topográficas moderadas. Se concluye que la integración de métricas ambientales constituye un avance metodológico esencial que democratiza el acceso a herramientas de análisis urbano avanzado y proporciona bases empíricas robustas para estrategias de mejoramiento urbano adaptadas a contextos específicos.

Palabras clave: Índice de Caminabilidad, Confort térmico, SIG, Países en Desarrollo, Centros Históricos

1 Introducción

La evaluación de la caminabilidad urbana ha evolucionado significativamente en las últimas décadas, transitando desde enfoques morfológicos tradicionales hacia metodologías integrales que consideran las condiciones ambientales como factores determinantes de la experiencia peatonal. En el contexto de ciudades intermedias latinoamericanas, particularmente aquellas ubicadas en regiones andinas, las condiciones topográficas extremas y climáticas adversas imponen limitaciones adicionales que no son capturadas adecuadamente por los índices convencionales de caminabilidad basados exclusivamente en métricas de conectividad, proximidad, mixticidad de uso de suelo y densidad comercial.. Esta problemática refleja una limitación metodológica más amplia en la evaluación de caminabilidad en contextos de ciudades en desarrollo, donde las herramientas de análisis tradicionales no consideran las particularidades climáticas y topográficas que condicionan estructuralmente los patrones de movilidad peatonal. Por ello se formula la pregunta de investigación de ¿Qué métricas deben integrarse al modelo convencional de evaluación de la caminabilidad que considere las particularidades climáticas y topográficas de ciudades en desarrollo con herramientas de análisis limitadas?

El objetivo de esta investigación es desarrollar y aplicar una metodología integral para la evaluación de caminabilidad que incorpore métricas de confort ambiental derivadas de datos de observación territorial, específicamente pendiente, radiación y temperatura de superficie terrestre (LST), como factores de corrección al índice convencional. La metodología implementa un enfoque híbrido que combina el análisis morfológico tradicional con el procesamiento de datos satelitales térmicos y topográficos de alta resolución, utilizando plataformas de Big Data como Landsat 8 y modelos digitales de elevación para generar métricas ambientales sistemáticas a escala metropolitana usados por su carácter abierto, gratuito y replicable especialmente en ciudades donde la información geoespacial es escasa o de difícil acceso por barreras administrativas, para estas metricas se toman como caso de estudio el area metropolitana y el centro histórico de la ciudad de Arequipa. Los resultados revelan discrepancias significativas entre la caminabilidad teórica calculada mediante métricas convencionales y las condiciones ambientales reales, evidenciando que el centro histórico, pese a presentar valores óptimos de conectividad y densidad comercial, enfrenta limitaciones ambientales asociadas a temperaturas superficiales elevadas (efecto isla de calor urbano) y pendientes moderadas que reducen la caminabilidad efectiva. La principal conclusión establece que la integración de métricas de confort ambiental constituye un avance metodológico fundamental que permite una evaluación más precisa de las condiciones reales que enfrentan los peatones, proporcionando una base empírica robusta para el desarrollo de estrategias de mejoramiento urbano adaptadas a las particularidades climáticas y topográficas y democratizando el acceso a herramientas de análisis urbano avanzado mediante el aprovechamiento de tecnologías de observación territorial de acceso abierto.

2 Marco Conceptual

2.1 Índice de Caminabilidad y Métricas de Confort Ambiental

La evolución conceptual del índice de caminabilidad ha experimentado una transformación significativa en las últimas décadas, incorporando variables ambientales que reflejan la experiencia real del peatón en el espacio urbano [1]. Este avance responde a la necesidad de superar las limitaciones de enfoques tradicionales centrados exclusivamente en aspectos morfológicos y funcionales del entorno construido, sin considerar las condiciones microclimáticas que influyen directamente en la decisión de caminar [2]. La integración de métricas ambientales representa un progreso hacia modelos más comprensivos que reconocen la caminabilidad como un fenómeno multidimensional, influenciado por factores físicos, sociales y ambientales. Sin embargo, la aplicación de estos índices en ciudades en desarrollo enfrenta desafíos relacionados con la disponibilidad de datos y la adaptación de metodologías diseñadas para contextos del norte global [3].

En este sentido, la incorporación de tecnologías de observación territorial y sensores remotos ofrece oportunidades para superar dichas limitaciones, permitiendo la evaluación sistemática de variables ambientales a escalas urbanas amplias [4]. Esta aproximación tecnológica facilita la generación de métricas de confort ambiental que complementan los indicadores tradicionales, proporcionando una base empírica sólida para la planificación de infraestructuras peatonales sostenibles. Estudios específicos en ciudades del Consejo de Cooperación del Golfo (GCC), como Dubái, Doha, Mascate y Riad, mediante un método Delphi modificado, identificaron 69 indicadores de caminabilidad, destacando la pendiente y el confort como factores significativos bajo climas cálidos y áridos [5]. Estos hallazgos fundamentan la presente propuesta metodológica que integra pendiente y variables de confort ambiental en la evaluación de la caminabilidad.

2.2 Confort ambiental urbano y sus métricas

El confort ambiental urbano se define a partir de la experiencia de las personas con su entorno [6], e incluye variables relevantes para los peatones como temperatura del aire, humedad relativa, presión atmosférica, radiación solar, iluminación, velocidad y dirección del viento, concentración de CO2 [7], ausencia de ruido, asoleamiento [8], pendiente, condiciones de la superficie y obstáculos [9]. Sin embargo, en ciudades en desarrollo, la evaluación de muchas de estas variables se ve limitada por la disponibilidad de datos y herramientas. No obstante, la creciente accesibilidad a bases de datos de observación territorial como Copernicus, LandSAT y Google Earth Engine, que adoptan políticas de datos abiertos, ha impulsado un cambio disruptivo en el uso de macrodatos (Big Data) para el análisis ambiental urbano [10].

Entre los factores que afectan la caminabilidad, la pendiente es crucial, ya que su aumento reduce la velocidad de desplazamiento peatonal y la elección de rutas, especialmente para usuarios vulnerables como adultos mayores y niños [11,12]. Estudios indican que un incremento del 1% en la pendiente puede disminuir en aproximadamente un 10% el atractivo de caminar una ruta [13,14]. Otro factor relevante es la radiación

solar, que influye en la preferencia de los peatones por caminar bajo el sol en lugar de buscar sombra, afectando su salud. La temperatura de superficie terrestre (LST), influenciada por la pendiente y la radiación solar, representa el intercambio de energía entre la radiación solar, la atmósfera y la superficie terrestre, y su variación espacial está afectada por la temperatura del aire, la pendiente y el índice de vegetación normalizado (NDVI) [15]. Además, métricas de textura derivadas de la LST describen mejor la heterogeneidad ambiental que las métricas tradicionales de paisaje [16].

2.3 Impacto de la topografía en patrones de movilidad peatonal

La topografía urbana constituye uno de los determinantes más significativos en la configuración de patrones de movilidad peatonal, ejerciendo una influencia directa sobre la accesibilidad y la percepción de esfuerzo asociada al desplazamiento a pie [17]. Las variaciones altimétricas del terreno no solo afectan la velocidad de desplazamiento y el gasto energético de los peatones, sino que también influyen en la selección de rutas y la frecuencia de uso de diferentes corredores urbanos [18]. Esta relación se intensifica en contextos de centros históricos, donde las características topográficas originales del asentamiento se mantienen inalteradas, creando condiciones particulares que requieren enfoques metodológicos específicos para su evaluación. La cuantificación del impacto de las pendientes en la caminabilidad ha evolucionado desde aproximaciones empíricas basadas en observación directa hacia metodologías que integran tecnologías de información geográfica y modelos predictivos [19].

Los estudios contemporáneos demuestran que la percepción de la pendiente varía significativamente entre diferentes grupos demográficos, siendo particularmente relevante para poblaciones vulnerables como adultos mayores, personas con discapacidad y niños [20]. Esta variabilidad en la percepción sugiere la necesidad de desarrollar métricas diferenciadas que consideren las capacidades físicas específicas de distintos usuarios del espacio peatonal. La integración de datos topográficos de alta resolución obtenidos mediante tecnologías LiDAR y modelos digitales de elevación derivados de sensores remotos ha revolucionado la capacidad de análisis de la influencia topográfica en la movilidad urbana [21]. Estas tecnologías permiten la generación de métricas precisas de pendiente a escalas de detalle que facilitan la identificación de barreras topográficas específicas y la evaluación de alternativas de diseño urbano orientadas a mejorar la accesibilidad peatonal [22]. La disponibilidad creciente de estos datos a través de plataformas de acceso abierto democratiza el acceso a herramientas de análisis avanzado, especialmente relevante para ciudades en desarrollo con recursos técnicos limitados.

2.4 Temperatura de superficie terrestre y confort térmico urbano

La temperatura de superficie terrestre (LST) emerge como una métrica fundamental para la evaluación del confort térmico en espacios urbanos, proporcionando información crítica sobre las condiciones microclimáticas que influyen en la experiencia peatonal [23]. La LST refleja la interacción compleja entre la radiación solar incidente, las propiedades térmicas de los materiales urbanos y las condiciones atmosféricas locales, generando patrones espaciales de temperatura que varían significativamente a escalas

intraurbanas [24] Esta variabilidad espacial de la temperatura superficial constituye un factor determinante en la percepción de confort térmico de los peatones y, consecuentemente, en sus decisiones de movilidad.

Los avances en tecnologías de sensores remotos térmicos han facilitado el monitoreo sistemático de la LST a escalas urbanas, proporcionando datos temporalmente consistentes que permiten el análisis de patrones estacionales y diurnos de variación térmica [25]. La integración de datos LST derivados de sensores como MODIS, Landsat y Sentinel-3 con información de uso de suelo y características morfológicas urbanas permite la identificación de islas de calor urbano y la evaluación de estrategias de mitigación térmica [26]. Esta capacidad de análisis resulta particularmente relevante para la planificación de infraestructuras peatonales que consideren el confort térmico como criterio de diseño.La relación entre LST y caminabilidad se manifiesta a través de múltiples mecanismos que incluyen la percepción directa de calor por parte de los peatones, la influencia en la temperatura del aire a nivel peatonal y los efectos sobre la demanda energética del desplazamiento a pie [27]. Los estudios empíricos demuestran que temperaturas superficiales elevadas pueden reducir significativamente la disposición a caminar, especialmente durante períodos de alta radiación solar, generando patrones temporales de uso del espacio peatonal que reflejan estrategias de adaptación térmica [28]. Esta evidencia subraya la importancia de incorporar métricas térmicas en los índices de caminabilidad, particularmente en contextos climáticos caracterizados por altas temperaturas ambientales.

2.5 Integración de datos de observación territorial en estudios urbanos

La revolución en el acceso a datos de observación territorial ha transformado las metodologías de análisis urbano, permitiendo la evaluación sistemática de variables ambientales a escalas espaciales y temporales amplias [29]. Plataformas de datos abiertos como Google Earth Engine, Copernicus y Landsat han democratizado el acceso a información satelital de alta calidad, superando barreras técnicas y económicas que limitaban la investigación en contextos con recursos restringidos [30]. Esta disponibilidad facilita enfoques metodológicos que integran múltiples fuentes para generar métricas comprensivas de calidad ambiental urbana, mejorando la precisión en la caracterización del entorno urbano.

La integración de datos multiespectrales, térmicos y topográficos derivados de sensores remotos permite crear índices compuestos que capturan la complejidad ambiental urbana con mayor precisión que métodos tradicionales basados en mediciones puntuales [31]. Los algoritmos de Big Data posibilitan el análisis de series temporales extensas, revelando patrones estacionales y tendencias a largo plazo, información crucial para la planificación adaptativa [32]. Sin embargo, desafíos metodológicos como la validación de productos satelitales, la armonización de resoluciones espaciales y temporales, y la interpretación en contextos urbanos heterogéneos persisten [33]. Estrategias como la validación con datos de campo y la fusión de datos multiresolución son prometedoras para superar estas limitaciones, consolidando una frontera activa de investigación con impacto en la planificación urbana sostenible [34].

3 Metodología

3.1 Caso de Estudio

El Área Metropolitana de Arequipa, segunda ciudad más poblada del Perú con aproximadamente 1.4 millones de habitantes según el último censo nacional [35] presenta características urbanas particulares que la convierten en un caso de estudio relevante para el análisis de caminabilidad en ciudades intermedias latinoamericanas. Su configuración urbana se caracteriza por un crecimiento expansivo no planificado que ha generado una estructura policéntrica con alta dependencia del transporte motorizado, donde el 46% de los viajes diarios se realizan en transporte público informal y solo el 17% corresponde a desplazamientos peatonales [36]. Esta dinámica de movilidad se ve agravada por la topografía accidentada de la ciudad, ubicada a 2,335 metros sobre el nivel del mar en un valle rodeado por volcanes, lo que genera pendientes pronunciadas que constituyen barreras físicas significativas para la movilidad peatonal, especialmente en los sectores periféricos donde se concentra la población de menores recursos [37]. La falta de datos sistemáticos sobre condiciones ambientales urbanas y patrones de movilidad peatonal en Arequipa refleja una problemática común en ciudades intermedias del sur global, donde la limitada capacidad técnica e institucional para la recolección y análisis de información urbana dificulta la implementación de políticas de movilidad sostenible basadas en evidencia empírica [38].

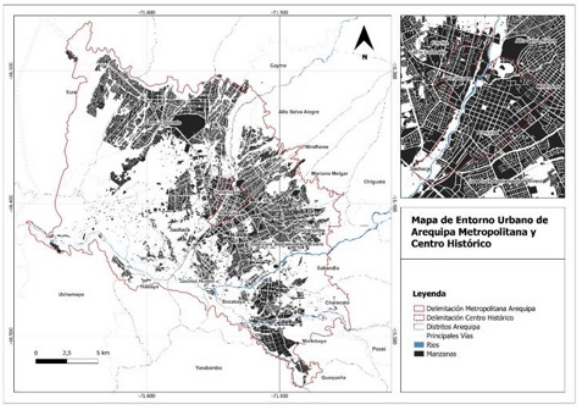


Figura 1: Entorno urbano de Arequipa metropolitana. y su centro histórico

3.2 Materiales y Herramientas

La metodología implementada en esta investigación adopta un enfoque híbrido que integra dos aproximaciones complementarias para la evaluación de la caminabilidad en el centro histórico de Arequipa: la metodología convencional basada en métricas morfológicas urbanas y la metodología de observación territorial integral que incorpora variables ambientales derivadas de sensores remotos. Esta estrategia metodológica dual permite la comparación sistemática entre enfoques tradicionales y aproximaciones

innovadoras que consideran el confort ambiental como factor determinante de la caminabilidad urbana.

3.21.1. Metodología Convencional: La primera fase metodológica implementa el índice de caminabilidad convencional desarrollado para ciudades patrimonio mundial, basado en cuatro métricas fundamentales que caracterizan la estructura urbana y funcional del área de estudio: conectividad, proximidad, mixticidad y densidad comercial.

$$Conectividad = \frac{1}{(SR^d)^2}\sum_{i=1}^{n}\cdot\left[\frac{3}{\pi}\cdot nL_i\left(1-\left(\frac{dist_i}{SR}\right)^2\right)^2\right]\cdot X_{norm} \quad \text{(Eq. 1)}$$

$$Proximidad = \frac{1}{(SR^w)^2}\sum_{i=1}^{n}\cdot\left[\frac{3}{\pi}\cdot\left(1-\left(\frac{dist_i}{SR^w}\right)^2\right)^2\right]\cdot X_{norm} \quad \text{(Eq. 2)}$$

$$Índice\ de\ Entropia\ para\ Mixticidad = \frac{-\sum_j P_j ln(P_j)}{\ln N} \quad \text{(Eq. 3)}$$

$$Densidad\ Comercial = \frac{\sum.(STl\ W)}{Area\ of\ circle} \quad \text{(Eq. 4)}$$

Estas métricas permiten cuantificar la accesibilidad y diversidad funcional del tejido urbano, estableciendo la base analítica para la siguiente fase.

3.2.2 Metodología de Observación Integral: La segunda fase metodológica amplia el análisis tradicional incorporando métricas de confort ambiental derivadas de datos de observación territorial, específicamente pendientes y temperatura de superficie terrestre (LST). Estas variables se integran como factores de corrección al índice convencional, con el fin de reflejar la influencia de las condiciones físicas y ambientales sobre la caminabilidad real. Las pendientes se extraen de modelos digitales de elevación de alta resolución procesados mediante algoritmos de análisis espacial que calculan la inclinación del terreno en grados, considerando que incrementos del 1% en la pendiente reducen aproximadamente 10% el atractivo peatonal según la literatura especializada. La LST se obtiene mediante el procesamiento de imágenes térmicas satelitales Landsat 8 y Sentinel-3, aplicando algoritmos de corrección atmosférica y conversión radiométrica que generan mapas de temperatura superficial con resolución espacial de 30 metros.

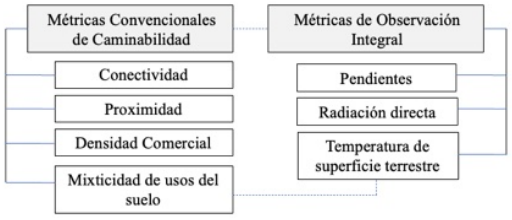


Figura 2: Cuadro Metodológico

4 Resultados

4.1. Métricas Convencionales de Caminabilidad

El mapa de conectividad metropolitana de Arequipa(Figura 3a) muestra una distribución espacial heterogénea que refleja las características morfológicas y topográficas del área. Se observa una concentración significativa de niveles altos de conectividad en el centro histórico y áreas consolidadas, contrastando con niveles bajos en las zonas periféricas y de expansión urbana reciente. Esta configuración evidencia un patrón de crecimiento radiocéntrico disperso, donde las áreas periféricas presentan deficiencias en la articulación vial que limitan la accesibilidad peatonal. Además, se identifican corredores de conectividad intermedia a lo largo de las principales vías metropolitanas, lo que sugiere oportunidades para intervenir y mejorar la integración entre el centro histórico y las periferias.

El mapa de proximidad metropolitana de Arequipa(Figura 3b) evidencia una distribución desigual de destinos y servicios urbanos, con alta concentración en el centro histórico y áreas comerciales consolidadas, y valores bajos en las zonas periféricas. Esta configuración refleja el modelo urbano monocéntrico, donde el núcleo histórico actúa como principal concentrador de actividades comerciales, institucionales y de servicios, favoreciendo la accesibilidad peatonal dentro de su radio de influencia. En contraste, las áreas de expansión reciente presentan deficiencias significativas en equipamientos y servicios de proximidad. Asimismo, se identifican corredores de proximidad intermedia a lo largo de arterias comerciales y centros de barrio consolidados, indicando subcentralidades emergentes que, aunque limitadas en diversidad funcional, ofrecen accesibilidad básica a las poblaciones periféricas.

El mapa de densidad comercial metropolitana de Arequipa (Figura 3c) evidencia una distribución espacial altamente concentrada, reflejando la estructura económica monocéntrica típica de ciudades intermedias latinoamericanas. El núcleo de máxima intensidad comercial se localiza en el centro histórico, con una extensión radial limitada hacia las periferias, donde predominan valores bajos de densidad comercial. La métrica aplicada revela una concentración extrema en el área patrimonial, con densidades significativamente superiores al resto del territorio, generando un gradiente de intensidad que decrece exponencialmente con la distancia al centro. Esta configuración refleja patrones coloniales de organización urbana, donde la Plaza de Armas y sus alrededores mantienen su función como principal distrito comercial, favoreciendo flujos peatonales por la alta concentración de destinos dentro del radio caminable. Además, se identifican núcleos de densidad comercial intermedia en centralidades barriales como Yanahuara y Cayma, que actúan como subcentralidades locales con intensidades menores, configurando un sistema urbano jerárquico que influye en los patrones de movilidad y la distribución de oportunidades de caminabilidad.

El mapa de mixticidad de usos de suelo metropolitana de Arequipa (Figura 3d) evidencia una distribución espacial polarizada que refleja la segregación funcional propia del crecimiento urbano no planificado. El índice de entropía muestra alta diversidad

funcional concentrada en el centro histórico y corredores comerciales consolidados, contrastando con extensas áreas periféricas dominadas por usos residenciales monofuncionales. El núcleo histórico presenta una trama urbana compacta que integra funciones residenciales, comerciales, institucionales y de servicios, favoreciendo la caminabilidad. En cambio, las zonas periféricas, especialmente en distritos de reciente expansión, exhiben especialización funcional que limita el desplazamiento peatonal multipropósito. Se identifican corredores de mixticidad intermedia en arterias urbanas y centros de barrio, indicando procesos incipientes de diversificación funcional, aunque con menor intensidad que el centro.

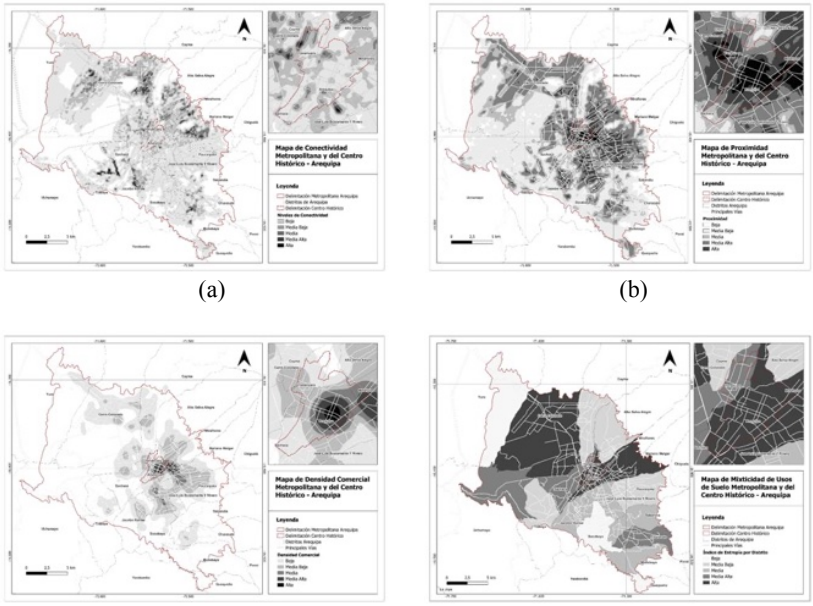


Figura 3:Mapas de las métricas convencionales (a. conectividad, b. proximidad, c. densidad comercial, d. mixticidad de usos del suelo) del área metropolitana y del centro histórico de Arequipa

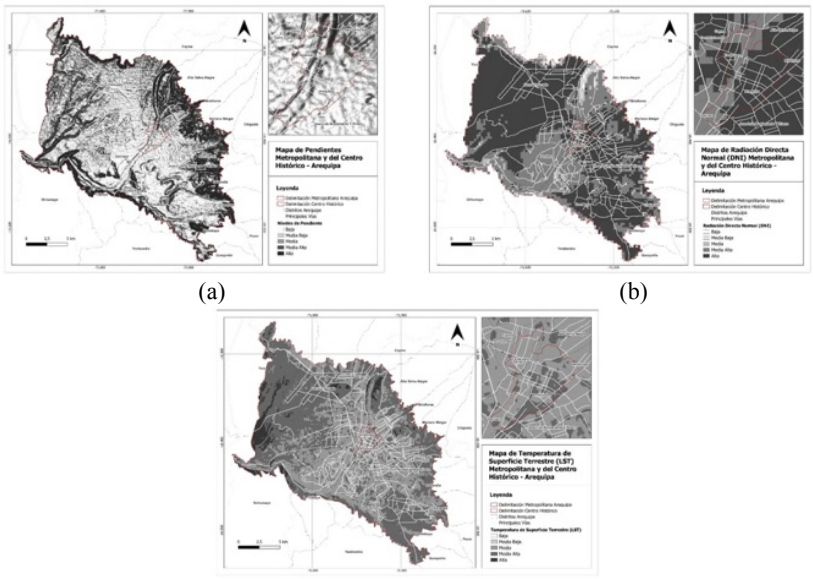
4.2 Métricas de Observación Integral

El mapa de pendientes metropolitanas de Arequipa (Figura 4a) revela una topografía volcánica compleja y heterogénea, con pendientes altas concentradas en las laderas de conos volcánicos y quebradas, y pendientes bajas en el fondo del valle donde se ubica el núcleo histórico. El análisis espacial de modelos digitales de elevación de alta resolución muestra que el centro histórico se encuentra en una zona relativamente plana, favoreciendo la movilidad peatonal, mientras que las áreas de expansión urbana en las laderas presentan inclinaciones significativas que, según la literatura, disminuyen el

atractivo peatonal en aproximadamente un 10% por cada 1% de aumento en la pendiente. Además, se identifican corredores topográficos de pendiente intermedia que coinciden con las principales vías metropolitanas, evidenciando la adaptación de la infraestructura vial a la geomorfología natural. Esta configuración topográfica impone barreras físicas que limitan la caminabilidad, especialmente en las zonas periféricas de mayor crecimiento demográfico, donde las pendientes pronunciadas y la insuficiente infraestructura peatonal generan entornos urbanos con baja accesibilidad, afectando principalmente a poblaciones de menores recursos.

El mapa de radiación directa normal (DNI) en la metrópoli de Arequipa (Figura 4b) muestra una distribución espacial homogénea con valores altos y muy altos, reflejando las condiciones áridas y la elevada altitud (2,335 msnm) que favorecen la incidencia solar directa con mínima atenuación atmosférica. El procesamiento de datos satelitales térmicos revela variaciones sutiles, donde las laderas volcánicas presentan los valores más elevados debido a la menor densidad atmosférica, mientras que el fondo del valle urbano exhibe valores ligeramente menores por sombreado topográfico y aerosoles urbanos. Esta radiación intensa genera estrés térmico significativo para la movilidad peatonal, especialmente entre las 10:00 y 16:00 horas, cuando la alta radiación y baja humedad relativa del clima desértico de altura crean un ambiente térmico adverso que reduce la disposición a caminar. La uniformidad de la radiación entre el centro histórico y las áreas periféricas indica la necesidad de estrategias de mitigación térmica generalizadas, como sombra e arborización urbana, para mejorar el confort térmico peatonal y promover políticas de movilidad sostenible adaptadas a las condiciones extremas del altiplano andino.

El mapa de temperatura de superficie terrestre (LST) de Arequipa (Figura 4c) revela un patrón térmico espacialmente heterogéneo, resultado de la interacción entre geomorfología, cobertura del suelo y procesos radiativos. Se identifican temperaturas altas concentradas en áreas urbanas consolidadas y superficies sin vegetación, contrastando con valores bajos en zonas con mayor cobertura vegetal y cuerpos de agua. El análisis de imágenes térmicas Landsat 8, corregidas atmosféricamente, muestra que el centro histórico presenta un efecto isla de calor urbano, con temperaturas superficiales 5-8 °C superiores a las áreas periféricas, debido a la alta densidad edificatoria y materiales térmicamente activos como piedra volcánica y concreto. Esta configuración térmica evidencia la influencia directa del uso del suelo en las microclimas urbanos, donde superficies pavimentadas actúan como acumuladores térmicos, elevando la temperatura ambiente peatonal, especialmente en horas de máxima radiación solar. Además, se observan gradientes térmicos entre el núcleo urbano y las zonas periféricas, donde la menor densidad y espacios abiertos favorecen condiciones térmicas más confortables para la movilidad peatonal. Esta variabilidad espacial de la LST es clave para evaluar el confort térmico peatonal, dado que las temperaturas superficiales elevadas se asocian con aumento de temperatura del aire y estrés térmico, factores esenciales para comprender la influencia ambiental en la caminabilidad y calidad de vida en climas áridos de altura.



(a) (b)

(c)

Figura 4: Mapas de métricas de observación integral (a. pendientes, radiación directa, c. temperatura de superficie terrestre) del área metropolitana y del centro histórico de Arequipa

5 Conclusiones

Los resultados de esta investigación confirman la hipótesis sobre la necesidad de integrar métricas de confort ambiental en la evaluación de la caminabilidad urbana, evidenciando que los enfoques convencionales basados únicamente en variables morfológicas ofrecen una visión incompleta de las condiciones reales que enfrentan los peatones en ciudades latinoamericanas. La metodología integral de observación territorial desarrollada revela discrepancias significativas entre la caminabilidad teórica, calculada con métricas tradicionales de conectividad, proximidad, mixticidad y densidad comercial, y las condiciones ambientales reales, caracterizadas por pendientes pronunciadas y temperaturas superficiales elevadas que limitan la movilidad peatonal. Esta divergencia es especialmente notable en el centro histórico de Arequipa, donde, a pesar de la alta concentración comercial y conectividad vial, el estrés térmico y las pendientes moderadas reducen la caminabilidad efectiva.

La incorporación de datos de sensores remotos en el análisis de caminabilidad representa un avance metodológico crucial. La temperatura de superficie terrestre (LST) se identifica como un indicador clave para evaluar el confort térmico peatonal, mostrando correlaciones espaciales con patrones de uso de suelo y densidad urbana que afectan la experiencia térmica. La aplicación de algoritmos de Big Data facilita la generación de métricas ambientales de alta resolución que capturan la variabilidad microclimática intraurbana, orientando la identificación de áreas prioritarias para intervenciones. No obstante, persisten desafíos metodológicos relacionados con la

validación de datos satelitales en contextos urbanos complejos y la necesidad de modelos integradores que consideren interacciones no lineales entre variables ambientales y comportamiento peatonal, lo que subraya la importancia de complementar con mediciones de campo y estudios perceptivos.

Finalmente, esta investigación contribuye al desarrollo de metodologías urbanas que aprovechan las oportunidades emergentes del acceso a datos de observación territorial, democratizando herramientas avanzadas de análisis para ciudades en desarrollo con recursos técnicos limitados. La propuesta metodológica ofrece una base empírica robusta para la planificación urbana sostenible, orientada a mejorar la calidad de vida y la movilidad peatonal en contextos climáticos y topográficos complejos, y establece un marco para futuras investigaciones que integren datos cuantitativos y cualitativos en la evaluación integral de la caminabilidad urbana.

6 Discusión

La aplicación de tecnologías de sensores remotos y plataformas de Big Data representa un avance significativo para superar las limitaciones tradicionales de disponibilidad de datos en ciudades en desarrollo, permitiendo la implementación de metodologías avanzadas de análisis urbano que antes requerían recursos técnicos y económicos elevados [29,30]. Los resultados confirman la utilidad de la temperatura de superficie terrestre (LST) como un indicador proxy robusto del confort térmico peatonal, mostrando correlaciones espaciales consistentes con patrones de uso de suelo y densidad urbana que facilitan la evaluación sistemática de condiciones microclimáticas a escala metropolitana [15,31]. No obstante, persisten desafíos metodológicos relacionados con la validación de productos satelitales en contextos urbanos heterogéneos y la necesidad de modelos integradores que consideren interacciones complejas entre variables ambientales, demográficas y comportamentales, lo que resalta la importancia de complementar la observación territorial con estudios de campo y análisis perceptivos [33,34].

Las implicaciones de esta investigación trascienden el caso de Arequipa, aportando a la construcción de marcos metodológicos replicables en ciudades intermedias latinoamericanas que enfrentan desafíos similares de crecimiento urbano no planificado, condiciones climáticas extremas y limitaciones topográficas [3,4]. La evidencia empírica subraya la necesidad de adaptar los enfoques de evaluación de caminabilidad a las particularidades ambientales locales, desarrollando estándares de confort específicos para contextos de altura y clima árido que consideren las adaptaciones fisiológicas y comportamentales de las poblaciones andinas [5,11,12]. Además, se abren oportunidades para futuras investigaciones en modelos predictivos de comportamiento peatonal que integren variables ambientales, evaluación de intervenciones de mitigación térmica y adaptación metodológica a diversos contextos climáticos y geográficos, incluyendo variables complementarias como seguridad y percepción para consolidar un modelo integral de caminabilidad [13,14,16]. Finalmente, esta investigación contribuye a democratizar el acceso a herramientas avanzadas de análisis urbano, aprovechando datos abiertos de observación territorial para fortalecer la capacidad técnica de planificación en ciudades con recursos limitados.

Agradecimientos.

Especial mención y agradecimientos a nuestra facultad y alma mater, la Facultad de Arquitectura y Urbanismo de la Universidad Nacional de San Agustín de Arequipa.

Declaración de Intereses. Los autores declaran no tener conflicto de intereses.

Referencias

[1] L. D. Frank *et al.*, "The development of a walkability index: Application to the neighborhood quality of life study", *Br J Sports Med*, vol. 44, núm. 13, pp. 924–933, 2010, doi: 10.1136/bjsm.2009.058701.

[2] J. A. Cook, M. Bose, W. E. Marshall, y D. S. Main, "How Does Design Quality Add to our Understanding of Walkable Communities?", *Landsc J*, vol. 32, núm. 2, 2013.

[3] A. Adkins, C. Makarewicz, M. Scanze, M. Ingram, y G. Luhr, "Contextualizing Walkability: Do Relationships Between Built Environments and Walking Vary by Socioeconomic Context?", *Journal of the American Planning Association*, vol. 83, núm. 3, pp. 296–314, jul. 2017, doi: 10.1080/01944363.2017.1322527.

[4] P. Gong, X. Huang, C. Huang, y M. White, "MACHINE LEARNING-BASED WALKABILITY MODELING IN URBAN LIFE CIRCLE", 2022.

[5] B. F. Alkrides, "Promoting Sustainable Urban Walkability: A Modified Delphi Study on Key Indicators for Urban Walkability in Gulf Cooperation Council Urban Streets", *Sustainability (Switzerland)*, vol. 17, núm. 3, feb. 2025, doi: 10.3390/su17031179.

[6] S. Ataee, M. Lopes, y H. Relvas, "Environmental comfort in urban spaces: A systematic literature review and a system dynamics analysis", *Urban Clim*, vol. 60, mar. 2025, doi: 10.1016/j.uclim.2025.102340.

[7] I. Pigliautile, G. Marseglia, y A. L. Pisello, "Investigation of CO2 variation and mapping through wearable sensing techniques for measuring pedestrians' exposure in urban areas", *Sustainability (Switzerland)*, vol. 12, núm. 9, may 2020, doi: 10.3390/su12093936.

[8] T. Al Shammas y F. Escobar, "Comfort and time-based walkability index design: A GIS-based proposal", *Int J Environ Res Public Health*, vol. 16, núm. 16, ago. 2019, doi: 10.3390/ijerph16162850.

[9] M. J. Campbell, P. E. Dennison, y M. P. Thompson, "Predicting the variability in pedestrian travel rates and times using crowdsourced GPS data", *Comput Environ Urban Syst*, vol. 97, oct. 2022, doi: 10.1016/j.compenvurbsys.2022.101866.

[10] J. Wagemann, S. Siemen, B. Seeger, y J. Bendix, "Users of open Big Earth data – An analysis of the current state", *Comput Geosci*, vol. 157, dic. 2021, doi: 10.1016/j.cageo.2021.104916.

[11] M. J. Campbell, P. E. Dennison, y B. W. Butler, "A LiDAR-based analysis of the effects of slope, vegetation density, and ground surface roughness on travel rates for wildland firefighter escape route mapping", *Int J Wildland Fire*, vol. 26, núm. 10, pp. 884–895, 2017, doi: 10.1071/WF17031.

[12] N. Basu, O. Oviedo-Trespalacios, M. King, M. Kamruzzaman, y M. M. Haque, "What do pedestrians consider when choosing a route? The role of safety, security, and attractiveness perceptions and the built environment during day and night walking", *Cities*, vol. 143, dic. 2023, doi: 10.1016/j.cities.2023.104551.

[13] M. Meeder, T. Aebi, y U. Weidmann, "The influence of slope on walking activity and the pedestrian modal share", en *Transportation Research Procedia*, Elsevier B.V., 2017, pp. 141–147. doi: 10.1016/j.trpro.2017.12.095.

[14] M. Tomasi, M. Nikolopoulou, R. Giridharan, M. Löve, y C. Ratti, "Dynamic analysis of a pedestrian network: The impact of solar radiation exposure on diverse user experiences", *Sustain Cities Soc*, vol. 112, oct. 2024, doi: 10.1016/j.scs.2024.105631.

[15] X. Zhang, A. Kasimu, H. Liang, B. Wei, y Y. Aizizi, "Spatial and Temporal Variation of Land Surface Temperature and Its Spatially Heterogeneous Response in the Urban Agglomeration on the Northern Slopes of the Tianshan Mountains, Northwest China", *Int J Environ Res Public Health*, vol. 19, núm. 20, oct. 2022, doi: 10.3390/ijerph192013067.

[16] E. Rahimi, S. Barghjelveh, y P. Dong, "Quantifying how urban landscape heterogeneity affects land surface temperature at multiple scales", *J Ecol Environ*, vol. 45, núm. 1, dic. 2021, doi: 10.1186/s41610-021-00203-z.

[17] M. J. Koohsari *et al.*, "Traditional and novel walkable built environment metrics and social capital", *Landsc Urban Plan*, vol. 214, oct. 2021, doi: 10.1016/j.landurbplan.2021.104184.

[18] Z. Guo y B. P. Y. Loo, "Pedestrian environment and route choice: Evidence from New York City and Hong Kong", *J Transp Geogr*, vol. 28, pp. 124–136, abr. 2013, doi: 10.1016/j.jtrangeo.2012.11.013.

[19] E. Leslie, N. Coffee, L. Frank, N. Owen, A. Bauman, y G. Hugo, "Walkability of local communities : Using geographic information systems to objectively assess relevant environmental attributes", vol. 13, pp. 111–122, 2007, doi: 10.1016/j.healthplace.2005.11.001.

[20] D. Rosenberg *et al.*, "Neighborhood Environment Walkability Scale for Youth (NEWS-Y): Reliability and relationship with physical activity", *Prev Med (Baltim)*, vol. 49, núm. 2–3, pp. 213–218, ago. 2009, doi: 10.1016/j.ypmed.2009.07.011.

[21] S. Hajna, K. Dasgupta, M. Halparin, y N. A. Ross, "Neighborhood walkability: Field validation of geographic information system measures", *Am J Prev Med*, vol. 44, núm. 6, pp. e55–e59, 2013, doi: 10.1016/j.amepre.2013.01.033.

[22] H. Y. Chan, D. Cheng, y A. Chen, "Routes with roots: Pedestrian route choices and sense of place of an urban university community", *J Transp Geogr*, vol. 118, jun. 2024, doi: 10.1016/j.jtrangeo.2024.103943.

[23] J. A. Voogt y T. R. Oke, "Thermal remote sensing of urban climates", *Remote Sens Environ*, vol. 86, núm. 3, pp. 370–384, ago. 2003, doi: 10.1016/S0034-4257(03)00079-8.

[24] M. Weng *et al.*, "The 15-minute walkable neighborhoods: Measurement, social inequalities and implications for building healthy communities in urban China", *J Transp Health*, vol. 13, pp. 259–273, jun. 2019, doi: 10.1016/j.jth.2019.05.005.

[25] N. Sobrino y R. Arce, "Understanding per-trip commuting CO2 emissions: A case study of the Technical University of Madrid", *Transp Res D Transp Environ*, vol. 96, jul. 2021, doi: 10.1016/j.trd.2021.102895.

[26] M. Santamouris, C. Cartalis, A. Synnefa, y D. Kolokotsa, "On the impact of urban heat island and global warming on the power demand and electricity consumption of

buildings - A review", *Energy Build*, vol. 98, pp. 119–124, jun. 2015, doi: 10.1016/j.enbuild.2014.09.052.

[27] N. Nasrollahi, A. Ghosouri, J. Khodakarami, y M. Taleghani, "Heat-mitigation strategies to improve pedestrian thermal comfort in urban environments: A review", el 1 de diciembre de 2020, *MDPI*. doi: 10.3390/su122310000.

[28] S. Thorsson, T. Honjo, F. Lindberg, I. Eliasson, y E. M. Lim, "Thermal comfort and outdoor activity in Japanese urban public places", *Environ Behav*, vol. 39, núm. 5, pp. 660–684, sep. 2007, doi: 10.1177/0013916506294937.

[29] P. Gong *et al.*, "Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM+ data", *Int J Remote Sens*, vol. 34, núm. 7, pp. 2607–2654, 2013, doi: 10.1080/01431161.2012.748992.

[30] N. Gorelick, M. Hancher, M. Dixon, S. Ilyushchenko, D. Thau, y R. Moore, "Google Earth Engine: Planetary-scale geospatial analysis for everyone", *Remote Sens Environ*, vol. 202, pp. 18–27, dic. 2017, doi: 10.1016/j.rse.2017.06.031.

[31] M. A. Wulder *et al.*, "Fifty years of Landsat science and impacts", el 1 de octubre de 2022, *Elsevier Inc.* doi: 10.1016/j.rse.2022.113195.

[32] Z. Zhu *et al.*, "Benefits of the free and open Landsat data policy", *Remote Sens Environ*, vol. 224, pp. 382–385, abr. 2019, doi: 10.1016/j.rse.2019.02.016.

[33] D. P. Roy *et al.*, "Landsat-8: Science and product vision for terrestrial global change research", *Remote Sens Environ*, vol. 145, pp. 154–172, abr. 2014, doi: 10.1016/j.rse.2014.02.001.

[34] C. Atzberger, "Advances in remote sensing of agriculture: Context description, existing operational monitoring systems and major information needs", febrero de 2013. doi: 10.3390/rs5020949.

[35] Instituto Nacional de Estadística e Informática del Perú, "Sistema de Consulta de Base de Datos REDATAM". Consultado: el 4 de diciembre de 2024. [En línea]. Disponible en: https://censos2017.inei.gob.pe/redatam/

[36] Instituto Municipal de Planeamiento IMPLA, "Plan de Desarrollo Metropolitano de Arequipa 2016-2025". Consultado: el 27 de marzo de 2024. [En línea]. Disponible en: https://impla.gob.pe/publicaciones/pdm-2016-2025/

[37] M. Dammert-Guardia, R. Cavagnoud, M. F. Barrios, y A. Calderón Rojas-Sandoval, "Woman traders and cross-border circuits: articulations between care and mobility on the borders of Peru, Bolivia and Chile", *Estudios Fronterizos*, vol. 25, 2024, doi: 10.21670/ref.2421157.

[38] L. S. Valverde-Caballero, L. M. Mendoza-Salazar, C. L. Butron-Revilla, E. Suarez-Lopez, y J. S. Aguilar-Ruiz, "Walkability index for world heritage cities in developing countries", *Environ Plan B Urban Anal City Sci*, vol. 52, núm. 1, pp. 76–96, ene. 2025, doi: 10.1177/23998083241250265.

Innovative technologies for governance and citizenship

Data Literacy for Smart Governance: An Educational-Analytical Framework for Ibero-American Cities

Mailyn Moreno-Espino[1,2][0000−0002−7613−3382], Jorge Alberto Pachecho-Sernard[3][1111−2222−3333−4444], Yahima Hadfeg-Fernández[4][0009−0001−1992−1040], Gerardo Acevedo-Sánchez[3][0000−0003−3076−8590], and Nayma Cepero-Pérez[5][0000−0003−3808−8135] Cornelio Yáñez-Márquez[3][0000−0002−6250−4728]

[1] Facultad de Informática, Universidad Complutense de Madrid mailymor@ucm.es
[2] Instituto de Tecnología del Conocimiento, Universidad Complutense de Madrid Laboratorio de Cómputo Inteligente, Centro de Investigación en Computación, Instituto Politécnico Nacional {jpachecos2024,gacevedos2024,cyanez}@cic.ipn.mx
[3] Facultad de Economía y Administración, Universidad Católica del Norte yahima.hadfeg01@ucn.cl {abc,lncs}@uni-heidelberg.de
[4] Facultad de Ingeniería Informática, Universidad Tecnológica de La Habana "José Antonio Echeverría" vric@tesla.cujae.edu.cu

Abstract. The development of smart cities necessitates two fundamental components: robust digital infrastructure and the cultivation of data literacy as a civic and governmental competence. These competencies are imperative for inclusive policy design, evidence-based decision-making, and socially sustainable governance. We hereby present an educational-analytical framework tailored to Ibero-American cities that integrates the following: The proposed initiative entails the following three components: The development of learning pathways tailored to meet the needs of citizens, public servants, and policymakers. The creation of a reproducible urban data analytics toolkit that addresses critical issues such as data quality, traceability, and impact metrics. The implementation of a digital participation and e-government model designed to foster informed citizenship, transparency, and accountability. The framework is illustrated through urban use cases in mobility, energy management, and land-use planning. We hereby propose a set of indicators to assess informed participation, equity of access, data quality and usage, administrative response times, and policy outcomes. In doing so, we also consider ethical dimensions such as privacy, bias, and fairness. This contribution strengthens the link between data literacy, smart governance, sustainable urban development, and inclusive citizenship, offering a pathway for empowered and resilient smart cities across Ibero-America.

Keywords: Data literacy · Smart governance · Sustainable cities · GIS · Urban informatics.

1 Introduction

In the contemporary era, the digital revolution has transformed nearly every aspect of urban life. The generation and utilization of data are now fundamental to the development of smart cities, encompassing smartphones, social media, and the integration of intelligent systems in domains such as homes, transportation networks, and public administration [1, 19]. The proliferation of urban data, generated by sensors, connected devices, and citizen activities, is increasing exponentially. These data influence critical decision-making processes in mobility planning, energy management, environmental monitoring, and citizen engagement. Consequently, they shape the governance and daily experience of cities.

Notwithstanding the pervasive presence of smart technologies, a conspicuous lacuna persists in educational and training endeavors concerning the effective management, interpretation, and application of data for evidence-based urban governance [2, 17]. The dearth of data literacy among the general public and public servants imposes limitations on the capacity of cities to reap the full benefits of urban analytics. This includes the potential for enhanced service delivery, infrastructure optimization, and participatory decision-making. Furthermore, a paucity of awareness regarding data collection and usage exposes individuals and institutions to risks such as misinformation, privacy breaches, algorithmic bias, and inequitable policy outcomes [3, 18].

Data literacy is not confined to a specific demographic; rather, it is a cross-generational necessity. While children and adolescents frequently exhibit proficiency in digital tools, they often lack a comprehensive understanding of how their data is utilized within digital ecosystems. Adults and senior citizens face distinct challenges. For example, adults and senior citizens must navigate e-government platforms and assess the reliability of information sources [4]. Addressing these gaps necessitates the implementation of customized educational strategies, meticulously tailored to address the diverse needs of urban stakeholders, including residents, policymakers, planners, and technical staff. [5].

These challenges are particularly salient in Ibero-American cities, where socio-economic inequalities and disparate access to digital infrastructure serve to exacerbate the risks of exclusion from data-driven governance [16, 17]. In this context, data literacy is not merely a technical proficiency but rather a strategic catalyst for smart governance, sustainable urban policies, and inclusive citizenship.

The objective of this study is to establish data literacy as a foundational skill for smart governance and socially sustainable urban development. We propose an educational-analytical framework that empowers multiple generations and stakeholder groups to interpret, manage, and apply data in ways that enhance transparency, foster citizen participation, and strengthen urban resilience. By integrating data literacy into the governance ecosystem of Ibero-American cities, this study builds upon previous research and contributes a framework tailored to the region's distinctive socio-economic context, facilitating more equitable and sustainable decision-making in an era marked by pervasive data collection and analysis [20].

2 Importance of Data Literacy in Smart Governance and Urban Informatics

Data has become an indispensable component of decision-making processes across various sectors. However, its strategic value is particularly evident in the governance of modern cities. In the Ibero-American context, the exponential growth of urban data—generated through sensors, mobile devices, social platforms, and e-government systems—offers unprecedented opportunities for evidence-based, inclusive, and sustainable policy-making [6, 17, 19].

Through the analysis of patterns and correlations across datasets related to mobility, energy, and the environment, decision-makers can predict outcomes, optimize public services, and allocate resources in a more equitable manner [7, 16]. In the domain of mobility management, for instance, data facilitates the optimization of traffic flow and the scheduling of public transportation. In the context of environmental governance, data supports the monitoring of air quality and the development of strategies for climate adaptation. Similarly, in the field of social policy, data plays a crucial role in identifying vulnerable communities and ensuring the effective targeting of support programs.

The integration of Geographic Information Systems (GIS), Big Data, and Artificial Intelligence (AI) constitutes the analytical foundation for sustainable land-use planning and infrastructure development. These technologies have the capacity to predict the impact of proposed interventions, thereby facilitating more participatory and transparent governance processes [18].

However, the implementation of data-driven governance is not without its challenges. A dearth of data literacy among policymakers, public servants, and citizens can result in decisions predicated on incomplete, misleading, or biased information. Misinterpretation of geospatial or demographic data, for example, may result in inequitable resource allocation or unintended harm to marginalized groups. Furthermore, a lack of awareness regarding privacy and security concerns exposes municipal systems and residents to the risk of identity theft, cyberattacks, and the misuse of sensitive information [9, 14, 18].

In order to address these challenges, it is essential to cultivate urban data literacy as a civic and governmental competence. Beyond technical training, this necessitates cultivating critical thinking, ethical awareness, and the capacity to evaluate credible data sources. The integration of data literacy education into public administration, citizen engagement initiatives, and community-based programs holds promise for enhancing urban resilience and fostering trust in institutions.

In essence, cultivating a culture of data-literate governance endows cities with the capacity to utilize data as a catalyst for innovation, transparency, and sustainability. When managed responsibly, data can contribute to more inclusive and resilient smart cities, enhance service delivery, and support equitable policy outcomes aligned with sustainable urban development goals [6, 9, 19].

3 The Current State of Urban Data Literacy

Despite its growing importance, data literacy remains significantly underrepresented in the governance of Ibero-American cities. A considerable number of training programs for public servants and stakeholders concentrate predominantly on fundamental digital skills, such as the utilization of office software or the navigation of online platforms. However, these programs frequently neglect to incorporate critical competencies in data analysis, visualization, interpretation, and ethical management [10, 17]. This phenomenon frequently gives rise to a fragmentation of knowledge and an impaired capacity to implement data insights within the domain of public policy.

The deficiencies are evident across a variety of demographic groups. Municipal authorities possess extensive knowledge of urban data, yet they frequently lack the skills necessary to validate, integrate, and communicate results effectively. Residents of neighborhoods experiencing socioeconomic disadvantage frequently encounter digital exclusion, which impedes their capacity to interact with e-government platforms and engage in data-driven decision-making processes. The prevailing perspective regarding younger generations, often termed "digital natives," is that they utilize applications and social media extensively. However, there is a notable lack of awareness among this demographic concerning the collection and monetization of their mobility, location, and consumption data. Conversely, older adults may encounter significant challenges in accessing digital platforms, thereby increasing their vulnerability to misinformation, privacy breaches, and exclusion from participatory processes [11, 16].

These disparities are further exacerbated by the absence of standardized frameworks for urban data literacy education. While certain municipalities and universities have integrated advanced analytics and GIS training into their curricula, many others lack the resources, expertise, or political determination to do so [12, 19]. This phenomenon engenders disparate access to fundamental competencies, thereby perpetuating socio-economic disparities and amplifying the urban digital divide. Infrastructure that is deficient in quality or quantity, as well as connectivity that is inadequate, poses significant obstacles to the adoption of smart governance initiatives in communities.

Addressing these deficiencies necessitates a multifaceted approach. Urban data literacy must be systematically embedded into training programs for policymakers, planners, and technical staff, as well as into citizen-oriented initiatives. This suggests a progression from fundamental ICT training to the cultivation of analytical competencies, encompassing the capacity to critically evaluate data sources, interpret geospatial and statistical outputs, and comprehend the ramifications of algorithmic decision-making in urban environments.

The expansion of lifelong learning opportunities through accessible platforms such as municipal data labs, open-data hackathons, community workshops, and online courses is also imperative. Public awareness campaigns are equally important for highlighting the relevance of data literacy in everyday urban life. For example, such campaigns can demonstrate how open data can support trans-

parent governance or how citizens can safeguard personal data when interacting with digital services.

The necessity of addressing the urban data literacy gap is not merely an educational challenge; it is a societal imperative. The enhancement of these competencies among all stakeholder groups facilitates the advancement of Ibero-American cities toward more transparent, participatory, and sustainable governance, thereby ensuring that the advantages of smart city initiatives are accessible to all citizens, not solely the digitally privileged [12, 20].

4 Proposed approach to urban data literacy

We propose an approach to data literacy that is not limited to technical training but framed as a civic competence essential for smart governance and sustainable urban life. Rather than being confined to classrooms, data literacy must be cultivated across all stages of life, in schools, communities, workplaces, and government institutions. This ensures that citizens, public servants, and decision-makers can engage meaningfully with urban data and contribute to inclusive governance.

4.1 When to start urban data literacy education

Urban data literacy should begin at the primary school level, where students can be introduced to simple, real-life problems such as measuring air quality in their neighborhood or visualizing public transport usage with graphs. These activities not only teach mathematical and scientific principles but also connect children to their urban environment, fostering early awareness of sustainability and collective responsibility [10, 12].

During adolescence, the curriculum should expand to include critical issues such as data privacy, digital rights, and the ethical use of urban data. For example, students could analyze how mobility apps or social networks influence daily choices, or explore how algorithmic bias may affect access to educational or social opportunities. This stage builds the critical perspective necessary for democratic participation in data-driven societies.

In higher education and professional development, data literacy should emphasize interdisciplinary applications: urban planning students may learn to work with geospatial datasets; public administration students may analyze indicators of service delivery; health sciences students could model epidemiological data. Such approaches connect technical proficiency with social impact, reinforcing data literacy as a transferable skill for governance, sustainability, and innovation.

For adults and older citizens, urban data literacy should focus on practical engagement with e-government platforms, digital participation tools, and the protection of personal information. Community workshops, public libraries, and civic associations can serve as hubs for inclusive education, empowering citizens to understand their digital footprint, participate in policy consultations, and hold administrations accountable [11, 12].

4.2 Why urban data literacy is essential

Data literacy equips individuals and communities with the ability to critically interpret and apply information for everyday urban decisions—ranging from managing personal energy consumption to understanding environmental alerts. At the societal level, it strengthens democratic processes by enabling informed debates, evidence-based policymaking, and transparent governance.

During the COVID-19 pandemic, for instance, the ability to understand infection curves, vaccination data, and public health recommendations was fundamental for collective resilience. Similarly, in the context of climate change, interpreting environmental data (air quality, water usage, or emissions) empowers citizens to support and co-design sustainable urban policies. Moreover, awareness of data privacy, misinformation, and algorithmic bias strengthens public trust and protects vulnerable groups from inequitable policy outcomes.

Thus, data literacy must be recognized not only as an individual skill but as a social infrastructure that underpins governance, sustainability, and equity in smart cities [13].

4.3 How to implement urban data literacy

Implementation requires context-sensitive and inclusive pedagogical approaches. For children, experiential learning—such as measuring school energy usage or mapping safe walking routes—provides both analytical and civic skills. For adolescents, project-based learning (e.g., designing community surveys or analyzing mobility data) links data analysis to social impact. For professionals, interdisciplinary courses and training should connect data tools with governance and sustainability challenges. For older adults, workshops on digital rights, e-government access, and online security ensure that no group is excluded from digital participation.

Integration into both formal education and community-based initiatives is essential. Schools should embed data modules into existing subjects; universities should promote interdisciplinary data literacy programs; municipalities should support ongoing adult learning. Partnerships between governments, academia, and civil society can ensure scalability, accessibility, and sustainability of these efforts.

Proposed Flow of Tasks for Implementing Urban Data Literacy

- Early Education (Primary School)
  - Introduce basic urban data concepts (e.g., measuring air quality, counting public transport usage).
  - Integrate activities into mathematics and science to develop analytical and civic awareness.
- Secondary Education
  - Teach digital rights, ethical data use, and algorithmic awareness.
  - Encourage project-based learning linked to urban issues (mobility, environment, inclusion).

- Higher Education and Professional Development
  - Offer interdisciplinary courses connecting data science with governance, sustainability, and health.
  - Promote real-world experiences such as internships in municipal data labs or civic-tech projects.
- Adult and Community Education
  - Provide workshops on e-government tools, privacy, and digital participation.
  - Leverage libraries, civic centers, and NGOs as inclusive hubs for lifelong learning.
- Continuous Monitoring and Feedback
  - Evaluate the effectiveness of programs through indicators (e.g., participation rates, service usage).
  - Adapt content to emerging urban challenges and technological developments.

5 Challenges and ethical considerations

The implementation of urban data literacy programs faces multiple challenges of a technological, economic, and socio-cultural nature. Technologically, the persistent digital divide across Ibero-American cities restricts equitable access to the tools required for learning and applying data skills. Unequal connectivity and limited access to devices remain significant obstacles for marginalized communities. Economically, the deployment of comprehensive data literacy initiatives demands substantial investment in training, infrastructure, and human resources—an effort often constrained by budget limitations in municipalities and public institutions. Culturally, barriers include distrust in government data practices, lack of awareness of the value of open data, and resistance to adopting new technologies [14].

Ethical considerations are equally critical. Data literacy must encompass instruction in privacy, data protection, transparency, and informed consent. Citizens should be able to understand how their personal and collective data are collected, processed, and utilized by both public and private actors. A data-literate population should also be capable of critically evaluating the fairness and inclusivity of data-driven policies, ensuring that algorithmic governance does not exacerbate social inequalities. Promoting awareness of digital rights and algorithmic accountability is fundamental for fostering trust and legitimacy in smart governance [15].

—

6 Potential impact of a data-literate urban society

A society equipped with strong urban data literacy has the potential to transform governance, sustainability, and social inclusion. In the field of urban planning,



Fig. 1. Proposed flow of tasks for implementing urban data literacy.

data-literate policymakers and citizens can co-create more resilient and sustainable cities by interpreting environmental and mobility data to inform equitable land-use decisions. In governance, data literacy strengthens transparency and accountability, enabling citizens to evaluate government dashboards, budget allocations, and policy outcomes. In social sustainability, it empowers vulnerable communities to use data in advocating for fair access to resources such as housing, transport, and healthcare [9, 10].

The COVID-19 pandemic demonstrated the importance of data-literate societies, where citizens capable of understanding infection trends and vaccination data contributed to collective resilience. Beyond health, urban data literacy can enhance participatory democracy by enabling citizens to interpret and question policy indicators in real time. Similarly, municipalities that cultivate data literacy among public servants can improve service delivery and accelerate evidence-based innovation in mobility, waste management, and energy efficiency.

The private sector also benefits: start-ups and enterprises operating in smart cities can leverage urban analytics responsibly to drive sustainable innovation. However, the emphasis must remain on balancing efficiency with equity, ensuring that technological progress contributes to inclusive urban development rather than deepening divides [11, 13].

—

7 Conclusions

In the context of the exponential growth of urban data, data literacy emerges as a foundational competence for smart governance and sustainable cities. A lack of such competence not only increases vulnerabilities—such as exposure to misinformation, privacy breaches, and inequitable policy outcomes—but also limits the ability of citizens and governments to harness data for inclusive development.

This work highlights the necessity of adopting a lifelong and multi-stakeholder strategy, integrating data literacy across schools, universities, municipal administrations, and community initiatives. The approach must be inclusive, addressing the needs of children, youth, professionals, and older adults, while also bridging gaps between citizens, governments, and private organizations.

Challenges related to the digital divide, cultural resistance, and ethical concerns such as privacy and algorithmic bias must be carefully addressed. Nonetheless, the potential benefits of urban data literacy are transformative. From improving health and education outcomes to enhancing participatory governance and resilience in the face of global challenges such as climate change, data literacy can enable Ibero-American cities to become more transparent, inclusive, and sustainable.

In essence, investing in urban data literacy is investing in the future of equitable and intelligent cities. It safeguards citizens from the risks of the digital age while empowering them to actively shape policies and decisions, thereby reinforcing democracy, social sustainability, and urban resilience.

References

1. Raffaghelli, J.E., Romero Carbonell, M., Romeu-Fontanillas, T.: In my opinion, the TOS... Situating personal data literacy interventions. Inf. Learn. Sci. **125**(5-6), 367–388 (2024). https://doi.org/10.1108/ILS-06-2023-0086
2. Pörtner, L., Riel, A., Klaassen, V., Sezgin, D., Kievits, Y.: Data literacy assessment – measuring data literacy competencies to leverage data-driven organizations. Procedia CIRP **128**, 78–83 (2024). https://doi.org/10.1016/j.procir.2024.07.047
3. Hannigan, A., Garry, F., Byrne, C., Phelan, H.: The role of the arts in enhancing data literacy: A scoping review protocol. PLOS One **18**(2), e0281749 (2023).
4. Miller, K.M., Yoon, S.A., Shim, J.: Surfacing collective pedagogical content knowledge for data literacy: An exploratory study with high school science teachers. In: Int. Conf. of the Learning Sciences (ICLS), pp. 1037–1040 (2022).

5. Puteri, S.A.: The alpha generation perspective on the use of technology in filtering actual information through social media. PRIMACY J. Engl. Educ. Lit. **2**(1), 41–53 (2023).
6. Liu, Y., Zeng, N., Papadonikolaki, E., Maritshane, K., Chan, P.W.: The future of digitalized project practices through data-savvy talent: A digital competence formation perspective. Proj. Lead. Soc. **5**, 100120 (2024). https://doi.org/10.1016/j.plas.2024.100120
7. Hojas-Mazo, W., Maciá-Pérez, F., Berná Martínez, J.V., Moreno-Espino, M., Lorenzo Fonseca, I., Pavón, J.: Framework based on simulation of real-world message streams to evaluate classification solutions. Algorithms **17**(1), 47 (2024). https://doi.org/10.3390/a17010047
8. Cepero-Pérez, N., Moreno-Espino, M., García-Borroto, M., Morales, E.F.: Progressive forest: An early stopping criteria for building ensembles. Comput. Sist. **27**(1), 89–97 (2023). https://doi.org/10.13053/CyS-27-1-4224
9. Bakken, S., Dreisbach, C., Kisselev, S., Turchioe, M.R.: Advancing data science competencies for nursing PhD students. In: Innovation in Applied Nursing Informatics, pp. 515–519. IOS Press (2024).
10. Sheriff, N., Sevukan, R.: Exploration of data literacy research using a network of cluster mapping approach. J. Scientometric Res. **12**(1), 130–143 (2023). https://doi.org/10.5530/jscires.12.1.002
11. Cui, Y., Chen, F., Lutsyk, A., Leighton, J.P., Cutumisu, M.: Data literacy assessments: A systematic literature review. Assess. Educ. **30**(1), 76–96 (2023).
12. Henderson, J., Corry, M.: Data literacy training and use for educational professionals. J. Res. Innov. Teach. Learn. **14**(2), 232–244 (2021). https://doi.org/10.1108/JRIT-11-2019-0074
13. Ming, W., Hui, H.: Data literacy education design based on needs of graduate students in UCAS. In: Kurbanoglu, S., Boustany, J., Spiranec, S., Grassian, E., Mizrachi, D., Roy, L. (eds.) Inf. Lit. Workplace, pp. 158–168. Springer, Cham (2018).
14. López-González, A., Moreno-Espino, M., Román, A.C.M., Fernández, Y.H., Pérez, N.C.: Ethics in artificial intelligence: An approach to cybersecurity. Int. Artif. **27**(73), 38–54 (2024). https://doi.org/10.4114/intartif.vol27iss73pp38-54
15. Ferretti, A., Ienca, M., Sheehan, M., Blasimme, A., Dove, E.S., Farsides, B., Friesen, P., Kahn, J., Karlen, W., Kleist, P., Liao, S.M., Nebeker, C., Samuel, G., Shabani, M., Velarde, M.R., Vayena, E.: Ethics review of big data research: What should stay and what should be reformed? BMC Med. Ethics **22**(1), 51 (2021). https://doi.org/10.1186/s12910-021-00616-4
16. Alexander, D., Døhl Diouf, L., Prescod, K.: Digital inclusion in Caribbean digital transformation frameworks and initiatives: a review. Stud. Perspect. – ECLAC Subregional Headquarters for the Caribbean **48652** (2023). Naciones Unidas, CEPAL
17. CAF Development Bank of Latin America: OECD Digital Government Studies. Digital Government Review of Latin America and the Caribbean: Building Inclusive and Responsive Public Services. OECD Publishing, Paris (2023)
18. UNESCO: Recommendation on the ethics of artificial intelligence. United Nations Educational, Scientific and Cultural Organization, Paris (2022)
19. UN-Habitat: World Cities Report 2022: Envisioning the Future of Cities. United Nations Human Settlements Programme, Nairobi (2022). https://unhabitat.org/world-cities-report-2022
20. Bettinger, K., Ferrari, J.Z., Lähteenoja, V.: Empowered data societies: A human-centric approach to data relationships. In: World Economic Forum, Geneva, vol. 2, pp. 2021 (2021)

Implementing data governance program in a Colombian public institution.

Gina Maestre-Góngora[1[0000-0002-2880-9245]], Diana López[1[0000-0003-0950-9068]], Fernando Osorio Giraldo[1], Osvaldo Cifuentes Restrepo[1], Juan Aragon-Mendoza[2[0000-0001-8832-0209]] and Myriam Patricia Cifuentes [2[0000-0002-2302-123X]]

[1] Universidad de Antioquia, Medellín, Colombia
[2] Superintendencia Nacional de Salud, Bogotá, Colombia

Abstract. This paper presents a case study of the implementation of a data governance program in a Colombian public institution, framed within the broader digital transformation agenda. The study explores the design, deployment, and evaluation of a data governance model based on maturity models and international best practices. Using a mixed-methods approach, the research combines documentary analysis, expert validation, and fieldwork conducted in collaboration with a public institution in Colombia. Key findings highlight the relevance of context-aware frameworks, stakeholder engagement, and adaptive capabilities for institutionalizing data governance. The study contributes to the literature by offering empirical insights into challenges and enablers in the public sector, particularly in Latin American settings. It also proposes a replicable methodology for assessing and enhancing data governance maturity. These results are particularly relevant for governments aiming to transition toward more intelligent and data-driven public management.

Keywords: Data governance, public sector, maturity models.

1 Introduction

Digital revolution has deeply transformed the way governments work, allowing for speeding up developing and implementing projects where data is the core goal of emerging technologies such as Artificial Intelligence (AI), internet of things (IoT), Big Data and Blockchain. These advances have significantly redefined the interactions between the State and the citizens, fostering even more intelligent governments. However, one of the main challenges is to fully understand the real potential data have, to leverage them effectively and to maximize their utility making them a strategic asset with high economic, social and political value. The ability to use data as a digital economy engine requires sound regulatory frameworks, translated into strategic initiatives involving key actors to guarantee their governance and effective exploitation [1]

The case of Colombia is an example for countries classified as “upper middle income” by the World Bank, as well as also, in the Latin America & the Caribbean region [2]. According to the 2017 Report on Digital Government [3], the Organization for Economic Cooperation and Development (OECD) spotlighted Colombia urgently needed to strengthen and formalize an integral data governance framework, recommending it must gather regulatory policies, technical standards and institutional

directives to align national goals with the development of a data-based culture inside public institutions. Critical topics include fostering interoperability, strengthening institutions' capability for efficiently managing the data life cycle as well as implementing exploitation mechanisms.

Correspondingly, a diagnosis made by the National Department of Planning in Colombia (DNP, acronym in Spanish) in 2017, showed the Colombian public sector had a low level of readiness for data-based decision making, as just 3.3% of government institutions had the necessary conditions to adopt and implement big data initiatives, while 97.6% needed specific interventions to overcome barriers and develop social and economic value from digital data. These results show an even deeper problem: the lack of a rooted data culture, which hinders recognizing data as a strategic national asset.

The Colombian Ministry of Information and Communication Technologies (MinTIC) has developed policy directives and tools to fill these by fostering adoption of best practices in data governance, and therefore, positioning data as a valuable resource to create public value [4–7]. Nevertheless, despite these efforts, adoption inside institutions is still unequal among public entities.

One of the main challenges for implementing data governance in the Colombian government is the institutions' uneven capabilities. Factors such as organization size, operative complexity, financial and human resources, and technological infrastructure set significant capability differences to adopt data governance principles.

Former research has flagged a scarce number of systematic studies on implementing and continuous monitoring of data governance strategies [8]. Notwithstanding international frameworks offer conceptual guidance, they do not provide practical guidance tailored to the needs and limitations of government organizations, even more at local level [9].

From a practical perspective, institutionalizing data governance inside government entities is a multifaceted challenge. For public organizations it is usually difficult to recognize the tangible benefits of cross-organizational data governance due to the scarce cooperation between areas. Diverse local practices and political considerations make it even harder to establish standardized data governance models.

Benfeldt et al. [10] found these barriers appear in different ways, such as hurdles to understand the available data, resistance to shared governance principles and fragmented efforts due to other responsibilities in conflict. In this way, implementing a data governance framework in public institutions requires a holistic scope gathering organizational, technical and political dimensions.

In this context, our objective was to develop a pragmatic framework to assess the level of readiness in Colombian public institutions for projects implementing data governance, and beyond, with the purpose of ease a more systematic and effective adoption of the data governance principles.

We developed a data governance maturity model to objectively record the technical observed present state directly from the public servants, which also answered the expected state. Observed and expected state allowed to identify the gap, a strategic action plan will fill. Therefore, beyond assessing maturity of data governance, our framework enables the design of implementation strategies adapted to the government institutions'

necessities and limitations to improve the public sector capacity to exploit data for decision making, service delivery and economic development in the digital era.

We show the results of applying our systematic framework and maturity model in a national public institution, The national public institution is part of the Colombian government's digital transformation, which is reflected in various regulations and initiatives, but mainly in Decree 1389-2022, which establishes the general guidelines and governance model for Colombia's data infrastructure [4].

2 Background

2.1 Data Governance

Data governance has become a foundational cornerstone for modern public management, because of its effects on decision making, policy development and government service delivery. The growing use of digital data demands State entities to prioritize the structured management of data assets to ensure transparency, efficiency and accountability. In general terms for technological purposes, governance involves the mechanisms and frameworks for decision making about information technology and data, over decision and responsibility domains [11]. However, data governance is not an exclusive topic of IT area. Instead, it is a strategic process over management, which requires straightforward authority, control mechanisms and a participatory scope for decision making at all government levels [12].

[10] highlights that data governance is further than government servants, as it impacts multiple actors, including also citizens and private sector workers. Public administrators rely on data to orient their daily tasks, while those responsible for policies use them for strategic decisions. In this context, [3] pleads for a structured data governance framework, which integrates regulatory, institutional, technical and political measures to effectively manage data life cycle. This involves data gathering and storage, as well as security, interoperability, reuse, and finally, maximizing their public value.

A growing number of studies outline the relevance of thinking data governance as a collective effort instead as an isolated institutional function. Several research results emphasize the need for politically accountable actors, researchers and public servants share a common understanding about data governance scope and challenges as the basis to develop feasible solutions [10]. Cerrillo-Mártinez [13] offer valuable insights of how governments approach data governance, identifying three main models: (i) Centralized model: a single organism has absolute control over data management. (ii) Coordination model: Foster decision making between multiple organisms. (iii) Decentralized model: Different organisms control their independent data management.

Liu [14] introduced the concept of social data governance, stating that data governance mirrors the sociopolitical context it came from, and has effects on policy development and regulatory practices. This perspective emphasizes interdependency between governance structures and broad social norms. An extensive review by [15] stresses the need for adopting a holistic and adaptative approach in the public sector where bureaucratic restrictions and different institutional capabilities have unique challenges.

From the technical point of view, [16] reinforces key factors for data governance effective in public administration, focusing on delivering intersectoral public services. This includes setting metadata repositories and the use of ontologies to standardize data classification and improve interoperability between government entities. Without these tools, efforts to integrate and manage public sector data will still be fragmented and ineffective.

Recent discussions suggest that just improving public data availability is not enough to foster innovation and achieve public policy objectives. These approaches look to ensure ethical data handling, fostering collaborative ecosystems where data are used equally to benefit the common good. His study contributes to debates about innovative governance models, by exploring tools and frameworks allowing sustainable and responsible data exchange. Simultaneously, [17] identifies three big challenges to work on for effective data governance at the local level government: (i) Data comprehension and valuing: Several government agencies do not have clear data inventories nor the mechanisms to assess data value for decision making. (ii) Improving collaboration and data exchange: Persistence of institutional data silos makes data integration difficult and limits more coordinated public policy development. (iii) Strengthening technical and political capabilities: Capabilities for effectively using data relies on available technological capabilities and political will to implement data-based strategies.

Despite advances in conceptualizing data governance, the critical gap to empirically validate data governance frameworks remains. Several proposed models in the academic literature stay at an abstract level, without tangible applications in real settings. Additionally, research on how implement, monitor and evaluate these governance frameworks in practice is still scarce. To advance in this field, strengthening empirical research and developing evidence-based methodologies is key. To maximize profit from data for economic and social development, countries must engage in: Guarantee data quality and integrity, ensure transparency and open access to data, implement solid measures to protect data privacy, build public trust in data management and use, strengthen technological infrastructure and human capital in government agencies to improve their capabilities for data analysis and exploitation.

Beyond national initiatives, international regulatory frameworks define the data governance future: European Union Data Act, backbone of the European digital economy strategy, looks to position Europe as the leader for equal use and access to industry data from all economic sectors [18]. Also, the Data Governance Law enacted directives for voluntary data exchange, promoting a structured approach for fairly and innovatively sharing information. Legislative updates during 2023 defined even more the rights and responsibilities related to data access and use in the EU, reinforcing the data ecosystem to be more cohesive and dynamic.

2.2 Maturity Models

Born in the technological field, maturity models have become widely accepted tools to assess the degree of development of diverse processes, comparing the observed performance in a topic to a grid predefined dimensions over a growing scale usually represented by Likert items. Information technology (IT) maturity models have become

established as diagnostic and strategic guidance tools that enable organizations to evaluate and improve their technological processes, capabilities, and competencies [19]. However, their development and application face challenges due to characteristics of the IT field, such as its dynamism, diversity, and dependence on the organizational context. Hence, the design and construction of maturity models require a critical and flexible approach [23].

The changing nature of IT presents one of the greatest challenges. The constant emergence of new technologies, agile methodologies, and reference frameworks creates an environment that is difficult to capture in a static model. In this sense, continuous updating becomes essential to prevent models from becoming obsolete [20].

Another challenge is the lack of consensus around definitions, processes, and expectations associated with maturity. Each organization develops different interpretations of what "technological maturity" means based on its strategic, sectoral, and cultural context [21]. This limits the possibility of applying a universal model, hence requires adapting each instrument to the particularities of the environment where it is implemented.

Maturity models are often based on self-assessments, which introduces a high level of subjectivity. Factors such as the perceptions of IT managers, organizational pressure, and the lack of clear criteria lead to inconsistencies. This limits the reliability of the results obtained. It also makes it difficult to compare organizations or different time periods [22].

Adopting a maturity model involves new processes, practices, or technologies, which often generates internal resistance. Organizational culture has a decisive influence on the acceptance of the model. Negative attitudes toward innovation reduce the effectiveness of its implementation. Therefore, change management is key to achieving sustainable results[23].

Developing and maintaining a maturity model requires investment in time, money, and specialized talent. Not all organizations have sufficient resources. This limits the ability to design robust and sustainable models over time. As a result, implementation may be incomplete or ineffective [19].

The effectiveness of a maturity model depends on the organizational and national context in which it is applied. Cultural, regulatory, and structural factors condition its adoption. What works in an Anglo-Saxon environment may not be viable in Latin America or other contexts. These differences reduce the universal applicability of maturity models [20].

Despite the challenges outlined above, maturity models remain valuable tools for organizations, as they facilitate the identification of gaps, the prioritization of investments, and the definition of improvement plans [22]. However, it is crucial to recognize their limitations, adapt the approach to each context, and ensure a continuous maintenance process that allows them to remain valid in a changing technological environment.

Table 1 summarizes multiple maturity models for assessing data governance progress, from initial levels of awareness to advanced and continuous integration into business processes.

Table 1. Data Governance maturity models in literature.

Model	Maturity Levels	Key Dimensions	Source
Gartner EIM	6 levels	Vision, strategy, metrics, governance, organization and roles, lifecycle, infrastructure	Structured and strategic model for all business information management
Stanford DGMM	5 levels	Awareness, formalization, metadata, stewardship, quality, master data	Measures basic capabilities and program resources in institutional environments
CMMI DMM	5 levels	Governance, data quality, strategy, architecture, operation, support processes	Formal and detailed approach for organizations seeking to improve their data management practices
IBM Data Governance	5 levels	Organizational awareness, stewardship, policy, value, risk, security, architecture, quality, metadata, lifecycle, audit	Broad coverage of critical elements including risk, value, and architecture
Oracle DGMM	6 levels:	Processes, policies, quality, organizational structure	Clear evolution from data chaos to institutionalized strategic governance

3 Methods

The figure 2 presents a five-phase structured process for developing and applying a maturity model. In Phase 1, the model is defined by reviewing existing frameworks. Phase 2 consists of developing tools for assessing the current and future status, as well as a pilot test to validate their effectiveness. In Phase 3, data collection and initial analysis are carried out. Then, in Phase 4, a gap analysis is performed comparing the current state with the desired state. Finally, Phase 5 culminates in the creation of a strategic platform that includes specific actions and a roadmap to achieve the defined objectives. This sequential approach allows for a comprehensive and planned assessment of organizational or process maturity.

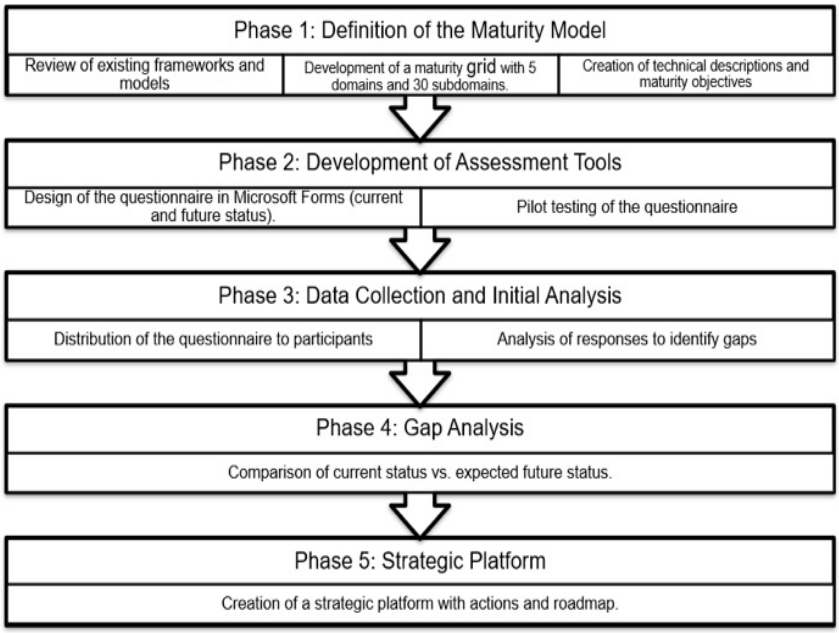


Figure 2. Methodology for data governance programs

From the review of several data governance frameworks such as the DGI [24], in the DMBOK [25], and related maturity models in Table 1**Table** , we designed a maturity model grid with five domains, detailed by subdomains (30) as shown in Table , over five levels of development of (1) absent, (2) reactive, (3) emerging, (4) established, and (5) embedded.

Following methodological steps [20, 26] we populated the intersection cells between domains/subdomains and maturity levels with descriptors of capacities [27], which, beyond the usual perception statements, introduced full technical descriptions of the progressive achievements expected for each subdomain topic to improve objectivity through explicit technical understanding. We used this grid to develop a questionnaire in Microsoft Forms, to explore the present observed state of data governance from public servants, as well as their expected state after 1 year of actions. We used these two ends (observed/present and future) for gap analysis, that actions will fill in a strategic plan, prioritized in a roadmap. Directors would approve it as other elements of the strategic platform of the wide institutional strategic data governance program.

After a broad institutional call, we obtained answers from 45 participants, mainly working with data (44) as users and some of them also with data management roles (12).

4 Results and Discussion

Table shows the basic average scales of the Likert-like items at each domain and subdomain, and Figs. 1 to 5 present the comparative visualization of average Likert-like scales by subdomains. The total average by observed domains was 2.70 (SD 0.22) below the midpoint of the scales, but the future expected values were optimistic with an average of 4.38 (SD 0.13).

Table 2. Data governance maturity model (DGMM) domains and subdomains, and results of Likert like scale observed and future expected results.

Data Governance domains and subdomains		Observed Avg	Future Avg
1.	Data governance policies	2.7	4.3
a.	Policy formulation	2.5	4.1
b.	Assessment of policy implementation	2.6	4.3
c.	Standards adoption	2.6	4.2
d.	Compliance with national and international regulations	2.6	4.3
e.	Adopting new Regulations and Legislations aligned with DG	2.6	4.3
f.	Communicating new Regulations and Legislations aligned with DG	3.2	4.6
g.	Adoption of best practices	2.7	4.3
h.	Processes and procedures to manage data governance operations	2.6	4.3
i.	Definition of data management processes by data governance	3.0	4.4
j.	Definition of SOP to operations with data, by data governance	2.4	4.2
2.	Data governance strategy	2.6	4.4
a.	Strategic planning of data governance	2.5	4.3
b.	Financial investment (financial capital for data governance)	2.9	4.6
c.	Technological infrastructure (technological capital for data governance)	2.7	4.4
d.	Human resources (human capital for data governance)	2.9	4.5
e.	Risk due to data governance absence	2.4	4.2
f.	Digital initiatives directed by data governance	2.2	4.4
3.	Roles and responsibilities for data governance	2.7	4.5
a.	Data governance structured roles	2.7	4.4
b.	Collaboration between data governance roles	2.6	4.5
c.	Data governance training	2.7	4.6
4.	Culture and ability to adapt to change to adopt data governance	2.7	4.4
a.	Awareness and cultural compromise about data governance	2.7	4.4
b.	Leaders' commitment to DG	2.8	4.4
c.	Communications about data governance	2.7	4.4
d.	Adoption of data governance practices	2.5	4.4
5.	Data management implementing the data governance areas	2.8	4.4

a.	Authentication and authorization for data access	3.1	4.5
b.	Mechanisms allowing and controlling data access	2.9	4.5
c.	Data protection	2.6	4.4
d.	Data quality	2.9	4.4
e.	Data use and usability	3.0	4.5
f.	Data integrity	2.7	4.4
g.	Data integration	2.6	4.4

The data governance maturity model applied to the Colombian public entity identified that, in general terms, the domains and subdomains evaluated are at an emerging level (≈2.6–2.9 on average), with a projection toward an established to integrated level (≈4.3–4.6) in future scenarios. This shows that, although there has been some initial progress, the organization still operates reactively on several fronts and needs to consolidate more systematic and sustainable practices.

The policy domain has an observed average of 2.7, which places it at an emerging level. Policy formulation (2.5) and evaluation of its implementation (2.6) show initial efforts, but without consolidation, reflecting a still reactive approach to regulatory requirements. Although the entity communicates some regulations (3.2, the highest value in this domain), the adoption of standards and good practices remains at an emerging level. The future projection (≈4.3–4.6) suggests a transition to an established level, where policies are more consistently integrated into institutional operations.

In the strategy domain, the average score is 2.6, also at an emerging level. Strategic planning (2.5) and financial investment (2.9) show the absence of a solid framework for prioritizing data governance as a cross-cutting issue. Governance-led digital initiatives (2.2) reflect an incipient state, where reliance on reactive decisions limits institutional projection. However, the expected scenario (4.2–4.6) points to an established to integrated level, with investments in technological infrastructure and human capital that strengthen strategic deployment.

Role domain has an average score of 2.7, placing it at an emerging level. The existence of minimal role structures (2.7) and collaboration between them (2.6) shows that the organization recognizes the need to formalize responsibilities, although it still lacks an integrated framework. Governance training (2.7) provides an opportunity to consolidate internal capacities. The projection towards values close to 4.5 reflects the intention to move to an established level, where governance is based on solid organizational capacities.

The cultural dimension has an average score of 2.7, corresponding to an emerging level. Although there is initial commitment from leaders (2.8) and a moderate level of communication (2.7), the adoption of governance practices (2.5) reveals that the organizational culture still fluctuates between reactive and emerging, with resistance to change and a low level of ownership. The projected jump to 4.4 indicates that the entity plans to consolidate an established level, where a pro-governance culture is an integral part of its operations.

Data management has the highest average observed (2.8), although still at an emerging level. Subdomains such as authentication and authorization (3.1) and use/usability (3.0) show relative progress compared to other areas, positioning themselves on the border between emerging and established. However, aspects such as protection (2.6)

and data integration (2.6) remain behind. The future trend toward values of 4.4–4.5 suggests consolidation at an established level, with standardized security, quality, and integration practices that contribute to institutional reliability.

To close the gap identified in the assessment, the entity designed a strategic platform to implement a data governance program that promotes, adapts, and adopts best practices from industry benchmarks and the guidelines of Colombia's Digital Government policy. The strategic platform that addresses the domains and subdomains evaluated is presented in figure 3.

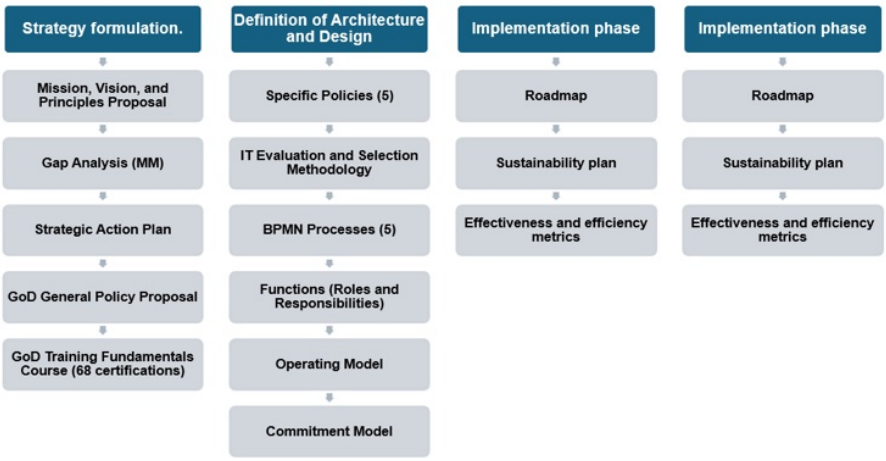


Figure 3. Strategic platform for designing data governance programs.

Overall, the results indicate that the entity is in a predominantly emerging phase, with isolated elements in a reactive state and some approaching an established level. The transition to higher levels will depend on the institutional capacity to: Formalize policies and processes beyond reactive initiatives, ensure sustained financial, technological, and human capital investment, develop roles with clear functions and effective collaboration mechanisms and foster an organizational culture that internalizes data governance as part of its daily work.

The gap between observed and projected values highlights not only the potential for development, but also the challenges in terms of policy continuity, resources, and cultural appropriation. Moving toward an embedded/integrated level will require governance practices to transcend specific projects and consolidate as a long-term institutional policy in the Colombian public sector.

5 Conclusions

While the public sector becomes even more dependent on data for decision making, improvement of services and economic development, data governance consolidates its role as a core element of government administration. However, implementation, standardization and interinstitutional collaboration challenges persist.

Overall, the results indicate that the entity is in a predominantly emerging phase, with isolated elements in a reactive state and some approaching an established level. The transition to higher levels will depend on the institutional capacity to: Formalize policies and processes beyond reactive initiatives; Ensure sustained financial, technological, and human capital investment; Develop roles with clear functions and effective collaboration mechanisms and foster an organizational culture that internalizes data governance as part of its daily work.

The gap between observed and projected values highlights not only the potential for development, but also the challenges in terms of political continuity, resources, and cultural appropriation. Moving toward an embedded/integrated level will require governance practices to transcend specific projects and consolidate as a long-term institutional policy in the Colombian public sector.

Our findings showed that related to the development of a sense of belonging, the involvement of public servants in building data governance in a public institution induces the tacit but robust incorporation of the multiple elements of the context these complex entities belong to. Therefore, providing technical knowledge to attain an informed assessment through introducing clear concepts by a detailed and improved maturity model, also by adding to the evaluation of the present situation, projecting the expected future state, allowed for finding patterns, which make clear how to fill the gap and stressing the relevance of the human factor and the related change management approach in data governance.

In the future, closing the gap between theory and practice through empirical research, inducing adaptative strategies and more flexible data governance models will be essential to allow for effective, sustainable data governance, and being able to produce real benefits for citizens.

Acknowledge: This study is funded and forms part of contracts 131-2024 (SNS-CD-118-2024) and 116-2025 (SNS-CD-104-2025) signed between the National Health Superintendency and the University of Antioquia between 2024 and 2025. The authors would like to thank the officials of the participating public entity, the National Health Superintendency of Colombia, for their willingness to apply the data governance maturity model and for facilitating access to the information necessary for the development of this study.

References

1. Góngora, G.P.M., Bernal, W.N.: Factores clave en la gestión de tecnología de información para sistemas de gobierno inteligente. Journal of Technology Management and Innovation. 10, (2015). https://doi.org/doi: https://doi.org/10.4067/S0718-27242015000400012.
2. WB: World Bank Country and Lending Groups – World Bank Data Help Desk, https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups, last accessed 2025/09/17.
3. OECD: Development Co-operation Report 2017. OECD (2017). https://doi.org/10.1787/dcr-2017-en.
4. MinTIC: Decreto 1389-2022, https://infraestructuradatos.gov.co/798/articles-195193_recurso_1.pdf, (2022).
5. MinTIC: Res 460-22 Plan Nacional de Infraestructura de datos, (2022).
6. DNP: Documento CONPES 3920. Consejo Nacional de Política Económica y Social - Departamento Nacional de Planeación, República de Colombia (2018).
7. MinTIC: Guía para el uso y aprovechamiento de Datos Abiertos en Colombia, (2016).
8. Alhassan, I., Sammon, D., Daly, M.: Data governance activities: an analysis of the literature. J Decis Syst. 25, 64–75 (2016). https://doi.org/10.1080/12460125.2016.1187397.
9. OECD: Estudios Económicos de la OCDE: Colombia 2019. , Paris (2019). https://doi.org/10.1787/805f2a79-es.
10. Benfeldt, O., Persson, J., Madsen, S.: Data Governance as a Collective Action Problem. Information Systems Frontiers. 22, 299–313 (2020). https://doi.org/10.1007/s10796-019-09923-z.
11. Niemi, E.: Designing a Data Governance Framework-Working Paper. (2013).
12. Mahanti, R.: Data Governance Implementation: Critical Success Factors. Software Quality Professional. 20, 4–21 (2018).
13. Cerrillo-Martínez, A., Casadesús-de-Mingo, A.: Data governance for public transparency, (2021).
14. Liu, J.: Social data governance: Towards a definition and model. Big Data Soc. 9, 20539517221111350 (2022). https://doi.org/10.1177/20539517221111352.
15. Nielsen, O.B., Stouby Persson, J., Madsen, S.: Why Governing Data is Difficult: Findings from Danish Local Government. In: International Working Conference on Transfer and Diffusion of IT (TDIT). Springer International Publishing, Portsmouth, UK (2018). https://doi.org/10.1007/978-3-030-04315-5_2.
16. Grimstad, T., Myrseth, P.: Information governance as a basis for cross-sector e-services in public administration. In: 2011 International Conference on E-Business and E-Government (ICEE). pp. 1–4 (2011). https://doi.org/10.1109/ICEBEG.2011.5887109.
17. Nielsen, O., Persson, J., Madsen, S.: Why Governing Data Is Difficult: Findings from Danish Local Government BT - Smart Working, Living and Organising. Presented at the (2019).
18. Europe Union: The European Data Act, https://www.eu-data-act.com/.

19. Becker, J., Knackstedt, R., Pöppelbuß, D.-W.I.J.: Developing maturity models for IT management. Business & Information Systems Engineering. 1, 213–222 (2009).
20. de Bruin, T., Freeze, R., Rosemann, M.: Understanding the Main Phases of Developing a Maturity Assessment Model. Australasian Conference on Information Systems. (2005).
21. Institute, I.T.G.: COBIT 5.0. The IT Governance Institute (ISACA) (2013).
22. Luftman, J.: Strategic Alignment Maturity. In: vom Brocke Jan and Rosemann, M. (ed.) Handbook on Business Process Management 2: Strategic Alignment, Governance, People and Culture. pp. 5–43. Springer Berlin Heidelberg, Berlin, Heidelberg (2015). https://doi.org/10.1007/978-3-642-45103-4_1.
23. Wendler, R.: The maturity of maturity model research: A systematic mapping study. Inf Softw Technol. 54, 1317–1339 (2012). https://doi.org/https://doi.org/10.1016/j.infsof.2012.07.007.
24. DGI: DGI Data Governance Framework - The Data Governance Institute, https://datagovernance.com/the-dgi-data-governance-framework/, last accessed 2024/05/13.
25. DAMA International: DAMA-DMBOK Data management Body of Knowledge. DAMA International, NJ, USA (2017).
26. García-Mireles, G., Moraga, M., Garcia, F.: Development of maturity models: A systematic literature review. In: IET Seminar Digest. pp. 279–283 (2012). https://doi.org/10.1049/ic.2012.0036.
27. Logframer: Likert type scale, https://www.logframer.eu/content/likert-type-scale, last accessed 2025/09/11.

Urban agriculture in Morocco – Systematic Review

Ouiam Fatiha Boukharta[1[0000-0001-5403-5207]], Loïc Sauvée[2 [0000-0002-9763-0126]], Francisco Tomatis[1 [0000-0002-5075-2310]], Leticia Chico Santamarta [1,3[0000-0003-2695-8471]] and Luis Manuel Navas Gracia [1[0000-0002-7895-925X]]

[1] University of Valladolid, Spain
[2] Polytechnic institute of UniLaSalle, France
[3] Harper Adams University, United Kingdom
ouiamfatiha.boukharta@uva.es

Abstract. Morocco's rapid development, with urban and agricultural expansion, is putting pressure on ecosystem services. This study assesses the implementation of urban agricultural practices in Morocco, through a systematic literature review, which analyzes the most relevant articles and results of studies related to this field. Indeed, it begins with the identification of a main research question, followed by sub-questions to make the analysis complete and precise. This methodology led to the identification of 19 relevant articles from the Scopus and Web of Science, which were analyzed and evaluated through coding in NVivo software. The results showed that urban agriculture is not widespread in Morocco, and that its implementation only began a few years ago, and only in the largest cities. Furthermore, this analysis has shown that this implementation identifies many benefits all related to sustainable development, and where various national and international projects seek to achieve the same goals, although farmers claim they need additional financial support. Furthermore, in future planning, urban planners should be more committed to including urban areas, to maintain the provision of these services for current and future generations.

Keywords: Urban Agriculture, Sustainable Development, Morocco, NVivo Software, Ecosystem services.

1 Introduction

Mediterranean regions are among the areas of the world most exposed to the combined effects of urban growth and climate change [1], and where Morocco faces significant pressure on its water supply due to limited water resources and growing demand from various sectors, including agriculture, industry, and domestic use [2]. Indeed, in Morocco, a country deeply affected by metropolitan development, agricultural areas are located on the outskirts of large cities and are undergoing rapid economic, social, and landscape transformations [3]. In addition, urbanization and urban growth are becoming increasingly important in the country, creating serious competition for land use,

particularly between urban and agricultural needs [4], with the urban population growing from nearly 9 million to over 16 million between 1982 and 2004 [5].

From this perspective, the debate on potential open space systems for urban growth centers takes on a new dimension, not only in terms of their contribution to recreation, ecological compensation, and food supply for the city through urban agriculture, but also because of their link to broader territorial dynamics, such as balancing urban expansion, the prevention of territorial fragmentation, and the promotion of social inclusion [6].

Urban agriculture is generally understood to mean any agricultural practice in an urbanized area. This includes vegetable gardens, collectively organized community gardens, individual plots, rooftop gardens, greenhouses, and other innovations in urban farming practices [7]. Indeed, the importance of urban agriculture has been highlighted due to its extensive social functionality and intensive use, its contribution to strengthening food resilience, and its importance for human well-being through the provision of various ecosystem services, all of which are aligned with sustainable development goals [8]. This study therefore assesses the implementation of urban agricultural practices in Morocco through a systematic review of the literature, analyzing the most relevant articles and study results in this field.

2 Material and Methods

2.1 Research question and methodology:

Considering the various considerations outlined above, this research consists of an in-depth literature review on urban agriculture in Morocco. For this purpose, the most relevant articles in the literature were selected and analyzed using NVivo software, one of the most advanced qualitative analysis tools currently available [9]. The research question defined is as follows:

What is the current situation of Urban Agriculture in Morocco?

This research question has generated several research sub-questions, which can be summarized in the following points, and which this systematic review will address:

RQ1: In which cities is urban agriculture most established in Morocco?

RQ2: What are the main objectives that authorities seek to achieve by implementing such projects?

RQ3: What are the main challenges and limitations that Morocco is facing while implementing urban agriculture projects?

Indeed, the framework is based on an initial general question, accompanied by specific sub-questions, which not only effectively structures the results but also optimizes the analysis process. This makes it easier to identify the methodological steps to be followed, organize data collection and processing, and ensure a thorough evaluation.

2.2 Data collection

The methodology used and followed in this literature review consisted of collecting relevant articles on the subject under review, conducting in-depth research and using key terms derived from our main question, then sorting them according to their relevance to the specific subject of this study and other criteria that will be described in more detail in the following subsections.

This involved extracting articles from various research engines, assessing their eligibility and exclusion criteria, applying all steps related to the analysis process (identification, selection, and eligibility), and the abstracting and analyzing of data. This methodological approach is consistent with the systematic literature review process recommended in 2016 by James et al. It also allows for better identification and analysis of articles [10].

Browsing articles using bibliographic databases

The databases used in this systematic review included the internationally renowned bibliographic databases Web of Science and Scopus. The search terms used were developed from keywords related to our specific research question. In addition, Boolean operators such as "AND" and 'OR' were used for our keyword combinations for this search, resulting in the following main string: ((((urban AND (agricul* OR farm*)) OR "community garden*" OR "school garden*" OR allotment*) AND Morocco*).

Data Extraction and Analysis

The search and selection process identified a total of 219 articles from the Web of Science and Scopus search engines. After removing 54 duplicates, the total number of articles selected was 165. After applying the inclusion criteria, selecting all published and accepted articles (open or non-open access) from all years, and applying the exclusion criteria, excluding articles that did not address our research question and/or did not concern Morocco, 23 articles were deemed eligible for mapping the results. During the coding and analysis of each article, four articles were excluded: two because they included urban agriculture projects outside Morocco, and two because they were not accessible. This resulted in a final selection of 19 articles for analysis, 13 from Scopus and six from Web of Science. The diagram presented in Figure 1 of PRISMA details the selection process and results at all stages.

Coding Categories for Analysis

In this systematic review, the coding process has been done by establishing various categories and sub-categories responding to our main question and sub-questions.
Indeed, it is relevant to present the different coding categories used in this systematic review, as clarity and precision in their definition are important for a robust analysis. These categories were created with the research objectives into consideration. For example, since the main objective was to analyze the existence of urban agriculture in

Morocco, the category "Presence of urban agriculture" was designed. Then, as one of the objectives was to identify where urban agriculture is located, another category was created, entitled "Location of urban agricultural areas," and other categories were created in the same way. This greatly improves the accuracy of our research, allowing for more precise extraction of information and an even more advanced analytical approach.

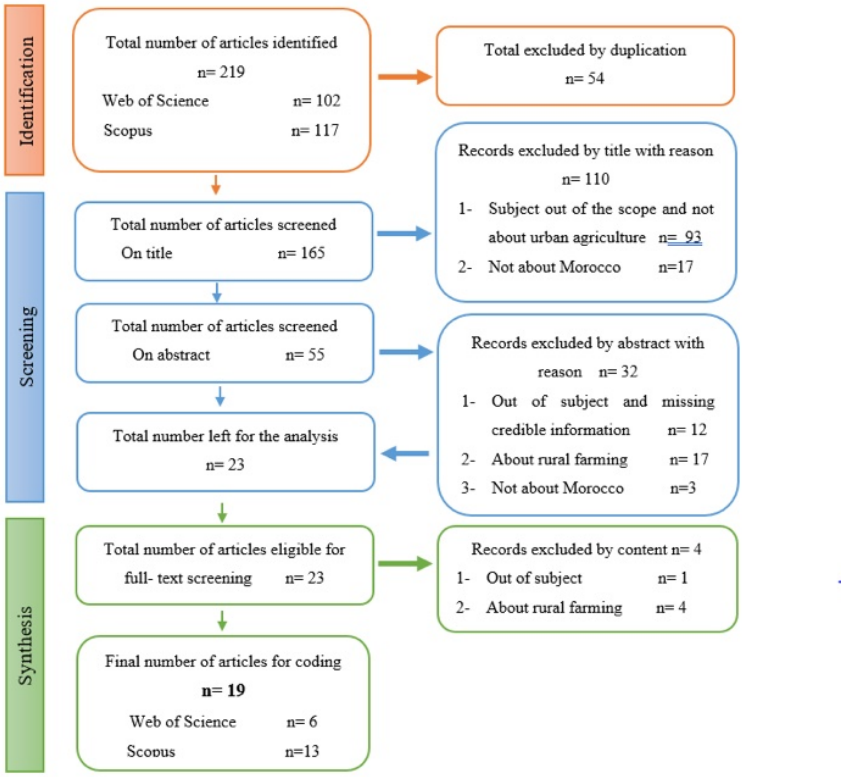


Figure 1: Prisma Diagram on the selection of articles related to Urban Agriculture in Morocco. Source: own work

Software employed

A qualitative synthesis approach was adopted to map the results of the included articles. To this end, the software used in this analysis is NVivo, one of the most widely used qualitative data management programs. NVivo has features such as character coding, rich text capabilities, and multimedia functions that are crucial for qualitative data management [9]. In addition, it allows researchers to process large amounts of data with greater transparency and offers the possibility of verifying the reliability of coding by members of the research team [11].

3 Results and Discussion

Morocco, located in North-West of Africa, is a Mediterranean country stretching along the northern coast of the Atlantic Ocean, covering an area of approximately 710,850 square kilometers [12]. Morocco is divided into 12 Administrative Regions and had a population of approximately 33.8 million in 2022, more than half of whom lived in urban areas [13]. Morocco is experiencing a notable trend toward urbanization, with migration from rural to urban areas. Indeed, between 2004 and 2016, the urbanization rate rose from 55% to 60.2% [14].

3.1 Urban Agriculture over the years:

This systematic review was conducted without any exclusion criteria, without any year restrictions, and without any restrictions on the type of article. However, the total number of articles relevant to our research question was 19. It is nevertheless necessary to assess the year in which urban agriculture practices first appeared in Morocco and its evolution over the years (Figure 2).

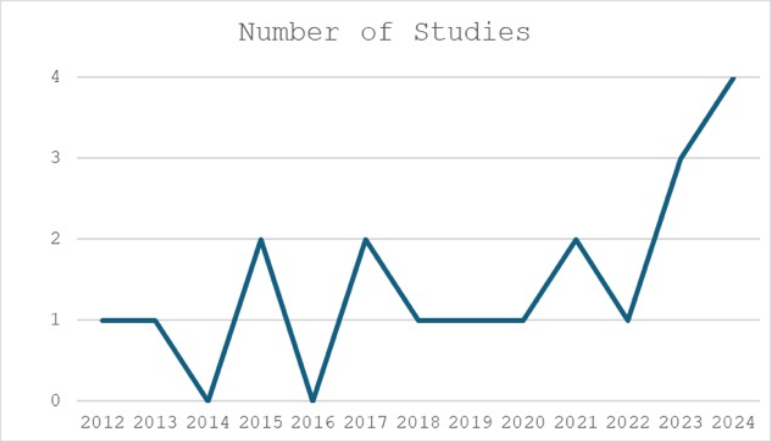


Figure 2: Number of studies related to Urban Agriculture in Morocco over the years. Source: own work

Figure 2 shows that urban agriculture practices first appeared at university level in Morocco in 2012, which means that this concept is new and has not been developed for many years. Furthermore, over the years, the line of research has been unstable, until 2022, when urban agriculture received very strong support compared to previous years, and when it became increasingly important until 2024. This clearly shows us that urban agriculture in Morocco is not very developed and that, until recent years, it has been becoming increasingly involved in the country.

3.2 Distribution of Urban Agriculture in Morocco

Based on the NVivo software coding process applied to the articles selected for this systematic review, the results show that urban agriculture is not very widespread in Morocco, where it accounts for fewer than 16 cases in the whole country. In addition to these cases, four of the 19 studies evaluated urban agriculture in Morocco, but in a general context and without focusing on specific urban cases. These articles explain in more detail the conditions that must be met to ensure the effective and well-structured implementation of these practices.

Urban agriculture is most developed in the Rabat-Salé-Kénitra Region, followed by the Marrakech-Safi Region and the Fès-Meknès Region. The Rabat-Salé-Zemmour-Zaer Region is considered a large community of producers and consumers who are very active in organic farming or agroecology thanks to local initiatives, including community farming, farmers' markets, etc., particularly in Rabat, the Moroccan capital [15]. On the other hand, other cities belonging to the same region, such as Kenitra, Salé, etc., are also characterized by a high level of activity [16, 17]. As for the Marrakech-Tensift-Al Haouz Region, located in south-west of Morocco, in the foothills of the Atlas Mountains, it is considered one of Morocco's main agricultural areas and is widely known for its extensive agricultural activities. Meanwhile, Casablanca is the largest urbanized region in the Kingdom of Morocco, home to 60% of Moroccan industry, and is also the largest port in North Africa [18].

3.3 Challenges and limitations regarding Urban Agriculture in Morocco

Morocco faces various challenges and constraints related to the integration of urban agricultural practices in cities. First, social and environmental issues, where population growth is putting increased pressure on resources, with challenges such as adequate housing and infrastructure in the face of climate change [15, 6]. Second, the complexity of land use, where Morocco's diverse landscapes make it difficult to classify land [19] and where small private plots complicate urban development and increase the cost of land, with the influence of commercial sellers complicating the acquisition process [20]. Thirdly, the lack of information hinders the expansion of the organic market, as certified products are sold at high prices due to production costs [21].

Another point to highlight is the participation of farmers, which is essential for the proper functioning and implementation of urban practices, with participants demanding adequate compensation for their commitment and development. Finally, there is the issue of legal recognition, where participatory guarantee systems encounter difficulties, as well as competition with public regulations on organic farming [15].

Overall, for the development of urban agricultural spaces in the Moroccan cities, it is essential to address the challenges related to land acquisition and support producers in their transition for more sustainable production patterns and practices [1].

3.4 Moroccan projects related to Urban agriculture

In Morocco, many projects are being implemented at the national and international levels to support agriculture and ecosystem services, ensure sustainable development goals, etc., and above all to support urban agriculture in cities. The main national project mentioned by several authors is the Green Morocco Plan, adopted in 2008 and set to continue until 2020, which has become a comprehensive policy encompassing all agricultural actions, aimed at providing joint support to small and medium-sized farmers [22]. The Green Morocco Plan has therefore been in force long enough to allow for analysis of the link between its objectives and the actions undertaken [23].
In addition, in 2004, a program entitled "Cities without Slums" was launched throughout Morocco to address a widespread lack of social housing and lead to the eradication of slums and the development of housing construction programs involving partners from the public and private sectors. This interdisciplinary and transdisciplinary research project is being conducted jointly by Moroccan and German researchers and practitioners, and its main objective is to develop energy- and climate-efficient structures in urban growth centers [6].

4 Conclusion

Urban and peri-urban agriculture involves complex interactions between the three aspects of sustainable development [6]. Agriculture is considered one of the Moroccan government's top priorities, accounting for 17% of the country's gross domestic product and employing around 40% of the national workforce [24]. The results of this systematic review revealed that urban agriculture is not widespread in Morocco, but its implementation and application are increasing over the years. However, most of the cities where urban agriculture is practiced are the largest ones, such as Rabat, Casablanca, and Marrakesh, which can be explained by their economic development, university research, and encouragement from local authorities. Nevertheless, in other cities, many forms of agriculture are practiced, such as rural and peri-urban agriculture, which are useful for local consumption but are not classified as urban.

Another important aspect that was evaluated is the amount of funds allocated to the implementation of urban agricultural projects in Morocco to facilitate their execution. Indeed, for the future implementation, it is important to launch studies on the effects of environmental and landscape factors on reproductive performance, to analyze in depth the spatial and temporal dynamics of nest installation, diet, and composition, and to assess landscape dynamics and attributes. This would help guide future urban development actions in favor of this species [25].

Several limitations were identified in this systematic review, through analysis of the final selected articles, such as climate change, land availability, etc., as indicated above. However, it is essential to mention some very important dimensions, namely the political and institutional aspects. Indeed, as already pointed out in 2019 by Lall et al., the urbanization and institutional reform program currently being pursued in Morocco is

characterized by significant gaps in the formulation of certain policies and in the general institutional frameworks aimed at exploiting the potential of urban growth and ensuring spatial equity [26]. In addition, several authors have highlighted the lack of adequate institutional commitment, which has long hampered the integration of urban agriculture into urban development plans [22]. However, recent research shows that Morocco has undertaken a structural reorientation of its agri-food systems, demonstrating a gradual consolidation of policy and institutional coordination in the urban agriculture sector, thus providing an essential reference point for the integration of urban agriculture into broader sustainable development programs [27].

Finally, it should be noted that this study is one of the first to analyze this aspect of urban agriculture implementation in Morocco, as well as to assess its distribution, advantages, limitations, and recommendations for future studies. However, it could serve as a basis for future research and a means to effectively implement urban agricultural practices.

Acknowledgments. This research was supported by FUSILLI Project, which is based on food and natural resources, and which is funded by the European Union's Horizon 2020 Research and Innovation program under grant agreement No.101000717 (https://fusilli-project.eu/). This work has also been funded by the European Union through the Horizon Europe program (HORIZON-CL6-2022-FARM2FORK-01), under the project 'Agro-ecological strategies for resilient farming in West Africa (CIRAWA)', with ID 101084398.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Soulard, C. T., Valette, E., Perrin, C., Abrantes, P. C., Anthopoulou, T., Benjaballah, O., ... & Paoli, J. C. Peri-urban agro-ecosystems in the Mediterranean: diversity, dynamics, and drivers. Regional Environmental Change, 18, 651-662. (2018)
2. Kammoun, A., Ouazzani, N., El Fels, A. E. A., Hejjaj, A., & Mandi, L. Enhancing pollutant removal efficiency in urban domestic wastewater treatment through the hybrid multi-soil-layering (MSL) system: A case study in Morocco. *Water Science & Technology*, *89*(10), 2685-2702. (2024)
3. Rousseau, M., Amarouche, M., & Salik, K. Requiem for a Rural Hinterland: The Contradictions of Laissez-Faire Regulation and the Urbanisation of Grey Areas in Rabat (Morocco). *disP-The Planning Review*, *57*(3), 68-82. (2021)
4. Debolini, M., Valette, E., François, M., & Chéry, J. P. Mapping land use competition in the rural–urban fringe and future perspectives on land policies: A case study of Meknès (Morocco). Land use policy, 47, 373-381. (2015)

5. HAUT COMMISSARIAT AU PLAN. Statistiques Environnementales au Maroc. 2006. Available online: https://www.hcp.ma/ downloads/Statistiques-environnementales-au-Maroc_t22451.html (accessed on 18 September 2021).
6. Kasper, C., Giseke, U., & Han, S. M. Designing multifunctional spatial systems through urban agriculture: the Casablanca case study. In Sustainable food planning: evolving theory and practice Wageningen Academic. (pp. 495-506). (2012)
7. Hammelman, C. Challenges to supporting social justice through food system governance: examples from two urban agriculture initiatives in Toronto. *Environment and Urbanization*, *31*(2), 481-496. (2019)
8. Camps-Calvet, M., Langemeyer, J., Calvet-Mir, L., & Gómez-Baggethun, E. Ecosystem services provided by urban gardens in Barcelona, Spain: Insights for policy and planning. *Environmental Science & Policy*, *62*, 14-23. (2016)
9. Zamawe, F. C. The implication of using NVivo software in qualitative data analysis: Evidence-based reflections. Malawi Medical Journal, 27(1), 13-15. (2015)
10. James, K.L.; Randall, N.P.; Haddaway, N.R. A methodology for systematic mapping in environmental sciences. Environ. Evid., 5, 7. (2016)
11. Ozkan, B.C. Using NVivo to Analyze Qualitative Classroom Data on Constructivist Learning Environments. 2004. Available online: http://www.nova.edu/ssss/QR.QR9-4/ozkan.pdf (accessed on 29 December 2023).
12. Berkat, O., & Tazi, M. Country pasture/forage resource profiles. *MOROCCO: FAO*. (2006)
13. Cohn, T. H., & Hira, A. A. International Development. In Global Political Economy (pp. 368-415). Routledge. (2025)
14. Thaiki, M., Bounoua, L., & Cherkaoui Dekkaki, H. Using satellite data to characterize land surface processes in Morocco. *Remote Sensing*, *15*(22), 5389. (2023)
15. Romagny, B., Aderghal, M., Auclair, L., Ilbert, H., & Lemeilleur, S. From rural to urban areas: new trends and challenges for the commons in Morocco. *The Journal of North African Studies*, *28*(1), 57-74. (2023)
16. Zouahri, A., Dakak, H., Douaik, A., El Khadir, M., & Moussadek, R. Evaluation of groundwater suitability for irrigation in the Skhirat region, Northwest of Morocco. *Environmental Monitoring and Assessment*, *187*, 1-15. (2015)
17. Rousseau, M., Amarouche, M., & Salik, K. Requiem for a Rural Hinterland: The Contradictions of Laissez-Faire Regulation and the Urbanisation of Grey Areas in Rabat (Morocco). *disP-The Planning Review*, *57*(3), 68-82. (2021)
18. UN-Habitat. World Cities Report: Urbanization and Development, Emerging Futures; UN-Habitat: Nairobi, Kenya, (2016)
19. Kusi, K. K., Khattabi, A., & Mhammdi, N. Analyzing the impact of land use change on ecosystem service value in the main watersheds of Morocco. *Environment, Development and Sustainability*, *25*(3), 2688-2715. (2023)
20. Okacha, A., Salhi, A., Abdelrahman, K., Fattasse, H., Lahrichi, K., Bakhouya, K., & Mondal, B. K. Balancing Environmental and Human Needs: Geographic Information System-Based Analytical Hierarchy Process Land Suitability Planning for Emerging Urban Areas in Bni Bouayach Amid Urban Transformation. *Sustainability*, *16*(15), 6497. (2024)
21. Marzouk H., Gbemenou B.S. Consommation des produits biologiques: analyse de marchés et comportement du consommateur marocain. Rabat: Institut Agronomique et Vétérinaire Hassan II. (2014)
22. Chehbouni, G. Assessing the Alignment and Implementation Challenges of the Sustainable Development Goals (SDGS) 2030 in Morocco: A Relevant Inquiry. *Journal of Sustainable Development Law and Policy (The)*, *15*(2), 186-223. (2024)

23. Faysse, N. The rationale of the Green Morocco Plan: missing links between goals and implementation. *The Journal of North African Studies*, *20*(4), 622-634. (2015)
24. Schilling, J., Freier, K. P., Hertig, E., & Scheffran, J. Climate change, vulnerability and adaptation in North Africa with focus on Morocco. Agriculture, Ecosystems & Environment, 156, 12–26. https:// doi. org/ 10. 1016/j. agee. 2012. 04. 021 (2012)
25. Hmamouchi, M.-J., Agharroud, K., Dahmani, J., Hanane, Saâ., Landscape and coloniality are robust predictors of White Stork nest habitat selection in a coastal urban environment, Estuarine, Coastal and Shelf Science, doi: https://doi.org/10.1016/j.ecss.2020.106835. (2020)
26. Lall, S., Mahgoub, A., Maria, A., & Touati, A. Leveraging Urbanization to Promote a New Growth Model While Reducing Territorial Disparities in Morocco: Urban and Regional Development Policy Note. World Bank Publications. (2019)
27. de Souza, Poliana M., Aziza Zeghmane, Dadaie Daihi, Doriane D. Ouedraogo, Fadwa ElKindoussy, Fidelia Gandiya, Hammed O. Rasaq et al. "Morocco's agrifood transformation: A regional model for sustainable agricultural development." Outlook on Agriculture: https://doi.org/10.1177/00307270251360431. (2025)

Artificial Intelligence in Digital Public Policies in Latin America: Tensions, Gaps, and Transferable Lessons

Felipe Bonilla-Sánchez[1][0009-0001-8674-8209],
Adrián Cabral Martínez[1][0009-0008-5314-6333], and
Elmer Iván Sánchez Rabadán[1][0009-0006-0071-3168]

[1]Facultad de Contaduría, Administración e Informática,
Universidad Autónoma del Estado de Morelos, Cuernavaca, México
{fbonilla, adrian.cabral, elmer.sanchez}@uaem.mx

Abstract. The application of artificial intelligence in the development of public policies in Latin America has become a field of increasing academic and practical significance, given its implications for transparency, equity, and the democratic legitimacy of governments. However, the literature presents fragmented results that make it challenging to identify regional patterns and lessons applicable to specific contexts such as Mexico. This article presents an integrative theoretical review that combines conceptual synthesis with contrasts between approaches and regional empirical evidence, organizing the field into axes and subthemes that interact with one another. In conclusion, five realistic fronts are proposed, directly linked to the identified gaps. Based on the identified gaps and tensions, five research lines are proposed regarding public trust and the quality of decision-making.

Keywords: Digital Public Policies; Artificial Intelligence in Government; Algorithmic Governance; Transparency and Accountability; Latin America.

1 Introduction

Artificial intelligence (AI) has become one of the most significant topics for the design and evaluation of public policies over the last decade, due to its potential to transform governmental processes, optimize resource allocation, and enhance the quality of public services. In Latin America, the emergence of national AI strategies reflects both the influence of international frameworks and the region's unique characteristics. In this context, Correa [4] introduces a qualitative approach that recovers the situated narratives of users, developers, and decision-makers in Chile, revealing the importance of incorporating local perceptions into AI governance, beyond formal regulatory frameworks. Faúndez-Ugalde et al. [2] examine the Brazilian experience in the use of AI for taxation, highlighting regulatory implications and risks of bias that illustrate tensions between revenue efficiency and the protection of rights. This way, recent studies converge in highlighting the centrality of ethics and transparency as guiding principles, as well as the need to strengthen institutional capacities and foster social trust in the use of algorithms. However, recent studies diverge in terms of implementation mechanisms, with some prioritizing the adoption of solid and anticipatory regulatory mechanisms.

Governance and Design of AI Policies in Latin America. National artificial intelligence policies in Latin America reveal a persistent dilemma between the pursuit of global standards and the diverse social and economic realities. Comparing several cases, Filgueiras [7] argues that the policy formulation processes deepen static pieces of bureaucracy.

Multilevel Governance and Multi-Actor Arrangements. Public participation legitimizes algorithmic governance and reduces substantive informational asymmetries. National strategies describe consultations and co-creation, but their performativity depends on how governmental narratives articulate public values, risks, and benefits as perceived by heterogeneous key actors across multiple levels and territories [10, 11]. Regional evidence suggests that innovation. Even when multisectoral roundtables are announced, the evidence shows limited traceability between citizen contributions and regulatory decisions, reinforcing perceptions of tokenism and weakening trust in the promise of participatory algorithmic governance [10, 18]. In Chile, institutional closures of regulatory controversies reveal dependence on political timing, which hinders the stabilization of deliberative processes.

Public Participation, Legitimacy, and Narratives. Public participation is crucial in legitimizing the use of algorithms in administrative processes. Guenduez and Mettler [10] demonstrate that governmental narratives establish legitimacy through metaphors of progress, modernization, and innovation, thereby shaping perceptions of technological risks and benefits within institutional processes. Hjaltalín and Sigurðarson [11] argue that public value is strengthened when narratives integrate transparency, equity, and rights protection rather than focusing solely on efficiency and economic competitiveness. Correa [4] documents situated narratives in Chile, where local actors highlight concerns about bias, surveillance, and the cultural appropriation of artificial intelligence in fragile democratic contexts. These findings reinforce the notion that public participation cannot be limited to formal consultations but must reflect concrete experiences and local perceptions. As Correa [3] maintains, narrative becomes a constitutive component of algorithmic governance, beyond abstract regulatory frameworks. Ricaurte [21] warns of algorithmic governmentality in Latin America, where the discourse of innovation may mask social control practices and reinforce persistent authoritarian structures.

The research concurs in affirming that narrative is not an embellishment but a constitutive element of AI governance and its political legitimacy. When participation is genuine, official narratives reflect social demands and foster trust in the algorithmic decision-making processes employed by government. However, if consultations are formal but non-binding, credibility gaps are reproduced and public confidence in responsible institutions is eroded. Consequently, narratives and participation form an inseparable pair, determining the degree of legitimacy that AI policies achieve. Evidence shows that without traceability, narratives risk becoming symbolic devices of superficial legitimation, without substantive effects on policy design. Official narratives emphasize inclusion and ethics, although in practice, they are often limited to non-binding

formal consultations. The contrast between aspirational discourse and concrete participation mechanisms reveals tensions between tokenistic legitimacy and procedural legitimacy [4, 10, 18].

In Mexico, the lack of traceability mechanisms means that citizen contributions are not reflected in final policies, weakening public trust and increasing perceptions of political exclusion. An alternative explanation is that narrative is used as a tool of symbolic governance, substituting solid institutional practices with promises of ethics and formal participation. This gap undermines the ability of citizens to accurately assess the impacts of algorithmic systems adopted by the public sector. Although there is consensus on the need for social legitimacy, the studies diverge on the means to achieve it: discursive narratives, legal frameworks, or binding deliberative participation. The controversy reflects the structural tension between political communication and democratic governance in the adoption of artificial intelligence in Latin America. In summary, public involvement in artificial intelligence necessitates moving beyond symbolic narratives and advancing toward binding mechanisms of effective procedural legitimacy within Latin American institutional contexts. For Mexico, the lesson is to institutionalize binding deliberative processes and ensure narrative traceability, avoiding legitimacy being reduced to rhetoric without real effects on public policy.

In this context, this article presents a structured search conducted in indexed databases, such as Scopus, Web of Science (WoS), SciELO, and Redalyc, which cover the social sciences and public administration. The search captures relevant literature in the field of digital government and public policy.

In this context, the present study aims to provide a comprehensive understanding of how IA is being integrated into public policy frameworks across Latin America. The article is organized into four main sections. The first presents the methodological design and criteria used to select and analyze relevant literature. The second examines governance and policy design experiences in the region. The third discusses sectoral cases and the global–local tensions that shape AI adoption. Finally, the fourth section offers conclusions and transferable lessons for the Mexican institutional context.

2. Methodology of Review

This article is based on an integrative theoretical review with a systematic orientation, following transparency standards established for literature reviews in the social sciences and public policy fields. The methodological approach follows a qualitative and analytical perspective, aiming to identify conceptual patterns, theoretical frameworks, and emerging trends across recent studies on AI governance and public administration. The proposed approach seeks to analyze and synthesize research results to construct an articulated understanding of the field []. The main sources of information were the scientific databases Scopus, Web of Science (WoS), SciELO, and Redalyc, covering the domains of public administration, governance, and applied social sciences. The period analyzed ranged from 2020 to 2025, in order to capture the most recent strategies and regulatory frameworks concerning IA in Latin America.

The search strategy combined keywords in Spanish and English—*"inteligencia artificial," "políticas públicas," "gobernanza," "transparencia algorítmica," "América*

Latina," and *"public policy."* The inclusion criteria were: (i) articles published in peer-reviewed journals; (ii) focus on governance, public policy, or artificial intelligence; and (iii) regional relevance for Latin America. Gray literature, non–peer-reviewed reports, chapters lacking empirical or conceptual evidence, and documents without full-text access were excluded. From the initial set of results obtained across the four databases, 26 studies met all eligibility criteria. The final thematic classification grouped the selected texts into three axes:(a) governance and AI policy design;(b) sectoral experiences and institutional transferability; and (c) global–local tensions in the adoption of regulatory frameworks.

The classification is consistent with other regional systematic reviews that apply similar thematic coding processes [24], [25]. The analytical coding was conducted through full reading and extraction of key variables (author, country, methodological approach, findings, and contributions). Three analytical categories emerged from this process: legitimacy, transparency, and multilevel governance, contrasted against international frameworks such as those of the OECD, ECLAC, and UNESCO. Although this review does not follow the complete PRISMA protocol, it adopts its principles of transparency, traceability, and interpretive validity, aligning with recent practices in integrative theoretical reviews in Latin America [1], [25].

The proposed methodology ensures a verifiable systematization of the review process, reinforcing confidence in the findings and their relevance to public policy analysis.

3 State Capacities, Transparency, and Data Management

3.1 Institutional Capacities and Maturity for Governmental AI

Institutional capacities determine the maturity of AI in public administration. Valle-Cruz et al. [26], and Luna-Reyes and Harrison [15] remark that maturity depends on politics. Regional evidence indicates persistent limitations in funding, digital skills, and interinstitutional coordination, which hinder the effectiveness of policies and impede the consolidation of capacities. Case studies, such as that of the State of Mexico, confirm that, although promising pilots exist, institutional capacities remain lagging in strategic areas. Millán-Vargas et al. [16] remark that institutional capacities are necessary but still insufficient in the region. In contrast with several European contexts, where regulatory frameworks set minimum standards for data, funding, and talent, Latin American strategies rely on personal leadership and unstable external financing. Hjaltalín and Sigurðarson [11] stress that without institutionalized public values, technical capacities do not guarantee legitimacy or equity in the practical implementation of algorithmic systems. Olsen and Christensen [17] warn that the lack of algorithmic transparency reflects organizational weaknesses, given the absence of clear records of decisions or mandatory independent audits. An alternative explanation lies in path dependency from previous digitalization processes, which limits institutional modernization and reinforces bureaucratic inertia that is difficult to reverse. The Mexican case demonstrates that sectoral pilots operate as "isolated innovation islands," but the scarcity of hybrid talent hinders scaling solutions and consolidating a coherent and sustainable na-

tional ecosystem. For Mexico, the key lesson is to invest in hybrid talent, data governance, and standardized maturity metrics, ensuring sustainability beyond technological pilots.

3.2 Algorithmic Transparency, FOI, and Accountability

Algorithmic transparency is a basic condition for legitimacy in democratic digital government. Olsen and Christensen [17] note that freedom of information regimes were designed for documents, not for models, and therefore require extensions that incorporate explainability, data traceability, and effective rights to human review. Zuiderwijk et al. [28] argue that an accountability architecture must document the entire system lifecycle, encompassing data, design, validation, and performance across subpopulations, while also defining organizational responsibilities and requiring periodic, independent audits. Regional evidence suggests creating public registries of algorithms, explainability guidelines, and tiered supervision procedures according to sectoral risk, aligned with international standards adapted to the Latin American context [18, 28], stressing that meaningful transparency requires data lineage, model cards, versions, disaggregated metrics, and thresholds of minimum acceptable performance for automated administrative decisions. Likewise, Olsen and Christensen [17] insist that clear avenues must exist to challenge outcomes, with timely human intervention and understandable justifications accessible to non-experts.

In Chile, Humeres [13] showed that the political closure of regulatory controversies evidenced how the absence of procedural obligations can weaken substantive accountability in high-impact systems. Ricaurte [21] warns that in Latin America, formal transparency without social control may normalize surveillance practices and reinforce institutionalized power asymmetries. OECD [18] recommends ex ante and ex post algorithmic impact assessments. Guenduez and Mettler [10] caution that without these mechanisms, transparency remains declarative and operates as a legitimizing narrative rather than an effective guarantee of citizens' rights. The region faces persistent technical opacity and legal weaknesses in the operation of public algorithms. Olsen and Christensen [17] explain that documentary transparency frameworks often conflict with trade secrets and proprietary information, referred to as "black boxes."

Zuiderwijk [28] demonstrates that numerous agencies publish open data but often omit key details, including lineage, cleaning criteria, sampling methods, and model versions, rendering it impossible to reconstruct decisions for genuine administrative redress. Olsen and Christensen [17] note that the explanations generated are often technical and targeted at experts, failing to meet standards of comprehensibility for citizens. Luna-Reyes and Harrison [15] argue that the scarcity of hybrid talent, limited budgets, and institutional heterogeneity obstruct the operationalization of meaningful transparency and continuous auditing. To address this, Zuiderwijk [28] and Olsen and Christensen [17] propose a “minimum viable” transparency package, which includes a public registry, impact assessment, model cards, periodic blind testing, and mandatory independent external audits. Zuiderwijk [28] concludes that when such obligations are adopted, public trust increases and externalities of error, bias, and arbitrariness in algorithmic administration are reduced. For Mexico, the transferable pathway involves an

operational regulatory package, comprising an algorithm registry, impact assessments, model cards, external audits, and procurement with verifiable obligations of access and explainability.

3.3 Data Management, Bias, and Equity

Data management has a significant impact on the public perception and legitimacy of artificial intelligence. Zuiderwijk [28] emphasizes that data, design, validation, and performance must be thoroughly documented, including disaggregated metrics by subpopulations and reproducible versions of the complete system. The OECD [18] recommends national data strategies, open standards, and trustworthy data-sharing agreements across all levels of government and strategic sectors in Latin America. Ricaurte [21] warns that without adequate social controls, algorithmic governmentality may normalize surveillance practices, reinforce asymmetries, and erode the autonomy of historically marginalized groups. In the same vein, [24] documents the precariousness of data labor sustaining annotation and cleaning in the region, with direct implications for bias, quality, and distributive justice in public systems. Complementarily, Tubaro et al. [24] compare experiences in other Latin American countries. These conditions impact linguistic representativeness, territorial coverage, and the updating of administrative series, which are essential for evidence-based policymaking. As Grohmann [9] points out, data governance in Latin America requires recognition of the historicity of epistemic and material inequalities; ignoring them reinforces systemic biases and limits states' ability to design inclusive public policies. In Latin America, open data portals are abundant; however, they often lack a documented lineage, clear cleaning criteria, and links to auditable administrative decisions by independent third parties.

Public procurement can require datasheets and model cards, guaranteeing traceability, access to testing, and replication rights by independent oversight authorities. These measures integrate equity principles with institutional infrastructure, reducing structural biases and strengthening social trust in governmental use of algorithms. Critical gaps persist in documentation, representativeness, and rights to meaningful algorithmic access. Olsen and Christensen [17] explain the importance of the freedom of access to information. The OECD [18] warns that public procurement without access clauses perpetuates technological dependence, thereby limiting external audits and democratic control over systems that affect fundamental rights. Zuiderwijk [28] emphasizes the importance of comprehensive documentation, yet regional practice lacks standards for datasheets and model cards that align with sectoral risk. Data representativeness is unequal across territories, languages, and groups, which biases models and decisions; without mandatory disaggregated metrics, inequity remains invisible and is reproduced bureaucratically. Tubaro et al. [25] warn that precarious data labor undermines quality and exacerbates bias, demanding labor policies and contracting standards that protect those sustaining public informational infrastructure. An alternative explanation is the lack of hybrid capacities in agencies, which prevents translating abstract equity principles into verifiable procedures and concrete operational objectives. The OECD [18] recommends a national data strategy with trustworthy data-sharing agreements, man-

datory lineage, and pre-contractual and periodic bias testing as part of public procurement. Olsen and Christensen [17] and Zuiderwijk [28] argue that this package should include model cards, disaggregated metrics, external audits, and access clauses, moving beyond formal transparency toward material guarantees of equity. For Mexico, effective data management requires a mandatory documentary infrastructure, disaggregated metrics, and independent audits that translate principles into practical practices of algorithmic governance.

3.4 Networked Governance and Risk Regulation

In addition to the approaches reviewed, a conceptual framework is integrated that articulates networked governance and risk regulation regimes to interpret the identified tensions and gaps critically. The analysis of AI in digital public policies in Latin America requires not only describing experiences but also situating them within conceptual frameworks that enable the interpretation of their tensions and gaps. In this regard, networked governance and risk regulation regimes provide valuable tools for engaging with qualitative literature and strengthening the interpretation of findings obtained from the systematic review. Networked governance is based on the premise that public decisions no longer depend solely on hierarchical state structures but emerge from interactions among governments, international organizations, technology companies, civil society, and academia. Rhodes [20] defined it as a model that partially replaces hierarchical control through relationships of interdependence, negotiation, and cooperation. Subsequently, Bevir and Rhodes [1] developed the decentred theory variant, which highlights that the meanings attributed by actors determine how governance networks are configured and operate.

In the Latin American context, this approach allows for a reinterpretation of several findings from the review. For example, the strong dependence on international regulatory frameworks (UNESCO, OECD, IDB) in contrast to limited national capacities reflects that governance networks operate asymmetrically, with global actors exerting greater influence than local institutions. For their part, Hood et al. [12] developed the framework of risk regulation regimes, distinguishing between information and monitoring, standard-setting, and behavior modification. However, Sánchez and Nyabola [22] point out that attempts to harmonize regulatory frameworks in Latin America with international models reflect tensions like those experienced in Africa, where the uncritical transfer of global standards without local adaptation undermines both the regulatory relevance and social legitimacy of AI policies.

At the level of information, the reviewed studies show that many national strategies lack verifiable mechanisms for ex ante and ex post evaluation, limiting the ability to measure real impacts. Regarding standard-setting, the evidence reveals an uncritical transfer of international guidelines with limited adaptation to local contexts, compromising the relevance of regulatory frameworks. Finally, in terms of behavior modification, documented cases indicate that most policies remain declarative, without producing substantive changes in administrative practices or in adequate guarantees of algorithmic transparency.

This framework thus enables the reinterpretation of the findings not merely as descriptive shortcomings, but as deficits in the regulatory capacity of Latin American states to address AI-related risks. For Mexico, the transferable lessons suggest designing an operational regulatory package that includes monitoring with verifiable indicators, standards tailored to the federal reality, and binding compliance mechanisms. The integration of both approaches offers a solid analytical framework: networked governance sheds light on the dynamics of participation and legitimacy, while risk regulation structures tensions around monitoring, standard-setting, and compliance. Together, they provide a conceptual framework that enables a critical interpretation of AI policies in Latin America and the derivation of transferable pathways for Mexico, oriented toward strengthening democratic legitimacy, ensuring adequate transparency, and enhancing institutional capacity.

4. Sectoral Experiences and Transferability to Mexico

4.1 Tax Administration and Smart Auditing

Smart tax auditing has become a priority field for Latin American governments. Valle-Cruz et al. [26] demonstrate how artificial intelligence facilitates the optimization of tax collection processes by identifying patterns of evasion and tax fraud in large volumes of administrative data**.** In Brazil, the OECD [18] documents significant advances in the automation of tax audits, integrating predictive analytics and blockchain to reduce structural evasion and improve the auditing of large taxpayers. These cases show how artificial intelligence not only increases state revenues but also strengthens public trust in perceived tax fairness. The revenue potential of artificial intelligence contrasts with substantive risks of bias and opacity**. Olsen and Christensen** [17] note that, compared to European countries with independent algorithmic audits, Latin America is characterized by proprietary black boxes that lack citizen oversight and significant institutional counterbalances. **Zuiderwijk** [28] warns that the lack of comprehensive documentation of data and models makes it difficult to reconstruct fiscal decisions, thereby undermining the fundamental rights of taxpayers affected by automation. Ricaurte [21] warns that algorithmic governmentality may transform AI in taxation into a mechanism of social control, reinforcing inequalities in the application of tax sanctions. The OECD [18] emphasizes that without clear equity rules, algorithms may focus auditing on small sectors, leaving prominent corporate tax evaders unaddressed. The primary gap is the absence of mandatory regional guidelines that regulate explainability, bias, and appeal rights in contemporary algorithmic taxation. For Mexico, this gap entails risks of arbitrariness and social distrust in a particularly sensitive domain for governmental legitimacy and revenue stability.

In conclusion, AI in taxation must advance in tandem with external audits, clear equity rules, and effective mechanisms for institutionalized human review. For Mexico, the central lesson is to balance revenue efficiency with social legitimacy through frameworks that guarantee auditing, explainability, and verifiable tax equity.

4.2 Organizational Adoption and Barriers in Subnational Governments

Subnational adoption depends on organizational capacities and stable coordination arrangements. Millán-Vargas et al. [16] show in the State of Mexico that AI pilots improve risk identification, although they face shortages of talent, interoperable data, and sustained operating budgets. Valle-Cruz et al. [26] propose embedding artificial intelligence into the public policy cycle. At the state and municipal levels, procurement processes limit learning by privileging packaged solutions, thereby reducing public control over models and relevant strategic data. In summary, subnational adoption progresses when capacities, procurement processes, and data governance are designed as a coherent system, rather than as isolated initiatives disconnected from inter-institutional coordination and comparative evaluation. These conditions either consolidate or undermine the ability to scale and institutionalize AI practices in subnational governments. Dominant barriers are institutional rather than merely technological in Latin American subnational governments.

In contrast with European frameworks that establish standards for data, financing, and talent, Latin American subnational governments face resource asymmetries, fragmentation, and procurement processes with restrictive clauses on technical transparency in the text "isolated innovation islands" without cross-sector integration, making comparative evaluation and systematic organizational learning among local agencies difficult [17, 16]. An alternative explanation lies in path dependency from prior digitalization, which conditions rules, processes, and organizational cultures toward inertia resistant to responsible experimentation. Catalogs of use cases and regulatory sandboxes can reduce legal uncertainty but require metrics and external audits with binding effects on procurement processes. Without intergovernmental cooperation frameworks and earmarked budgets, pilots often expire with the end of political cycles, resulting in the loss of institutional knowledge accumulated within teams and repositories.

Evidence converges that data governance, results-based contracting, and ex ante and ex post evaluation are enabling conditions for solid and scalable adoption. However, debate remains over their optimal sequencing, as institutional capacities, budgetary resources, and local digital maturity condition possible pathways, rhythms, and prioritization of components within multi-year programs. Consequently, subnational adoption requires moving from opportunistic projects to programs with standards, open contracts, and controls that ensure legitimacy, scalability, and cumulative learning. For Mexico, the transferable pathway combines data standards, open contracting with mandatory documentation, and subnational capacity centers that institutionalize learning, evaluation, and the scaling of public AI solutions.

4.3 Contextual Adaptation: Global–Local Tensions and Policy Pathways

The adaptation of artificial intelligence policies faces tensions between global frameworks and local realities. Correa [3]shows that international influence on Latin America generates regulatory hybrids, where global standards are combined with local practices. Complementarily, Isaza [14] compares AI policies across countries in the region, highlighting how strategic documents tend to replicate international frameworks. Vivares and Rodríguez-Ortiz [27] argue that AI can reproduce asymmetric power relations, as countries in the Global South depend on narratives and technologies developed in the

Global North. Humeres [13] demonstrates in Chile that the political closure of regulatory controversies shows how local institutional timing reshapes the adoption of international guidelines into national regulatory frameworks. Ricaurte [21] warns that algorithmic governmentality may consolidate inequalities. Luna-Reyes and Harrison [15] argue that policy pathways must strike a balance between international cooperation and the strengthening of domestic capacities, ensuring institutional sustainability and strategic autonomy. In summary, contextual adaptation involves translating global principles into verifiable operational rules, enabling institutions to sustain legitimacy and democratic resilience in the application of artificial intelligence.

Global–local tensions in artificial intelligence reflect structural asymmetries of power and capacities. Correa [3] explains that regulatory hybrids import external legitimacy but may lack social and cultural grounding, weakening their real effectiveness in Latin American institutional environments. Vivares and Rodríguez-Ortiz [27] note that global multi-actorism tends to perpetuate colonial hierarchies, thereby limiting the capacity of governments and local communities to influence the design of international rules. The OECD [18] acknowledges that the uncritical adoption of global standards generates local tensions, recommending instead the adaptation of frameworks to national and territorial institutional contexts. Humeres [13] observes that in Chile, regulatory controversies are closed politically rather than through citizen deliberation, revealing the limits of transferring global guidelines into local practices. [28]**,** warn that the lack of metrics and datasheets prevents the evaluation of whether international standards are materially met in day-to-day governmental practice. Tubaro et al. [25] argue that precarious labor in data-related tasks reproduces inequity. Luna-Reyes and Harrison [15] suggest that policy pathways should incorporate international technical cooperation, complemented by domestic capacity-building, to mitigate structural dependence. Consequently, contextual adaptation requires recognizing tensions, redistributing capacities, and articulating international standards with local social control mechanisms and inclusive practices of democratic governance. For Mexico, the key lesson is to translate international frameworks into verifiable national rules, strengthen domestic capacities, and avoid dependence on external standards without adapting them to local contexts.

5 Discussion

This article aims to critically analyze the experiences of designing public policies with artificial intelligence in Latin American countries, to identify lessons transferable to Mexico's institutional and social contexts.

5.1 Comparison with Previous Reviews

This review differs from earlier surveys that addressed artificial intelligence in public policy from general or descriptive perspectives. Studies such as those by Zuiderwijk [28]**,** and Valle-Cruz et al. [26] analyzed conceptual frameworks or specific applications of digitalization in government, but without integrating the three dimensions pro-

posed in this article. Furthermore, unlike approaches focused exclusively on international guidelines, such as those of OECD**,** this review incorporates comparisons among regional experiences (Chile, Brazil, Mexico, Uruguay), highlighting not only normative principles but also transferable operational mechanisms. In this way, it provides an integrative and critical approach that complements more normative reviews or those focused solely on the ethics of artificial intelligence [10]**.**

5.2 Critical Synthesis of Findings

To synthesize the regional contrasts identified in the reviewed literature, Table 1 summarizes the main gaps observed in Latin American AI policy implementation and highlights the transferable lessons for Mexico's institutional context.

Table 1: Regional gaps and transferable lessons in IA public policy

Dimension	Observed Regional Gap	Transferable Lesson for Mexico
Institutional capacity	Fragmented institutional structures and limited technical expertise hinder the implementation of AI strategies in the public sector.	Strengthen inter-institutional coordination and develop specialized units for AI governance within public administration.
Transparency and accountability	Algorithmic systems often lack traceability and citizen oversight, leading to weak procedural legitimacy.	Incorporate mechanisms for algorithmic auditability and ensure open data standards in public decision-making.
Regulatory frameworks	Most Latin American countries still rely on soft-law instruments without binding mechanisms.	Promote the development of enforceable regulatory frameworks aligned with OECD and UNESCO ethical standards.
Public participation	Civic engagement remains symbolic or tokenistic, with limited deliberative participation.	Institutionalize participatory mechanisms in the design and evaluation of AI policies.
Digital inequality	Persistent gaps in infrastructure and access to digital education limit inclusive AI governance.	Integrate digital literacy programs and equity-driven policies to prevent algorithmic exclusion.

Note. The table summarizes key findings from the reviewed literature and highlights lessons applicable to Mexico's public sector.

There is consensus that data governance is a structural condition for the legitimate use of IA. Zuiderwijk [28] emphasizes the importance of documenting the entire system lifecycle, which aligns with Olsen and Christensen [17], who propose adapting freedom of information regimes to algorithms by incorporating public registries, impact

assessments, and human review. Their ideas align with OECD [18], which suggests regional standards for openness and interministerial coordination.

There is also convergence that legitimacy does not depend solely on technical efficiency. Guenduez and Mettler [10] emphasize the relevance of public narrative and social participation in the acceptance of artificial intelligence, while Ricaurte [21] warns that, without democratic counterbalances, algorithmic governmentality may result in surveillance mechanisms and social control.

Contradictions arise around global–local transfer. Humeres [13] shows that in Chile, regulatory frameworks are gradually consolidated through participatory processes rather than through direct transfer of external models. Divergences also persist in the way transparency is conceived. Olsen and Christensen [17] advocate algorithmic registries and audits, while some regional frameworks still limit accountability to the publication of open data without full traceability.

The present review contributes by emphasizing the subnational scale. The document "isolated innovation islands" in Mexican local governments that improve tax auditing but lack cross-sectoral integration. Estevez et al. [6] note that these limitations are attributed to the absence of hybrid profiles and weak inter-institutional coordination.

5.3 Theoretical and Practical Implications

The findings allow for a redefinition of theoretical categories. Rather than conceiving of AI ethics in abstract terms, the reviewed studies emphasize the need for distributed accountability throughout the entire sociotechnical cycle. Transparency ceases to be a normative principle and becomes an operational package that includes data lineage, external audits, model cards, and disaggregated metrics [28,17].

In practical terms, regional evidence shows that Mexico must advance toward five operational minimums: (i) a public registry of algorithms; (ii) *before* and *after* impact assessments; (iii) comprehensive and mandatory documentation; (iv) open public procurement with clauses of access and non-opposition to auditing; and (v) multilevel coordination with external audit offices and independent ombuds institutions [18,16]. The identified operational lessons are transferable and can strengthen the country's institutional legitimacy.

5.4 Limitations of the Review

First, temporality, the 2020–2025 period captured recent developments but may have excluded important precedents. Second, source selection was limited to articles with DOIs in indexed journals, which ensures quality but excludes gray literature that may contain applicable practices. Third, the review focused on literature in English, Spanish, and Portuguese, which may have limited its perspectives to other Latin American countries. These limitations were mitigated through the use of explicit criteria, critical classification, and a structured extraction table.

Based on the identified gaps and tensions, five research lines are proposed:

1. Empirically validate minimum transparency and audit packages in high-impact sectors (taxation, social policy, security).
2. Study subnational scaling and interoperable data governance in federations with high institutional heterogeneity.
3. Analyze distributive effects of algorithmic models with disaggregated metrics in vulnerable populations.
4. Compare contractual arrangements and access clauses that reduce dependence on international vendors.
5. Deepen participatory anchoring and the right to human review as effective mechanisms of democratic legitimacy.

In sum, this review distinguishes itself from previous surveys by combining regional comparative analysis with critical conceptual frameworks that reinterpret structural tensions in Latin American algorithmic governance. While studies such as Pereira and Vásquez [19]and Olsen and Christensen [7] have documented the initial advances of AI policies in the region, the present work articulates operational and normative dimensions that transcend the descriptive level. Moreover, the incorporation of evidence on digital labor and social justice in countries such as Argentina, Brazil, and Venezuela [25], along with the comparative synthesis of national strategies conducted by [23], reinforces the need for integrative approaches that link multilevel governance with the material conditions of implementation [23]. In this sense, the central contribution of this study lies in situating the discussion on artificial intelligence in Latin America.

6 Conclusions

The reviewed evidence highlights three decisive factors for designing public policies with AI in Latin America. First, governance and design: national strategies have served as devices of legitimation and orientation, but their effectiveness depends on closing the gap between discourse and execution through multilevel and multi-actor arrangements with clear mandates, metrics, and budgets.

The review addresses its objective by critically analyzing design experiences in the region and identifying lessons transferable to Mexico. Ultimately, its central contribution is to translate abstract principles (such as ethics, transparency, and participation) into verifiable mechanisms that articulate governance, data, and organizational adoption. The cross-reading of evidence and global–local tensions enables the proposal of policy pathways tailored to Mexican federalism and institutional heterogeneity, moving beyond the mere transfer of international frameworks.

This is an integrative theoretical review that combines conceptual synthesis with contrasts between approaches and regional empirical evidence, thereby organizing the field into axes and subthemes that interact with one another. Its utility is threefold: research (clarifying categories and observable variables for comparative studies), teaching (offering a curricular map of algorithmic governance in government), and professional development (translating principles into an implementation checklist for decision-makers and oversight bodies).

In conclusion, five realistic fronts are proposed, directly linked to the identified gaps: (*i*) empirical validation of minimum transparency and audit packages in high-impact sectors; (*ii*) longitudinal studies on subnational scaling, with comparable performance and legitimacy metrics; (*iii*) distributive analyses with disaggregated metrics to assess biases and effects on vulnerable populations; (*iv*) comparison of contractual arrangements and access/audit clauses that reduce dependence on vendors; and (*v*) evaluation of participatory devices—binding deliberation, traceability of inputs, and the right to human review—regarding public trust and decision-making quality. As a final synthesis, the region already possesses the normative components; the challenge is to institutionalize them into procedures, capacities, and contracts that operationalize the promise of AI as a tool for the public interest in Mexico.

References:

[1] Bevir, M., & Rhodes, R. A. W. (2016). Decentred theory, change and network governance. In E. Sørensen & J. Torfing (Eds.), Theories of Democratic Network Governance (pp. 77–91). Palgrave Macmillan.
[2] Faúndez-Ugalde, A., Mellado-Silva, R., & Aldunate-Lizana, E. (2020). Use of artificial intelligence by tax administrations: An analysis regarding taxpayers' rights in Latin American countries. Computer Law & Security Review, 38, 105441.
[3] Correa, M. L. (2025). The shaping of AI policies in Latin America: A study of international influence and local realities. In G. Zavattaro & N. de Kool (Eds.), Public Governance and Emerging Technologies (pp. 221–240). Springer.
[4] Correa, T. (2025). Reimagining AI in Latin America: Situated narratives of stakeholders in Chile. AI & Society.
[5] Criado, J. I., Alcaide-Muñoz, L., & Liarte, I. (2025). Two decades of public sector innovation: Building an analytical framework from a systematic literature review of types, strategies, conditions, and results. Public Management Review, 27(3), 623–652.
[6] Estevez, E., Janowski, T., & Luna-Reyes, L. F. (2024). Governing artificial intelligence in the public sector: Towards a research agenda. Digital Government: Research and Practice, 5(2), 1–19.
[7] Filgueiras, F. (2023). Designing artificial intelligence policy: Comparing experiences in Latin America. Latin American Policy, 14(1), 99–114.
[8] Flórez Rojas, M. L. (2025). The shaping of AI policies in Latin America: A study of international influence and local realities. In G. Zavattaro & N. de Kool (Eds.), Public Governance and Emerging Technologies (pp. 221–240). Springer.
[9] Grohmann, R. (2025). Latin American critical data studies. Big Data & Society, 12(1), 1–12.

[10] Guenduez, A. A., & Mettler, T. (2023). Algorithmic government: Public sector legitimacy and acceptance of decision-making algorithms. Government Information Quarterly, 40(1), 101719.
[11] Hjaltalín, M., & Sigurðarson, H. (2024). Governing with algorithms: Public sector adoption and legitimacy in Iceland. Government Information Quarterly, 41(1), 101914.
[12] Hood, C., Rothstein, H., & Baldwin, R. (2001). The Government of Risk: Understanding Risk Regulation Regimes. Oxford University Press.
[13] Humeres, H. (2025). Algorithmic regulation and political controversies in Chile: Closing debates through institutional shortcuts. Big Data & Society, 12(1), 1–14.
[14] Isaza, C. (2023). What to do on AI policy and why? A Latin American comparative paper. SSRN Electronic Journal.
[15] Luna-Reyes, L. F., & Harrison, T. M. (2024). Building institutional capacity for artificial intelligence in the public sector. Digital Government: Research and Practice, 5(2), 1–15.
[16] Millán-Vargas, G., Sandoval-Almazán, R., & Valle-Cruz, D. (2024). Artificial intelligence in Mexican public administration: Pilots, barriers and institutional capacity. In Proceedings of the 25th Annual International Conference on Digital Government Research (pp. 112–121). ACM.
[17] Olsen, A. L., & Christensen, J. (2024). Artificial intelligence and public sector accountability. Communications of the ACM, 67(3), 64–71.
[18] Organisation for Economic Co-operation and Development (OECD), Economic Commission for Latin America and the Caribbean (ECLAC), Development Bank of Latin America (CAF), & European Commission. (2022). Latin American Economic Outlook 2022: Towards a Green and Just Transition. OECD Publishing.
[19] Pereira, G., & Vásquez, I. (2021). Artificial intelligence public policies in Latin America: Disruption or more of the same? In Proceedings of the 22nd Annual International Conference on Digital Government Research (pp. 403–412). ACM.
[20] Rhodes, R. A. W. (1996). The new governance: Governing without government. Political Studies, 44(4), 652–667.
[21] Ricaurte, P. (2024). Data epistemologies and the politics of artificial intelligence in Latin America. Big Data & Society, 11(1), 1–12.
[22] Sánchez, A., & Nyabola, N. (2024). Comparative legal frameworks for regulating artificial intelligence: A model for harmonizing AI laws in Latin America and Africa. Information & Communications Technology Law, 33(2), 145–163.
[23] Sistema Económico Latinoamericano y del Caribe (SELA). (2025). Artificial intelligence and public policy in Latin America and the Caribbean: Comparison of national strategies. SELA Publications.
[24] Castro-Díaz, L., Nahuelhual, L., & Ochoa-Huespe, J. (2022). A systematic review of social participation in ecosystem services studies in Latin America. Science of the Total Environment, 851, 158–172. https://www.sciencedirect.com/science/article/pii/S0048969722016163

[25] Sancho, J. R., Pineda, D. A., & Torres, L. (2023). University social responsibility in Latin America: A systematic review using PRISMA methodology. Science, Culture and Technology Journal, 6(2), 85–99. https://sct.ageditor.ar/index.php/sct/article/view/855
[26] Alvarenga, J. M., Garcia, M. E., & Santos, F. R. (2024). The integrative review in studies of educational public policy. Revista Brasileira de Educação, 29(1), 1–15. https://www.scielo.br/j/rbedu/a/CYXLcVCCkWLBQD7H45ZbWxI

From Digital Government to Smart and Sustainable Cities in Latin America

Juárez Merino, Miguel Ángel[1][0000-0001-5562-7677]

[1] Universidad Nacional Autónoma de México, Cuernavaca, Mor 62210, Mexico
dr.migueljm@gmail.com

Abstract. In the current global scenario, the development of smart and sustainable cities is one of the most relevant avenues for finding solutions to the problems arising from the climate crisis and environmental degradation. In this context, local governments seek to protect the economic, social, and cultural rights linked to these phenomena using innovative digital technologies, placing citizens at the center of action, and fostering environmental awareness.

Building on this, the present study seeks to identify the most advanced smart cities in Latin America and to understand which strategies are proving effective within the regional context, as well as the critical factors that may be delaying the transformation process. Based on the analysis of international indicators and indices, the necessary data were extracted to conclude that the level of maturity in smart and sustainable urbanization in Latin America is, on average, low. Consequently, relevant reflections are presented to propose criteria that could raise the effectiveness of the transition toward sustainability.

Keywords: Sustainability, Smart Cities, Latin America.

1 Introduction

This research aims to link the design of smart cities with sustainable development. To that end, the theoretical foundations that encompass the idea of using digital technologies as a tool to improve public management and strengthen integration between citizens and local governments are established, including other actors such as the private and social sectors.

Likewise, the foregoing is related to economic, social, cultural, and environmental rights, with particular emphasis on the human right to a healthy environment, for which the State is obliged to guarantee access to a dignified life and to provide the necessary conditions for the care of the environment and its natural resources.

Regarding the object of study, this work focuses on Latin America and investigates progress in smart urbanization and sustainability, all using data provided by the main global measurement indices and by comparing the methodologies employed.

To increase the operational precision of this study, the guiding research question is to identify, based on indices and specialized literature, which cities in Latin America show greater maturity toward smart and sustainable models, to observe the institutional

strategies that have supported that advancement and to recognize the structural factors that act as barriers. To aid analytical clarity, the main conceptual variables used are defined so that smart and sustainable city maturity is understood as the institutional and technical capacity to deploy digital solutions aimed at sustainability, digital government level refers to the degree of digitization of services and administrative processes and governance quality encompasses aspects of transparency, interinstitutional coordination, and public management. These definitions situate the comparative analysis carried out using the selected indices and documentary sources.

Finally, the results are discussed, and the resulting conclusions are presented to provide an overview of the degree of progress the Latin American region has made in the urban transformation process, and to reflect on the causes that may limit this process, as well as the advances that have been achieved.

2 Theoretical framework

The emergence of smart cities around the world is a phenomenon resulting from diverse causes that have converged to give rise to a type of urbanization that proposes the use of digital technologies as a basis for developing plans and solutions that effectively address citizens' needs. In this sense, the two main factors that have influenced the construction of smart spaces are, on the one hand, the growing urbanization trend of recent decades and, on the other hand, the rise of digital technologies stemming from the Fourth Industrial Revolution.

From this it is possible to argue that the accelerated growth of urban concentrations has modified population and economic distribution worldwide, exacerbating problems linked to poverty, increasing inequalities, and widespread discontent—especially in regions such as Latin America, where these vulnerabilities were already preexisting (Olivera 2020).

According to United Nations estimates, by 2050 the global urban population will have doubled; therefore, economic, social, and cultural activities, as well as their environmental and humanitarian repercussions, will increasingly be concentrated in cities, which raises sustainability challenges in terms of housing, infrastructure, basic services, food security, health, education, employment, safety, and natural resources (UN-Habitat 2017).

For his part, according to Schwab (2016), the world is undergoing a Fourth Industrial Revolution characterized by intensive use of the internet, more powerful and cheaper digital tools, and a surge in the use of artificial intelligence, robotics, and machine learning. In this context, two problems inherent to public administration arise: the functioning of bureaucracy and the relationship with the community (Montecinos 2021).

The Fourth Industrial Revolution represents a great opportunity for public administration and governments to transform toward agile, effective, and transparent management (Llanes-Font et al. 2020). This new digital era provides a platform for governments to focus their efforts on citizens, better understanding their needs and preferences through data analysis and citizen participation. This not only represents a pathway for modernizing public administration but also tends to strengthen the relationship between

government and citizens, fostering greater collaboration and trust in the management of public affairs.

Criado (2016) defines the above as smart public governance, a concept that refers to an approach in which governments rely on the most innovative technologies to create new forms of interaction with the private and social sectors, with the aim of sharing resources and responsibilities and seeking solutions to public problems.

On this basis, different theories have sought to explain how governance has acquired new nuances. Following the criterion of Luna, Gil and Sandoval (2015), the theoretical approaches that best fit the explanation of digital government are structuration theory (DeSanctis & Poole 1994; Giddens 1984; Orlikowski 1992), the sociotechnical approach (Kling 2000), and institutional theory (Fountain 2001).

In structuration theory, technology itself becomes an additional institutional element that facilitates or hinders the performance of activities by actors in the system; the sociotechnical approach, for its part, understands how different factors impact the use of digital technologies in government and what effects they produce; thirdly, institutional theory explains that digital government is not based solely on the inclusion of technological infrastructure but must include organizational, institutional, and contextual aspects (Luna, Gil & Sandoval 2015).

Therefore, in an increasingly urbanized and digitalized world, it is necessary to harness this transformative force to address global challenges and provide effective solutions that stem from local intervention. Cities are an essential instrument of sustainable development; however, without proper implementation of well-designed plans, achieving the expected results will be impossible (UN-Habitat 2020).

At this point, Alvarado (2018) proposes that smart and sustainable cities are a necessary alternative toward which metropolises must move to guarantee the provision of public services and social cohesion more efficiently without jeopardizing the environment.

The use of digital technologies in different urban activities proposes a new form of territorial management (Alvarado 2020), in which various aspects—from mobility to waste management—adopt an integrated approach focused on resource optimization, cost reduction, and improving the population's quality of life through care for the environment and reduction of inequalities.

The foregoing is directly linked to the human right to a healthy environment, through which people have an inherent right to live in a clean, safe, and ecologically balanced environment, while the State is also obliged to provide basic public services and to protect, preserve, and improve the environment (IACHR 2021); all of this in terms of what is established in Article 11 of the Additional Protocol to the American Convention on Human Rights in the Area of Economic, Social and Cultural Rights (also known as the Protocol of San Salvador) (OAS 1988).

This category of human right forms part of the evolution of Economic, Social and Cultural Rights (ESCR), from which new concepts such as the rights to water, housing, health, and the aforementioned healthy environment have been incorporated; thus arise the Economic, Social, Cultural and Environmental Rights (DESCA), which seek to guarantee the necessary conditions for a dignified life, both individually and collectively (National Human Rights Commission [CNDH] 2023).

Therefore, the State and local governments, now more than ever, have the obligation to establish mechanisms that satisfy the comprehensive needs of human beings in the pursuit of improving their quality of life and addressing the challenges posed by economic, social, and environmental phenomena. The design of a smart city is not merely a matter of urban planning or public management; it must be a response to the demand for building an inclusive and sustainable society, harmonious with the environment and forward-looking.

3 Methodology

This research was developed from a theoretical and legal approach regarding the meaning of sustainable development and its connection with the process of smart urbanization in contemporary cities.

The study adopts a comparative and descriptive approach grounded in the synthesis of international indices and in a critical review of secondary official and academic sources. No primary data collection or surveys were conducted, therefore the conclusions stem from the integration and critical assessment of the metrics and evidence available in the consulted indices and in the literature. Aware of methodological differences across sources, the manuscript explains the criteria used to select indicators and delimits the scope of the analysis so that the results are presented as analytical approximations that contribute to policy debate without claiming to replace original empirical studies.

4 Sustainable Development

The concept of sustainable development first appeared in 1980 in the report by the International Union for Conservation of Nature. Later, in 1987, the report Our Common Future by the United Nations World Commission on Environment and Development provided the first definition, stating that sustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs (United Nations 1987).

In this same vein, three pillars were defined to achieve this goal in a balanced way: economic development, social development, and environmental protection. Following this, subsequent agreements discussed the means to put sustainable development into practice, notably the 1992 Earth Summit in Rio, the 2002 World Summit on Sustainable Development, and the 2012 Earth Summit (United Nations 1992; United Nations 2002; United Nations 2012).

In 2015, 193 United Nations Member States adopted the 2030 Agenda for Sustainable Development, a plan of action for people, the planet, and prosperity, aimed at strengthening universal peace and access to justice. The Agenda sets out 17 Sustainable Development Goals (SDGs) with 169 integrated and indivisible targets covering the economic, social, and environmental spheres. By adopting it, States commit to mobilizing the means necessary for its implementation through partnerships focused especially on the needs of the poorest and most vulnerable (United Nations 2015).

The main objective of the SDGs is to move toward a more equal society. To this end, they aim to eradicate extreme poverty, combat all forms of inequality, and promote inclusive growth, sustainable cities, peace, justice, and the strengthening of institutions. Policy improvement initiatives to advance toward the proposed goals are based on global, regional, and national strategies, as well as on strengthening partnerships between developing and developed countries (Turra & Fernandes 2021).

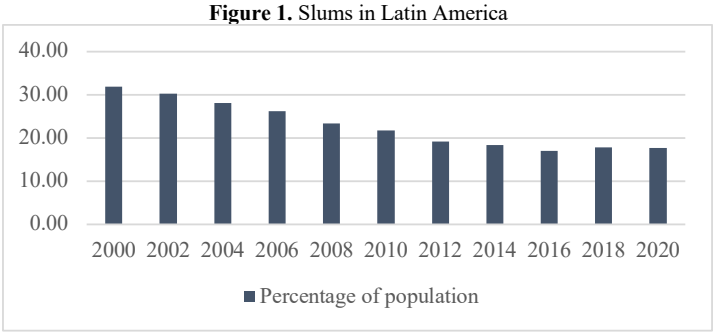
Specifically, SDG 11, entitled Sustainable Cities and Communities, aims to make human settlements inclusive, safe, resilient, and sustainable (United Nations 2015). To achieve this, the city must be able to provide basic services to its citizens, including energy, housing, transportation, education, and access to equitable growth opportunities for all.

"Cities and metropolitan areas are engines of economic growth and contribute about 60% of global GDP. However, they also account for around 70% of global carbon emissions and more than 60% of resource consumption" (United Nations 2019). If we add to this the growing global trend toward urbanization, there is an urgent need to prioritize sustainable development.

In parallel, it is important to note that accelerated urban expansion results in the emergence of marginal areas with insufficient infrastructure and services, or those exceeding the planning capacity of local governments. Issues such as wastewater treatment, waste collection, water stress, and mobility directly impact the environment and exacerbate existing inequalities.

According to data from ECLAC, the percentage of the urban population living in slums in Latin America showed a downward trend throughout the present century; however, since 2016 this indicator has reversed, and current data show an increase in slums or informal settlements—meaning those in which the population lives in inadequate housing and lacks basic services.

Figure 1. Slums in Latin America



CEPAL, 2024

According to Gálvez (2022), this is the result of unlimited urban expansion toward the outskirts of cities, where people with fewer economic resources settle, causing wealth to concentrate in the central areas and relegating the majority to informal zones. In this way, not only are preexisting gaps deepened, but conditions are also created for city infrastructure to become overburdened, slowing sustainable development.

For its part, the World Bank considers that, although urbanization is driving the global economy, growing inequality and exclusion in cities can hinder progress in development; however, if managed properly, the urbanization process can create the conditions to improve people's quality of life (World Bank 2024).

To achieve inclusive, safe, resilient, and sustainable cities, the organization proposes a roadmap that encompasses the following dimensions:

- Spatial inclusion: Provide affordable housing, water, and sanitation services.
- Social inclusion: Equal rights and participation for all.
- Economic inclusion: Job creation and economic growth (World Bank 2024).

On this basis, it is possible to design strategies and public policies aimed at closing the gaps resulting from current urbanization. Addressing issues such as slums, access to public services, economic growth, and the reduction of social inequalities would create more equitable and sustainable urban spaces.

In addition, it is estimated that by the end of 2023, only 15 percent of the SDG targets are on track to be achieved by 2030 at the global level (ECLAC 2023). This reflects a sense of urgency to take bold, fair, and effective actions in response to the global commitment to creating sustainable spaces. The response of local governments and cities must be directed toward meeting these goals with a transformative vision capable of balancing urban growth and sustainability.

Based on the above, it is essential to focus on the case of Latin America to identify which measures are proving effective in the development of smart and sustainable cities, as well as the roadmap to follow to achieve harmonious regional growth. All of this will make it possible to view the problem from a comprehensive perspective and provide an opportunity to reflect in ways consistent with Latin America's particular characteristics.

5 Indicators in Latin America

In response to the risks arising from urbanization and human digitalization, smart and sustainable cities are spaces that seek to provide solutions to public problems using technological tools that serve as a channel of connection between local governments and their citizens.

To understand how the inclusion of technologies influences the construction of smart and sustainable urban spaces, it is useful to examine the main global indices that annually compare the development of this type of cities and narrow the focus to those that are best positioned in the region.

In this regard, the IMD Smart City Index 2023 indicates that the five best-rated smart cities in Latin America in the last year were Medellín, Colombia; Santiago, Chile; Mexico City; Buenos Aires, Argentina; and San José, Costa Rica, ranked 118th, 119th, 121st, 124th, and 127th respectively, out of a total of 141 cities worldwide (IMD 2023).

Furthermore, in terms of technology and sustainability, the instrument assessed the progress of these cities as shown in the following graphs.

Figure 2. Health and Safety

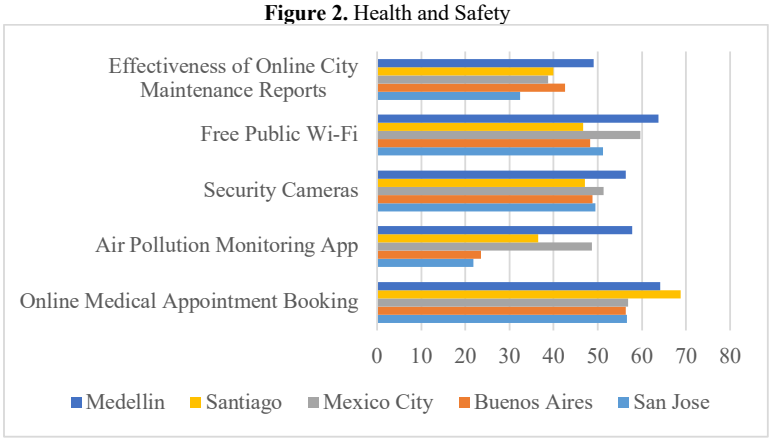


Figure 3. Mobility

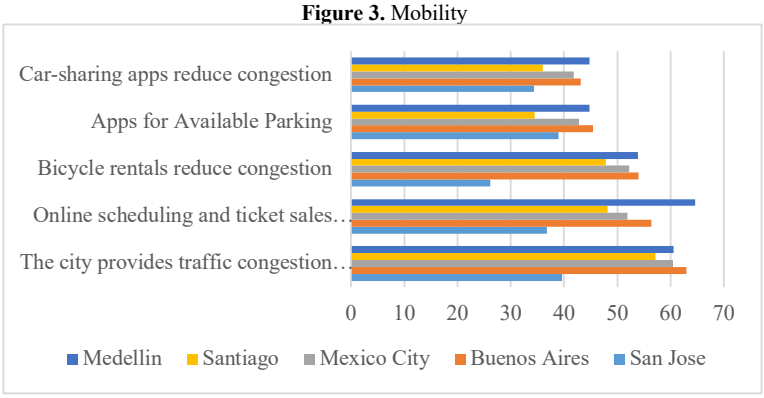


Figure 4. Work and Education

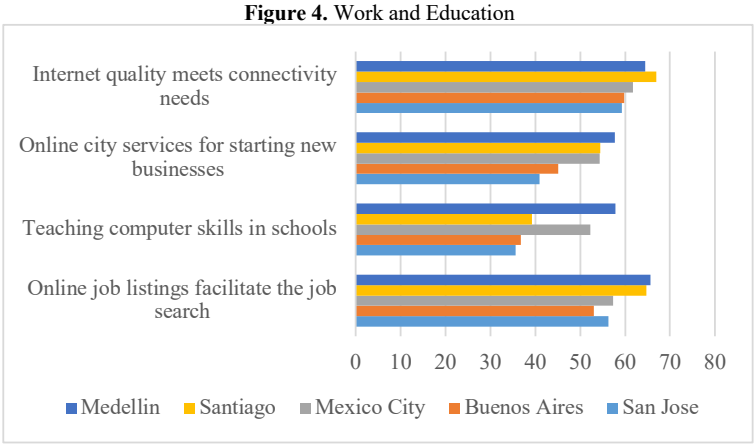
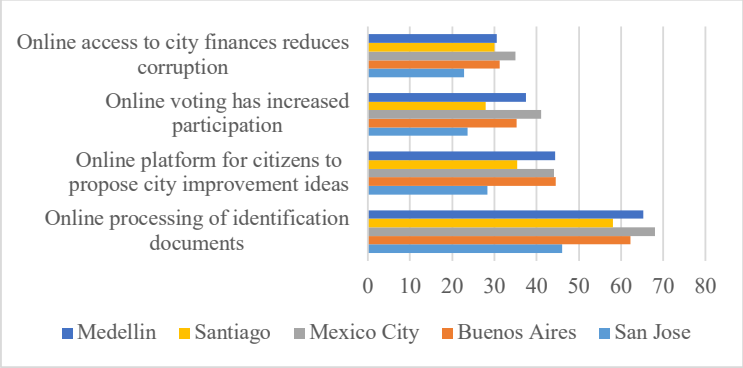


Figure 5. Governance



Figures 2 to 5 Author's elaboration based on IMD 2023 data.

Based on these scores, it is possible to observe that the most prominent smart and sustainable cities in Latin America show a medium-to-low level of progress. The indicators presented demonstrate that there are various challenges to overcome in terms of urban planning and development; furthermore, the potential offered by digital tools to facilitate access to public services, encourage citizen participation, and connect local governments with their residents is not being utilized effectively.

It is important to emphasize the governance dimension, as it has become evident that this is one of the greatest weaknesses in the region (see Figure 5). Efforts to combat corruption, promote transparency, and engage citizens using digital technologies have not achieved adequate penetration in Latin America's major cities, such as those referenced in the IMD Smart City Index 2023. On the contrary, rather than serving as a benchmark for maturity in smart and sustainable urbanization, these cities show resistance to using such technologies to improve governance and tackle issues like corruption.

On the other hand, the IESE Cities in Motion 2024 Index by the University of Navarra establishes a model of nine dimensions for smart urbanization that allows measuring each city's sustainability for the future. The criteria considered are human capital, social cohesion, economy, governance, environment, mobility and transportation, urban planning, international projection, and technology (Berrone & Ricart 2024).

Using this methodology, out of a total of 183 cities evaluated worldwide, the five Latin American cities with the highest progress were Santiago, Chile; Buenos Aires; Montevideo; Mexico City; and Panama City, as shown in the following table:

Table 1. Results on Smart and Sustainable Cities in Latin America

City	Global position	Human capital	Social cohesion	Economy	Governance	Environment	Mobility and Transportation	Urban planification	International proyection	Technology
Santiago	91	76	109	137	29	76	44	96	64	97
Buenos Aires	115	47	141	171	35	85	158	19	29	125
Montevideo	119	131	106	143	111	51	124	122	131	123
Mexico City	120	43	120	121	90	172	144	46	42	148
Panama	127	151	102	108	154	101	88	140	99	156

Author's elaboration based on Berrone & Ricart 2024.

As shown in the table, where green colors indicate greater progress and the scale toward red indicates the opposite, Latin America exhibits a medium-to-low level of smart and sustainable development. In fact, even though there are significant advances in certain dimensions, these are not integrated with the other indicators, suggesting disordered and limited growth that does not consider the integration of relevant sustainability aspects, and whose isolated approach prevents harmonious maturity.

For example, in the case of Buenos Aires, there is a strong position in urban planning, making it a regional reference; however, this is not linked to areas such as mobility and transportation, social cohesion, or the economy, where a significant crisis is evident, particularly in the economic sector. Similarly, Mexico City shows advanced indicators in areas such as international projection, human capital, and urban planning, but in terms of the environment, it faces a crisis that requires urgent intervention.

Regarding Santiago, Chile, this case is particularly relevant due to its focus on integrating digital solutions applied to public transportation and energy efficiency. The city has implemented intelligent systems for traffic control and mass transit management, contributing to improved urban mobility. Additionally, there is a noticeable trend by

the local government to promote citizen participation and implement transparency policies. Nevertheless, despite these advances, Santiago faces challenges like other Latin American metropolises, such as the need to strengthen social cohesion and improve environmental management in peripheral areas, where density and disordered urban growth place additional pressure on resources.

According to Berrone and Ricart (2024), this problem stems from Latin America's high urban density. The region's metropolises face increasingly large global challenges common to the area; furthermore, international crises directly affect countries in the region and exacerbate preexisting problems.

It is important to note that in the two measurement instruments presented, results regarding the development of smart and sustainable cities in Latin America vary in terms of ranking criteria. However, even though different methodologies were established for each index, both agree in evaluating Santiago, Chile; Mexico City; and Buenos Aires as among the most advanced cities in these terms, allowing us to affirm that these three serves as the region's main reference points.

Based on all the above, it is possible to build an argument regarding the observed results and present relevant conclusions on the topic of this research, to provide insights consistent with sustainable development in Latin America.

6 Discussion of Results

In line with the challenges faced by current local governments, the new urban agenda must highlight the links between sustainable urbanization and the improvement of quality of life. It is essential to emphasize the incorporation of all global actors into urban development or renewal policies and strategies (UN-Habitat 2020). This entails on one hand, close collaboration between the public sector, private sector, and civil society, and on the other, the integration of international knowledge and experiences that enrich local solutions.

It is important to note that innovation does not mean transitioning to a new technological paradigm, but rather making effective, practical, and novel use of available technologies to provide solutions to existing problems (Alvarado 2018). In this regard, the emphasis should be on the adaptation and application of tools that allow addressing concrete challenges such as vehicular congestion, environmental pollution, or inefficiencies in public service delivery. The goal is to enhance citizens' quality of life without compromising environmental balance, leveraging each technological resource to generate tangible and sustainable benefits.

The creation of smart urban spaces is based on the integration of society and the knowledge economy, where creativity, innovative ideas, scientific knowledge, and technological development shape the dominant profile of the city (Cabrero & Díaz 2013). This model also requires promoting citizen participation, as community involvement allows the identification of real needs and co-creation of solutions that effectively address local problems.

In the case of Latin America, local governments and civil society have yet to fully incorporate technology and its most innovative potential, such as artificial intelligence

and robotics, into the evaluation of state management and citizen participation. Local governments have mostly advanced in integrating the Internet, mobile applications, and digitization of procedures (Montecinos 2021); however, these initiatives often operate in isolation and do not constitute a multidimensional plan that comprehensively addresses the challenges of sustainable urbanization. Consequently, significant gaps are evident in crucial aspects such as social cohesion and economic development.

Some progress can be observed in the implementation of digital tools in public management that impact certain dimensions of smart urbanization; however, the progress is limited, as it does not constitute a multidimensional plan affecting the most important areas of sustainable city design. Therefore, the essential elements to improve citizens' quality of life in an integrated manner, aligned with environmental balance, are lacking.

The Latin American cities with the greatest growth potential are Santiago, Chile; Mexico City; and Buenos Aires. However, they still face very relevant issues to guarantee the human right to a healthy environment. Economic and social cohesion aspects are two dimensions with low or medium growth in these three cities, indicating significant remaining gaps such as poverty, access to better living conditions, and the promotion of citizen participation and its connection with other sectors.

For example, Santiago, Chile, has achieved notable advances in integrating digital technologies into its transport systems and implementing solutions to increase energy efficiency. Additionally, the city stands out for its governance focus and promotion of citizen participation, which has enabled initiatives connecting different levels of government and private sectors. Nevertheless, challenges remain in environmental management and inequality reduction, particularly in peripheral areas requiring greater investment in social infrastructure and basic services.

Mexico City, for its part, shows strengths in human capital, urban planning, and international projection; however, it also faces significant environmental challenges and deficiencies in social cohesion. Air pollution, congested traffic, and lack of integration in sustainable mobility policies highlight the need to rethink strategies that effectively link digitalization with sustainability actions.

Buenos Aires stands out for its advanced urban planning and capacity to design regeneration and transformation projects for public spaces, particularly led by the local government. However, it presents limitations in areas such as mobility and social integration, where additional efforts are needed to promote citizen participation and equitable development. One of its main challenges is to transform its urban planning advantages into comprehensive solutions addressing both economic and social needs, establishing a network of interconnected policies that strengthen its urban structure sustainably.

Beyond the previously mentioned aspects, substantial differences can be identified in how each city approaches the integration of digital solutions and management of urban growth. In Mexico City, efforts have focused on enhancing citizen participation platforms and real-time information networks, with strategies directed toward digitalizing processes and implementing early warning systems, allowing proactive emergency response and optimized resource allocation. Additionally, the city has promoted public-private partnerships to facilitate the transfer of emerging technologies and foster innovation in urban service management.

In Santiago, Chile, the focus has been on transforming urban infrastructure through pilot projects integrating IoT sensors and data-driven management systems. This approach aims to optimize water and energy resource use to enhance resilience to climate challenges. Strategic collaborations with universities and research centers have also been encouraged, allowing the development of solutions tailored to the specific needs of neighborhoods and improving the quality of public services.

Finally, Buenos Aires has focused on modernizing administrative processes and promoting transparency in public management. The city has incorporated geographic information systems for more effective monitoring of urban growth and territorial planning, complemented by programs promoting digital inclusion and technological training, seeking to actively engage citizens in decision-making and shaping urban space.

Therefore, it is reasonable to argue that each approach reflects a response adapted to its context and, together, offers a complementary view of how digitalization can serve as a catalyst for more efficient and sustainable urban development in the region.

Based on the above, although each city presents specific contexts and challenges, the comparative analysis highlights that innovation must be accompanied by a holistic strategy. Only through close coordination between different levels of government, the inclusion of global actors, and the adaptation of technological solutions to the specific needs of each environment can the existing gaps in poverty, access to services, and social cohesion be closed.

7 Conclusions

The main contribution of this article is to provide an integrated and critical reading of existing indices and evidence that helps identify trends, governance gaps and policy priorities for the transition to smart and sustainable cities in Latin America. The findings highlight the importance of institutional coordination and investment in digital infrastructure to progress in that transition, as well as the need for policies that protect rights and reduce social and digital divides. Future research should clearly pursue longitudinal designs and municipal administrative data to move from description toward assessments of causality and the effectiveness of public interventions.

Cities in Latin America are at a critical juncture in their development toward smart urbanization and, even more importantly, sustainability. Given the global crises caused by chaotic urban growth, the adoption of digital tools that enhance the design and implementation of public policies and local government plans is imperative to improve people's living conditions and safeguard environmental balance.

Implementing isolated programs that affect only certain aspects of sustainability is insufficient; what is required is a holistic strategy that simultaneously addresses economic, social, and environmental dimensions, in harmony with fundamental rights inherent to these areas, such as the human right to a healthy environment. In this regard, integrating digital technologies into urban management not only allows for resource optimization and cost reduction but also promotes greater transparency in public administration, improves efficiency in service delivery, and strengthens citizen participation.

Practically, the benefits of an integrated strategy manifest across various domains, highlighted in this research as follows:

1. Economic: The digitalization of procedures and the implementation of intelligent management systems can reduce bureaucracy and foster investment in modern infrastructure, resulting in a more dynamic economy and the strengthening of a more efficient state.
2. Social: By enhancing citizen participation through digital platforms, the creation of solutions is facilitated, and greater social inclusion is promoted. This is reflected in improved access to basic services, optimized public transportation, better public service provision, and reduced inequalities, allowing more people to benefit from a high-quality urban life.
3. Environmental: The implementation of technologies for waste management, energy consumption optimization, and sustainable mobility not only reduces cities' carbon footprints but also improves air quality and ensures efficient management of natural resources, protecting the environment and promoting environmentally respectful development.

Furthermore, the transition toward smart cities offers the advantage of transforming traditional management into a participatory and inclusive administration. By placing citizens at the center of action and connecting various actors—from government institutions to private companies and social organizations—collaborative networks are created that enhance innovation and comprehensively address urban challenges. In this scenario, technology becomes a catalyst for change, promoting transparency, efficiency, and a sense of belonging—essential elements for sustained transformation and progress.

Therefore, the path toward sustainable growth is directly linked to the design of smart cities. Current governments must transition from traditional management to a model in which citizen participation and the integration of global actors are central to transformation and innovation. In doing so, existing gaps in inequality will be closed, paving the way for a more prosperous, equitable, and environmentally respectful urban future.

References

1. Alvarado, R.: Ciudad inteligente y sostenible: hacia un modelo de innovación inclusiva [Smart and Sustainable City: Towards a Model of Inclusive Innovation]. Paakat Revista de tecnología y sociedad **7**(13), 1-17 (2018). http://dx.doi.org/10.32870/Pk.a7n13.299.
2. Alvarado, R.: Ciudades inteligentes y sostenibles: una medición a cinco ciudades de México [Smart and Sustainable Cities: Measuring Five Cities in Mexico]. Estudios sociales **30**(55), 2-28 (2020). https://doi.org/10.24836/es.v30i55.860.
3. Banco Mundial: Ciudades inclusivas [Inclusive cities]. Grupo Banco Mundial, Washington D.C. (2024). https://www.bancomundial.org/es/topic/inclusive-cities.
4. Berrone, P. & Joan, R: IESE Cities in Motion Index. Universidad de Navarra, Spain (2024). https://www.iese.edu/media/research/pdfs/ST-0649.
5. Cabrero, E. & Díaz, A: Introducción [Introduction]. In: Cabrero, E. (ed.) Retos de la competitividad urbana [Challenges of Urban Competitiveness], pp. 15-36. CIDE (2013).

6. CEPAL: La Cumbre de los ODS de 2023: países unidos en Nueva York para acelerar la acción hacia los Objetivos de Desarrollo Sostenible [The 2023 SDG Summit: Countries United in New York to Accelerate Action Toward the Sustainable Development Goals]. Naciones Unidas, Santiago (2023). https://www.cepal.org/es/notas/la-cumbre-ods-2023-paises-unidos-nueva-york-acelerar-la-accion-objetivos-desarrollo-sostenible.
7. CEPAL: Proporción de la población urbana que vive en barrios marginales, asentamientos informales o viviendas inadecuadas [Proportion of the Urban Population Living in Slums, Informal Settlements, or Inadequate Housing]. Naciones Unidas, Santiago (2024). https://consensomontevideo.cepal.org/es/indicadores/proporcion-de-la-poblacion-urbana-que-vive-en-barrios-marginales-asentamientos.
8. CIDH: Derechos Económicos, Sociales, Culturales y Ambientales. Estándares Interamericanos [Economic, Social, Cultural, and Environmental Rights: Inter-American Standards]. OEA, Washington, D.C. (2021). https://www.oas.org/es/cidh/informes/pdfs/compendio%20desca_esp_completo.pdf.
9. CNDH: Derechos Económicos, Sociales, Culturales y Ambientales [Economic, Social, Cultural, and Environmental Rights]. México (2023). https://www.cndh.org.mx/programa/39.
10. Criado, I.: Gobernanza inteligente, innovación abierta y tecnologías sociales en unas administraciones públicas colaborativas. ¿Hacia un cambio de paradigma en la gestión pública? [Smart Governance, Open Innovation, and Social Technologies in Collaborative Public Administrations: Towards a Paradigm Shift in Public Management?]. In: Criado, I. (coord.) Nuevas tendencias en la gestión pública [New Trends in Public Management], pp. 27-50. Instituto Nacional de Administración Pública, Madrid (2016).
11. DeSanctis, G. & Poole, S.: Capturing the Complexity in Advanced Technology Use: Adaptive Structuration Theory. Organization Science **5**(2), 121-147 (1994). http://www.jstor.org/stable/2635011.
12. Fountain, J.: Building the Virtual State. Washington, D.C.: Brookings Institution Press, Washington, D.C. (2001).
13. Gálvez, C.: Barrios marginales y discapacidad: ¿Hacia dónde van los asentamientos latinoamericanos? [Slums and Disability: Where Are Latin American Settlements Headed?]. Limaq (9), 15-24 (2022). https://doi.org/10.26439/limaq2022.n009.5384.
14. Giddens, A.: The Constitution of Society. University of California Press, L.A. (1984).
15. IMD: Smart City Index Report 2023. Lausanne: World Competitiveness Center, Lausanne (2023). https://imd.cld.bz/IMD-Smart-City-Index-Report-20231.
16. Kling, R.: Learning about Information Technologies y Social Change: The Contribution of Social Informatics. The Information Society **16**(3), 217–32 (2000). https://doi.org/10.1080/01972240050133661.
17. Llanes-Font, M., Salvador-Hernández, Y., Suárez-Benítez, M. & Solórzano-Benítez, R.: Cuarta revolución industrial y administración pública en América Latina y el Caribe [Fourth Industrial Revolution and Public Administration in Latin America and the Caribbean]. Ciencias Holguín **26**(3), 78-87 (2020). https://www.redalyc.org/articulo.oa?id=181563834007.
18. Luna, L., Gil, J. & Sandoval, R.: Avances y retos del gobierno digital en México [Advances and Challenges of Digital Government in Mexico]. Universidad Autónoma del Estado de México, México (2015).
19. Montecinos, E.: Cuarta revolución industrial y la Administración Pública en América Latina [Fourth Industrial Revolution and Public Administration in Latin America]. Revista venezolana de gerencia **26**(93), 10-25 (2021). https://www.redalyc.org/journal/290/29066223002/29066223002.pdf.

20. Naciones Unidas: Informe de la Comisión Mundial sobre el Medio Ambiente y el Desarrollo [Report of the World Commission on Environment and Development]. UN: New York (1987). https://digitallibrary.un.org/record/139811/files/A_42_427-ES.pdf?ln=es.
21. Naciones Unidas: Cumbre de la Tierra de Río [Earth Summit in Rio]. Asamblea General, Río de Janerio (1992). https://www.un.org/es/conferences/environment/rio1992.
22. Naciones Unidas: Cumbre Mundial sobre el Desarrollo Sostenible [World Summit on Sustainable Development]. Asamblea General, Johannesburgo (2002). https://www.un.org/es/conferences/environment/johannesburg2002.
23. Naciones Unidas: Conferencia de las Naciones Unidas sobre el Desarrollo Sostenible [United Nations Conference on Sustainable Development]. Asamblea General, Río de Janerio (2012). https://www.cepal.org/es/eventos/conferencia-naciones-unidas-desarrollo-sostenible-rio20.
24. Naciones Unidas: Transformar nuesrtro mundo: la Agenda 2030 para el Desarrollo Sostenible [Transforming our world: the 2030 Agenda for Sustainable Development]. Asamblea General, New York (2015). https://documents.un.org/doc/un-doc/gen/n15/291/93/pdf.
25. Naciones Unidas: Informe de los Objetivos de Desarrollo Sostenible 2019 [The Sustainable Development Goals Report 2019]. New York (2019). https://unstats.un.org/sdgs/report/2019/The-Sustainable-Development-Goals-Report-2019_Spanish.pdf.
26. OEA: Protocolo Adicional a la Convención Américanana sobre Derechos Humanos en materia de Derechos Económicos, Sociales y Culturales "Protocolo de San Salvador" [Additional Protocol to the American Convention on Human Rights in the Area of Economic, Social and Cultural Rights "Protocol of San Salvador"]. Departamento de Derecho Internacional, San Salvador (1988). https://www.oas.org/juridico/spanish/tratados/a-52.html.
27. Olivera, A.: Globalización, urbanización y salud: Impactos de la COVID-19 [Globalization, Urbanization and Health: Impacts of COVID-19. Architecture and Urbanism]. Arquitectura y Urbanismo **XLI**(3), 6-16 (2020). https://www.redalyc.org/articulo.oa?id=376865021002.
28. ONU-Hábitat: Nueva Agenda Urbana [New Urban Agenda]. ONU, Quito (2017). https://habitat3.org/wp-content/uploads/NUA-Spanish.pdf.
29. ONU-Hábitat: Nueva Agenda Urbana [New Urban Agenda]. ONU, Nairobi (2020). https://publicacionesonuhabitat.org/onuhabitatmexico/Nueva-Agenda-Urbana-Ilustrada.pdf.
30. Orlikowski, W.: The Duality of Technology: Rethinking the Concept of Technology in Organizations. Organization Science **3**(3), 398-427 (1992). http://www.jstor.org/stable/2635280.
31. Turra, C. & Fernandes, F.: La transición demográfica. Oportunidades y desafíos en la senda hacia el logro de los Objetivos de Desarrollo Sostenible en América Latina y el Caribe [Demographic Transition: Opportunities and Challenges on the Path to Achieving the Sustainable Development Goals in Latin America and the Caribbean]. CEPAL, Santiago (2021). https://repositorio.cepal.org/server/api/core/bitstreams/3f7b0487-2d48-453f-a083-e7d3278da358/content.

Integra Hub: innovación educativa territorial para ciudades inteligentes inclusivas

Submodelo universitario de articulación tecnológica y social en contextos regionales

Integra Hub: Territorial Educational Innovation for Inclusive Smart Cities A University

Submodel for Technological and Social Articulation in Regional Contexts

Vergara Berra, Alejandro [1][0009-0001-3642-0870] and
Ortega Cerecedo, José David [2][0009-0007-0406-0384]

[1] Research director. Universidad Interamericana. 72828. Puebla, México
[2] Rector. Universidad Interamericana. 72828. Puebla, México
alexvergaraberra@hotmail.com, jdortega@lainter.edu.mx

Abstract. This presentation introduces the design and preliminary validation of Integra Hub as a university-based model for territorial innovation, aimed at strengthening digital inclusion, situated learning, and multisectoral articulation within regional contexts.

The research was conducted in the state of Puebla, Mexico, and enabled the identification of structural limitations in traditional innovation hubs, particularly their weak connection to local dynamics and limited adaptability in rural and semi-urban areas. In contrast, Integra Hub emerges as a proactive proposal grounded in social co-creation, collaborative governance, and the strategic application of emerging technologies to foster regional development.

The methodology employed a mixed-methods approach—comprising expert interviews, documentary analysis, and structured surveys—which allowed for a comprehensive understanding of the underlying challenges. The findings reveal broad cross-sectoral acceptance of the model and affirm its potential for integration into socially oriented smart urban ecosystems.

In conclusion, the Integra Hub emerges as an innovative pathway for democratizing innovation processes within higher education, fundamentally contributing to sustainable impacts across educational, productive, community, and social domains.

Keywords: Smart cities, territorial innovation, digital inclusion, collaborative governance, higher education.

1 Introducción

La ciudad de Puebla, reconocida como Capital de la Tecnología y la Sostenibilidad en el marco del ICSC-CITIES 2025, se configura como un escenario privilegiado para examinar la pertinencia y los límites de los modelos de innovación territorial.

Su complejidad geográfica -donde convergen tanto centros urbanos consolidados, como zonas semiurbanas en expansión, y también comunidades rurales con arraigo histórico- refleja, en escala regional, los retos globales de las ciudades inteligentes, entre lo que puede destacarse de la conciliación del dinamismo tecnológico con la inclusión social y la sostenibilidad ambiental. La presencia de instituciones de educación superior agrega un componente estratégico, ya que permite la vinculación de la producción de conocimiento con la transformación de realidades locales.

En este contexto, los hubs de innovación tradicionales evidencian deficiencias expuestas. A pesar de ser modelos exitosos en la generación de capital tecnológico en entornos urbanos densos, siguen reproduciendo esquemas de gobernanza verticales y cerrados, con débil conexión hacia actores sociales y comunitarios. Su orientación, se centra en la competitividad global, que en ciertos casos puede traducirse como acciones de exclusión digital, baja participación ciudadana y escasa capacidad de adaptación a territorios periféricos o rurales. De esta forma, se percibe un desfase entre diseño institucional y las dinámicas territoriales reales, limitando su potencial como entes de transformación social.

Frente a estas limitaciones presentadas, resulta imperativo un enfoque territorial que trascienda la lógica tecnocrática dominante en el paradigma de la ciudad inteligente. Un enfoque que además de concebirse como un fin en sí mismo, también represente un medio que amplíe derechos, reduzca brechas y sirva como fortalecedor de la cohesión social. En esta perspectiva es como se logra concebir a la ciudad inteligente inclusiva -que enfatiza la participación ciudadana, la equidad en el acceso a las tecnologías y la sostenibilidad contextual- y a la gobernanza territorial colaborativa.

El modelo Integra Hub, impulsado por la Universidad Interamericana en Puebla, se plantea como una respuesta a este desafío. Concebida como una plataforma de convergencia entre educación superior, tecnologías emergentes y desarrollo regional; la propuesta es llevar a cabo una articulación entre la cocreación social, la gobernanza participativa y el accionar multisectorial con un doble impacto: por una parte, sensibilizar al contexto inmediato, y por otro, lograr una alta capacidad de replicabilidad en otros territorios.

En suma, la presente ponencia defiende el postulado de que el Integra Hub constituye una alternativa viable y necesaria para reconfigurar el concepto de ciudad inteligente: no solo como un espacio tecnológicamente avanzado, sino como un territorio donde la innovación se democratiza, la inclusión digital se convierte en derecho y la sostenibilidad se traduce en prácticas compartidas. Esta es la tesis que guía la reflexión y el análisis aquí presentados.

2 Marco conceptual

2.1 Innovación educativa territorial

La innovación educativa territorial va más allá de incorporar herramientas tecnológicas en el aula; es un proceso más complejo que se orienta a la transformación de vínculos entre educación, sociedad y territorio. (García-Hernández, 2021; UNESCO, 2020). En este contexto, la educación superior asume un rol destacado en dos vertientes: productora de conocimiento especializado y mediadora entre saberes académicos, demandas sociales y dinámicas productivas locales.

El modelo Integra Hub se anota en esta lógica, ya que propone espacios de formación situada, combinando tecnologías emergentes con metodologías participativas, con un diseño que responde a las condiciones específicas de contextos urbanos, semiurbanos y rurales. De esta forma, la innovación educativa va más allá de su presentación como un proceso vertical impuesto desde instituciones centrales, elevándose al nivel de ser una construcción colectiva fuerte, capaz de solidificar las capacidades locales, potenciando el desarrollo regional desde la base.

2.2 Inclusión digital como condición habilitante

La inclusión digital constituye un prerrequisito indispensable para el despunte de ciudades inteligentes con vocación social. No es suficiente el aseguramiento al acceso a dispositivos o infraestructura de conectividad. Se debe priorizar la garantía de que individuos y comunidades se apropien de estas herramientas para participar activamente en los procesos de innovación, aprendizaje y toma de decisiones conscientes.

En espacios territoriales como el estado de Puebla, donde las brechas digitales evidencian desigualdades históricas entre zonas urbanas y rurales, la inclusión digital adquiere un carácter estructural sin precedente (CEPAL, 2021; INEGI, 2022). El Integra Hub responde a este reto con programas de alfabetización tecnológica, provisión equitativa de recursos educativos y diseño de soluciones ajustadas a las capacidades locales. Y de esta forma contribuye al surgimiento de una ciudadanía digital crítica y activa, que sea capaz de incidir en la construcción de ecosistemas inteligentes desde sus propios horizontes culturales y sociales.

2.3 Gobernanza colaborativa en ciudades inteligentes

La noción de gobernanza colaborativa se consolida como un principio clave en la configuración de ciudades inteligentes inclusivas. Frente a esquemas de gestión centralizados y dominados por lógicas tecnocráticas, se plantea una gobernanza abierta, participativa y multisectorial, donde universidades, instituciones, gobiernos, empresas y organizaciones sociales, de forma colaborativa diseñan estrategias de desarrollo adaptadas al territorio.

El Integra Hub logra este enfoque, ya que articula actores diversos en torno a metas compartidas, como la innovación, la inclusión y la sostenibilidad. Este tipo característico de gobernanza, además de ampliar la legitimidad de las iniciativas, también incrementa su eficacia, integrando conocimientos situados, recursos locales y redes de

colaboración intersectorial. Para con ello, alinearse a las tendencias internacionales, que son las promotoras de ecosistemas de innovación abiertos y resilientes, orientados al crecimiento tecnológico, y sobre todo a la generación de impactos sociales sostenibles (Nam & Pardo, 2011; OECD, 2020).

3 Metodología

La investigación se desarrolló bajo un enfoque mixto secuencial exploratorio, combinando técnicas cualitativas y cuantitativas de manera complementaria. Este diseño metodológico respondió a la necesidad de comprender la innovación educativa territorial en toda su complejidad, abarcando dimensiones estructurales, pedagógicas, tecnológicas y comunitarias. Tal como advierten Vázquez-Pérez (2024) y Forni & De Grande (2020), los estudios aplicados en contextos regionales requieren esquemas flexibles, capaces de integrar evidencia empírica con validación territorial.

3.1 Técnicas aplicadas

El estudio empleó tres técnicas principales:

- *Análisis documental*: se revisaron 47 fuentes académicas, institucionales y técnicas. El material fue procesado y codificado con NVivo, lo que permitió identificar patrones temáticos y marcos normativos clave.
- *Entrevistas semiestructuradas*: se realizaron 29 entrevistas a actores del ecosistema de innovación en Puebla -académicos, líderes comunitarios y representantes del sector tecnológico- con el fin de conocer perspectivas diversas.
- *Encuesta estructurada*: aplicada a 120 participantes, validada estadísticamente mediante el coeficiente alfa de Cronbach ($\alpha = 0.87$), lo cual garantizó la fiabilidad del instrumento.

Tabla 1. Diseño metodológico del estudio aplicado en Puebla

Componente	Técnica / Instrumento	Resultados clave / Indicador
Documental	**Revisión sistemática / Codificación NVivo**	**r = 0.81 (coocurrencia temática)**
Casos internacionales	**Matriz de similitud / Jaccard**	**J = 0.67 (convergencia estructural)**
Entrevistas cualitativas	**Codificación abierta, axial y selectiva**	**Saturación teórica en entrevista n.º 25**
Encuesta cuantitativa	**Cuestionario estructurado / Alfa Cronbach**	**$\alpha = 0.87$ / x^2 con $p < 0.05$ en dos cruces**

Elaboración propia con base en Vázquez-Pérez (2024), Forni & De Grande (2020) e instrumentos aplicados en campo.

La Tabla 1 resume la lógica operativa del diseño metodológico, evidenciando la coherencia entre técnicas, instrumentos y resultados esperados. Esta estructura permitió triangular hallazgos y fortalecer la validez interna del estudio.

3.2 Validación territorial y ética

La validación territorial se realizó en localidades del estado de Puebla, seleccionadas por criterios de representatividad geográfica y diversidad socioeconómica. Se incluyeron zonas urbanas, semiurbanas y rurales, lo que permitió contrastar la aplicabilidad del modelo Integra Hub en distintos entornos.

Tabla 2. Cobertura territorial del estudio por tipo de zona

Tipo de zona	Localidades incluidas
Urbana	**Ciudad de Puebla, San Andrés Cholula, San Pedro Cholula**
Semiurbana	**Atlixco, Huejotzingo**
Rural	**Tepeaca, San Juan Atenco, Santa María Coapan**

Fuente: Elaboración propia con base en el diseño territorial del estudio (2024).

En términos éticos, se garantizó el consentimiento informado, el anonimato de los participantes, la participación voluntaria y la aprobación institucional por parte del Comité de Ética de la Universidad Interamericana.

4 Resultados clave

Los resultados del estudio se organizaron en cuatro ejes temáticos, lo que permitió evaluar la pertinencia, viabilidad y alineación estratégica del modelo Integra Hub en el contexto territorial de Puebla.

4.1 Percepción intersectorial del modelo

Los tres sectores consultados valoraron positivamente el modelo; los académicos destacaron su potencial formativo, los líderes comunitarios su enfoque inclusivo, y el sector tecnológico su capacidad de articulación territorial.

Tabla 3. Percepción intersectorial del modelo Integra Hub

Grupo estratégico	Alta (%)	Media (%)	Baja (%)
Académicos	**74.2**	**21.7**	**4.1**
Expertos locales	**68.5**	**26.3**	**5.2**
Sector tecnológico	**62.8**	**32.4**	**4.8**

Fuente: Elaboración propia con base en encuesta estructurada aplicada en Puebla (n=120).

Como se observa en la Tabla 3, la percepción del modelo es mayoritariamente positiva en todos los sectores, lo que refuerza su viabilidad como plataforma colaborativa de innovación educativa.

4.2 Brecha digital en zonas rurales

El estudio identificó una brecha estructural en conectividad y acceso a dispositivos en comunidades rurales. El 68% de estas reportó velocidades de conexión inferiores a 5 Mbps, lo que limita el acceso a servicios digitales básicos.

Tabla 4. Acceso a conectividad en comunidades rurales

Velocidad de conexión	Porcentaje de comunidades rurales
< 5 Mbps	**68%**
5–10 Mbps	**22%**
> 10 Mbps	**10%**

Elaboración propia con base en trabajo de campo y datos de INEGI (2022).

La Tabla 4 evidencia la necesidad de modelos adaptativos como Integra Hub para cerrar las brechas digitales en territorios con infraestructura limitada.

4.3 Viabilidad como plataforma multisectorial

El modelo fue evaluado como viable por el 84% de los participantes, quienes reconocieron su capacidad para articular esfuerzos entre universidad, gobierno, sector privado y sociedad civil.

Tabla 5. Viabilidad del Integra Hub como plataforma multisectorial

Valoración	Porcentaje
Viable	84%
Duda	11%
Poco viable	5%

Elaboración propia con base en resultados de encuesta intersectorial (2024).

Este respaldo cuantitativo valida la pertinencia del modelo como estrategia de desarrollo regional con enfoque colaborativo..

4.4 Alineación con ODS y políticas públicas

El modelo Integra Hub se articula con los Objetivos de Desarrollo Sostenible (ODS) 4, 9 y 11, así como con los planes estatal (Puebla) y nacional (México) de desarrollo, especialmente en los ejes de equidad educativa, inclusión digital y transformación social.

Tabla 6. Alineación del modelo Integra Hub con ODS y políticas públicas

ODS / Política Nacional	Alineación estratégica
ODS 4: Educación de calidad	Fortalece capacidades educativas en zonas vulnerables
ODS 9: Innovación e infraestructura	Promueve hubs tecnológicos inclusivos
ODS 11: Ciudades sostenibles	Articula actores para desarrollo urbano sostenible
Plan Estatal de Desarrollo	Vinculación educativa y tecnológica
Plan Nacional de Desarrollo	Impulso a la ciencia y equidad digital

Fuente: Elaboración propia con base en análisis documental de CEPAL (2021), UNESCO (2020) y Plan Nacional de Desarrollo (2020–2024).

La alineación del modelo con los ODS y planes nacionales, como se muestra en la Tabla 6, refuerza su potencial como instrumento operativo de política pública territorial, con capacidad de incidir en procesos de transformación estructural desde lo local.

5 Discusión

La implementación del Integra Hub en Puebla revela contrastes significativos frente a las tendencias nacionales en innovación educativa territorial. Mientras que, a nivel nacional, las iniciativas se concentran en zonas metropolitanas con alta densidad tecnológica -como Ciudad de México, Nuevo León o Jalisco-, en Puebla el despliegue se orienta a entornos heterogéneos que incluyen localidades rurales con notorias limitaciones estructurales.

Tabla 7. Comparativa nacional-local en innovación educativa territorial (2024)

Indicador	Promedio nacional (México)	Estado de Puebla
Cobertura de conectividad escolar	78.20%	62.40%
Presencia de hubs universitarios	14 entidades federativas	1 modelo piloto (Integra Hub)
Participación comunitaria formal	41.60%	58.3% (Integra Hub)
Vinculación multisectorial activa	36.90%	72.5% (Integra Hub)

Fuente: Elaboración propia con base en INEGI (2022), Mejoredu (2023) y CEIDE-UNAM (2024).

Como se aprecia en la Tabla 7, Puebla exhibe indicadores más bajos en infraestructura digital, pero destaca en participación comunitaria y articulación multisectorial. Este desbalance permite posicionar al Integra Hub como una alternativa viable para territorios que enfrentan brechas estructurales, y a su vez cuentan con capital social y vocación de colaboración intersectorial.

5.1 Comparación con modelos internacionales

El Integra Hub puede ser analizado en relación con referentes internacionales, como los living labs europeos (ej. Helsinki Urban Lab) o los hubs tecnológicos promovidos en Norteamérica y Asia (ej. MIT Media Lab, Singapore Smart Nation Initiative). Dichos

modelos han probado eficacia en la construcción de ecosistemas de innovación sofisticados, pero su campo de acción suele estar circunscrito a entornos urbanos de alta conectividad, con abundancia de capital humano especializado y acceso a financiamiento público-privado sostenido.

Tabla 8. Comparación entre Integra Hub y modelos internacionale

Dimensión	Living Labs / Hubs globales	Integra Hub (Puebla, México)
Infraestructura tecnológica	**Alta (urbana, conectividad avanzada)**	**Mixta (urbana, semiurbana y rural)**
Financiamiento	**Sostenido, público-privado**	**Limitado, institucional y comunitario**
Participación ciudadana	**Moderada, en fases de validación**	**Alta, desde el diseño y ejecución**
Replicabilidad	**Condicionada a entornos urbanos**	**Flexible, adaptada a contextos diversos**
Inclusión digital	**Implícita, no siempre prioritaria**	**Eje central del modelo**

Fuente: Elaboración propia con base en Nam & Pardo (2011), OECD (2020) y análisis comparativo del estudio

Este análisis expuesto en la Tabla 8, pone de relieve que el Integra Hub no pretende replicar mecánicamente los esquemas internacionales, sino aportar una perspectiva complementaria, centrada en la inclusión digital y la cocreación social. Con ello, responde a una de las limitaciones más señaladas de los modelos globales: su escasa capacidad de adaptación a contextos con brechas profundas de conectividad e infraestructura.

De este modo, el Integra Hub se perfila como una propuesta estratégica para articular innovación educativa, inclusión digital y gobernanza colaborativa desde una lógica territorial. Su valor radica menos en competir con experiencias internacionales y más en demostrar que es posible edificar ecosistemas de innovación desde la base, con la participación activa de comunidades y con sensibilidad a la diversidad de contextos.

Así es como la experiencia poblana contribuye al debate latinoamericano sobre ciudades inteligentes inclusivas, donde los desafíos estructurales exigen modelos flexibles, socialmente comprometidos y capaces de traducir la innovación en beneficios tangibles para todos los sectores de la sociedad.

6 Conclusiones

6.1 Potencial de replicabilidad

Los resultados del estudio evidencian que el Integra Hub posee un alto potencial de replicabilidad en contextos territoriales diversos de México y América Latina. Su arquitectura modular, sustentada en principios de inclusión digital, gobernanza colaborativa y formación situada, permite la adaptación de sus componentes a escenarios con distintos niveles de infraestructura tecnológica y cohesión social. A diferencia de modelos internacionales que dependen de ecosistemas urbanos altamente tecnificados, el Integra Hub se construye desde la heterogeneidad territorial, lo que le otorga flexibilidad operativa y pertinencia social.

Tabla 9. Factores que favorecen la replicabilidad del modelo Integra Hub

Factor estructural	Condición observada en Puebla	Implicación para replicabilidad
Participación comunitaria	**Alta**	**Favorece la apropiación local**
Infraestructura básica	**Mixta**	**Requiere adaptación tecnológica**
Vinculación institucional	**Activa**	**Permite escalamiento regional**

Fuente: Elaboración propia con base en resultados del estudio territorial (2024).

La Tabla 9 sintetiza los factores estructurales observados en Puebla que favorecen la replicabilidad del modelo Integra Hub. Destaca cómo la alta participación comunitaria, la infraestructura mixta y la vinculación institucional activa permiten adaptar el modelo a otros territorios con condiciones similares, reforzando su viabilidad como estrategia escalable de innovación educativa territorial.

6.2 Recomendaciones para política pública

A partir de la validación territorial realizada, es posible delinear recomendaciones estratégicas que contribuyan a consolidar políticas públicas de innovación educativa y territorial:

1. Diseñar modelos híbridos que integren conectividad, formación situada y participación comunitaria en programas estatales de innovación educativa.
2. Impulsar financiamiento específico para hubs territoriales en zonas semiurbanas y rurales, priorizando aquellas con baja cobertura digital, pero con fuerte capital social.
3. Generar marcos normativos flexibles que promuevan la cocreación entre universidades, gobiernos locales, sector privado y organizaciones comunitarias.

4. Incorporar métricas de impacto territorial en los sistemas de evaluación, complementando los indicadores tecnológicos con dimensiones sociales y comunitarias.

Tabla 10. Recomendaciones estratégicas para política pública

Eje de política	Recomendación clave	Nivel de implementación sugerido
Educación	**Formación situada y conectividad inclusiva**	**Estatal / municipal**
Innovación	**Financiamiento de hubs territoriales**	**Federal / estatal**
Gobernanza	**Articulación multisectorial**	**Local / interinstitucional**
Evaluación	**Indicadores de impacto territorial**	**Nacional / institucional**

Fuente: Elaboración propia con base en alineación con ODS y planes nacionales (2024).

La Tabla 10 presenta recomendaciones clave para orientar políticas públicas que fortalezcan la innovación educativa territorial. Cada eje -educación, innovación, gobernanza y evaluación- se vincula con acciones concretas y niveles de implementación diferenciados, lo que permite articular estrategias desde lo local hasta lo nacional, en coherencia con los ODS y los planes de desarrollo vigentes.

6.3 Integración en ecosistemas urbanos inteligentes

El Integra Hub muestra capacidad para insertarse de manera estratégica en ecosistemas urbanos inteligentes a través de tres rutas principales:

- Nodos periféricos de innovación, orientados a reducir las brechas digitales en comunidades marginadas cercanas a los centros urbanos.
- Laboratorios de cocreación ciudadana, que funcionen como espacios de colaboración entre universidades, gobiernos y ciudadanía para diseñar soluciones urbanas inclusivas.
- Plataformas de datos abiertos, que permitan integrar los resultados del modelo en sistemas de monitoreo urbano y facilitar la toma de decisiones basadas en evidencia.

Tabla 11. Vías de Integración del Integra Hub en ecosistemas urbanos inteligentes

Estrategia de integración	Objetivo principal	Ejemplo de aplicación urbana
Nodo periférico de innovación	Reducir brechas digitales en zonas marginadas	Zonas limítrofes de Puebla capital
Laboratorio de cocreación	Diseñar soluciones urbanas participativas	Espacios universitarios abiertos
Plataforma de datos abiertos	Monitorear impacto territorial en tiempo real	Sistemas municipales de innovación

Fuente: Elaboración propia con base en experiencias urbanas en Helsinki, Medellín y Ciudad de México

En conjunto, los hallazgos permiten afirmar que el Integra Hub constituye una propuesta innovadora y socialmente pertinente para avanzar en la construcción de ciudades inteligentes inclusivas. Su capacidad de replicabilidad, su alineación con los Objetivos de Desarrollo Sostenible y su énfasis en la participación comunitaria lo posicionan como un instrumento estratégico para orientar políticas públicas y prácticas institucionales hacia un desarrollo regional más equitativo y sostenible.

Esta orientación se refuerza en la Tabla 11, que presenta tres vías concretas de integración del modelo en ecosistemas urbanos inteligentes: nodos periféricos, laboratorios de cocreación y plataformas de datos abiertos, cada una con objetivos específicos y ejemplos de aplicación contextualizados.

El modelo no solo responde a las necesidades locales de Puebla, sino que ofrece una alternativa viable para territorios que, enfrentando limitaciones estructurales, busquen construir ecosistemas de innovación desde la base social y no únicamente desde la infraestructura tecnológica.

References

1. CEPAL. (2021). Transformación digital para una recuperación sostenible e inclusiva. Comisión Económica para América Latina y el Caribe. https://www.cepal.org/es/publicaciones/46623-transformacion-digital-una-recuperacion-sostenible-inclusiva.
2. Forni, F., & De Grande, P. (2020). Innovación territorial y gobernanza colaborativa: Modelos emergentes en América Latina. FLACSO.
3. García-Hernández, M. (2021). Educación digital y brechas tecnológicas en México: Un enfoque territorial. Revista Latinoamericana de Educación, 21(2), 45–62. https://doi.org/10.5281/edu.lat.2021.02.
4. Ibarra-Sáiz, M. S., Rodríguez-Gómez, G., & Boud, D. (2023). Evaluación participativa en educación superior: Perspectivas internacionales. Narcea Ediciones.

5. INEGI. (2022). Encuesta Nacional sobre Disponibilidad y Uso de Tecnologías de la Información en los Hogares (ENDUTIH). Instituto Nacional de Estadística y Geografía. https://www.inegi.org.mx/programas/dutih/2022/
6. Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). Qualitative data analysis: A methods sourcebook (3rd ed.). SAGE Publications.
7. Nam, T., & Pardo, T. A. (2011). Conceptualizing smart city with dimensions of technology, people, and institutions. Proceedings of the 12th Annual International Digital Government Research Conference, 282–291. https://doi.org/10.1145/2037556.2037602.
8. OECD. (2020). Innovation and digitalisation in regional development. Organisation for Economic Co-operation and Development. https://www.oecd.org/publications/innovation-and-digitalisation-in-regional-development.htm.
9. Strauss, A., & Corbin, J. (2008). Basics of qualitative research: Techniques and procedures for developing grounded theory (3rd ed.). SAGE Publications.
10. UNESCO. (2020). Education in a post-COVID world: Nine ideas for public action. United Nations Educational, Scientific and Cultural Organization. https://unesdoc.unesco.org/ark:/48223/pf0000373717.
11. Vázquez-Pérez, L. (2024). Modelos híbridos de innovación educativa en territorios vulnerables. Universidad Autónoma Metropolitana.

Assessing the Impacts of Mega Projects on Urban Thermal Dynamics in Riyadh, Saudi Arabia

Ayman Imam [0000-0003-0555-7601]

Department of Urban and Regional Planning, Faculty of Architecture and Planning, King Abdulaziz University, Jeddah, Saudi Arabia.

Abstract. Riyadh's Vision 2030 transformation has triggered a wave of large-scale and giga-level urban projects whose thermal impacts have not been systematically quantified. This study analyzes urban thermal dynamics in Riyadh between 2016 (pre-implementation stage) and 2024 (consolidation stage) using multi-temporal Landsat 8/9 imagery, single-channel land surface temperature (LST) retrieval, and GIS overlays of official mega-project footprints. The results show a citywide increase of about 2–3°C in LST, with the highest warming in airport, industrial, and newly suburbanized areas where impervious surfaces expanded and vegetation declined. By contrast, landscape- and heritage-oriented projects such as King Salman Park and Diriyah recorded localized cooling (−2 to −11°C), confirming the moderating role of green infrastructure, shading, and water-sensitive design. Comparative insights from Dubai, Doha, and Phoenix indicate that these outcomes are design-dependent rather than inevitable. Accordingly, the study recommends embedding minimum greening ratios, high-albedo materials, and continuous green–blue corridors into Riyadh's mega-project controls to align with Green Riyadh and Vision 2030 livability and climate-resilience targets.

Keywords: Urban Heat Island (UHI); Riyadh; Mega Projects; Sustainability; Remote Sensing; Climate-Sensitive Design; Green Infrastructure

1 Introduction

Urbanization is a defining force shaping the 21st-century landscape, with cities serving as epicenters for economic diversification, innovation, and national transformation strategies. However, in arid environments like the Arabian Peninsula, urban expansion brings elevated environmental risks—including increased energy consumption, resource depletion, and altered microclimates.

Among these, the Urban Heat Island (UHI) effect remains one of the most pressing urban environmental challenges, where densely built-up zones exhibit significantly higher temperatures than surrounding rural areas. This effect is amplified in desert cities due to high solar radiation, impervious land cover, limited vegetation, and the use of heat-retaining materials such as asphalt and concrete [1, 2].

Riyadh, the capital of Saudi Arabia, exemplifies these dynamics. Over the past five decades, it has transformed from a compact desert town into a sprawling metropolis of over 7 million residents [3, 4]. Central to this transformation are the ambitious mega

projects launched under Saudi Arabia's Vision 2030, including King Salman Park, Diriyah Gate, the Riyadh Metro, and New Murabba. These initiatives aim to enhance mobility, restore cultural identity, and improve livability [5, 6].

Yet, despite their potential, these mega projects also risk exacerbating thermal stress by expanding impervious surfaces, reducing vegetated land, and increasing energy-intensive activities. Previous research has evaluated city-wide UHI trends or broad land-use transitions, but there remains a critical knowledge gap in understanding how specific mega projects influence thermal patterns at a micro-urban scale [4, 7].

This study addresses that gap by evaluating land surface temperature (LST) changes in Riyadh between 2016 (pre-project implementation) and 2024 (post-development). It leverages geospatial analysis and remote sensing to assess how distinct project typologies—ranging from cultural to industrial to residential—affect UHI intensities. It also contextualizes Riyadh's case within international best practices to derive policy-relevant insights for other arid-region cities undergoing rapid development.

The primary objective of this study is to assess how Riyadh's mega projects—implemented under Vision 2030—have influenced urban thermal dynamics, particularly the Urban Heat Island (UHI) effect, between 2016 and 2024. By isolating the thermal signatures of individual projects, the research seeks to offer insights into the relationship between urban design typologies and local climate performance in an arid environment.

To fulfill this objective, the following research questions are posed:

1. What are the spatial and temporal changes in land surface temperature (LST) in Riyadh between 2016 (pre-development) and 2024 (post-development)?
2. How do specific mega projects contribute to either intensifying or mitigating the UHI effect across different districts of Riyadh?
3. What role do green infrastructure, design typology, and climate-sensitive interventions play in shaping LST and thermal comfort outcomes?
4. How do Riyadh's UHI dynamics compare with international case studies from other rapidly urbanizing desert cities such as Dubai, Doha, and Phoenix?
5. What evidence-based policy and planning recommendations can be developed to enhance the sustainability and climate resilience of future mega projects in Riyadh?

By answering these questions, the study aims to contribute both to urban climate resilience discourse and to the practical design and governance of sustainable mega projects in hot-arid urban regions.

2 Materials and Methods

This section details the study area, data sources, satellite image processing, LST derivation techniques, GIS analyses, and comparative approaches used to assess the impacts of Riyadh's mega projects on urban thermal dynamics.

2.1 Study Area

Riyadh is located in central Saudi Arabia (24.31°–25.07° N, 46.49°–46.87° E), positioned atop the Najd plateau at approximately 600 meters above sea level. The city experiences a hot desert climate (BWh classification), characterized by intensely hot summers and mild, dry winters. Temperatures during peak summer frequently exceed 45°C.

Over the past 50 years, Riyadh has evolved from a compact desert city into a sprawling metropolitan region composed of traditional neighborhoods, high-rise corridors, and newly developed suburbs [3]. This transformation has been accelerated by Saudi Vision 2030, under which a suite of mega projects has been launched to promote livability, economic diversification, and cultural revitalization.

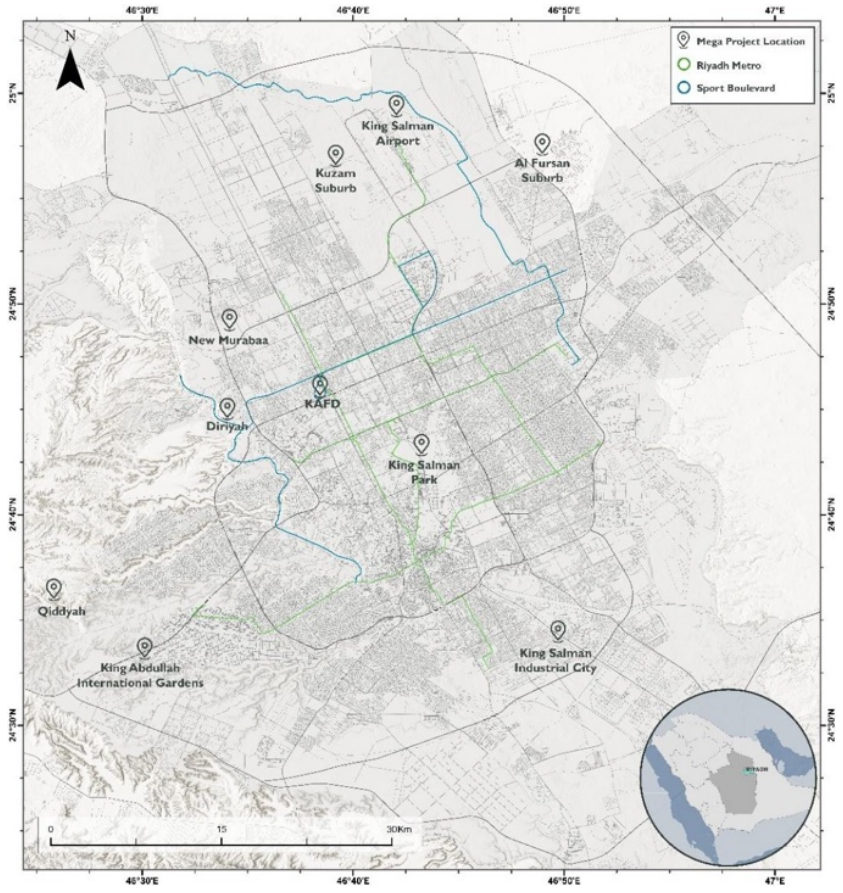


Fig. 1. Study area map of Riyadh showing the locations of mega projects.

Mega Projects Overview. Riyadh's mega projects form the cornerstone of its transformation under Vision 2030, focusing on cultural revival, economic diversification, enhanced mobility, and environmental sustainability. These initiatives are designed to improve Riyadh's infrastructure, enrich its cultural identity, and promote environmental resilience (Figure 2).

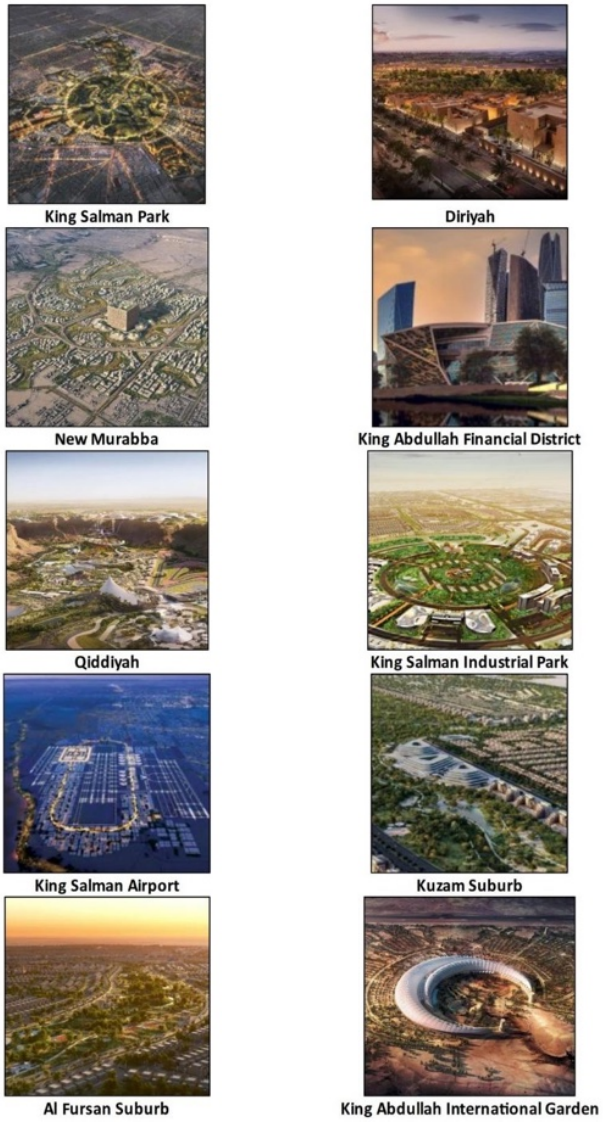


Fig. 2. Riyadh's mega projects

Among the most significant projects is King Salman Park, a 13-square-kilometer urban green space envisioned as one of the largest globally. It aims to improve air quality, increase biodiversity, and provide recreational and cultural amenities, serving as a vital green lung for the city. Similarly, Diriyah restores one of Saudi Arabia's most iconic heritage sites, blending traditional Najdi architecture with modern infrastructure to create a world-class cultural and tourism hub.

The New Murabba Development integrates residential, commercial, and cultural spaces within a sustainable framework, emphasizing pedestrian-friendly designs, energy efficiency, and green infrastructure. Meanwhile, the King Abdullah Financial District (KAFD) positions Riyadh as a leading financial hub with sustainable architecture, walkable spaces, and advanced transportation systems, promoting economic growth and urban livability.

Qiddiyah, Riyadh's entertainment, sports, and leisure destination, incorporates extensive green spaces and sustainable design principles to support both environmental and social well-being. Industrial projects, such as King Salman Industrial Park, foster eco-friendly practices, including renewable energy integration and waste management, to advance sustainable industrial ecosystems.

Infrastructure developments, such as King Salman Airport, enhance Riyadh's connectivity and economic activity while prioritizing renewable energy and green landscaping to mitigate environmental impact. Suburban projects such as Kuzam Suburb and AlFursan Suburb focus on providing modern, sustainable housing with integrated green spaces and community-oriented designs, improving urban livability.

Lastly, the King Abdullah International Garden stands as a beacon of environmental education and biodiversity conservation, combining innovative landscaping techniques with renewable energy systems. Together, these transformative initiatives aim to position Riyadh as a leading urban hub, balancing rapid growth with sustainability, cultural preservation, and quality-of-life enhancements.

These projects serve varied urban functions and offer a range of design and thermal outcomes, making them suitable for spatial comparisons of UHI effects.

2.2 Data Acquisition and Preprocessing

This study employed satellite-based thermal imagery from the Landsat program, renowned for its long temporal coverage and suitability for urban climate research. Landsat 8 and Landsat 9 scenes covering the Riyadh metropolitan area were acquired for two representative periods: 2016 (pre-mega project expansion) and 2024 (post-development). The years were selected not only for the availability of cloud-free imagery but also because 2016 corresponds to the pre-implementation/baseline stage of Riyadh's Vision 2030 mega and giga projects, while 2024 captures the consolidation phase in which major project footprints (e.g., King Salman Park, Diriyah, New Murabba, KAFD expansion) become thermally detectable in medium-resolution imagery. Thermal infrared (TIR) bands were extracted from the Landsat scenes using ESRI's Landsat Explorer platform, which provides straightforward access to atmospherically corrected and georeferenced datasets. Preprocessing steps included radiometric calibration and

atmospheric correction to convert digital numbers to at-sensor brightness temperatures. Surface emissivity values, essential for accurate land surface temperature (LST) retrieval, were derived from NDVI-based emissivity estimation methods. Ancillary data—such as land-use/land-cover (LULC) maps, meteorological records from the Saudi National Center for Meteorology, and official shapefiles or digitized footprints of the mega projects obtained from Riyadh planning and development authorities—were incorporated for spatial and contextual analysis.

2.3 Land Surface Temperature Derivation

Accurate LST calculation from Landsat Thermal Infrared (TIR) data is fundamental for UHI studies. This study employs the Single-Channel Algorithm (SCA) method, which utilizes top-of-atmosphere brightness temperature, surface emissivity, and atmospheric parameters to estimate LST. The methodological workflow is summarized in Figure 3.

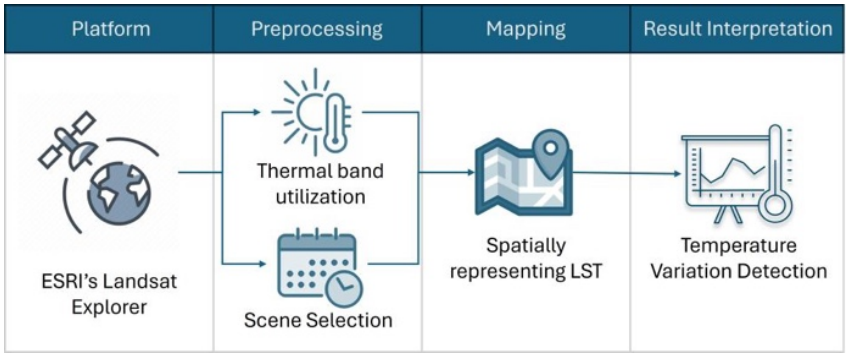


Fig. 3. The framework of the proposed methodology.

Single-Channel Algorithm (SCA) Overview. The SCA method involves the following steps:

1. **Brightness Temperature Calculation:** Digital numbers from the thermal bands are converted to at-sensor radiance, then to brightness temperature using sensor-specific calibration coefficients.
2. **Emissivity Correction:** Surface emissivity is estimated based on NDVI, which differentiates between vegetated and non-vegetated surfaces. Emissivity values typically range from 0.95 for urban surfaces to 0.98–0.99 for vegetated areas [8].
3. **LST Calculation:** Adjust brightness temperature for emissivity to derive true LST, accounting for atmospheric absorption and emission effects.

While ground-based validation was limited, we compared LST estimates with available ground station temperature records and historical climate data. The consistency checks indicated that the derived LST values reasonably approximated measured conditions, though future research could incorporate portable radiometers or in-situ sensors for improved validation.

2.4 Spatial Analysis and Indicators

GIS-based analyses were performed using ArcGIS Pro. LST rasters from 2016 and 2024 were co-registered and differentiated to map spatio-temporal changes. We overlaid mega project footprints, identified from official government plans and development authority data, onto LST maps to isolate temperature trends within project areas. To understand the drivers of LST changes, we computed land cover indices:

- **Normalized Difference Vegetation Index (NDVI):** To assess vegetation distribution and its influence on LST.
- **Normalized Difference Built-up Index (NDBI):** To identify the built-up intensity and correlate it with rising surface temperatures.
- **Normalized Difference Water Index (NDWI):** To detect surface water presence and relate it to localized cooling effects.

By correlating these indices with LST changes, we identified patterns linking vegetation reduction, construction density, and the emergence of new impervious surfaces to thermal shifts [2, 8].

2.5 Statistical and Comparative Analyses

We computed basic descriptive statistics (mean, median, standard deviation) for LST values during the pre- and post-development periods within mega-project zones and their neighboring areas. Additionally, box plots and histograms were used to visualize the distribution of temperature changes. Comparison with other arid and semi-arid cities—such as Dubai (UAE), Doha (Qatar), and Phoenix (USA)—provided a global context, referencing studies that highlight similar UHI intensification from large-scale infrastructural interventions [9-12]. Such comparative insights can help understand whether the observed changes in Riyadh are unique or follow broader regional and climatic patterns.

3 Results

This section presents the key findings of the study, analyzing changes in Land Surface Temperature (LST) between 2016 and 2024 across Riyadh's major mega project zones. The analysis highlights both intensification and mitigation of UHI effects, depending on the design approach, land use, and environmental considerations of each project.

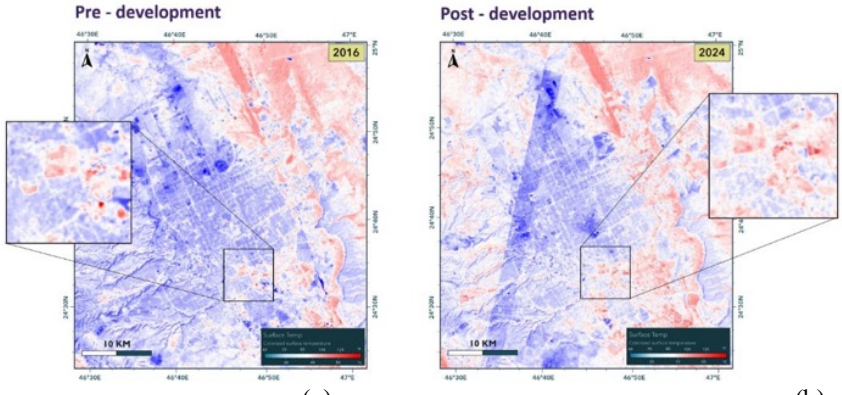
3.1 Spatiotemporal Variations in Land Surface Temperature

LST maps from 2016 and 2024 show clear spatial shifts in thermal intensity across Riyadh. Figure 4a and 4b display the distribution of surface temperatures in the pre- and post-development years.

- In 2016, cooler areas were concentrated around Wadi Hanifa and central green corridors, with LST values generally ranging between 32°C and 42°C.

- By 2024, significant urban hotspots emerged, particularly in the city's north-east and southwest quadrants, correlating with areas undergoing rapid infra-structure growth.

The mean LST across the city increased by approximately 2.8°C, with some zones recording temperature spikes of over 5°C. These changes are consistent with global literature on the link between impervious surfaces, reduced vegetation, and UHI for-mation [13, 14].



(a) (b)

Fig. 4. Distribution of LST in Riyadh: (a) Pre-development period (2016); (b) post-development period (2024).

Table 1. Mean land surface temperature (LST) for selected Riyadh mega-project polygons in 2016 and 2024 and the resulting variation (ΔLST).

Project name	LST 2016 (°C)	LST 2024 (°C)	ΔLST 2016–2024 (°C)
King Salman Park	44	39	**−5**
Diriyah	40	38	**−2**
New Murabba	43	34	**−9**
King Abdullah Financial District (KAFD)	33	35	**+2**
Qiddiyah	39	28	**−11**
King Salman Industrial Park	44	48	**+4**
King Salman Airport	45	49	**+4**
Kuzam Suburb	42	34	**−8**
AlFursan Suburb	42	49	**+7**
King Abdullah International Garden	35	38	**+3**
King Salman Park	44	39	**−5**
Diriyah	40	38	**−2**
New Murabba	43	34	**−9**

The above values were derived from spatially clipped LST rasters averaged over each project boundary using zonal statistics in ArcGIS. Surface temperature increases were especially prominent in highly built-up, vegetatively sparse areas, while decreases were observed in projects with green design strategies and landscape integration.

Citywide Temperature Trends. As shown in Figure 4a (2016), LST patterns indicated relatively moderate temperatures, with cooler zones primarily concentrated around Wadi Hanifa and other green corridors. These areas benefited from natural shading, evapotranspiration from vegetation, and reduced impervious surface coverage. However, by 2024 (Figure 4b), numerous urban hotspots will emerge in newly developed districts characterized by high-density buildings, extensive road networks, and reduced green cover.

The mean LST across Riyadh increased by approximately (2–3°C) over the eight-year period, with the highest temperature increments (up to 5°C) observed in areas undergoing significant land cover transformation. These findings align with established urban climate theory, which posits that increased building density and impervious surfaces contribute to higher surface temperatures and intensified UHI effect.

Seasonal and Diurnal Patterns. While the study primarily focused on annual LST changes, preliminary analyses of seasonal and diurnal patterns suggest that UHI effects are more pronounced during the summer months and at nighttime. In the summer, the intense solar radiation exacerbates heat absorption by urban materials. At the same time, nighttime UHI intensity remains elevated due to the slow release of stored heat from buildings and pavement. These seasonal variations highlight the compounded stress of extreme heat on urban populations, infrastructure, and energy systems.

3.2 Influence of Mega Projects on UHI Intensity

The study's focal point—assessing the specific impacts of mega projects on UHI intensity—revealed a heterogeneous landscape of thermal responses across different project areas. Table 1 summarizes LST variations across key mega-project zones between 2016 and 2024, illustrating contrasting effects driven by design and implementation strategies.

Projects Exacerbating UHI. Several mega projects were identified as significant contributors to UHI intensification due to their reliance on conventional construction materials and designs lacking sufficient green infrastructure:

- AlFursan Suburb experienced the highest LST increase (+7°C). The area underwent extensive transformation from semi-natural land into densely packed housing with limited tree cover or cooling strategies.
- King Salman Airport and Industrial Park each recorded +4°C increases, driven by vast paved zones, reflective roofs, and a lack of microclimatic buffers.
- Even KAFD, despite being a "smart" financial district, showed a modest increase (+2°C), likely due to dense built-up mass and vertical heat storage.

These results underscore the design-dependence of thermal outcomes. Projects that ignored nature-based solutions or high-albedo surfaces exacerbated urban heat, as observed in similar cases in Dubai and Phoenix.

Projects Mitigating UHI. Conversely, several projects demonstrated strong thermal moderation:

- Qiddiyah, with its extensive landscape and water-sensitive urban design, showed the highest cooling effect (−11°C).
- New Murabba recorded a significant −9°C drop, attributed to high green plot ratios, walkable layouts, and building orientation minimizing heat gain.
- King Salman Park and Diriyah Gate also saw measurable LST reductions of −5°C and −2°C, respectively. These results confirm that green infrastructure and vernacular cooling principles can effectively mitigate UHI in mega developments [15, 16].

These examples highlight that mega projects can play a dual role—either exacerbating or mitigating UHI effects—depending on their design strategies. Projects prioritizing green infrastructure, reflective materials, and sustainable urban design principles can effectively counterbalance the heat-retaining tendencies of urbanization, contributing to more resilient and livable cities.

3.3 Comparative International Perspectives

To contextualize Riyadh's UHI dynamics, the study compared its findings with assessments conducted in other rapidly urbanizing desert cities, such as Dubai (UAE), Doha (Qatar), and Phoenix (USA). These comparisons highlight common challenges and successful mitigation strategies, providing valuable lessons for Riyadh.

Dubai, United Arab Emirates. Dubai's UHI challenges mirror those of Riyadh. However, the adoption of vertical gardens, linear green corridors, and high-albedo pavements has led to a gradual reduction in surface temperatures [17-19]. Lessons include the importance of mandated greening thresholds and permeable infrastructure. In response, Dubai has implemented several mitigation strategies, including:

- **Green Corridors:** Development of linear green spaces along highways and urban arteries to provide shading and promote airflow.
- **Green Roofs and Vertical Gardens:** Incentives for building owners to install green roofs and vertical gardens, enhancing urban greenery without requiring additional land.
- **Reflective Pavements:** Use of high-albedo materials for pavements and roads to reduce heat absorption and lower surface temperatures.

These initiatives have demonstrated measurable localized temperature reductions and improved microclimates, and they offer an operative model for Riyadh's linear mega projects, such as the Sports Boulevard and Metro corridors.

Doha, Qatar. Doha introduced urban canopy programs and energy-efficient building codes that reduced UHI intensities in suburban districts [20-22]. Their multimodal transport expansion also lessened heat emissions from traffic, a major secondary contributor to UHI. In response, Doha has prioritized sustainable urban design through:

- **Urban Canopy Development:** Strategic planting of trees and creation of shaded public spaces to provide natural cooling.
- **Energy-Efficient Building Designs:** Adoption of building codes that mandate energy-efficient materials and designs, reducing heat gain and enhancing indoor thermal comfort.
- **Public Transportation Systems:** Expansion of public transit networks to decrease vehicular emissions and traffic congestion, indirectly mitigating UHI effects by reducing heat emissions from vehicles.

These measures illustrate how combining building-level standards with citywide mobility reforms can slow UHI growth in a Gulf capital.

Phoenix, USA. Phoenix applies a multi-pronged strategy—urban forestry, cool roofs, and permeable pavements—that has shown substantial impact on mitigating peak LST during summer months [23-25]. Phoenix has implemented a variety of UHI mitigation strategies, including:

- **Urban Forestry Programs:** Extensive planting of native trees to provide shade and promote evapotranspiration.
- **Cool Roof Initiatives:** Promotion of reflective roofing materials to decrease heat absorption by buildings.
- **Permeable Pavements:** Installation of permeable surfaces to reduce heat retention and improve water infiltration, contributing to cooler urban surfaces.

Phoenix's experience shows that UHI mitigation can be mainstreamed into municipal codes and capital works even under sustained urban growth.

Lessons for Riyadh. Drawing from these international examples, Riyadh can adopt a multi-faceted approach to UHI mitigation in its mega projects:

1. Integrate green infrastructure as primary, not decorative, elements (tree-lined boulevards, pocket parks, green roofs) in King Salman Park edges, Diriyah extensions, and Sports Boulevard segments.
2. Mandate reflective/cool materials for roofs, plazas, and widened road sections in airport, industrial, and suburban projects where our results showed +4 to +7°C warming.
3. Apply water-sensitive urban design (WSUD)—shaded water features, treated greywater for irrigation—to leverage evaporative cooling where potable water is constrained.
4. Align building energy codes with thermal targets so that new mixed-use clusters (e.g. New Murabba precincts) do not add to anthropogenic heat loads.
5. Plan green–blue corridors that connect mega projects (King Salman Park → Sports Boulevard → Wadi Hanifa), creating a city-scale cooling network.

These measures would translate the study's LST evidence into implementable controls and align Riyadh's mega projects with regional best practice in UHI management.

1 Discussion

The thermal impact analysis of Riyadh's mega projects provides valuable evidence for integrating climate resilience into future urban development, particularly in arid environments. The findings demonstrate that while some projects successfully mitigated UHI effects through nature-based design, others contributed to significant temperature increases due to conventional construction practices and a lack of vegetation.

1.1 Policy Recommendations

1. **Mandatory Green Space Requirements:** Introduce regulations that mandate minimum green space allocations for all new developments, particularly mega projects. This could include requirements for urban parks, green roofs, and vegetated public spaces to ensure that development does not come at the expense of essential greenery.
2. **Reflective Material Standards:** Establish building and pavement codes that require the use of reflective or high-albedo materials in construction. Incentivize developers to adopt cool roofing systems and reflective pavements to reduce heat absorption and lower surface temperatures.
3. **Urban Form Guidelines:** Develop comprehensive urban form guidelines that promote compact, mixed-use development patterns. Encourage the creation of pedestrian-friendly zones to reduce reliance on vehicular transport and minimize the expansion of impervious surfaces.
4. **Water-Sensitive Urban Design (WSUD):** Implement policies that promote the integration of water-sensitive designs in mega projects. This includes incorporating water features, constructed wetlands, and permeable pavements to enhance evaporative cooling and support microclimatic regulation.
5. **Energy-Efficient Building Codes:** Strengthen building codes to incorporate energy-efficient designs and materials. Encourage the use of passive cooling strategies, natural ventilation, and energy-efficient appliances to reduce overall energy consumption and associated heat emissions.
6. **Comprehensive Monitoring and Evaluation:** Establish a robust monitoring framework to track thermal changes and UHI intensity over time. Utilize remote sensing data, ground-based sensors, and GIS tools to continuously assess the effectiveness of implemented strategies and inform future policy adjustments.

1.2 Planning Strategies

1. **Green Corridors and Urban Forestry:** Design and implement green corridors that connect major parks, water bodies, and residential areas. Urban forestry programs should prioritize the planting of native, drought-tolerant species to ensure sustainability and reduce water consumption.
2. **Integrated Water Management:** Develop integrated water management systems that support both human use and ecological functions. This includes

rainwater harvesting, wastewater recycling, and the creation of green infrastructure that leverages water for cooling and vegetation support.

3. **Climate-Responsive Design Principles:** Incorporate climate-responsive design principles in all new developments. This involves designing buildings and public spaces that maximize natural shading, enhance airflow, and minimize heat gain through strategic orientation and material selection.
4. **Public Transportation Enhancement:** Expand and enhance public transportation networks, such as the Riyadh Metro, to reduce the dependency on private vehicles. Efficient public transit systems can decrease traffic congestion, lower vehicular emissions, and contribute to cooler urban environments.
5. **Community Engagement and Awareness:** Engage with local communities to raise awareness about the importance of UHI mitigation and sustainable urban practices. Encourage community participation in greening initiatives, such as tree planting drives and the maintenance of public green spaces.
6. **Cross-Sectoral Collaboration** Foster coordination between **urban planning, environment, housing, and transport authorities** under a unified UHI mitigation framework. Incentivize private sector developers through expedited approvals for sustainable projects.

2 Conclusions

This study assessed the thermal impacts of mega projects in Riyadh between 2016 and 2024, using Landsat-derived Land Surface Temperature (LST), GIS analysis, and land use indices (NDVI, NDBI, NDWI). The research aimed to determine how large-scale urban interventions influenced Urban Heat Island (UHI) patterns, and what factors contributed to either intensifying or mitigating thermal stress in different project zones. The results showed that Riyadh's urban expansion, while transformative in scope and scale, produced divergent thermal outcomes:

- Projects with conventional, impervious urban form—such as King Salman Airport, AlFursan Suburb, and King Salman Industrial Park—were associated with significant LST increases, ranging from +4°C to +7°C. These projects typically featured minimal green space, dense infrastructure, and limited climate-responsive design.
- Projects integrating green infrastructure, water-sensitive elements, and vernacular urbanism—such as Qiddiyah, New Murabba, and King Salman Park—demonstrated strong localized cooling, with LST reductions between −5°C and −11°C. These interventions align with international best practices in UHI mitigation.
- The overall citywide LST increased by approximately 2–3°C, indicating an intensification of UHI effects in peripheral districts experiencing high growth. However, the study also confirmed that UHI trajectories are not inevitable; they are contingent on design decisions, land use composition, and environmental integration.

- International comparisons with Dubai, Doha, and Phoenix further validated that climate-sensitive mega project design—particularly in arid regions—can play a pivotal role in moderating urban temperatures and promoting sustainability.

In conclusion, the study highlights the need for evidence-based planning and regulations for mega projects in growing desert cities. It shows urban development and thermal resilience can coexist, and Riyadh can be a global model for climate-adaptive urbanization if sustainability is prioritized from the start.

References

1. Oke, T. R. (1982). The energetic basis of the urban heat island. Quarterly Journal of the Royal Meteorological Society, 108(455), 1–24. https://doi.org/10.1002/qj.49710845502
2. Nuruzzaman, M. (2015). Urban Heat Island: Causes, Effects and Mitigation Measures—A Review. International Journal of Environmental Monitoring and Analysis, 3, 67–73. https://doi.org/10.11648/j.ijema.20150302.15
3. Garba, S. B. (2004). Managing urban growth and development in the Riyadh metropolitan area, Saudi Arabia. Habitat International, 28(4), 593–608. https://doi.org/10.1016/j.habitatint.2003.10.008
4. Alajizah, S. M., & Altuwaijri, H. A. (2024). Assessing the Impact of Urban Expansion on the Urban Environment in Riyadh City (2000–2022) Using Geospatial Techniques. Sustainability, 16(11), 4799. https://doi.org/10.3390/su16114799
5. King Salman Park Foundation. (2021). The Destination for Investment. https://www.kingsalmanpark.sa
6. RCRC (Royal Commission for Riyadh City). (2024). Green Riyadh Program. https://www.rcrc.gov.sa/en/news/55236
7. Ahn, C., Lee, S., Peña-Mora, F., & Abourizk, S. (2010). Toward environmentally sustainable construction processes: The U.S. and Canada's perspective on energy consumption and GHG/CAP emissions. Sustainability, 2(1), 354–370. https://doi.org/10.3390/su2010354
8. Bala, R., Prasad, R., Yadav, V. P., and Sharma, J. (2018). A Comparative Study of Land Surface Temperature with Different Indices on Heterogeneous Land Cover Using Landsat 8 Data, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLII-5, 389–394, https://doi.org/10.5194/isprs-archives-XLII-5-389-2018
9. Zhao, L., Lee, X., Smith, R. B., & Oleson, K. (2021). Strong contributions of local background climate to urban heat islands. Nature, 511(7508), 216–219. https://doi.org/10.1038/nature13462
10. Kim, J. -H., Gu, D., Sohn, W., Kil, S. -H., Kim, H., & Lee, D. -K. (2016). Neighborhood Landscape Spatial Patterns and Land Surface Temperature: An Empirical Study on Single-Family Residential Areas in Austin, Texas. International Journal of Environmental Research and Public Health, 13(9), 880. https://doi.org/10.3390/ijerph13090880
11. Santamouris, M., & Vasilakopoulou, K. (2023). Recent progress on urban heat mitigation technologies. Science Talks, 5, 100105. https://doi.org/10.1016/j.sctalk.2022.100105
12. Chow, W. T. L., & Brazel, A. J. (2012). Urban heat island research in Phoenix, Arizona: Theoretical contributions and policy applications. Bulletin of the American Meteorological Society, 93(4), 517–530. https://doi.org/10.1175/BAMS-D-11-00011.1

13. Arnfield, A. J. (2003). Two decades of urban climate research: A review of turbulence, exchanges of energy and water, and the urban heat island. International Journal of Climatology, 23(1), 1–26. https://doi.org/10.1002/joc.859
14. Feng, L., Xie, N., & Cui, N. (2022). Materials to Mitigate the Urban Heat Island Effect for Cool Pavement: A Brief Review. Buildings, 12(8), 1221. https://doi.org/10.3390/buildings12081221
15. Balany, F., Ng, A. W., Muttil, N., Muthukumaran, S., & Wong, M. S. (2020). Green Infrastructure as an Urban Heat Island Mitigation Strategy—A Review. Water, 12(12), 3577. https://doi.org/10.3390/w12123577
16. Vujovic, S., Haddad, B., Karaky, H., Sebaibi, N., & Boutouil, M. (2021). Urban Heat Island: Causes, Consequences, and Mitigation Measures with Emphasis on Reflective and Permeable Pavements. CivilEng, 2, 459–484. https://doi.org/10.3390/civileng2020026
17. Al-Hajri, S., Al-Ramadan, B., Shafiullah, M., & Rahman, S. M. (2025). Microclimate Performance Analysis of Urban Vegetation: Evidence from Hot Humid Middle Eastern Cities [Review of Microclimate Performance Analysis of Urban Vegetation: Evidence from Hot Humid Middle Eastern Cities]. Plants, 14(4), 521. Multidisciplinary Digital Publishing Institute. https://doi.org/10.3390/plants14040521
18. Korkut, T. B., & Rachid, A. (2024). Numerical Investigation of Interventions to Mitigate Heat Stress: A Case Study in Dubai. Energies, 17(10), 2242. https://doi.org/10.3390/en17102242
19. Wen, J., Abuhani, D. A., Mazzarello, M., Duarte, F., Norford, L. K., & Ratti, C. (2024). Optimizing Pedestrian Comfort in Arid Cities: A Shade-Oriented Pathfinding Approach in Dubai. Research Square (Research Square). https://doi.org/10.21203/rs.3.rs-5379740/v1
20. Alfarsi, Z. M. A., & Üzümcüoğlu, D. (2025). Integrating Biophilic Design and Smart Technologies for Sustainable Urban Development in Muscat, Oman – Advancing SDG 11: Sustainable Cities and Communities. Journal of Lifestyle and SDGs Review, 5(7). https://doi.org/10.47172/2965-730x.sdgsreview.v5.n07.pe07354
21. Najjar, M., Indraganti, M., & Furlan, R. (2025). The Role of Building Geometry in Urban Heat Islands: Case of Doha, Qatar. Designs, 9(3), 77. https://doi.org/10.3390/designs9030077
22. Rajeswari, J. R., Fountoukis, C., Siddique, A., Moosakutty, S. P., Mohieldeen, Y. E., Ayoub, M., & Alfarra, M. R. (2024). Urban heat island phenomenon in a desert, coastal city: The impact of urbanization. Urban Climate, 56, 102016. https://doi.org/10.1016/j.uclim.2024.102016
23. Park, Y., Zhao, Q., Guldmann, J., & Wentz, E. A. (2023). Quantifying the cumulative cooling effects of 3D building and tree shade with high resolution thermal imagery in a hot arid urban climate. Landscape and Urban Planning, 240, 104874. https://doi.org/10.1016/j.landurbplan.2023.104874
24. Schneider, F. A., Ortiz, J. C., Vanos, J., Sailor, D. J., & Middel, A. (2023). Evidence-based guidance on reflective pavement for urban heat mitigation in Arizona. Nature Communications, 14(1). https://doi.org/10.1038/s41467-023-36972-5
25. Ünal, M., & Middel, A. (2025). Improving Thermal Comfort in Hot-Arid Phoenix, Arizona Courtyards: Exploring the Cooling Benefits of Ground Surface Cover and Shade. Building and Environment, 113001. https://doi.org/10.1016/j.buildenv.2025.113001

Reframing Urban Mobility: Assessing Land Use, Accessibility, and Equity through the Riyadh Metro

Abdulrhman M Gbban [0000−0002−1057−277X]

Department of Urban and Regional Planning, Faculty of Architecture and Planning,
King Abdulaziz University, Jeddah 80200, Saudi Arabia
amgabban@kau.edu.sa

Abstract. The riyadh Metro, part of Saudi Arabia's Vision 2030, aims to reduce car use, improve accessibility, and foster sustainable growth. This study explores how its impacts on land use, accessibility, and equity vary across station types. Using GIS land analysis, field audits, a Barrier Effect Index, and 25 stakeholder interviews, it examines six stations representing Riyadh's central, transitional, and peripheral areas. Results show central stations like Al-Olaya and KAFD experience commercial growth, better pedestrian access, and diverse land use, while peripheral stations like Al-Mashaal and Haramain remain disconnected due to zoning, last-mile issues, and weak governance. Applying Bertolini's Node–Place Model and Community Severance Theory, the study highlights that infrastructure alone can't ensure equitable transit outcomes without integrated land use, governance, and social inclusion. It recommends policies prioritizing equity-driven Transit-Oriented Development, coordination with the Metropolitan Transit Development Authority, and smart city strategies aligned with Vision 2030.

Keywords: Riyadh Metro, Transit-Oriented Development, Spatial Equity, Urban Mobility, Community Severance, Gulf Cities, Public Transport Planning

1 Introduction

Urban mobility transformation sits at the heart of Saudi Arabia's Vision 2030, which aspires to diversify the economy, enhance livability, and transition toward sustainable and smart city models. The Riyadh Metro, one of the largest infrastructure investments in the Middle East, is a central component of this transformation. With six operational lines and more than 80 stations, it aims to reduce private car dependency, mitigate congestion, and stimulate equitable urban growth.

Despite its strategic importance, systematic evaluations of the Metro's early spatial and social impacts remain limited. Global literature demonstrates that transit investments can yield mixed outcomes depending on their integration with urban form and governance mechanisms. In rapidly motorized Gulf cities, however, the challenge is twofold: ensuring that transit infrastructure fosters not only accessibility but also spatial justice and inclusive growth.

This study explicitly addresses this gap by asking the following research question:

How does the Riyadh Metro produce differentiated impacts on land use, accessibility, and spatial equity across central, transitional, and peripheral station typologies?
Three sub-questions guide the analysis:

1. How have land-use patterns and densities evolved within 200–800 m of selected metro stations since 2015?
2. To what extent do walkability and perceived accessibility correlate with metro usage and local vitality?
3. What governance, zoning, and social barriers constrain equitable transit-oriented development in peripheral areas?

By combining the Node–Place model and Community Severance Theory, this study evaluates the Riyadh Metro as both a mobility infrastructure and social-urban system. It analyzes six stations representing Riyadh's spatial diversity: Al-Olaya, KAFD, Al-Malaz, King Fahad Road, Al-Mashaal, and Haramain.

The article adds to Gulf-region TOD discourse with one of the first empirical assessments of metro-related urban change in Riyadh. It places its findings within the broader smart city planning context, emphasizing how equitable transit access supports Vision 2030's Quality of Life Program and Riyadh's Smart Mobility Strategy.

2 Literature Review

Transit-Oriented Development (TOD) promotes dense, mixed-use, and pedestrian-friendly areas near transit nodes [4]. Studies from Seoul, Portland, and Hong Kong show TOD can increase land value, reduce car dependency, and encourage sustainable behaviors [10]. However, Litman [7] warns that without car-use restrictions and pedestrian improvements, transit infrastructure might induce demand, offsetting TOD benefits.

In the Gulf region, the TOD concept faces unique implementation challenges. Gbban et al. [6], Aljoufie et al. [1], and Schwedes [12] argue that mono-functional zoning, fragmented planning institutions, and prioritization of capital-intensive development limit the transformative potential of metro systems in cities such as Riyadh, Doha, and Dubai. Aldosary (2020) and MoMRAH policy documents (2019–2023) further note that Saudi cities still operate within regulatory frameworks that emphasize land subdivision and automobile-based access over integrated transport–land-use models.

In this context, the Riyadh Metro offers a unique testbed for assessing how TOD principles are being locally interpreted—and whether they can deliver on promises of equity and livability embedded in Saudi Vision 2030.

Recent debates on "smart cities" highlight the need for transport infrastructure supporting physical, digital, inclusive, and green urban systems. Riyadh's Smart Mobility Strategy (RCRC, 2023) envisions a multi-modal, tech-enabled network, but car-centric infrastructure and inadequate last-mile planning persist. This paper evaluates whether the Riyadh Metro helps create a bright, just, and livable city.

2.1 Node–Place Theory and Station Area Performance

The Node–Place model, developed by Bertolini (Fig. 1) [3], evaluates the dual role of transit stations as transportation nodes (connectivity) and urban places (land-use diversity, pedestrian quality). High-performing station areas are both well-connected and socially vibrant. If node strength does not match place quality, stations become "mobility islands"—technically accessible but disconnected from daily life [13].

The model has been applied in cities such as Lisbon, Melbourne, and Amsterdam, helping identify mismatches between investment priorities and local contexts. However, its use in Gulf cities remains limited, making this study one of the first to operationalize Node–Place evaluations in Riyadh.

Figure 1 shows how a transit station must succeed both as a highly connected network node and a vibrant urban place (left). The right panel shows community severance, with transport corridors creating physical, psychological, and functional barriers.

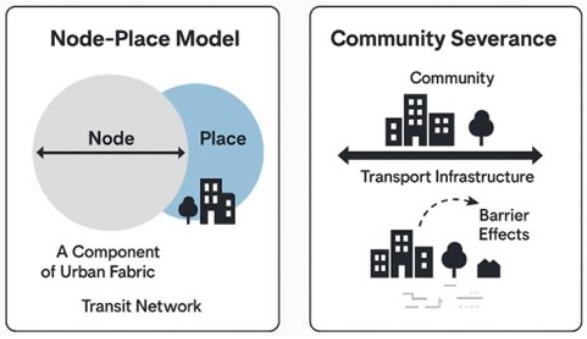


Fig. 1. The Node–Place model

To structure the analysis, this study defines ex-ante performance criteria per station typology:

- Strong Node–Strong Place: Integrated transport and urban vibrancy (e.g., Al-Olaya)
- Moderate Node–Weak Place: Accessible but underutilized urban form (e.g, Al-Malaz)
- Weak Node–Weak Place: Disconnected infrastructure and poor land use (e.g., Haramain)

This typology allows for systematic comparison across Riyadh's urban spectrum and links theoretical expectations to field-measured performance.

2.2 Community Severance and Barrier Effects

Community Severance Theory highlights how transport infrastructure can become a barrier rather than a connector [2, 11]. This includes:

- Physical severance (e.g., highways, viaducts, fencing),
- Psychological severance (e.g., discomfort, fear, disorientation), and
- Functional severance (e.g., poor crossings, lack of shade/signage).

Recent work in Melbourne and London shows that removing physical barriers can improve connectivity—but may also trigger displacement or gentrification without inclusive policies [16]. In the Gulf, Gbban et al. [6] show how Riyadh's arterial roadways create socio-spatial disconnection, particularly for youth, elderly, and low-income populations. This study introduces a calibrated Barrier Effect Index, combining spatial indicators (walkway continuity, pedestrian crossings) and user perceptions (safety, accessibility) to assess severance severity around metro stations.

2.3 Advancing Equity and Institutional Alignment

Infrastructure-led growth without institutional integration often exacerbates inequalities. As Lucas [8] and Pereira et al. [10] argue, transportation justice requires evaluating how infrastructure serves people of different genders, income levels, and abilities—not just whether it exists. In Riyadh, multiple entities, including RCRC, MoMRAH, and the Public Transport Authority (PTA), share overlapping responsibilities, with no unified planning body. This fragmentation impedes TOD implementation, as found in the literature on Gulf governance gaps [15, 5].

The study thus examines:

- Equity of access across station areas, using walkability scores and interview quotes,
- Institutional performance, including coordination failures and planning misalignments.
- Recommendations for establishing a Metropolitan Transit Development Authority (MTDA), akin to Victoria's Level Crossing Removal governance model [9].

2.4 Conceptual Model: Transit-Led Urban Transformation

To integrate these insights, a conceptual model is proposed that aligns with both TOD and smart city frameworks. As shown in Fig. 2, successful metro-led transformation requires convergence across four domains:

1. Infrastructure Quality (frequency, intramodality, reliability)
2. Urban Form Integration (mixed-use density, public space, walkability)
3. Governance and Institutional Capacity (zoning overlays, policy alignment)
4. Social Accessibility (affordability, safety, inclusivity).

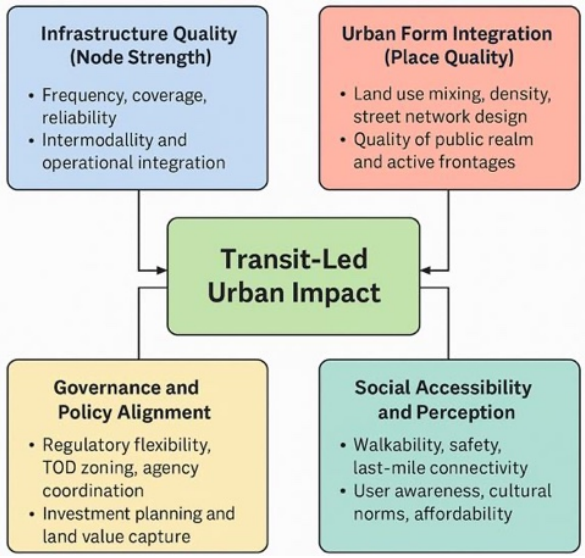


Fig. 1. The proposed conceptual model

This model operationalizes global TOD theory through a Saudi-specific lens, embedding Vision 2030 priorities, including urban humanization, digital inclusion, and spatial equity. It also provides a diagnostic framework for assessing which domains enable or constrain transformation at each station.

3 Methodology

This study assesses the Riyadh Metro's spatial and social impacts using a mixed-methods approach, including GIS analysis, barrier-effect audits, walkability surveys, and stakeholder interviews, conducted from January to March 2024. The methodology comprises six linked parts (Fig3).

3.1 Case Study Selection

Six station areas were purposively selected to represent a spectrum of spatial typologies and socio-economic conditions in Riyadh:

- **Al-Olaya** – Central Business District (CBD); high density and commercial value
- **KAFD** – New financial hub; mixed-use high-rise enclave
- **Al-Malaz** – Historic core; transitioning mixed-use zone
- **King Fahad Road** – Linear corridor; fragmented land uses
- **Al-Mashaal** – Peripheral low-income district
- **Haramain** – Edge city station near highway networks

This stratified sample ensures comparative insight into how metro infrastructure interacts with urban form, governance, and accessibility across typologies.

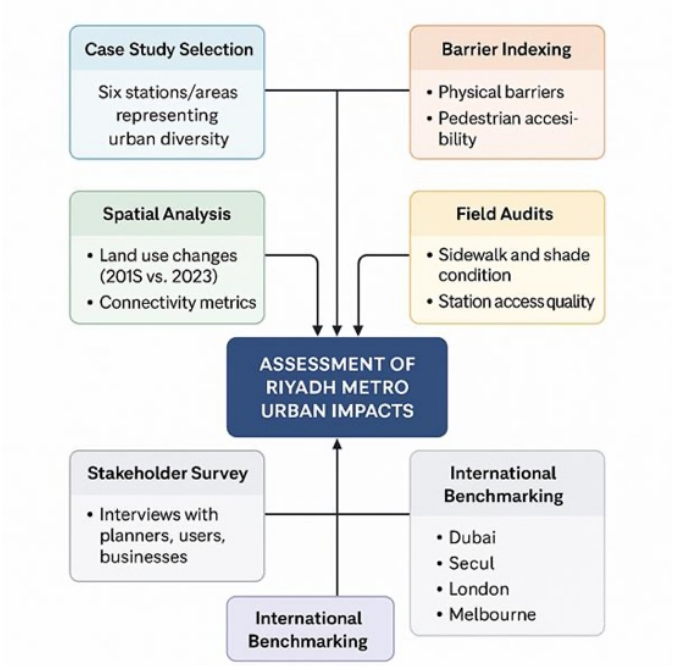


Fig. 2. The components of assessments of Riyadh metro urban impacts

3.2 Spatial and Land Use Analysis

Remote sensing analysis used Sentinel-2 imagery (10m) via Google Earth Engine, comparing 2015 (pre-metro) and 2023 (post-implementation). Land use classifications (residential, commercial, open space, transport, institutional) were derived using supervised classification (Random Forest algorithm) and validated using field ground-truthing (n = 60 points per site) with a classification accuracy of 92.6% (Kappa = 0.89). GIS buffer zones were delineated at 200m, 400m, and 800m radii around each station, following international TOD assessment norms [10]. The following metrics were extracted:

- Commercial and Residential Land Use (m²)
- Floor Area Ratio (FAR) change
- Open/Public Space share (%)
- Intersection Density (per km²)
- Network Connectivity Index (Beta and Alpha indices)

3.3 Barrier Effect Index (BEI)

To capture both objective and subjective severance, we developed a composite Barrier Effect Index (0–5 scale) for each station, drawing from Gbban et al. [6] and van Eldijk et al. [14]. The index comprises five components (Table 1):

Table 1. Components and Scoring Criteria of the Barrier Effect Index (BEI)

Component	Description	Scoring (0–1)
B_1	Physical barriers (walls, fences, highways)	0 = none; 1 = multiple
B_2	Crossing frequency/quality	0 = high quality; 1 = few or unsafe
B_3	Average distance to station entrance	0 = <300m; 1 = >600m
B_4	Shading, sidewalk continuity	0 = continuous/shaded; 1 = fragmented/exposed
B_5	Perceived safety/access (from interviews)	0 = high; 1 = low
$BEI = (B_1 + B_2 + B_3 + B_4 + B_5) \rightarrow$ Total range: 0 (no severance) to 5 (high severance)		

Each index was calculated by two independent coders and cross-validated with an inter-rater reliability of 93.5% (Cohen's Kappa = 0.91). Calibration was done using walkability benchmarks adapted from the ITDP TOD Standard (2017).

3.4 Field-Based Walkability Audits

On-site urban audits were conducted between January 20 and February 28, 2024. The audit tool was adapted from the ITDP TOD Standard and included:

- Sidewalk quality and continuity
- Presence of shading, lighting, seating, and greenery
- Legibility of signage and directional wayfinding
- Cleanliness and station entrance design
- Observation of informal economic activity (vendors, kiosks)

Data were geo-tagged using ArcGIS Collector and documented via a structured checklist and photographic log (n = 162 audit points across 6 sites). These audits informed both spatial scores and qualitative user experience narratives. Fig. 3. illustrates the relative severity of severance at the six stations, with peripheral zones like Al-Mashaal and Haramain showing the most significant connectivity disruptions and deficits.

3.5 Stakeholder Interviews

A total of 25 semi-structured interviews were conducted (**Table 2**), divided across user and institutional categories. Sample guiding questions included:

- "What changes have you observed in your area since the metro opened?"
- "How do you reach the station, and what challenges do you face?"

- "Do you feel safer or more mobile due to the metro?"
- "Has your business improved or worsened?"

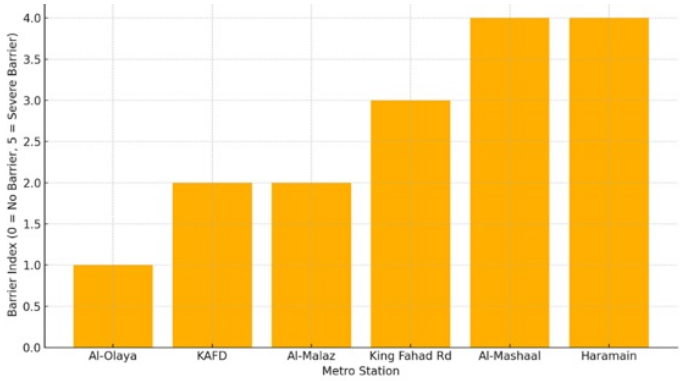


Fig. 4. Barrier Effect index by Station.

Table 2. Stakeholder Groups and Interview Sample for Metro Accessibility Assessment

Group	Number	Description
Daily commuters	10	Male/female, aged 18–55, 5+ trips/month
Business owners	6	Small shops/cafes near stations
Elderly residents	3	Aged 60+, using feeder transport or walking
Urban planners & RCRC officials	4	Strategic/operational roles
Transit staff (Riyadh Metro)	2	Line and station-level operations

Language: Interviews were conducted in Arabic (n=19) and English (n=6) and translated by the author. Verbatim responses were coded thematically using NVivo 14 across six themes: access, safety, equity, economic change, land use, and user satisfaction.

3.6 Quantitative Analysis and Benchmarking

Correlation analysis was conducted between walkability scores (0–10) and monthly ridership growth (%) at each station, using Pearson's r (n = 6 sites; 12-month averages). The result was $r = 0.77$, $p < 0.05$, indicating a strong positive relationship between pedestrian infrastructure quality and transit uptake.

Note: While this correlation suggests an association, causal claims are avoided given the limited sample size and observational design.

Additionally, the study benchmarks Riyadh against:

- Melbourne's Level Crossing Removal Project (governance + severance)
- Seoul's Integrated Rail Corridors (land use + TOD)
- Dubai's Automated Metro and TOD pilot zones
- London's Jubilee Line Extension (accessibility + urban renewal)

Comparative indicators include:

- Transit–land use integration (planning instruments)
- Feeder service availability and quality
- Station ridership vs. population catchments
- Equity and affordability strategies

4 Results and Discussion

This section presents the integrated analysis of the Riyadh Metro's early-stage urban impacts, drawing on geospatial change detection, walkability audits, the Barrier Effect Index, and stakeholder interviews. Results are organized by four thematic domains, aligned with the conceptual model: land-use transformation, mobility and accessibility, economic–environmental shifts, and equity/severance.

4.1 Land-Use Transformation: Central Uptake vs. Peripheral Inertia

Spatial analysis using Sentinel-2 imagery and GIS buffers (200m radius) revealed strong commercial intensification in central stations, especially Al-Olaya (+15,200 m²) and KAFD (+12,400 m²). These areas also showed positive shifts in Floor Area Ratio (FAR), rising by +0.32 and +0.28, respectively, between 2015 and 2023 **(Table 3)**.

In contrast, peripheral stations like Al-Mashaal and Haramain exhibited minimal growth or even loss in residential stock (–1,000 m² and –1,500 m², respectively). These sites remain encumbered by rigid zoning, poor market demand, and low planning prioritization.

Table 3. Net Land-Use and FAR Change within 200 m Buffers (2015–2023)

Station	Δ Commercial (m²)	Δ Residential (m²)	FAR Change (2015–23)
Al-Olaya	+15,200	−500	+0.32
KAFD	+12,400	+1,000	+0.28
Al-Malaz	+8,100	+2,000	+0.14
King Fahad Rd	+6,300	−200	+0.11
Al-Mashaal	+1,900	−1,000	+0.03
Haramain	+1,100	−1,500	+0.02

These patterns confirm the hypothesis that metro-induced land-use transformation is highly conditional on pre-existing urban form, planning alignment, and land-market responsiveness. As one planner from RCRC noted:

"*The metro creates potential, but unless the zoning supports vertical growth, not much will happen around it.*"

This outcome mirrors findings in Tehran and Kuala Lumpur, where TOD benefits accrued unevenly across the urban fabric [5, 10].

4.2 Mobility Patterns and Accessibility Gaps

Walkability audits and interview data reveal a sharp divide between centrally located, pedestrian-accessible stations and peripheral nodes that are disconnected (Table 4).

Field interview insights:

"*The metro is great, but I still need to take a taxi even though it's a 10-minute walk—there's no shade, and it feels unsafe.*" — Female commuter, Al-Mashaal

"*There's no clear signage. Tourists get confused. Locals avoid it*." — Vendor, Haramain Station.

These insights reinforce the importance of last-mile design, especially for vulnerable populations, and validate the use of walkability audits and barrier scoring as proxies for TOD performance.

Table 4. Walkability, Feeder Services, and Accessibility Characteristics by Station

Station	Walkability Score	Feeder Services	Accessibility Notes
Al-Olaya	8.5	Yes	Shaded, legible paths and continuous sidewalks
KAFD	7.8	Yes	Ongoing path development, good signage
Al-Malaz	6.4	Partial	Historic grid; poor shading
King Fahad Rd	6.1	No	Wide arterials, unsafe crossings
Al-Mashaal	4.2	No	No shade, lengthy detours, few crossings
Haramain	3.8	No	Disconnected from the pedestrian network

Correlation analysis ($r = 0.77$, $p < 0.05$) indicates that walkability is a strong predictor of ridership uptake, though not sufficient on its own.

4.3 Economic and Environmental Shifts

Economic vitality varied widely. Al-Olaya and KAFD recorded 20–30% increases in commercial footfall, attributed to better metro access and increased land desirability.

"We now have customers who never used to visit our shop before. Weekends are especially busier.".

— Retail operator, Al-Olaya

By contrast, Al-Mashaal and Haramain showed no measurable commercial uplift—highlighting the limits of infrastructure without surrounding urban activation. This supports the "mobility island" effect identified by Bertolini [3]. Fig. 4 shows the divergence in commercial rent index (base 100 = 2015), where central stations outpaced peripheral ones by over 20 points.

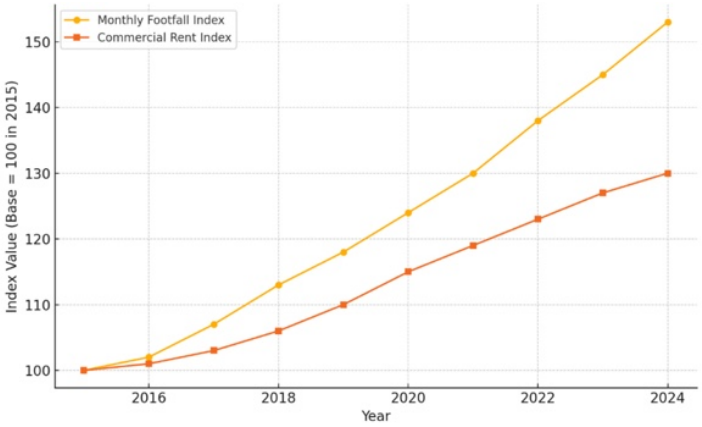


Fig. 4. Monthly Footfall vs. Commercial Rent Index (Base 100 = 2015)

Environmentally, interviewees reported a modest modal shift:

Average weekly car trips dropped from 4.8 to 3.2 among metro users, however, system-wide emissions remain high, due to a lack of:

- Park-and-ride facilities
- Feeder service integration
- Low off-peak demand

"I'd take the metro more often if it connected to my neighborhood directly. Driving is still faster." — Male commuter, King Fahad Road

This suggests that technical infrastructure must be embedded in a multimodal ecosystem to generate environmental benefits.

4.4 Barrier Effects and Spatial Justice

Using the Barrier Effect Index, Haramain and Al-Mashaal scored 4 out of 5 (Fig. 6 & Table 5)—indicating severe severance due to:

- Major roadways,
- Absence of crossings,
- Lack of shade,
- Poor signage and wayfinding.

The clustering of high barrier scores with low-income communities worsens spatial injustice. Residents in Al-Mashaal and Haramain frequently walk over 800 meters to access the station, often lacking shade, clear guidance, or safety features.

"We see the metro but can't reach it without walking along fast roads." — Elderly resident, Haramain

"It's designed for those with cars. People like us feel excluded." — Youth, Al-Mashaal

These insights show that smart city infrastructure risks reproducing inequalities unless equity is embedded into both design and governance. Peripheral users walk farther under more hostile conditions, despite being the most transit-dependent.

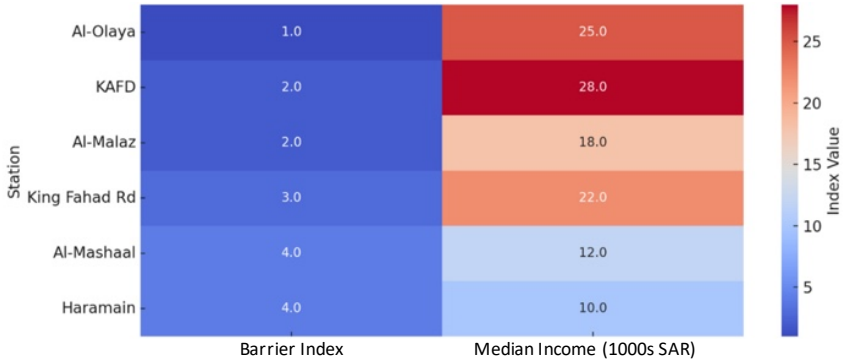


Fig. 5. Monthly Footfall vs. Commercial Rent Index (Base 100 = 2015).

Table 5. Barrier Effect Index (BEI) Scores and Dominant Severance Types by Station

Station	BEI Score (0–5)	Dominant Barrier Type
Al-Olaya	1	Minor visual clutter, but highly walkable
KAFD	2	Fencing and block scale interrupt paths
Al-Malaz	2	Few signalized crossings
King Fahad Rd	3	Heavy traffic, disconnected sidewalks
Al-Mashaal	4	Long distances, unshaded, unsafe paths
Haramain	4	Highway encirclement, no direct linkage

5 Recommendations and Policy Directions

This section outlines four pillars to leverage the Riyadh Metro as a catalyst for equitable, sustainable, and smart urban transformation. These recommendations support the proposed model and address spatial disparities across station types.es.

5.1 Spatial Integration: From Infrastructure to Inclusive Urban Fabric

Issue: Peripheral stations remain disconnected due to zoning rigidity, low density, and weak regeneration policies.

Recommendations:

- Introduce TOD Zoning Overlays within an 800-meter radius of all metro stations to allow mixed-use mid-rise development, reduce parking minimums, and incentivize vertical growth.

- Launch regeneration pilot programs in underserved areas such as Al-Mashaal and Haramain, linking transit investment with affordable housing, public realm upgrades, and economic anchors (e.g., municipal offices, health centers).
- Implement Land Value Capture (LVC) mechanisms (e.g., betterment levies, development rights transfer) to reinvest gains from central stations into peripheral upgrades.

5.2 Accessibility: Fixing the Last-Mile Deficit

Issue: Walkability, signage, and last-mile connectivity remain major barriers to effective metro access, particularly in edge districts.

Recommendations:

- Upgrade pedestrian infrastructure with shade trees, continuous sidewalks, safe crossings, and legible, bilingual wayfinding systems (Arabic–English).
- Deploy micromobility solutions such as bike- and scooter-sharing in high-potential but underserved areas.
- Redesign feeder loops with real-time integration between buses, taxis, and metro entry points—especially in low-mobility neighborhoods.

5.3 Social Equity: Embedding Inclusivity into TOD Practice

Issue: Current development patterns risk reinforcing exclusion—access is skewed toward middle-income users, while low-income residents face spatial and psychological severance.

Recommendations:

- Establish Equity-First TOD Demonstration Zones, embedding:
- Affordable housing mandates (e.g., 25% below-market units),
- Community amenities (e.g., clinics, youth centers),
- Gender-sensitive urban design (lighting, surveillance, inclusive signage).
- Create participatory planning platforms for residents and local businesses to co-design TOD strategies and feedback systems.
- Adopt equity indicators (walkshed quality, time-cost burden, user satisfaction) into metro performance KPIs.

5.4 Governance: Aligning Institutions through a Metropolitan Transit Development Authority (MTDA)

Issue: Institutional fragmentation hinders integrated TOD implementation across land use, housing, and transport agencies.

Recommendations:

- Establish a Metropolitan Transit Development Authority (MTDA) to coordinate strategic planning across:
  - RCRC (metro infrastructure),
  - MoMRAH (zoning and urban development),
  - NHC (housing and affordability).

- Adopt performance-based planning instruments, where agencies are held accountable to cross-sector KPIs (e.g., TOD coverage, equity index, emissions avoided).
- Develop an interagency knowledge-sharing platform, including best practices from Melbourne, Seoul, and Dubai.

These recommendations advance Riyadh's ambitions to become a leading smart city by 2030. By integrating mobility, land use, digital systems, and social equity, the metro can evolve from a transportation project into a platform for inclusive innovation. Embedding equity and spatial justice into smart city planning is no longer optional—it is essential to sustainable urban governance.

6 Conclusion

This study assessed the Riyadh Metro's impact on land use, accessibility, and spatial equity across central and peripheral station zones using GIS, walkability audits, the Barrier Effect Index, and interviews. Three key conclusions emerged:

1. Infrastructure Alone is Insufficient: Central stations show commercial growth and better pedestrian access, while peripheral stations remain disconnected. This indicates that successful transit-oriented development (TOD) requires supporting policies, last-mile solutions, and coordinated planning beyond just infrastructure.
2. Equity-First TOD is Essential: Unintentional metro expansion risks exacerbating urban disparities. The study proposes "Equity-First TOD" models, integrating affordable housing, community services, and participatory design, alongside the Barrier Effect Index to mitigate disconnection, as crucial for inclusive urbanism.
3. Governance Reform is Critical: Riyadh's fragmented planning limits the metro's potential. Establishing a Metropolitan Transit Development Authority (MTDA) would enable integrated land, transport, and housing strategies, shifting towards cohesive and equitable urban planning.

The research extends the Node-Place model and Community Severance theory to a Gulf megacity, offers an evaluative framework combining spatial metrics with user experiences, and provides evidence that TOD outcomes depend on governance and social design, not just infrastructure. These findings are applicable to other rapidly urbanizing cities in the Middle East, North Africa, and the Global South.

References

1. Aljoufie, M., Zuidgeest, M., Brussel, M., & van Maarseveen, M. (2011). Urban growth and transport: Understanding the spatial–temporal relationship. In C. A. Brebbia & A. Pratelli (Eds.), *Urban Transport XVII* (pp. 315–328). WIT Press.
2. Anciaes, P. R., Jones, P., & Mindell, J. S. (2016). Community severance: Where is it found and at what cost? *Transport Reviews*, 36(3), 293–317. https://doi.org/10.1080/01441647.2015.1077286
3. Bertolini, L. (1999). Spatial development patterns and public transport: The application of an analytical model in the Netherlands. *Planning Practice & Research*, 14(2), 199–210. https://doi.org/10.1080/02697459915724
4. Cervero, R., & Murakami, J. (2009). Rail and property development in Hong Kong: Experiences and extensions. *Urban Studies*, 46(10), 2019–2043. https://doi.org/10.1177/0042098009339431
5. Forouhar, A., & Kheyroddin, R. (2016). The impact of commercialisation on the spatial quality of residential neighbourhoods: Evidence from Nasr neighbourhood of Tehran. *Geographical Planning of Space Quarterly Journal*, 6(20), 63–84.
6. Gbban, A. M., Kamruzzaman, M., Delbosc, A., & Coxon, S. (2023). Understanding spatial change and social impact in Melbourne's level crossing removal project. *Journal of Transport and Land Use*, 16(1), 103–121. https://doi.org/10.5198/jtlu.2023.2341
7. Litman, T. (2017). Generated traffic and induced travel: Implications for transport planning. *Victoria Transport Policy Institute*. http://www.vtpi.org/induces.pdf
8. Lucas, K. (2012). Transport and social exclusion: Where are we now? *Transport Policy*, 20, 105–113. https://doi.org/10.1016/j.tranpol.2012.01.013
9. State Government of Victoria. (2022). *Project overview – Victoria's Big Build*. https://bigbuild.vic.gov.au/projects/level-crossing-removal-project/about/project-overview
10. Pereira, R. H. M., Schwanen, T., & Banister, D. (2017). Distributive justice and equity in transportation. *Transport Reviews*, 37(2), 170–191. https://doi.org/10.1080/01441647.2016.1208265
11. Roda, C., Charreire, H., Feuillet, T., Mackenbach, J. D., Compernolle, S., Glonti, K., & Oppert, J. M. (2016). Mismatch between perceived and objectively measured environmental obesogenic features in European neighborhoods. *Obesity Reviews*, 17(S1), 31–41. https://doi.org/10.1111/obr.12386
12. Schwedes, O. (2011). The field of transport policy: An introduction. *Transport Policy*, 18(1), 1–6. https://doi.org/10.1016/j.tranpol.2010.07.003
13. Vale, D. S., Viana, C. M., & Pereira, M. (2018). The extended node–place model at the local scale: Evaluating the integration of land use and transport for Lisbon's subway network. *Journal of Transport Geography*, 69, 282–293. https://doi.org/10.1016/j.jtrangeo.2018.05.014
14. van Eldijk, J., Gil, J., & Marcus, L. (2022). Disentangling barrier effects of transport infrastructure: Synthesising research for the practice of impact assessment. *European Transport Research Review*, 14(1), 1–19. https://doi.org/10.1186/s12544-021-00517-y
15. Woodcock, I. (2016). The design speculation and action research assemblage: 'Transit for All' and the transformation of Melbourne's passenger rail system. *Australian Planner*, 53(1), 15–27. https://doi.org/10.1080/07293682.2015.1132870
16. Woodcock, I., & Martin, S. (2016). Of skyrails and skytrains – Elevated rail in the Australasian urban transport environment. In *Proceedings of the 38th Australasian Transport Research Forum* (pp. 1–16). Australasian Transport Research Forum.

Urban Agriculture with an IAR-FAO Approach During COVID-19: A Socio-Educational Strategy from Cuenca, Ecuador

Fredi Portilla Farfán[1][0000-0002-1234-9758], Paola Duque-Sarango[2][0000-0003-4484-7273], María José Quizhpi Pintado[1], Tania Alexandra Sarango Guamán[1].

[1]Grupo de investigación en Comunicación, Educación y Ambiente GICEA, Universidad Politécnica Salesiana. Calle Vieja 12-30 y Elia Liut, Cuenca 010203, Ecuador.

[2]Grupo de Investigación en Recursos Hídricos (GIRH-UPS), Universidad Politécnica Salesiana. Calle Vieja 12-30 y Elia Liut, Cuenca 010203, Ecuador.

pduque@ups.edu.ec

Abstract. This paper describes a socio-educational strategy implemented in Cuenca, Ecuador, during the COVID-19 pandemic to address food insecurity through urban agriculture in public schools. The initiative was based on the FAO's Resilient Agroecological Intensification (IAR-FAO) framework and involved children and adolescents in the design, implementation, and maintenance of organic gardens. Activities included environmental education workshops, participatory learning sessions, and encouraging students and their families to replicate home gardens. The methodology followed a four-phase process: planning, training, action, and evaluation. A total of 229 students from five schools participated, with measurable impacts on environmental awareness, food self-sufficiency, and gender empowerment. Results indicated that 80% of students demonstrated a strong understanding of agroecological principles and food security concepts. Moreover, school gardens served not only as educational spaces but also as tools to strengthen community resilience during crises. This experience highlights the potential of inclusive, school-based urban agriculture to support social sustainability by enhancing food sovereignty, promoting environmental stewardship, and building community solidarity in vulnerable urban areas. It contributes to the Smart Communities concept by integrating nature-based education and citizen engagement into urban resilience strategies.

Keywords: Urban agriculture, Food security, COVID-19, IAR-FAO principles, Environmental education, Social sustainability, ODS6.

1 Introduction

The COVID-19 pandemic significantly disrupted food systems worldwide, exposing structural weaknesses in urban food supply chains and worsening socio-economic inequalities. In Latin American cities, including Cuenca (Ecuador), this crisis increased

the vulnerability of children and families living in peripheral urban neighborhoods, especially regarding food insecurity and limited access to fresh, nutritious produce. In this context, urban agriculture re-emerged as a practical strategy for self-sufficiency and resilience.

Urban agriculture refers to the cultivation of food within or around cities, often on small plots, rooftops, or community spaces. Beyond food production, it serves social, educational, and environmental purposes by fostering sustainable habits, green infrastructure, and community cohesion [1,2]. Within academic institutions, school gardens can be leveraged as experiential learning environments that promote environmental awareness, healthy eating, and inclusive participation [3,4].

Between 2020 and 2021, a participatory project was carried out in five public schools in Cuenca to improve food security and social inclusion through agroecological practices. The initiative followed the Principles for Responsible Investment in Agriculture and Food Systems (IAR-FAO), which promote fair, inclusive, and sustainable food systems [5]. Children, families, and teachers worked together to create and manage organic school gardens, supported by environmental education workshops and gender-sensitive activities.

This paper outlines the methodology, results, and lessons learned from this intervention. It emphasizes how school-based urban agriculture can serve as a powerful tool to promote social sustainability by tackling food justice, empowering youth, and strengthening community resilience in crises. Urban agriculture is aligned with the Smart Cities paradigm by promoting citizen participation, local data generation, and sustainable urban governance. In this sense, school gardens represent a bridge between community engagement and environmental monitoring, supporting inclusive and adaptive urban systems.

2 Theoretical Framework

2.1 Urban Agriculture

Urban agriculture involves a variety of productive activities conducted in and around cities, including growing vegetables, fruits, medicinal plants, and sometimes small-scale animal husbandry. It transforms vacant lots, backyards, rooftops, or schoolyards into sites for food production and community involvement. In cities of the Global South, it acts as a response to food shortages, economic crises, and social marginalization [1].

Besides providing fresh, nutritious produce, urban agriculture also offers various co-benefits: it boosts urban greenery, recycles organic waste, improves air quality, cuts down food transportation emissions, and encourages environmental education [2]. In schools, gardens serve as practical teaching tools for science, nutrition, teamwork, and environmental care [3]. These advantages are significant in post-pandemic urban areas, where mental health, well-being, and local resilience have become urgent concerns.

2.2 IAR-FAO Principles

The Principles for Responsible Investment in Agriculture and Food Systems (IAR-FAO), endorsed by the Committee on World Food Security (CFS) in 2014, provide a rights-based framework to guide agricultural development by the Sustainable Development Goals (SDGs). These principles highlight food security and nutrition, economic inclusion, environmental sustainability, and fair access to resources [4].

Among the nine principles, this study focuses on five: (1) contribute to food security and nutrition, (2) foster sustainable and inclusive economic development, (3) promote gender equality, (4) empower youth, and (8) ensure safe and healthy agricultural systems. These principles were translated into practical educational actions and activities throughout the project. Their integration into school-based urban agriculture enhances not only technical knowledge but also civic and ethical awareness among students and educators.

2.3 Food Security and Educational Inclusion

Food security is defined as reliable, consistent access to enough, safe, and nutritious food to support an active and healthy life. In school communities, food security includes both the availability of healthy meals and the ability to make informed food choices [5]. The COVID-19 pandemic disrupted food supply chains and household economies, making educational institutions important platforms for restoring food access and promoting long-term food sovereignty.

At the same time, education systems play a vital role in tackling social inequality and encouraging inclusive participation. Incorporating food production into school curricula enhances students' understanding of sustainability, nutrition, and community health [6]. Additionally, practical agricultural education can inspire students to become proactive change-makers, sharing knowledge and practices within their homes and neighborhoods.

3 Methodology

This process started with the collaborative design of activities and the selection of participating institutions and students. It then moved on to implementing school gardens and conducting agroecological training workshops. These two parts were developed simultaneously to strengthen both practical and theoretical knowledge. Finally, the project concluded with an evaluation and analysis phase that used both quantitative tools and qualitative insights to assess educational and social outcomes. The entire methodological sequence is summarized (see Fig. 1).

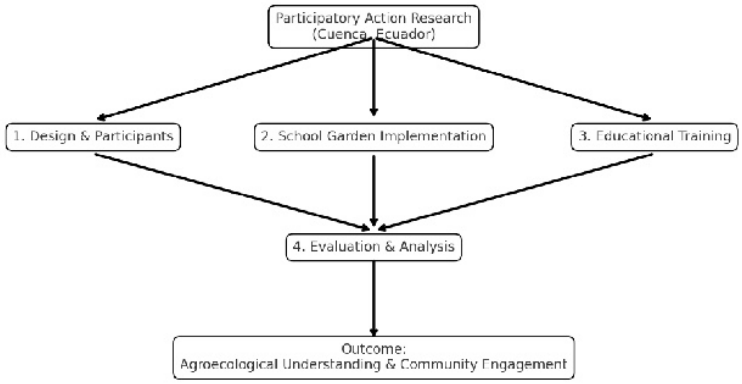


Fig. 1 Experimental design of the participatory school garden project showing four phases: (1) design and student selection, (2) garden implementation, (3) agroecological training based on IAR-FAO principles, and (4) mixed-method evaluation.

3.1 Design and Participants

The planning phase involved identifying suitable schools, selecting student groups, and establishing collaboration frameworks with teachers and families. The target groups consisted of fifth- and sixth-grade students (ages 9–12), selected for their cognitive and emotional readiness to engage in agricultural learning and for the feasibility of long-term follow-up. A total of 229 students (124 boys and 105 girls) from five institutions were involved.

Each school designated a physical space for establishing a garden, either in open soil or raised beds, depending on available infrastructure. Educators and project facilitators collaboratively developed activity schedules and crop plans.

3.2 Implementation of School Gardens

The implementation phase involved direct student participation in preparing, maintaining, and harvesting school gardens. Local black soil was mixed with vermicompost to boost fertility. Guided by their teachers, students engaged in soil conditioning, crop sowing, irrigation, and organic pest control. The chosen crops included lettuce (*Lactuca sativa*), broccoli (*Brassica oleracea*), cabbage (*Brassica viridis*), and aromatic herbs. A randomized block design (RBD) was used for planting: 48 plants were arranged into four treatments with three replications each per school.

Irrigation was carried out manually, using harvested rainwater or tap water depending on local availability. Pest and weed control relied on organic methods, such as manual removal and natural repellents. After three months, students and their families

collaborated to harvest the produce, strengthening community involvement and food self-sufficiency (see Fig. 2).

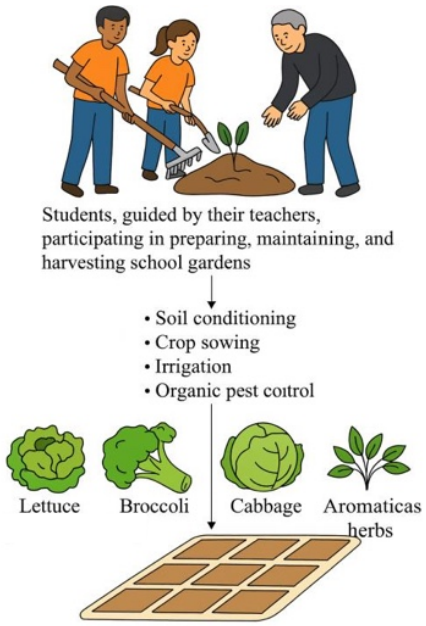


Fig. 2 Participatory experimental design for school garden implementation.

Meanwhile, a series of educational workshops were held to enhance agroecological knowledge and responsible food production principles. Table 1 summarizes the training sessions provided at each school.

Table 1. Educational Workshops Conducted in Each School

Workshop	**Topics Covered**
Introduction to Urban Agriculture	•Concept of urban agriculture and school gardens •Benefits of urban and home gardens •Importance of growing food at home •Types of home gardens •Stages of a garden: -soil preparation - sowing - irrigation - harvesting
Good Agricultural Practices (GAP) and Cultivation Techniques	• Basic cultivation techniques • Cultural practices and crop care • Good Agricultural Practices (GAP)

		• Pillars of GAP: - food safety - food security
Responsible Investment Principles in Agriculture (IAR-FAO)		• Key IAR-FAO Principles: - Principle 1: Contribute to food security and nutrition - Principle 2: Promote sustainable and inclusive economic development - Principle 3: Support gender equality and women's empowerment - Principle 4: Empower youth - Principle 8: Ensure safe and healthy food systems

3.3 Activities to Promote IAR-FAO Principles

To strengthen students' understanding of responsible and inclusive agriculture, a series of interactive and play-based activities was carried out. These actions promoted key IAR-FAO principles, including food security, gender equality, and sustainable food systems. The goal was to engage students in participatory learning and reinforce their environmental and nutritional awareness. Table 2 summarizes the core activities implemented during the project.

Table 2. Activities Conducted to Promote IAR-FAO Principles Among Students

Activity	Description
Wheel of Fortune	This game aimed to enhance learning by making urban agriculture concepts more accessible and engaging. Ten students (5 boys and 5 girls) each spun a colorful wheel. Wherever the pointer landed, the student had to explain or reflect on that topic (e.g., composting, planting, water use). This encouraged active recall and peer learning.
Word Search Puzzle	This activity supported cognitive development by stimulating concentration and vocabulary building. Students were asked to find specific agroecological terms hidden in a word grid, then write their definitions. This reinforced key concepts while promoting spelling, focus, and thematic understanding.
Educational Field Trip to San Joaquín Agroecological Garden	A guided field visit was organized to a large-scale agroecological garden located in San Joaquín. Students observed organic production techniques, interacted with local farmers, and connected classroom theory with real-world practice. Such experiences have been shown to enhance learning outcomes and foster environmental responsibility.
Creation of an Environmental	Each student contributed to a booklet structured around the five IAR-FAO principles emphasized in the project. Sections included key facts, fun illustrations, reflections, and personal

Education Booklet	commitments. The process stimulated creativity and reinforced the importance of responsible food systems.

3.4 Evaluation

Evaluation is a vital part of students' personal growth, as it offers insights into their progress or challenges in a specific subject over time. In this project, students' understanding of key agroecological and sustainability concepts was assessed using a structured crossword-based tool. Each item was rated on an ordinal scale designed to rank different levels of qualitative understanding. Scores ranged from 1 (poor) to 4 (excellent), as shown in Table 3.

Table 3. Student Comprehension Assessment Scale

Score	Description	Interpretation
1	The student did not demonstrate an adequate understanding of the topic.	Poor
2	The student demonstrated partial understanding of the topic.	Fair
3	The student demonstrated nearly satisfactory comprehension of the topic.	Good
4	The student demonstrated full and excellent comprehension of the topic.	Excellent

3.5 Data Collection and Analysis

Data from environmental education evaluations were compiled and analysed using descriptive statistics. Graphs were used to illustrate differences in performance between schools and between male and female students. Descriptive statistics (means, percentages and standard deviations) were employed to evaluate comprehension differences between schools and genders. Excel and cross-tabulation were used for data visualisation to explore potential associations between participation levels and comprehension outcomes, and field observations, photos and feedback from teachers and parents were also included as qualitative evidence to aid interpretation of the results. Using both quantitative and qualitative methods provided a thorough understanding of the educational, social and environmental impacts of the initiative.

4 Results and Discussion

This section presents the project's primary outcomes and interprets them in terms of the social sustainability goals: promoting food security, environmental education, youth empowerment, and community resilience.

4.1 Evaluation of Environmental Education Based on IAR-FAO Principles

After using the evaluation tool, data were recorded to assess students' understanding of key environmental education topics. It is essential to note the total number of students assessed, broken down by gender, as shown in Table 4.

Table 4. Sample Size by Educational Institution and Gender

Institution	Grade	Male Students	Female Students	Total
Cristo Rey Basic Education School	6th	5	4	9
Francisca Dávila de Muñoz School	5th "A"	22	19	41
Francisca Dávila de Muñoz School	5th "B"	20	19	39
Federico Proaño Basic Education School	8th "A"	36	19	55
Julio María Matovelle Francisca Dávila de Muñoz School	5th "A"	20	22	42
Tres de Noviembre Francisca Dávila de Muñoz School	6th "A"	21	22	43
Total	—	**124**	**105**	**229**

Following this demographic overview, Fig. 3 shows the percentage distribution of student comprehension levels by school, based on the four-tier scale introduced earlier (Table 4: Excellent, Good, Fair, Poor). These results provide valuable insight into the educational impact of the intervention at each school.

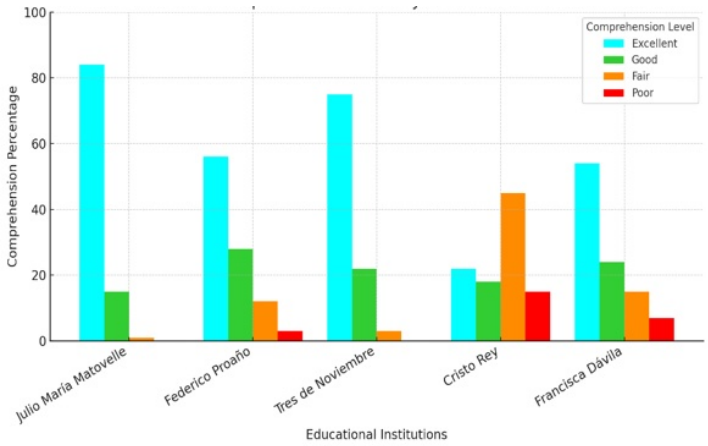


Fig. 3 Student comprehension levels across five schools based on IAR-FAO principles. *Julio María Matovelle* and *Tres de Noviembre* achieved the highest "Excellent" ratings, while *Cristo Rey* showed a broader distribution including more "Fair" and "Poor" results.

As shown in the chart, *Julio María Matovelle* and *Tres de Noviembre* schools achieved the highest levels of comprehension, with 84% and 75% of students rated as "Excellent," respectively. In contrast, *Cristo Rey School* had a lower proportion in the highest category and a more balanced distribution across all levels, including higher percentages of students in the "Fair" and "Poor" categories.

These differences may reflect varying degrees of school engagement, teacher involvement, and socio-environmental context. Overall, the data indicate a positive trend toward student understanding of agroecological concepts and responsible food systems. Differences in comprehension levels among schools revealed significant variations in the degree of institutional involvement and community participation. As shown in Fig. 3, *Julio María Matovelle* and *Tres de Noviembre* achieved the highest comprehension rates (84% and 75% rated as "Excellent," respectively), whereas *Cristo Rey* displayed a more dispersed performance, including higher percentages in the "Fair" and "Poor" categories.

These contrasts may be explained by contextual factors such as the level of teacher engagement, parental participation, and the socio-environmental setting of each institution. Schools that actively integrated the gardens into their curricular and extracurricular activities tended to generate stronger learning outcomes. This suggests that institutional commitment and continuity of practice are decisive factors in fostering meaningful environmental learning. Moreover, the gender-balanced participation and emphasis on peer collaboration appear to have reinforced students' motivation and sense of collective responsibility.

From an analytical standpoint, these findings demonstrate that educational performance improves when agroecological learning is embedded within a supportive community network. Schools that promoted continuous reflection, teacher facilitation, and student-led initiatives displayed not only higher comprehension scores but also greater enthusiasm for replicating the experience at home. This relationship highlights a key feature of Smart Communities: the co-production of knowledge and collective empowerment as drivers of sustainable behavior.

Therefore, higher comprehension levels in schools with strong teacher engagement suggest that institutional commitment is a critical variable in the success of community-based smart education models. The observed pattern aligns with participatory learning theories and with FAO's emphasis on responsible investment in education and food systems, where knowledge-sharing and social inclusion underpin resilience.

4.2 Learning Outcomes on Environmental and Food Education

Evaluation results indicated that the majority of students achieved a high level of comprehension across the nine thematic items. Specifically, 58% of students were rated "excellent" and 22% "good" in their understanding of concepts such as planting, food safety, pest control, and the importance of gender equality in agriculture. Schools with greater involvement from teachers and administrative staff (e.g., Federico Proaño and Julio Matovelle) reported the highest comprehension levels, with 75%–84% of students scoring in the "excellent" range.

These results suggest that integrating practical garden-based learning with theoretical content significantly enhances students' capacity to internalize and apply agroecological and sustainability principles. The FAO has reported similar findings in other Latin American school garden programs [1].

4.3 Urban Gardens as a Tool for Community Resilience

During the COVID-19 lockdown, several families replicated home gardens using the skills and materials learned from school-based activities. This spontaneous replication demonstrates the transferability of knowledge from institutional to domestic environments, turning households into micro-centers of sustainability. Families reported improvements in food variety, reduced expenses, and a renewed appreciation for local agroecological practices.

Such replication processes reflect the principles of Smart Communities, where citizen-led innovation, peer-to-peer learning, and adaptive knowledge networks enhance urban resilience. By transforming educational experiences into community action, the project nurtured a form of social intelligence—a bottom-up process through which communities collectively address food insecurity and environmental challenges.

Furthermore, this approach exemplifies how digital and participatory strategies can be integrated into future stages of similar initiatives. For instance, the use of mobile platforms to record crop data or monitor household garden growth could strengthen feedback loops between schools, families, and municipal programs. This transition from localized agroecological action to networked civic participation represents a tangible pathway for operationalizing Smart City goals through community-driven sustainability (Fig. 4).



Fig. 4 Family garden sowing during the experimental phase.

4.4 Gender, Empowerment, and Youth Engagement

The project explicitly promoted gender equality and youth empowerment, aligned with Principles 3 and 4 of the IAR-FAO framework. Activities such as group work, environmental education games, and the environmental booklet fostered inclusive participation among boys and girls. Teachers observed that girls showed strong leadership in organizing gardening tasks and decision-making.

Empowering youth to become actors in their communities supports long-term sustainability. By engaging students in hands-on agroecological activities, the initiative not only enhanced food security but also developed critical awareness, cooperative skills, and environmental responsibility—key dimensions of social sustainability [2,3].

5 Conclusions and Future Work

This study demonstrates that urban agriculture, when integrated into educational environments under a socially inclusive framework, can serve as a powerful tool for enhancing food security, environmental awareness, and community resilience in times of crisis. The implementation of school gardens based on IAR-FAO principles in Cuenca, Ecuador, contributed not only to the production of fresh food but also to the development of key values such as equity, cooperation, and sustainability among children and families. The study provides a model for incorporating social sustainability into smart city policies by promoting food sovereignty, digital education and participatory urban governance.

The project's participatory nature empowered students, especially girls and marginalized youth, to take active roles in food production and environmental stewardship. The initiative also stimulated home-based replication, indicating its potential for broader community impact. These outcomes align with the goals of social sustainability in smart cities, particularly by fostering inclusion, health equity, and citizen engagement through nature-based, low-cost strategies.

As a replicable and adaptable model, this experience could inform educational policy and municipal sustainability programs in other urban settings. Future work should focus on scaling the model to secondary schools, integrating digital tools (e.g., IoT-based irrigation or mobile monitoring apps), and evaluating long-term impacts on household nutrition and community resilience. Additionally, interdisciplinary collaborations between educators, agronomists, and social scientists could enhance the pedagogical and technical components of urban agriculture in school systems.

By reinforcing local food systems and cultivating socio-environmental responsibility from an early age, this model contributes meaningfully to the vision of inclusive, resilient, and sustainable cities.

Future projects could include mobile-based monitoring of garden performance, digital mapping of community gardens, and participatory data collection to inform urban food policy.

Acknowledgments. This study was supported by the Universidad Politécnica Salesiana of Ecuador, through its institutional programs on applied research and community engagement. The authors thank the participating schools, teachers, and families for their collaboration and commitment throughout the project.

Disclosure of Interests. Disclosure of Interests. The authors declare no competing interests relevant to the content of this article.

References

1. FAO. Urban and peri-urban agriculture: A strategy for food security and sustainable cities. Rome: Food and Agriculture Organization of the United Nations; 1999.
2. Moreno O. Agricultura urbana: nuevas estrategias de integración social y recuperación ambiental en la ciudad. Rev Electr Diseño Urbano Paisaje.
3. FAO. School gardens: Nutrition and education go hand in hand. Rome: Food and Agriculture Organization of the United Nations; 2005. http://www.fao.org/newsroom/en/news/2005/104116/index.html
4. Altieri MA, Toledo VM. The agroecological revolution of Latin America: Rescuing nature, securing food sovereignty and empowering peasants. J Peasant Stud. 2011;38(3):587–612.
5. Committee on World Food Security (CFS). Principles for Responsible Investment in Agriculture and Food Systems. Rome: FAO/IFAD/UNCTAD/World Bank; 2014. http://www.fao.org/3/a-au866e.pdf
6. FAO. Family farming in Latin America and the Caribbean: Policy recommendations. Santiago: FAO Regional Office for LAC; 2014.
7. Hernández L. La agricultura urbana y caracterización de sus sistemas productivos y sociales como vía para la seguridad alimentaria en nuestras ciudades. Cult Trop. 2006;27(2):13–25.
8. Moreno F. Huertos escolares de la región de Murcia. Murcia: Consejería de Educación, Formación y Empleo; 2013.
9. Merçon J, Escalona MA, Noriega MI, Figueroa II, Atenco A, González ED. Cultivando la educación agroecológica: El huerto colectivo urbano como espacio educativo. Rev Mex Inv Educ. 2012;17(55):1201–1224.
10. Rosales R, Flores H. Importancia del agua de riego para la producción sostenible de frijol en Durango. Instituto Nacional de Investigaciones Forestales, Agrícolas y Pecuarias (INIFAP); 2017.
11. Silva VM. Manual para el productor: El cultivo de las hortalizas. Proyecto Jatun Sach'a – UNODC Bolivia; 2017. https://www.unodc.org/documents/bolivia/DIM_Manual_de_cultivo_de_hortalizas.pdf
12. Gramunt N. Ejercicios y actividades para la estimulación cognitiva. Fundación "la Caixa"; 2010. https://fiapam.org/wp-content/uploads/2012/10/ejercicios_es.pdf
13. García D, Priotto G. Aportes políticos y pedagógicos en la construcción del campo de la Educación Ambiental. Secretaría de Ambiente y Desarrollo Sustentable, Argentina; 2009.
14. Tenorio KP, Fuenmayor MM. Salidas de campo como estrategia didáctica para la enseñanza-aprendizaje de las ciencias naturales. Univ Córdoba, Colombia; 2018. https://repositorio.unicordoba.edu.co/handle/ucordoba/930
15. Céspedes C. Producción hortofrutícola orgánica. Ministerio de Agricultura, Chile; 2012. http://biblioteca.inia.cl/medios/biblioteca/boletines/NR38258.pdf

El marco de las Ciudades Inteligentes como un reforzamiento de la Cohesión Urbana Inclusiva

Miroslaba Taide Bautista Flores[1] and Ryzard Edward Rozga Luter[2]

[1] Doctorante del Doctorado en Urbanismo de la Facultad de Planeación Urbana y Regional Universidad Autónoma del Estado de Mexico.
mbautistaf050@alumno.uaemex.mx

[2] Doctor en Ciencias Económicas; Profesor-Investigador de la Universidad Autónoma Metropolitana Unidad Lerma (UAM-L) y de la FaPUR UAEM;
r.rozga@correo.ler.uam.mx

Abstract. Las ciudades inteligentes se han consolidado como uno de los marcos más influyentes para enfrentar los desafíos urbanos del siglo XXI. Al integrar innovación tecnológica, digitalización y sostenibilidad, este modelo ha buscado transformar la gestión de los servicios públicos y mejorar la calidad de vida de la ciudadanía. Sin embargo, su desarrollo ha estado acompañado de tensiones entre eficiencia tecnológica e inclusión social, lo que plantea la necesidad de repensar la relación entre inteligencia urbana y cohesión social.

Este artículo analiza cómo el paradigma de las ciudades inteligentes pueden reforzar la cohesión urbana inclusiva, partiendo de una revisión teórica sobre la evolución de ambos conceptos y de los principales desafíos que enfrentan los municipios metropolitanos en México. Se recupera la noción de Cohesión Urbana Inteligente (CUI) como antecedente que articula integración socioespacial, planeación local e innovación social, mostrando que la verdadera inteligencia urbana no reside únicamente en la infraestructura digital, sino en su capacidad de fortalecer la equidad y la justicia espacial.

Sobre esta base, se presenta la propuesta de Cohesión Social Urbana Inteligente (CSUI), concebida como un modelo conceptual que integra factores sociales, ambientales, económicos, tecnológicos y de planeación urbana. La CSUI constituye un aporte original para trascender la visión instrumental de la ciudad inteligente, ofreciendo un marco de análisis y acción que permita orientar investigaciones académicas y diseñar políticas públicas más inclusivas y sostenibles en los municipios metropolitanos de América Latina.

Keywords: Ciudad Inteligente, Cohesión Social, Innovación Social, Gobernanza digital y Planeación Urbana.

1 Introducción

La acelerada urbanización de las últimas décadas ha consolidado a las áreas metropolitanas como los principales escenarios de innovación, producción y diversidad social. Sin embargo, este crecimiento también ha intensificado problemas de desigualdad, fragmentación espacial y exclusión digital que desafían la gobernanza urbana. Frente a estos retos, las ciudades inteligentes han emergido como un paradigma orientado a aprovechar las tecnologías de la información y la comunicación (TIC) para mejorar la gestión urbana, la sostenibilidad ambiental y la calidad de vida de la ciudadanía. Aun así, la experiencia internacional y el caso mexicano muestran que la transición hacia este modelo no está exenta de tensiones: los beneficios tecnológicos no siempre se traducen en inclusión, y persisten brechas sociales que limitan el acceso equitativo a la innovación.

En este contexto, la cohesión social se convierte en un elemento importante para repensar el papel de las ciudades inteligentes en América Latina. El debate contemporáneo evidencia que la inteligencia urbana no debe limitarse únicamente a la infraestructura digital o a la eficiencia operativa; debe incorporar valores como equidad, justicia espacial, participación ciudadana y sostenibilidad social. Diversas experiencias en Europa han demostrado que la integración de políticas de diversidad, inclusión y derechos sociales fortalece la resiliencia de las ciudades inteligentes, mientras que en México la fragmentación institucional y las desigualdades territoriales siguen siendo obstáculos significativos para avanzar en esa dirección.

Este artículo propone un marco teórico-conceptual que vincula la evolución de la idea de ciudades inteligentes con la cohesión social urbana, destacando sus desafíos en los municipios metropolitanos mexicanos. A partir del análisis de marcos normativos, experiencias internacionales y problemáticas locales, se presenta la propuesta de Cohesión Social Urbana Inteligente (CSUI) como un aporte original. La CSUI busca articular dimensiones sociales, económicas, ambientales, tecnológicas y de planeación urbana, constituyéndose en una herramienta conceptual para orientar tanto la investigación académica como la formulación de políticas públicas inclusivas y sostenibles. El objetivo de la investigación es evaluar el impacto de la implementación de algunos elementos de la idea de ciudad inteligente en la mejora de la cohesión social en los municipios metropolitanos del Estado de México durante el periodo 2014-2024, con el propósito de comprender los mecanismos que promueven la participación ciudadana, la sostenibilidad social, económica y ambiental, así como la gobernanza digital.

1.1 Ciudad inteligente: conceptualización, modelos y enfoques críticos

La noción de ciudad inteligente (CI), también denominada smart city en la literatura especializada, surge del reconocimiento de que las tecnologías de la información y la comunicación (TIC) son importantes para que las ciudades enfrenten de manera eficaz

los procesos contemporáneos de transformación social. Este concepto ha evolucionado, ampliando sus alcances y significados al integrar nuevas dimensiones (sociales, económicas, ambientales y políticas) con el objetivo de aprovechar de forma más integral los avances tecnológicos. Esta ampliación responde tanto a los desafíos propios de la urbanización acelerada como a los cambios sociales derivados del uso masivo de herramientas digitales.

Este enfoque comenzó a consolidarse en Europa durante la década de los noventa, impulsado por la aplicación de las tecnologías de la información y la comunicación (TIC) en diversos ámbitos, desde la atención a la salud y el diseño de entornos laborales innovadores, hasta la gobernanza participativa y la gestión energética [1]. Según Fernández [2], desde sus orígenes la noción de ciudad inteligente ha estado vinculada con el uso de soluciones tecnológicas para enfrentar retos urbanos como la eficiencia energética, la disminución de emisiones contaminantes y la mitigación del cambio climático.

En esta misma línea, Caragliu, Del Bo y Nijkamp [2] destacan que el crecimiento del interés por las ciudades inteligentes ha sido impulsado por una combinación de factores: la atención mediática, el debate en el ámbito de las políticas públicas, y el involucramiento de sectores empresariales y académicos. En este contexto, la innovación tecnológica aplicada a sectores como la energía, el transporte o la conectividad digital se ha promovido como motor del desarrollo urbano. Con el paso del tiempo, se ha producido una evolución en el entendimiento del concepto: se ha transitado de una visión centrada en dispositivos tecnológicos avanzados a una comprensión sistémica, donde la ciudad inteligente es un entramado complejo en el que la ciudadanía, junto con actores públicos y privados, comparte responsabilidades en la toma de decisiones [2].

Desde la perspectiva de la Asociación Española de Normalización y Certificación (AENOR, 2014), una ciudad inteligente es aquella que emplea las tecnologías de la información y la comunicación (TIC) para mejorar la calidad de vida y la accesibilidad de sus habitantes, al tiempo que promueve un desarrollo sostenible en sus dimensiones económica, social y ambiental [2]. Esta interacción tecnológica ocurre de forma multidisciplinar y en tiempo real, permitiendo responder a las necesidades cambiantes de la población.

La noción de ciudad inteligente debe entenderse como un concepto amplio y dinámico que integra múltiples dimensiones del entorno urbano. Su estructura se sostiene sobre cuatro componentes: el espacio físico de la ciudad, una red de infraestructuras, sistemas digitales interconectados y una ciudadanía que participa y articula su funcionamiento. El propósito de una smart city es lograr una gestión eficiente en áreas como el urbanismo, el transporte, la energía, la salud, la educación y la seguridad, satisfaciendo las necesidades de la ciudad como las expectativas de sus habitantes [3].

De acuerdo con Maestre [4], el concepto de ciudad inteligente ha cobrado relevancia en la agenda pública global debido al acelerado crecimiento de las zonas urbanas. En este contexto, las smart cities pueden surgir de dos maneras: como proyectos diseñados desde cero o como procesos de transformación gradual sobre ciudades preexistentes que incorporan soluciones tecnológicas de forma progresiva. En ambos casos, su desarrollo es el resultado de la interacción entre diversos factores sociales, políticos, económicos y tecnológicos, lo que implica que su implementación varía según el contexto, los objetivos, las capacidades financieras disponibles y la visión de desarrollo que se adopte.

Además, no puede pasarse por alto el papel que juega la ciudadanía en este proceso. Tal como lo plantea Cocchia [4], son los ciudadanos, a través del uso cotidiano de dispositivos digitales, quienes generan presión sobre los gobiernos locales para ofrecer servicios electrónicos eficientes. Esta dinámica impulsa una transformación urbana orientada a mejorar la calidad de vida. Así, la ciudad inteligente se concibe como un territorio que integra las TIC con el objetivo de fomentar la cooperación, la innovación y la sostenibilidad, promoviendo una gestión urbana eficiente y centrada en las personas.

En esta misma línea, Maroto [5], sostiene que las ciudades inteligentes buscan optimizar su funcionamiento mediante el aprovechamiento de tecnologías emergentes y el análisis de grandes volúmenes de datos generados en el entorno urbano. Este modelo se estructura a partir de seis ejes: movilidad, población, gobierno, calidad de vida, medio ambiente y economía. Todos estos componentes se articulan a través del uso estratégico de la información y de una participación ciudadana.

En cuanto a su evolución, la literatura especializada coincide en tres generaciones en la evolución del modelo de ciudad inteligente [6]. La primera generación de Ciudad Inteligente 1.0 se caracteriza por la iniciativa del sector privado, de las empresas tecnológicas, que promueven el uso de herramientas digitales entre las administraciones locales. En esta etapa, las ciudades suelen contar con escasa infraestructura tecnológica y existe cierto escepticismo por parte de las autoridades respecto a los beneficios de largo plazo que las TIC pueden generar en la calidad de vida. Además, la participación ciudadana es limitada, ya que la construcción de inteligencia urbana se orienta desde fuera, sin una articulación con las comunidades locales [6].

Posteriormente, en la segunda generación de Ciudad Inteligente 2.0 presenta una mayor implicación por parte de los gobiernos locales, quienes reconocen el valor estratégico de la tecnología para el desarrollo urbano. En esta fase, los proveedores de soluciones tecnológicas actúan como aliados que facilitan procesos orientados a mejorar la calidad de vida, mediante la incorporación de innovación en los servicios urbanos [6].

Finalmente, la tercera generación de Ciudad Inteligente 3.0, pone en el centro a la sociedad. Las ciudades de esta etapa logran involucrar a la ciudadanía en el diseño y gestión de soluciones urbanas. La participación social, la corresponsabilidad en la

planificación y la construcción de un sentido de pertenencia se vuelven factores determinantes para el éxito del modelo [6].

Este desarrollo no sigue una secuencia obligatoria para todas las ciudades. Rózga [6] propone el enfoque de la Ciudad Inteligente 3 – Escalable, que considera diferentes niveles de madurez tecnológica y social. En la primera fase, el énfasis está en la convergencia tecnológica, la competitividad y la seguridad. En la segunda, se incorporan elementos de cohesión social, favoreciendo procesos de convergencia entre lo tecnológico y lo comunitario. En la tercera etapa, el rol de las autoridades centrales es importante, y el eje prioritario se centra en la sostenibilidad, apoyada por programas y políticas públicas diseñadas desde un enfoque colaborativo y de largo alcance [6].

Por otro lado, las ciudades inteligentes de cuarta generación representan una etapa avanzada en la evolución del modelo smart city, caracterizada por la integración de tecnologías emergentes como la inteligencia artificial, el Internet de las Cosas (IoT) y las redes 5G. Gracias a estas herramientas, es posible configurar entornos urbanos más sostenibles, orientados a mejorar la calidad de vida de sus habitantes.

Una de las cualidades de estas ciudades es su capacidad para operar de forma interconectada y en tiempo real. A través de sistemas de monitoreo y análisis de datos, gestionan de manera eficiente sectores como el transporte, la energía, la seguridad y los servicios públicos. Un ejemplo es Medellín, donde se ha puesto en marcha un Plan Maestro de Ciudad Inteligente centrado en la conectividad y la sostenibilidad, impulsado mediante la colaboración entre actores públicos y privados [7].

Estas ciudades promueven una visión integral del desarrollo urbano, que abarca el uso eficiente de la energía, una mejor gestión de los residuos y estrategias para mitigar el impacto ambiental. Para ello, recurren a herramientas tecnológicas como sensores inteligentes y plataformas de análisis de grandes volúmenes de datos (big data), consolidando un modelo de gobernanza urbana orientado al futuro [8].

Las ciudades inteligentes enfrentan desafíos en su camino hacia la sostenibilidad, en la gestión de la energía, la movilidad urbana, la seguridad y el control de la calidad del aire. Frente a estas problemáticas, el uso de tecnologías emergentes se ha vuelto indispensable. Herramientas como la inteligencia artificial, el análisis de datos, las energías limpias y los sistemas de transporte innovadores mejoran la eficiencia operativa, ayudan a minimizar el impacto ambiental. Además de sus beneficios, estas tecnologías impulsan la generación de empleo, estimulan la inversión y fortalecen los ecosistemas de innovación. En conjunto, permiten avanzar hacia modelos urbanos sostenibles, resilientes, centrados en el bienestar de la ciudadanía [8].

Las dimensiones de las ciudades inteligentes son importantes en la transición hacia entornos urbanos sostenibles. Los avances tecnológicos han permitido el diseño de soluciones innovadoras para enfrentar desafíos como la gestión de residuos, la movilidad sustentable y la reducción de la contaminación. Tecnologías como el Internet

de las Cosas (IoT) y la inteligencia artificial hacen posible un uso eficiente de los recursos urbanos, contribuyendo a disminuir la huella ecológica de las ciudades. Así, la innovación en sostenibilidad busca optimizar los servicios, mejorar la calidad de vida de la población y preservar el entorno natural. Para alcanzar estos objetivos, es necesario la colaboración entre los gobiernos, el sector privado, la academia y la sociedad civil [8].

De forma complementaria, Díaz [9], advierte que, en el caso de México, la incorporación de tecnologías inteligentes al entorno urbano representa una oportunidad como un desafío. Las zonas metropolitanas del país, marcadas por una alta densidad poblacional y una creciente demanda de servicios públicos, requieren estrategias innovadoras que mejoren la gestión urbana de manera efectiva. En este contexto, las tecnologías emergentes pueden desempeñar un papel determinante frente a problemáticas como el tráfico vehicular, la contaminación ambiental, la escasez de agua y los altos índices de inseguridad [9].

Esta transición hacia un modelo de ciudad inteligente exige transformaciones en la manera en que se gobiernan las ciudades. La gobernanza metropolitana, muchas veces basada en esquemas jerárquicos y fragmentados, resulta insuficiente frente a la complejidad de los desafíos actuales. Ante esta realidad, Díaz [9], plantea la necesidad de avanzar hacia modelos de gobernanza más inclusivos, colaborativos y horizontales, en los que se integren de forma efectiva los distintos niveles de gobierno, así como los actores sociales en los procesos de planeación y toma de decisiones.

Desde otra perspectiva, Saskia Sassen [10], enfatiza que la tecnología en el entorno urbano debe concebirse como una herramienta al servicio del bienestar colectivo, y no solo como un medio para optimizar sistemas o procesos. Para ella, la cohesión social se fortalece cuando la infraestructura tecnológica se orienta hacia la inclusión, permitiendo que todos los sectores sociales puedan acceder y beneficiarse de sus ventajas. En esta línea, sostiene que la verdadera inteligencia de una ciudad radica en su capacidad tecnológica y en su potencial para generar entornos urbanos equitativos, habitables e inclusivos.

Por su parte, Soja [11] sostiene que es importante analizar cómo el poder y la organización espacial influyen en las dinámicas urbanas, especialmente a través del enfoque de la justicia espacial. A juicio del autor, los espacios urbanos reflejan las desigualdades sociales existentes, también pueden intensificarlas. Advierte que la incorporación de tecnologías en las ciudades corre el riesgo de ampliar estas brechas si se concentra únicamente en zonas privilegiadas, dejando al margen a comunidades con menores recursos [11].

Desde esta perspectiva, las ciudades deben entenderse como entornos físicos, como escenarios en los que se ejercen relaciones de poder. Por ello, promover la cohesión social en el marco de una ciudad inteligente implica asegurar que los beneficios de la transformación tecnológica se distribuyan de manera equitativa en todo el territorio

urbano. Integrar el principio de justicia espacial en la planificación urbana es necesario para prevenir nuevas formas de exclusión digital. Así, las tecnologías pueden convertirse en herramientas efectivas para disminuir desigualdades territoriales y avanzar hacia ciudades más justas, inclusivas y cohesionadas [11].

La gobernanza local se enfrenta a desafíos en el contexto urbano contemporáneo. El ideal de ciudad inteligente implica la incorporación de nuevas tecnologías y la articulación entre los actores, colocando al ciudadano en el centro del diseño urbano y tecnológico [12]. El concepto de Ciudad Inteligente ha evolucionado desde una visión tecnológica hacia un enfoque integral, en el que cobran relevancia la sostenibilidad social, la inclusión y la participación ciudadana. En este marco, la información geoespacial adquiere un papel importante, ya que permite anticipar escenarios de desarrollo y orientar mejor las decisiones urbanas [12].

Por otro lado, en países como China, el modelo de ciudad inteligente se ha impulsado con un fuerte respaldo del Estado, priorizando la eficiencia tecnológica. Este enfoque ha generado impactos contrastantes en la distribución de los servicios urbanos, lo que plantea interrogantes sobre la equidad y el equilibrio territorial en su implementación [13].

1.2 Cohesión social en el contexto urbano

Desde la década de los noventa, el concepto de cohesión social comenzó a incorporarse en el diseño de políticas públicas con el propósito de enfrentar los problemas derivados de la exclusión social. Esta noción tiene sus raíces en las ideas de Émile Durkheim sobre la integración social, las cuales plantean que una sociedad cohesionada debe garantizar la participación de la ciudadanía y asegurar condiciones equitativas de acceso a las oportunidades.

En ese sentido, el informe elaborado por la Comisión Económica para América Latina y el Caribe (CEPAL) en colaboración con las Naciones Unidas y la Agencia Española de Cooperación Internacional para el Desarrollo (AECID), señala que la cohesión social es un concepto abierto, de carácter flexible y abstracto, empleado para referirse a situaciones que requieren intervención desde lo público [14]. La definición, por tanto, varía según la institución que lo emplee, y está condicionada por el contexto político y social en el que se plantea. Como sostienen Jenson y Beauvais (2005), la forma en que se aplica el concepto depende del problema que se busca resolver y de las características del entorno [15].

Por su parte, en el contexto europeo, la cohesión social ha sido importante en el diseño de políticas públicas orientadas a promover el desarrollo, tanto en países con economías avanzadas como en aquellos con rezagos estructurales. De acuerdo con Mora, este concepto adquirió mayor relevancia a mediados de los años noventa, particularmente en el marco de integración regional consolidado en el Tratado de Maastricht de 1993. En dicho contexto, se estableció el Fondo de Cohesión Social, como un instrumento

financiero destinado a respaldar a naciones con menor nivel de desarrollo, como España, Portugal, Grecia e Irlanda, y así fomentar un crecimiento equilibrado dentro de la Unión Europea [16].

La cohesión social en el ámbito urbano se entiende como la capacidad de los diferentes actores sociales para construir y compartir un propósito común, ya sea en lo social, lo económico o lo político, a partir de necesidades específicas de barrios, colonias o grupos determinados. Esta articulación se posibilita mediante mecanismos de ajuste, diálogo y consenso.

En contextos urbanos y metropolitanos, la cohesión social hace referencia a la capacidad de las comunidades para fortalecer la integración entre diversos sectores de la población, al mismo tiempo que promueve la participación de la ciudadanía en la vida púbica. En regiones metropolitanas, caracterizadas por su complejidad y diversidad social, esta cohesión es necesaria para prevenir la fragmentación de la sociedad y el aislamiento de grupos vulnerables. Fomentar la cohesión implica garantizar la convivencia entre los habitantes, garantizar el acceso equitativo a servicios públicos, al empleo y a mecanismos de participación que permitan a la ciudadanía incidir en las decisiones que afectan su entorno urbano [17].

En la actualidad, el entorno urbano suele estar compuesto como un conjunto de fragmentos desconectados, lo que hace necesario promover una estructura urbana que favorezca la cohesión social y evite la desarticulación del espacio colectivo. Bravo y Rivas [18] destacan la importancia de corredores urbanos que conecten territorios, revitalicen espacios históricos y generen oportunidades para nuevas formas de habitar. En su análisis del área metropolitana al sur de Granada, identifican problemáticas como la pérdida de identidad local y la falta de integración territorial. Frente a ello, proponen el fortalecimiento de estructuras colectivas que respondan a las necesidades sociales mediante una mejor integración de equipamientos y servicios urbanos en la ciudad [18].

Por otro lado, en América Latina y el Caribe, la evaluación de dimensiones vinculadas con la cohesión social, como el reconocimiento, la participación ciudadana y la capacidad para resolver conflictos, representa un campo en desarrollo, en contextos marcados por la desigualdad. Uno de los factores que más incide es el sobreendeudamiento de los hogares, asociado a la precariedad laboral y la insuficiencia de ingresos. En ese contexto, las decisiones financieras tomadas por las familias no siempre responden a una lógica racional, ya que la incertidumbre económica genera una sensación de inseguridad, que deteriora su calidad de vida.

Durante los últimos quince años, se ha observado un incremento en el endeudamiento familiar en América Latina. Este fenómeno es visible en países como Brasil, Chile y Colombia, donde a pesar del crecimiento en el ingreso disponible, persisten desigualdades socioeconómicas. Las cifras agregadas provenientes de las cuentas nacionales ofrecen una visión general del comportamiento financiero de los hogares, resultan insuficientes para revelar las disparidades entre los distintos grupos sociales en

cuanto a su acceso a recursos, niveles de endeudamiento y vulnerabilidad financiera [15].

En este contexto, cobra relevancia el concepto de la Cohesión Urbana Inteligente (CUI) no sólo comprende la participación de la sociedad en los asuntos públicos urbanos como miembros de una comunidad urbana, contribuye en disminuir la exclusión, segregación y polarización -política, social, económica y ambiental-. Esto es, la CUI busca potenciar las políticas públicas urbanas mediante mecanismos, instrumentos y programas de participación e innovación público y social para conseguir objetivos de cohesión urbana. Por tanto, el reto de instrumentar cohesión urbana inteligente es para disminuir diferencias de inclusión y participación en el desarrollo urbano local, con formas de interacción social dinámicas [19]. El concepto de CUI está operacionalizado en tres dimensiones que lo comprenden: Integración socio-espacial, Integración inteligente y Planeación local e innovación social.

1.3 Propuesta: Cohesión Social Urbana Inteligente (CSUI)

El concepto de Cohesión Social Urbana Inteligente (CSUI) que se presenta en esta investigación es una construcción propia, elaborada a partir del análisis y la síntesis del Marco teórico-conceptual revisado. Es una propuesta con un enfoque integral de la aplicación de las tecnologías emergentes relacionados con los principios de sostenibilidad y gobernanza inclusiva, adaptándolos al contexto de los municipios metropolitanos del Estado de México. Tiene por objetivo fomentar comportamiento de la ciudadanía donde la inclusión social, la equidad en el acceso a servicios y la participación ciudadana estén alineadas con el desarrollo económico, ambiental y tecnológico.

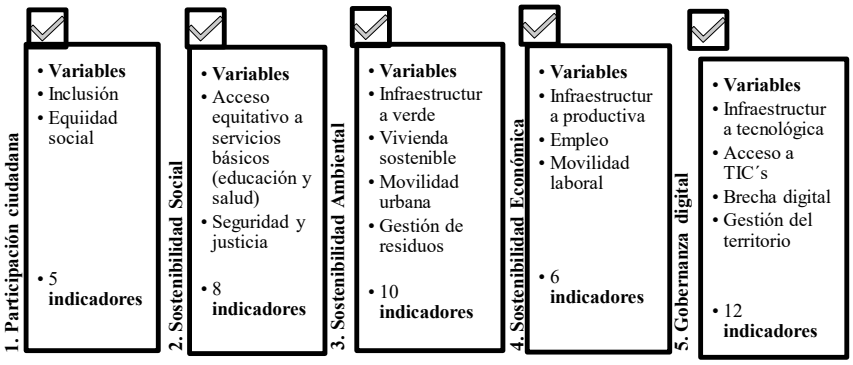
Este concepto surge de la necesidad de transformar los municipios metropolitanos en espacios de convivencia más cohesionados y sostenibles, donde la interacción entre los actores urbanos esté mediada por soluciones tecnológicas que promuevan la justicia social y la equidad en el acceso a los servicios en las ciudades contemporáneas.

La cohesión social urbana inteligente debe ser el resultado de la integración y el fortalecimiento de las comunidades urbanas a través del uso de tecnologías inteligentes que fomentan la participación ciudadana. La cohesión se logra promoviendo la sostenibilidad social mediante el acceso equitativo a servicios básicos, la sostenibilidad ambiental con prácticas responsables y protección del entorno, y la sostenibilidad económica impulsando el desarrollo y empleo locales. Además, la gobernanza digital al utilizar herramientas tecnológicas promueve la mejora en la gestión de la ciudad y reduce las brechas. Por consecuencia al implementar la idea de ciudad inteligente, se promueve un ambiente de unidad y cohesión social.

Para analizar la cohesión social urbana inteligente (CSUI), se han definido cinco categorías que permiten evaluar su desarrollo de manera integral. Estas categorías son: participación ciudadana, sostenibilidad social, sostenibilidad ambiental, sostenibilidad

económica y gobernanza digital. Cada una aborda componentes del entorno urbano, desde el involucramiento de la ciudadanía en los procesos de decisión, hasta el uso de la tecnología en la gestión pública. El concepto de cohesión social urbana inteligente (CSUI) se integra de cinco categorías, quince variables y cuarenta y seis indicadores, de manera parcial en la figura 1 y 2.

Figura 1. Categorias de Cohesión Urbana Inteligente (CSUI)



Fuente: Elaboración propia, 2025

Cada categoría se integra de las siguientes variables:

1. Participación ciudadana

1. Inclusión social: Refiere a la frecuencia y en el alcance de las consultas públicas, que permiten a los ciudadanos expresar sus opiniones sobre proyectos urbanos.
2. Equidad social: Participación justa y representativa de todos los grupos sociales en los procesos de toma de decisiones.

2.Sostenibilidad Social

3. Acceso equitativo a servicios básicos (educación y salud): Ciudadanos que tienen igualdad de oportunidades para acceder a educación y salud. seguridad y justicia.
4. Seguridad y justicia: Ciudadanos que tienen acceso a sistemas legales y de seguridad.

3.Sostenibilidad Ambiental

5. Infraestructura verde: Espacios verdes y soluciones sostenibles para mejorar la calidad ambiental y promover la biodiversidad.
6. Vivienda sostenible: Espacios habitacionales que minimizan el impacto ambiental y garantizan calidad de vida mediante el uso de tecnologías eficientes.
7. Movilidad urbana: Sistemas de transporte accesibles, eficientes y sostenibles que facilitan el desplazamiento de la ciudadanía en la ciudad.
8. Gestión de residuos: Prácticas sostenibles para la recolección, tratamiento y disposición de desechos, orientadas a minimizar impactos ambientales.

4.Sostenibilidad Económica

9. Infraestructura productiva: Impulsan el desarrollo económico, facilitando la producción, distribución y generación de empleo en un municipio.
10. Empleo: Oportunidades laborales, que permiten el desarrollo económico y social de la ciudadanía.
11. Movilidad laboral: Ciudadanos que se trasladan entre diferentes lugares de trabajo, se refiere al acceso a oportunidades laborales.

5. Gobernanza digital

12. Infraestructura tecnológica: Sistemas y equipos tecnológicos que soportan el acceso, intercambio y procesamiento de información, para facilitar la innovación y el desarrollo en la ciudadanía.
13. Acceso a TIC´s: Disponibilidad de la ciudadanía para utilizar tecnologías de la información y la comunicación, para promover la inclusión digital, el acceso a servicios y recursos en línea.
14. Brecha digital: Acceso, uso y aprovechamiento de las tecnologías de la información y la comunicación entre la ciudadanía, limitando las oportunidades de desarrollo.
15. Gestión del territorio: Planificación y organización del uso del suelo y los recursos naturales, que promueven el desarrollo sostenible, la equidad y la calidad de vida de los ciudadanos.

1.4 Instrumentos y marcos normativos de ciudades inteligentes en México

El desarrollo de ciudades inteligentes en México está condicionado por la existencia de marcos normativos e instrumentos de política pública que buscan vincular la transformación digital con objetivos de inclusión y sostenibilidad. En el ámbito internacional, destaca la Estrategia Digital 2022–2025 del Programa de las Naciones Unidas para el Desarrollo (PNUD), cuyo propósito es convertir la digitalización en una herramienta para mejorar la vida de las personas, fortalecer la equidad y proteger el medio ambiente. Sus líneas de acción incluyen potenciar el desarrollo a través de la tecnología, construir ecosistemas digitales más inclusivos y resilientes, y garantizar que los beneficios de la transformación digital lleguen a grupos históricamente vulnerables, como mujeres y personas con discapacidad [20]. Este esfuerzo se vincula con el Plan Estratégico del PNUD 2022–2025, que reconoce a la digitalización como un eje estructural para alcanzar los Objetivos de Desarrollo Sostenible (ODS) [20].

Un segundo referente internacional es el Índice de Ciudades Prósperas (CPI) de ONU-Hábitat, diseñado para medir la productividad, la infraestructura, la calidad de vida, la equidad, la inclusión social, la sostenibilidad ambiental y la gobernanza. Este índice permite evaluar y comparar el desempeño urbano en dimensiones estrechamente vinculadas con la cohesión social y constituye una guía para orientar políticas públicas hacia el desarrollo inclusivo [21, 22].

En el ámbito nacional, México ha diseñado diversas políticas con implicaciones directas para las ciudades inteligentes. Un ejemplo es el programa Aldeas Inteligentes,

Bienestar Sostenible 2023–2024 de la Secretaría de Infraestructura, Comunicaciones y Transportes (SICT), orientado a cerrar la brecha digital en zonas rurales mediante la incorporación de tecnologías de la información y comunicación. Más que una estrategia de conectividad, busca transformar comunidades al impulsar desarrollo sostenible y bienestar local, en línea con los Objetivos de Desarrollo Sostenible (ODS) de la Agenda 2030 [23].

Asimismo, la Estrategia Digital Nacional 2021–2024 constituye un marco relevante para impulsar la transformación digital en México. Esta política establece cinco objetivos: transformación gubernamental, economía digital, educación digital, salud universal y efectiva, e innovación cívica con participación ciudadana. Incluye temas como la protección de datos, la seguridad de la información y el comercio electrónico, con el propósito de garantizar un entorno digital seguro y eficiente que soporte la transición hacia ciudades inteligentes [24].

Finalmente, el Plan Nacional de Desarrollo 2019–2024 incorpora iniciativas como Internet para Todos, que busca universalizar el acceso a la conectividad y reducir desigualdades territoriales [25]. Estos instrumentos nacionales se complementan con políticas sectoriales, como la Política Nacional de Vivienda 2020–2024, que introduce criterios de sostenibilidad e inclusión en la planeación urbana [26]. Sin embargo, persiste el desafío de articular estas estrategias en un marco de gobernanza digital coherente y sostenible, que logre trascender los ciclos sexenales y consolidar una visión de largo plazo para las ciudades inteligentes en México.

1.5 Desafíos de las ciudades inteligentes para integrar la cohesión social

Las ciudades inteligentes no pueden limitarse a la infraestructura tecnológica y a la eficiencia operativa; deben integrar la cohesión social como un pilar importante para garantizar inclusión y sostenibilidad. Experiencias como las de Ámsterdam, Berlín y Estocolmo muestran que la verdadera inteligencia urbana se construye con políticas activas de diversidad, igualdad de género, integración de migrantes y respeto a la pluralidad social [27]. Estos casos evidencian que la cohesión social implica atender factores como la inmigración, la salud pública, la seguridad y el desarrollo comunitario.

A pesar de los avances globales en la reducción de la pobreza extrema, la urbanización acelerada ha incrementado la desigualdad salarial en tres de cada cuatro ciudades en las últimas dos décadas, incluso en países de la OCDE [27]. La segregación espacial es otro desafío: mientras barrios exclusivos concentran servicios de alto nivel, los asentamientos informales se expanden con acceso limitado a salud, educación y transporte. Estas dinámicas fragmentan las urbes y amplían la brecha entre riqueza y vulnerabilidad.

Autores como Sassen y Soja [10, 11] advierten que la digitalización, si se centra solo en áreas privilegiadas, puede reforzar desigualdades sociales y espaciales. Por ello, las ciudades inteligentes deben incorporar principios de justicia espacial y garantizar que

la infraestructura tecnológica beneficie a todos los grupos sociales. El desafío central radica en equilibrar innovación tecnológica con inclusión social, de modo que las ciudades inteligentes se conviertan en entornos equitativos, cohesionados y sostenibles.

La construcción de ciudades inteligentes en México enfrenta el reto de no quedar reducida a un proyecto tecnológico, sino de convertirse en una plataforma de integración social. Los municipios metropolitanos, al adoptar estrategias de ciudad inteligente, requieren garantizar que los beneficios de la digitalización y de la innovación urbana se distribuyan de manera equitativa. Sin políticas de inclusión digital, programas de acceso universal a servicios y estrategias de integración comunitaria, la ciudad inteligente corre el riesgo de profundizar las desigualdades históricas presentes en las zonas periféricas y marginadas. En este sentido, la cohesión social debe entenderse como una condición necesaria para que la ciudad inteligente cumpla su promesa de sostenibilidad y eficiencia.

Asimismo, el verdadero potencial de la ciudad inteligente en los municipios metropolitanos radica en su Capacidad de generar nuevas formas de gobernanza colaborativa. La fragmentación institucional y la falta de coordinación intermunicipal obstaculizan el despliegue de soluciones tecnológicas integradas. Una ciudad inteligente con cohesión social require mecanismos de cooperación entre gobiernos locales, universidades, empresas y ciudadanía, que permitan consolidar una gobernanza digital inclusiva. Solo bajo este enfoque, la innovación tecnológica se convierte en una herramientas para fortalecer la confianza social y crear ciudades más justas, resilientes y sostenibles.

Conclusiones y recomendaciones

El análisis desarrollado muestra que el marco de las ciudades inteligentes constituye una oportunidad estratégica para reforzar la cohesión urbana inclusiva en los municipios metropolitanos. No obstante, los beneficios de la digitalización y la innovación tecnológica no son automáticos: dependen de políticas públicas capaces de integrar lo social, lo económico y lo ambiental en un mismo horizonte de planeación. La revisión teórica y normativa evidencia que, a pesar de avances relevantes en materia de conectividad, estrategias digitales y programas de vivienda, la fragmentación Institucional y la persistencia de desigualdades territoriales siguen siendo obstáculos significativos.

En este contexto, la propuesta de Cohesión Social Urbana Inteligente (CSUI) aporta un marco conceptual que permite superar la visión meramente instrumental de la ciudad inteligente. La CSUI enfatiza que la verdadera inteligencia urbana no se limita a la eficiencia tecnológica, coloca la cohesión social en el centro de las dinámicas metropolitanas. Al integrar dimensiones sociales, ambientales, económicas, tecnológicas y de planeación urbana, esta propuesta se convierte en un referente para evaluar políticas públicas y orientar investigaciones futuras en América Latina.

De cara al futuro, es necesario fortalecer la gobernanza metropolitana mediante esquemas de cooperación intermunicipal que eviten la fragmentación y permitan desplegar soluciones tecnológicas integradas; garantizar la inclusión digital como política de Estado, asegurando acceso equitativo a conectividad, dispositivos y capacitación para los grupos más vulnerables; y vincular la innovación tecnológica con la cohesión social, de modo que los proyectos de ciudad inteligente se evalúen no solo por su eficiencia, sino también por su capacidad para reducir brechas y fortalecer la justicia espacial. Asimismo, se requiere dar continuidad a los marcos normativos nacionales e internacionales, trascendiendo los ciclos políticos de corto plazo y alineando las políticas locales con estándares globales de sostenibilidad. En suma, avanzar hacia ciudades inteligentes con cohesión social exige una visión integradora que articule tecnología, gobernanza e inclusión, y en este marco la Cohesión Social Urbana Inteligente (CSUI) constituye un aporte original para consolidar dicho enfoque, al ofrecer un marco de análisis y acción que refuerza el carácter inclusivo y sostenible de las ciudades inteligentes en los municipios metropolitanos.

References

1. Komninos, N., Schaffers, H., y Pallot. M. (2011). Developing a policy roadmap for smart cities and the future internet. In Cunningham, P., and Cunningham, M. (Eds) eChallenges e-2011 Conference Proceedings, IIMC International Information Management Corporation, pp. 1-
2. Fernández, J. M. (2015). Ciudades Inteligentes: La mitificación de las nuevas tecnologías como respuesta a los retos de las ciudades contemporáneas. *Economía Industrial* (395); pp. 17-28.
3. Enerlis, Ernst, E y Young, F., Ferrovial and Madrid Network. (2012). *Libro blanco Smart Cities.* Madrid: Enerlis; Ernst and Young; Ferrovial; Madrid Network.
4. Maestre, P., (2015) Revisión de literatura sobre ciudades inteligentes: una perspectiva centrada en las TIC. Ingeniare, Universidad Libre Barranquilla, 11(19), pp. 137-149.
5. Maroto, C. (2015). Big Data aquí y ahora 2015. Situación mundial y foco en el mercado de Colombia. Colombia: OBS.
6. Rózga, R. E. (2018). Modelos de Ciudad Inteligente (Smart City) y estrategias de su implementación. Ponencia presentada para el 23° Encuentro Nacional Sobre Desarrollo Regional en la Ciudad de Puebla del 16 al 19 de octubre de 2018, con el tema general: Agenda Pública para el Desarrollo Regional, la Metropolización y la Sostenibilidad. Benemérita Universidad Autónoma de Puebla (BUAP).
7. Alcaldía de Medellín. (s.f.). Medellín Ciudad Inteligente. Recuperado de [https://www.medellin.gov.co]
8. Tecno-Futuro. (2024). El Papel de las Ciudades Inteligentes en la Sostenibilidad del Futuro. Energías Renovables y Sostenibilidad. https://tecnofuturo.net/energias-renovables-y-sostenibilidad/papel-ciudades-inteligentes-sostenibilidad-futuro/
9. Díaz, A. A. (2023). Gobernanza metropolitana y ciudades inteligentes: Retos para la innovación del gobierno y la gestión del México urbano. Siglo XXI Editores.
10. Sassen, S. (2006). Cities in a World Economy (3rd ed.). SAGE Publications.
11. Soja, E. W. (2010). Seeking Spatial Justice. University of Minnesota Press.

12. Lechón Gómez, D. M., & Ramos Muñoz, D. E. (2020). ¿Es Internet un territorio? Una aproximación a partir de la investigación del hacktivismo en México. Economía, Sociedad y Territorio, 20(62), 273–301. https://doi.org/10.22136/est20201507
13. Morris, C. (2023). 'It would be smart to discuss this on telegram': China's digital territorialization project and its spatial effects on contentious politics. Territory, Politics, Governance, 11(6), 1081–1099. https://doi.org/10.1080/21622671.2023.2183894
14. Bernal Hidalgo, L. (2020). Los territorios digitales en el contexto del ciberespacio. Ar@cne. Revista Electrónica de Recursos de Internet sobre Geografía y Ciencias Sociales, 24(247), 1–18. https://doi.org/10.1344/ara2020.247.32468
15. CEPAL, Comisión Económica para América Latina y Caribe y Naciones Unidas (NU), Agencia Española de Cooperación Internacional para el Desarrollo (AECID). (2022). *Metodología de medición de la cohesión social en América Latina y el Caribe, Aprendizajes y desafíos futuros*. Santiago de Chile: CEPAL.
16. Haro, G., y Vázquez, J. (2018). La cohesión social desde una perspectiva no normativa: alternativa de un diseño instrumental. Tla-Melaua, revista de Ciencias Sociales. Facultad de Derecho y Ciencias Sociales, 11 (43), pp. 132-154. issn: 1870-6916 /
17. Berrone, P., Ricart, J. E., & Duch T-Figueras, A. I. (2017). *Cities and Social Cohesion: Designing More Inclusive Urban Areas.* Barcelona.
18. Bravo, B., y Rivas., J. L. (2011). Elementos de cohesión urbana en escenarios metropolitanos: intersticios y corredores. *IV jornadas internacionales sobre investigación en arquitectura y urbanismo*: Valencia 1,2,3, pp. 1-19. ISBN 9788493867508
19. Bautista, M. y Hoyos, G. (2021). Cohesión Urbana Inteligente: hacia una propuesta de innovación local en México. Colección: Recuperación transformadora de los territorios con equidad y sostenibilidad. Innovación, turismo y perspectiva de género en el desarrollo regional. (Vol. V). Universidad Nacional Autónoma de México, Instituto de Investigaciones Económicas y Asociación Mexicana de Ciencias para el Desarrollo Regional. México.
20. PNUD, Programa de las Naciones Unidas para el Desarrollo. (2022). Estrategia digital 2022–2025. Recuperado de https://digitalstrategy.undp.org/documents/Digital-Strategy-2022-2025-Full-Document_ES_Interactive.pdf
21. ONU HABITAT, Programa de las Naciones Unidas para los Asentamientos Humanos, (2016). Urbanización y desarrollo: Futuros emergentes: Reporte de ciudades del mundo 2016. ONU- Habitat.
22. ONU HABITAT, Programa de las Naciones Unidas para los Asentamientos Humanos. (2018). Índice Básico de las Ciudades Prósperas. Recuperado de https://infonavit.janium.net/janium/Documentos/68149.pdf
23. SICT, Secretaría de Infraestructura, Comunicaciones y Transportes. (2023). Aldeas inteligentes, Bienestar sostenible. Recuperado de https://coberturauniversal.gob.mx/aldeas-inteligentes
24. Presidencia de la República. (2021). Estrategia Digital Nacional. Gobierno de México. Recuperado de: https://www.dof.gob.mx/nota_detalle.php?codigo=5628886&fecha=06/09/2021#gsc.tab=0
25. Gobierno de México. (2019). Plan Nacional de Desarrollo 2019-2024. Gobierno de México. https://www.gob.mx/pnd/documentos/plan-nacional-de-desarrollo-2019-2024.
26. Gobierno de México. (2020). Política Nacional de Vivienda 2020-2024. Secretaría de Desarrollo Agrario, Territorial y Urbano (SEDATU). https://www.gob.mx/sedatu/documentos/politica-nacional-de-vivienda-2020-2024.
27. Berrios, E. (2022). El rol de los espacios colectivos en la cohesión social urbana. II Simposio Internacional de Doctorandos en Desarrollo Urbano Sustentaberale en Latinoamérica y el Caribe.

Identificación y análisis de factores relevantes del desempeño en matemáticas en México con datos PISA 2012

Esteban Paul Louis Nicolas[1][0009-0000-7000-5024], Gilberto Lorenzo Martínez Luna[2][0000-0002-0105-1112] y Adolfo Guzmán Arenas[2][0000-0002-8236-0469]

[1]École Nationale de l'Aviation Civile (ENAC), Toulouse, France.

[2]Centro de Investigación de Computación Ciudad de México, México.

Resumen. El presente artículo describe el proceso para modelar la relación de un conjunto de variables con el desempeño en un examen de matemáticas de tercer grado de secundaria, y realizar predicciones de este resultado, antes de desarrollar el examen. Las variables son de datos compartidos como datos abiertos por el estudio llamado PISA del año 2012, se describe la complejidad originada por diferentes causas: los formatos con que se almacena el archivo, la creciente cantidad de variables que describen el contexto de los estudiantes, los diferentes tipos de las variables, entre otras causas, las cuales relacionar con el desempeño.

Palabras clave: Evaluación Pisa, Estudiantes mexicanos, Desempeño en matemáticas, Selección de atributos.

1 Introducción

La calidad de la educación en México es un tema central en el diseño de políticas públicas y en la búsqueda de equidad social. El estudio PISA (Programme for International Student Assessment) [1], coordinado por la OCDE (Organización para la Cooperación y el Desarrollo Económicos), ofrece una evaluación comparativa del rendimiento de los estudiantes de 15 años en áreas clave como matemáticas, lectura y ciencias. México participa en PISA desde el año 2000, y los datos generados constituyen una amplia fuente de información para el análisis de los factores que inciden en el aprendizaje.

No obstante, la complejidad y el tamaño de estos conjuntos de datos requieren un cuidadoso proceso de limpieza, integración y transformación antes de que puedan ser utilizados para análisis significativos. Este proyecto se centra en el tratamiento de los datos PISA 2012 [2] para México, seleccionados como punto de partida; aunque se dispone de datos más recientes, el análisis inicial sobre 2012 permite establecer una base metodológica que puede extenderse a años posteriores. En particular, se busca identificar los factores escolares y contextuales que mejor explican el desempeño académico mediante técnicas de aprendizaje automático y selección de características. De esta manera, el estudio contribuye a comprender las relaciones entre las variables educativa.

2 Antecedentes

En México, diversas instituciones han aprovechado datos similares a los publicados por PISA para estudiar el desempeño educativo en contextos sociales complejos. Por ejemplo, el sistema ANCONE (Análisis de Contexto en Exámenes) [3], desarrollado en el CIC con apoyo de la Dirección General de Medición y Tratamiento de Datos del INEE (Instituto Nacional para la Evaluación de la Educación), utiliza los datos de los exámenes PLANEA y EXCALE, los cuales evalúan el logro académico de los estudiantes y la calidad educativa del sistema nacional. Esta herramienta facilita la validación de variables y la realización de estudios orientados a diseñar intervenciones que reduzcan desigualdades educativas.

El éxito académico en la educación secundaria es un fenómeno complejo influenciado por factores personales, familiares e institucionales. Estudios recientes han identificado que hábitos de estudio, organización del tiempo y apoyo familiar son determinantes clave del rendimiento escolar. Por ejemplo, Ramírez Lemus (2023) [4] destaca que la preparación académica desde la educación primaria, el establecimiento de metas claras y el apoyo emocional son esenciales para fomentar el éxito educativo en México.

La selección de variables relevantes en conjuntos de datos de alta dimensión ha recibido creciente atención en la literatura de aprendizaje automático. Estrategias híbridas de selección y ponderación de características, como la propuesta por Singh et al. (2024) [5], combinan métodos de filtrado estadístico con métodos basados en modelos para identificar atributos informativos y relevantes, demostrando eficacia en contextos de datos complejos y de alta dimensionalidad.

3 Metodología

Se trabajó con los archivos de PISA 2012 [2], que incluyen respuestas de exámenes, cuestionarios de contexto escolar, familiar y estudiantil. Los archivos originales en formato plano (.txt) fueron procesados a partir de sus sintaxis SPSS, lo que permitió extraer la estructura de variables, etiquetas y valores faltantes, para convertirlos en archivos CSV limpios y etiquetados. Durante este proceso también se eliminaron columnas sin variabilidad, se corrigieron inconsistencias en nombres y se homogenizaron los valores faltantes, para garantizar conjuntos de datos listos para el análisis.

Posteriormente, se integró la información de estudiantes y escuelas mediante el identificador SCHOOLID, generando una tabla unificada con datos personales, escolares y de contexto. Dado que el conjunto inicial contenía más de 900 variables. Para una preselección se descartaron aquellas con más del 60% de valores perdidos. Por ejemplo, “Father Age Arrival Test Country” o “Subjective Norms in Mathematics”, así como variables redundantes o poco relevantes para el estudio del rendimiento académico para México como “Family Country of Birth” (para los padres y abuelos) o “Language at home” que tiene más de 95% de respuestas de misma repuestas.

Finalmente, para medir el desempeño en matemáticas se utilizó el primer valor plausible (PV1MATH), clasificando a los estudiantes según los niveles de competencia definidos por la OCDE. Esta elección se fundamenta en que los valores plausibles no

representan estimaciones individuales exactas, sino simulaciones que mantienen la variabilidad poblacional. El uso de un único valor plausible permitió simplificar el análisis computacional y la interpretación de los resultados. Aunque la metodología de PISA recomienda considerar los cinco valores plausibles para estimaciones puntuales [6], diversos estudios [7] han mostrado que, a nivel agregado, las tendencias y asociaciones entre variables se mantienen estables cuando se emplea un solo valor plausible. Por lo tanto, esta decisión no compromete la validez estadística del análisis comparativo.

3.1 Variable del estudio

El dataset final contenía aproximadamente 33,000 estudiantes con casi 800 características combinadas de todas las fuentes. Se guardó en un archivo .csv listo para su análisis. PISA define oficialmente siete niveles de desempeño en matemáticas, cada uno acompañado de una descripción cualitativa de las habilidades que los estudiantes demuestran [8]. Además, para cada nivel, PISA establece los valores de puntuación correspondientes, lo que permite ubicar a los estudiantes según su desempeño real, ver Fig. 1. En nuestro estudio, los valores continuos de puntaje se transformaron a estos niveles oficiales, de manera que la clasificación de los estudiantes refleje los estándares de la evaluación y permita analizar la distribución del rendimiento de forma comparativa, ver Tabla 1.

Tabla 1. Nivel de matemáticas sobre el resultado del estudiante.

Nivel	Rango de puntuación	Descripción general
Nivel 6	≥ 669.30	Puede conceptualizar, generalizar y utilizar información basada en modelos complejos, además de aplicar pensamiento estratégico y visión crítica.
Nivel 5	607.04 – 669.29	Puede desarrollar y trabajar con modelos, seleccionar y aplicar estrategias de resolución de problemas en situaciones complejas.
Nivel 4	544.68 – 607.03	Puede trabajar eficazmente con modelos explícitos, razonar con cierta profundidad e interpretar información compleja.
Nivel 3	482.38 – 544.67	Puede ejecutar procedimientos claramente descritos, interpretar y usar representaciones, y razonar directamente a partir de los datos.
Nivel 2	420.07 – 482.37	Puede interpretar y reconocer situaciones en contextos que requieren razonamiento simple. Tiene una comprensión básica del contenido matemático.
Nivel 1	358.77 – 420.06	Puede responder preguntas utilizando procedimientos simples. Capacidad limitada para hacer inferencias.
Por debajo del Nivel 1	< 358.77	Solo puede realizar tareas matemáticas muy básicas en contextos conocidos.

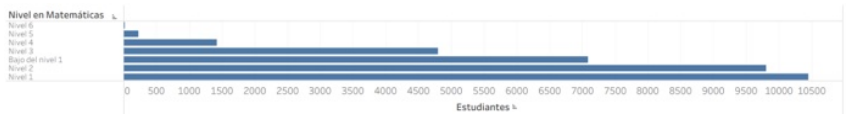


Fig. 1. Número de estudiantes por nivel en matemáticas

3.2 Primer acercamiento a la selección de variables

Para este proyecto se utilizó Weka [9], un entorno que ofrece implementaciones confiables de distintos métodos de selección de atributos, tanto estadísticos (filtros) como basados en modelos (wrappers y embebidos). Esto permitió aplicar enfoques complementarios bajo un mismo marco y diseñar un protocolo de selección más sólido, evitando depender de un único criterio evaluador.

Como primer acercamiento, se aplicó el evaluador CfsSubsetEval [10] en combinación con el método de búsqueda BestFirst [10] sobre la base de datos de estudiantes mexicanos de PISA 2012 (33,806 observaciones y 774 atributos), utilizando como variable objetivo el primer valor plausible en matemáticas. La evaluación se ejecutó en modo "usar todos los datos", lo cual es adecuado en la etapa de selección de atributos, ya que el interés no es validar la generalización de un modelo predictivo, sino identificar las variables más asociadas con el desempeño matemático en el conjunto completo. En este proceso se exploraron 79,701 subconjuntos de atributos, seleccionando el de mayor mérito (0.598), compuesto por 108 variables, ver Tabla 2.

El procedimiento de CfsSubsetEval [10] se basa en dos principios clave: relevancia y no redundancia. El primero busca maximizar la correlación de cada atributo con la variable objetivo, mientras que el segundo evita incluir variables demasiado correlacionadas entre sí, que aportarían información duplicada, por ejemplo "Class Size", "Mother qualification" y "Father qualification". De esta manera, el subconjunto final no solo conserva los factores más predictivos, sino que también garantiza diversidad informativa, lo que facilita la construcción de modelos más simples, robustos e interpretables. Además, también de reducir riesgos de sobreajuste.

Tabla 2. Resultado de las variables seleccionadas.

Variables seleccionadas
Math Career - Child study interest
Repeat Grade
Possessions - Internet
Father
Grade compared to modal grade in country
ICT Entertainment Use
Out-of-school - Obtain practical information from the Internet
Maths Self-Concept - Not Good at Maths
Maths Self-Concept - Learn Quickly
Maths Anxiety - Get Very Nervous
Maths Self-Efficacy - Calculating TV Discount
Maths Self-Efficacy - Solving Equation 1
Experience with Pure Maths Tasks - Solve Equation 1
Familiarity with Maths Concepts - Polygon
Attitudes - Not suitable for schoolwork

A pesar de su eficacia, este procedimiento tiende a favorecer variables directamente relacionadas con el contenido evaluado, dejando fuera otras que, aunque menos correlacionadas de forma inmediata, podrían ser clave para comprender factores contextuales, como el entorno familiar o escolar. Por ejemplo, aspectos como las posesiones en el hogar, el nivel educativo de los padres o la infraestructura escolar podrían quedar descartados, a pesar de su posible relevancia educativa, ver Tabla 2.

Para no perder esta información, se implementó una selección complementaria de variables agrupadas por contexto: escuela, familia y estudiante. Dentro de cada grupo se identificaron las variables con mayor influencia relativa en su ámbito, preservando así la capacidad explicativa de cada bloque. Este enfoque híbrido permite construir modelos predictivos más completos y, al mismo tiempo, analizar cómo distintos factores —académicos, familiares y escolares— interactúan para influir en el desempeño matemático. Gracias a este método, es posible estudiar no solo si un alto puntaje en matemáticas está asociado a las competencias previas del estudiante, sino también a

variables como el apoyo familiar, los recursos disponibles en el hogar o las características del centro escolar. De esta forma, el análisis no se limita a explicar "quién obtiene mejores resultados", sino qué aporta información sobre "por qué" y "en qué condiciones" se producen dichos resultados, proporcionando insumos valiosos para diseñar estrategias educativas más efectivas y equitativas.

3.3 Diseño y procedimiento

El protocolo se fundamenta en la idea de que cada técnica presenta sesgos y limitaciones: los métodos estadísticos identifican dependencias simples y son computacionalmente eficientes, mientras que los basados en modelos capturan interacciones complejas a mayor costo. Al integrarlos, se obtiene un proceso más robusto, que reduce el riesgo de sobreajuste y mejora la fiabilidad de las variables seleccionadas.

Los métodos filtro incluyeron CorrelationAttributeEval [10] (correlación lineal vía Pearson), GainRatioAttributeEval [10] (ganancia de información ajustada por número de categorías) y ReliefF (diferencias entre vecinos cercanos). Estos métodos son rápidos y útiles para descartar variables irrelevantes. Los métodos basados en modelos aplicados fueron CfsSubsetEval [10] + BestFirst [10], que maximiza relevancia y minimiza redundancia entre atributos, y WrapperSubsetEval [10] + GreedyStepwise [10] con J48 [11], que evalúa subconjuntos mediante el desempeño de un árbol de decisión.

Finalmente, se adoptó una estrategia de combinación en tres pasos: (1) normalizar los puntajes individuales (Ecuación 1); (2) calcular un puntaje promedio de los rankings (Ecuación 2); y (3) integrar estos resultados con los métodos subset y wrapper, asignando pesos específicos (Ecuación 3). Esto permitió construir un criterio equilibrado entre robustez estadística y capacidad predictiva.

$$\text{score}_{f,i} = \frac{\text{score}_{f-ranked,i}}{\max_i \text{score}_{f-ranked,i}} \tag{1}$$

- $\text{score}_{f,i}$: puntaje normalizado del atributo *i* según el método de filtro *f*.
- $score_{f-ranked,i}$: puntaje original del atributo *i* en el ranking *f*.

$$\text{score}_{avg\text{-}ranking,i} = \frac{1}{\text{card}(F)} \cdot \sum_{f=1}^{F} \text{score}_{f,i} \tag{2}$$

- $\text{score}_{avg\text{-}ranking,i}$: puntaje promedio del atributo *i* considerando todos los métodos de filtro.
- $\text{card}(F)$: número total de métodos de ranking utilizados.
- F : conjunto de todos los filtros

$$\text{score}_{final,i} = \text{score}_{avg\text{-}ranking,i} \cdot w_0 + \sum_{m=1}^{M} \delta_{i,m} \cdot w_m \tag{3}$$

- $\text{score}_{final,i}$: puntaje final del atributo *i*, combinando ranking promedio y selección por modelos.
- w_0: peso asignado al puntaje promedio de los rankings.
- M : número total de métodos de selección (subset y wrapper).
- $\delta_{i,m}$: 1 si el atributo *i* fue seleccionado por el modelo *m*, y 0 en caso contrario.
- w_m : peso asignado a la selección del atributo por el método *m*.

4 Resultado de las selecciones

4.1 Métodos de filtro

Clasificación de los atributos sobre su correlación

Para identificar las variables más relevantes que influyen en el desempeño en matemáticas, se aplicó en Weka el evaluador CorrelationAttributeEval junto con el método de búsqueda Ranker. Esta estrategia permitió calcular la correlación individual de cada atributo de los cuestionarios de estudiantes, padres y escuelas del PISA 2012 respecto al primer valor plausible en matemáticas. De esta manera, las variables se ordenaron jerárquicamente según su fuerza de asociación con el puntaje, sin considerar redundancia entre ellas. Para garantizar la comparabilidad entre los tres cuestionarios y evitar sesgos derivados de la diferencia en el número total de atributos, se decidió trabajar con los 85 primeros atributos de cada conjunto, tomando como referencia que el cuestionario de padres fue el que presentó menor número de variables seleccionadas. Esta selección ofrece una base homogénea y sólida para posteriores análisis interpretativos y modelos predictivos, permitiendo resaltar los factores más influyentes en el rendimiento académico de los estudiantes.

Los resultados generales muestran que los factores familiares destacan principalmente por la educación y ocupación de los padres, así como por el involucramiento en la vida escolar de los estudiantes. Las variables con mayor score incluyen “Repeat Grade”, expectativas educativas de los padres y participación parental, evidenciando que el capital educativo y socioeconómico del hogar influye de manera directa en el rendimiento matemático ver Tabla 3.

Tabla 3. Resultados de la selección de correlación en los tres contextos: Familiar, estudiantil y escolar (5 primeras).

Familiar		Estudiantil		Escolar	
Mérito (Score)	**Variables**	**Mérito (Score)**	**Variables**	**Mérito (Score)**	**Variables**
0.1139	Repeat Grade - At <ISCED 1>	0.12429	International Grade	0.1063	Student Computers - Computers with Internet
0.1061	Completed levels - < level 5A/6>	0.12429	Grade compared to modal grade in country	0.09887	Student Computers - Computers for Education
0.0983	Parental involvement in their child’s school	0.12014	Index of economic, social and cultural status	0.09401	Quality of school educational resources
0.095	Expected completed levels of education	0.11973	ISCED level	0.08939	Enrolment - Number of girls
0.0864	Father Qualification - < level 5A/6>	0.11657	Grade Repetition	0.08909	Total school enrolment

En el contexto estudiantil, los atributos más relevantes reflejan tanto la trayectoria académica como actitudes y autopercepción hacia las matemáticas ver Tabla 3. Variables como “International Grade”, “Mathematics Anxiety”, “Mathematics Self-Concept” y el índice socioeconómico del estudiante muestran que el desempeño está condicionado por el historial escolar, la posición socioeconómica y las creencias sobre las propias capacidades, así como el acceso a recursos educativos en el hogar.

Por último, en el ámbito escolar ver Tabla 3, los factores más destacados se relacionan con la disponibilidad y calidad de recursos educativos, la infraestructura

tecnológica, las actividades extracurriculares y la organización institucional. Aunque sus scores son menores que los individuales o familiares, indican que la infraestructura y la capacidad de gestión de las escuelas ejercen un efecto significativo sobre el desempeño.

Clasificación de los atributos sobre el gain y ratio

Para identificar los factores escolares más relevantes que predicen el desempeño en matemáticas de los estudiantes mexicanos en PISA 2012, se aplicó el método de selección de atributos basado en Gain Ratio. Este enfoque permite cuantificar la importancia de cada variable respecto a la variabilidad del puntaje en matemáticas, priorizando aquellas que aportan mayor información discriminativa sobre los resultados. A diferencia del análisis previo basado en correlación, que solo mide relaciones lineales entre variables y el puntaje, Gain Ratio considera también la relación de variables categóricas y la reducción de incertidumbre en los resultados, ofreciendo una perspectiva más completa sobre la relevancia de los factores escolares.

En el contexto familiar, los mayores méritos corresponden a la repetición de grado en primaria, las expectativas educativas de los padres y su nivel educativo, evidenciando la influencia del capital cultural y las aspiraciones familiares en el rendimiento matemático. También aparecen factores como las cualificaciones y la ocupación de los padres, así como la participación en actividades escolares, aunque con menor peso, reflejando un compromiso indirecto en el aprendizaje ver Tabla 4.

En el ámbito estudiantil, destacan la repetición escolar, el estatus migratorio y el nivel ISCED, mostrando que la trayectoria académica y el origen del alumno son determinantes clave. Recursos individuales como el acceso a computadora o internet en casa, junto con variables de contexto cultural (lengua hablada y país de nacimiento) y percepciones personales (auto-concepto en matemáticas), presentan una influencia intermedia.

En el contexto escolar, los valores más altos se observan en la responsabilidad de contratación docente por autoridades nacionales, los salarios iniciales de los profesores y la titularidad de la escuela. Otros factores, como las actividades extracurriculares, el uso de computadoras y las políticas de admisión, tienen menor peso, aunque inciden en la calidad del entorno educativo, ver Tabla 4.

Tabla 4. Resultados de selección de Gain Ratio en los tres contextos: familiar, estudiantil y escolar (5 primeras).

Familiar		Estudiantil		Escolar	
Mérito (Score)	**Variables**	**Mérito (Score)**	**Variables**	**Mérito (Score)**	**Variables**
0.123068	Repeat Grade - At < ISCED 1>	0.120196	Repeat - <ISCED 1>	0.0377112	Responsibility - Hiring teachers - National education authority
0.063555	Expected completed levels - <ISCED level 5A or 6>	0.119035	Grade Repetition	0.0325476	Responsibility - Teacher start salaries - <School governing board>
0.051334	Expected completed levels of education	0.084604	Immigration status	0.029583	Public or private
0.046547	Repeat Grade - At < ISCED 2>	0.076677	ISCED level	0.0286964	Responsibility - Salary increase - <School governing board>
0.040742	Father Qualification - <ISCED level 5A, 6>	0.064991	International Grade	0.0285811	School Ownership

Clasificación con el algoritmo RELIEF-F.

El algoritmo RELIEF-F, extensión de RELIEF, evalúa la calidad de los atributos considerando, para cada instancia, los vecinos más cercanos de distintas clases. De este modo asigna un peso según su capacidad discriminativa, incluso con ruido o valores faltantes, lo que lo hace más robusto que evaluadores basados solo en relaciones lineales o información aislada. En nuestro caso, combinado con Ranker, generó un ranking comparativo entre cuestionarios de estudiantes, padres y escuelas.

En el contexto estudiantil, los mayores pesos se concentran en factores emocionales y cognitivos como el autoconcepto negativo en matemáticas (0.0377) y la ansiedad matemática (≈0.026), confirmando que percepciones y emociones influyen más que recursos materiales. En el contexto familiar, destacan el interés del hijo por una carrera matemática (0.0196) y las expectativas educativas de los padres (≈0.018), lo que refleja la centralidad de las aspiraciones familiares sobre los factores socioeconómicos. En el contexto escolar, sobresale la gestión institucional, con variables como tamaño de clase (0.1428) y fijación de metas educativas (≈0.11–0.13), que superan en importancia a recursos materiales o condiciones específicas ver Tabla 5.

Tabla 5. Resultados de selección de RELIEF-F en los tres contextos: familiar, estudiantil y escolar (5 primeras).

Familiar		Estudiantil		Escolar	
Mérito (Score)	**Variables**	**Mérito (Score)**	**Variables**	**Mérito (Score)**	**Variables**
0.019617	Math Career – Child interest	0.0377108	Maths Self-Concept - Not Good at Maths	0.171353	Stratum ID
0.018198	Expected completed levels - level 5A/6>	0.0266882	Maths Self-Concept - Learn Quickly	0.142817	Class size
0.015794	Expected completed levels of education	0.0261368	Maths Anxiety - Get Very Nervous	0.130349	School Leadership - Informal observations
0.01193	Activities - Discuss behaviour teacher	0.0259894	Maths Anxiety - Feel Helpless	0.125954	School Leadership - Enhance reputation
0.011285	Education Cost	0.0259144	Maths Anxiety - Get Very Tense	0.123414	Internet schoolwork - Assignments

4.2 Métodos wrapper y embebido

Método Correlation-based Feature Selection

Para este análisis se utilizó el evaluador CFS (Correlation-based Feature Selection) en combinación con el método de búsqueda Best First. El evaluador CFS selecciona atributos considerando dos criterios: alta correlación con la variable objetivo y baja redundancia entre ellos, priorizando subconjuntos equilibrados y no redundantes. Los parámetros empleados fueron: -P 1, que indica que la evaluación se realiza en función del poder predictivo de los subconjuntos, y -E 1, que especifica la inclusión de atributos localmente predictivos, es decir, que aportan información en combinación, aunque no lo hagan de forma independiente. El algoritmo de búsqueda BestFirst se aplicó en dirección forward, partiendo de un conjunto vacío y añadiendo atributos de forma progresiva. El proceso se detuvo tras 5 expansiones sin mejora, garantizando eficiencia y evitando la exploración de combinaciones poco informativas.

Tabla 6. Resultados de la selección de atributos mediante CFS con el algoritmo BestFirst en el contexto familiar (18 valores).

Variables seleccionadas
Stratum ID 7-character (cnt + region ID + original stratum ID)
Father Qualification - <ISCED level 5A, 6>
Mother Qualification - <ISCED level 5A, 6>
Annual Household Income
Education Cost
Activities - Discuss behaviour teacher
Activities - Volunteer maintenance
Available Schooling
Math Career - In family
Math Career - Child interest
Repeat Grade - At < ISCED 1>
Expected completed levels - <ISCED level 5A or 6>
Parental report on mother's ISCO08 classification
Parental report on father's ISCO08 classification
Parental report on career expectations for their children expressed in ISCO08
Parental involvement in their child s school
Expected completed levels of education
PQ Highest Parental Occupation Status (ISEI)

El mejor subconjunto de atributos parentales alcanzó un mérito de 0.076, indicando que su capacidad predictiva sobre los puntajes de matemáticas es limitada frente a los subconjuntos de datos estudiantiles o escolares. Entre los 18 atributos seleccionados figuran el nivel educativo de los padres, la calificación profesional (ISEI), el ingreso anual, los costos educativos, la implicación parental en la escuela, la trayectoria matemática familiar y las expectativas educativas. Algunas de estas variables muestran correlaciones indirectas: por ejemplo, la participación en actividades escolares influye de manera mediada, mientras que la educación de los padres refleja tendencias generales del capital educativo sin predecir con precisión el rendimiento, ver Tabla 6. Como ilustración, se observa en la Fig. 2 que los estudiantes con bajo nivel matemático suelen tener pocas opciones de escolarización, lo que restringe su acceso y puede ubicarlos en entornos educativos menos favorables. En conjunto, estos atributos ofrecen una visión amplia del entorno parental y su influencia en la educación.

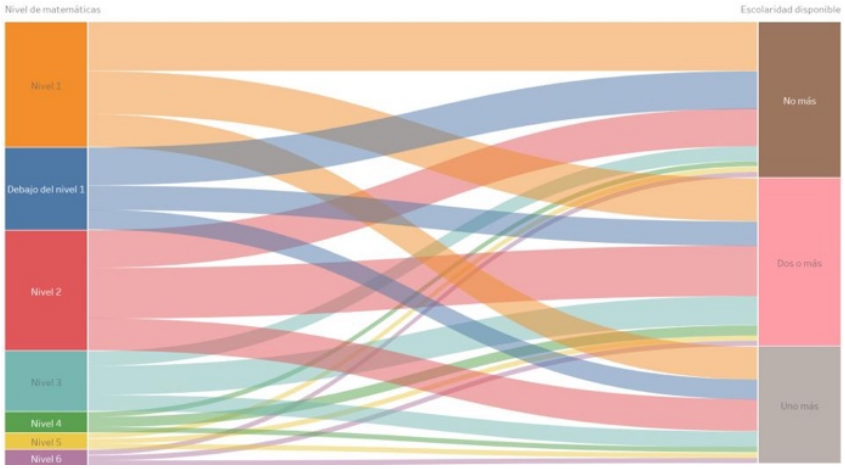


Fig. 2. Gráfico Sankey relacionando el nivel de matemáticas y opciones escolares

Tabla 7. Resultados de la selección de atributos mediante CFS con el algoritmo BestFirst en el contexto estudiantil (20 sobre los 88 valores).

Variables seleccionadas	Variables seleccionadas	Variables seleccionadas	Variables seleccionadas
Age at <ISCED 1>	At Home - Cell phone w/o Internet	At Home - Mother	At Home - USB (memory) stick
At school - Desktop Computer	At school - Internet connection	At school - Tablet computer	Attend <ISCED 0>
Attitude towards School - Does Little to Prepare Me for Life	Attitudes - Source of information	Attitudes - Too unreliable	Attitudes - Troublesome
Attitudes Towards Computers: Limitations of the Computer	Attributions to Failure - Bad Guesses	Attributions to Failure - Not Good at Maths Problems	Attributions to Failure – Unlucky
Country of Birth National Categories- Mother	Difference in Effort	Effort-real 2	Experience with Pure Mathematics Tasks at School

Se obtuvo un mérito de 0.115 para el mejor conjunto, que incluye 88 de un total de 411 variables. Destacan factores sociodemográficos y del hogar (nivel educativo y ocupación de los padres, recursos en casa, computadoras o internet), motivacionales y emocionales (autoeficacia, autoconcepto, ansiedad y atribuciones de fracaso), prácticas de estudio y exposición escolar (tiempo de tareas, tutorías, contenidos, y apoyo docente) y uso de tecnologías tanto en la escuela como fuera de ella, ver Tabla 7.

Aunque el mérito es bajo, la diversidad de variables refleja la complejidad multifactorial del rendimiento académico y ofrece una visión holística del aprendizaje, permitiendo identificar patrones transversales entre factores familiares, individuales, escolares y tecnológicos.

Tabla 8. Resultados de la selección de atributos mediante CFS con el algoritmo Best First en el contexto escolar (20 sobre los 42 valores).

Variables seleccionadas	Variables seleccionadas	Variables seleccionadas	Variables seleccionadas
Funding - Benefactors	School Location	Enrolment - Number of girls	No. of teachers - Total Full Time
No. of teachers - Total Part time	Student Computers - Total number of students	Student Computers - Computers for Education	Student Computers – Computers with Internet
Internet schoolwork - Homework	Shortage – Science lab equipment	Shortage - Computers for instruction	Shortage – Heating/cooling and lighting
Activities - School play	Activities - Volunteering	Additional maths lessons	Learning Hindrance – Students truancy
Parent Participation - Assistance building and grounds	Parent Participation – School government	Parent Participation – Fundraising	Teacher Monitoring – External observers

El subconjunto identificado en el contexto escolar alcanzó un mérito de 0.06, bajo frente a los contextos estudiantil o parental, pero los 42 atributos seleccionados ofrecen una visión diversa de las condiciones escolares que pueden influir en el rendimiento matemático ver Tabla 8. Incluye factores estructurales (infraestructura, número de docentes, equipamiento tecnológico), organizativos (autonomía escolar, contratación, políticas de admisión) y de clima escolar (participación de padres, monitoreo docente, actividades extracurriculares). Aunque no muestran correlaciones fuertes con los puntajes, reflejan cómo el entorno institucional condiciona las oportunidades de aprendizaje y ejerce influencia indirecta en el desempeño.

Método Wrapper

La selección de atributos se realizó con el Evaluador de Subconjuntos Wrapper en Weka, combinado con el método de búsqueda Greedy Stepwise, utilizando el clasificador J48 para medir el desempeño. J48 fue seleccionado por su balance entre precisión y eficiencia, adecuado dado que los métodos Wrapper son computacionalmente costosos al evaluar múltiples subconjuntos. La evaluación se realizó mediante validación cruzada de 5 pliegues con parámetros -C 0.25 (factor de confianza) y -M 2 (mínimo de instancias por hoja); -C controla la poda del árbol equilibrando complejidad y generalización, mientras que -M permite detectar patrones sutiles sin crear hojas inestables. Estos parámetros consideran el tamaño del conjunto de datos (más de 30,000 filas) y el número de atributos (85 a 412), asegurando que Wrapper evaluara eficientemente los subconjuntos y que J48 capturara patrones relevantes para predecir los puntajes matemáticos de los estudiantes.

Tabla 9. Resultados de wrapping conjunto con el algoritmo J48.

Variables Familiar seleccionadas	Variables Estudiantil seleccionadas	Variables Escolar seleccionadas
Stratum ID	Possessions – Internet	Enrolment – Number of girls
Mother Qualification - <ISCED level 5A, 6>	Maths Self-Efficacy – Solving Equation	No. of maths teachers – <ISCED5A> Qual Full Time
Math Career - Child interest	Maths Self-Concept – Not Good at Maths	Teacher morale – High Morale
Repeat Grade - At < ISCED 1>	Out-of-School Study Time – Homework	Responsibility – Teacher start salaries – Teachers
Repeat Grade - At < ISCED 2>	Familiarity with Concepts – Divisor	Responsibility – Salary increase – Teachers
Expected completed levels - <ISCED level 5A or 6>	Maths Lessons – Algebra	Student-Teacher ratio
	Unique National Study Programme Code	
	Attitudes Towards Computers: Limitations of the Computer as a Tool for School Learning	
	Mathematics Self-Efficacy	
	Mathematics Self-Concept	

Los subconjuntos seleccionados en los contextos familiar, estudiantil y escolar muestran consistencia y equilibrio, reflejando distintos tipos de factores que influyen en el rendimiento en matemáticas. En el contexto familiar, se identificaron variables como el “Stratum ID”, la formación académica de la madre, las expectativas educativas de los padres, el interés del niño por la carrera matemática y la repetición de grado en niveles tempranos, ver Tabla 9. Estas variables muestran cómo aspectos del capital cultural y la trayectoria académica temprana se combinan para influir en el desempeño, destacando que ningún factor por sí solo explica los resultados. En el contexto estudiantil, las variables más relevantes incluyen la autoeficacia y el autoconcepto matemático, la familiaridad con conceptos básicos como divisores y álgebra, el tiempo dedicado a tareas y el acceso a Internet, así como actitudes hacia las computadoras y el programa nacional de estudios. Esto evidencia que las competencias cognitivas, las percepciones del estudiante y los recursos disponibles actúan de manera complementaria en su aprendizaje. Finalmente, en el contexto escolar, destacan variables relacionadas con la estructura institucional y el liderazgo, como el número de alumnas matriculadas, la cantidad de docentes con formación completa, la moral del profesorado, la relación alumno-profesor y las responsabilidades salariales. Esto subraya la importancia del entorno organizativo y de los incentivos para la calidad educativa.

4.3 Conjuntos de los resultados

En esta sección se presentan los subconjuntos de atributos seleccionados a partir de los diferentes métodos de selección previamente discutidos. El objetivo es combinar los resultados de técnicas de ranking (correlación, gain ratio, relief-f) con métodos de selección por subconjunto y wrapper, utilizando una ponderación definida para cada tipo de método. Esto permite integrar la información de múltiples perspectivas, considerando tanto la relevancia individual de cada atributo como su contribución combinada a la predicción del desempeño matemático.

Tabla 10. 20 variables seleccionadas y puntuaciones finales de cada contexto combinando métodos estadísticos, subset y wrapper.

Familiar		Escolar		Estudiantil	
Score Final	**Variables**	**Score Final**	**Variables**	**Score Final**	**Variables**
0.956	Repeat Grade - At < ISCED 1>	0.874	Enrolment - Number of girls	0.866	Familiarity with Maths Concepts - Divisor
0.901	Math Career - Child interest	0.510	Teacher morale - High Morale	0.857	Mathematics Self-Concept
0.854	"Mother Qualification - <ISCED level 5A 6>"	0.497	No. of maths teachers - <ISCED5A> Qual Full Time	0.851	Maths Self-Efficacy - Solving Equation 1
0.802	Stratum ID	0.485	Responsibility - Salary increase - Teachers	0.846	Maths lessons - Algebra
0.514	Repeat Grade - At < ISCED 2>	0.484	Student-Teacher ratio	0.844	Out-of-School Study Time - Homework
0.471	Expected completed levels of education	0.480	Responsibility - Teacher start salaries - Teachers	0.843	Attitudes Towards Computers: Limitations of the Computer as a Tool for School Learning
0.422	"Father Qualification - <ISCED level 5A 6>"	0.455	Responsibility - Hiring teachers - National education authority	0.836	Mathematics Self-Efficacy
0.421	Activities - Discuss behaviour teacher	0.452	Student Computers - Computers with Internet	0.560	Possessions - Internet
0.411	Math Career - In family	0.449	Additional maths lessons	0.551	Unique national study programme code
0.408	Education Cost	0.447	Admission - Academic record	0.489	ISCED level
0.406	Parental involvement in their child s school	0.444	Student Computers - Computers for Education	0.485	Grade Repetition
0.400	Activities - Volunteer maintenance	0.442	School Location	0.471	Repeat - <ISCED 1>
0.395	PQ Highest Parental Occupation Status (ISEI)	0.438	Responsibility - Hiring teachers - Principals	0.440	International Grade
0.387	Available Schooling	0.436	School selectivity/student admission policies	0.440	Grade compared to modal grade in country
0.365	Annual Household Income	0.433	Shortage - Science lab equipment	0.423	Possessions - computer
0.355	Mother's ISCO08 classification	0.427	Responsibility - Teacher start salaries - <School governing board>	0.418	National Study Programme
0.347	Career expectations for their children	0.426	Stratum ID 7-character (cnt + region ID + original stratum ID)	0.411	Maths Self-Concept - Learn Quickly
0.341	Father s ISCO08 classification	0.425	Quality of school educational resources	0.407	"Index of economic social and cultural status"
0.198	Expected completed levels - <ISCED level 5A or 6>	0.423	Total school enrolment	0.407	Maths Anxiety - Get Very Nervous
0.106	Activities - Discuss progress teacher	0.412	Shortage - Heating/cooling and lighting	0.403	Maths lessons - Change in graphs

Además de los subconjuntos seleccionados, se exploraron patrones específicos en el ámbito estudiantil relacionados con factores emocionales, como la ansiedad frente a las matemáticas. En la Fig. 3 se observa que cerca del 50 % de los alumnos reportan ansiedad frente a las matemáticas, especialmente aquellos con bajo desempeño (nivel ≤1). Este hallazgo sugiere que la ansiedad puede ser un factor relevante para considerar al

interpretar los resultados académicos y que los alumnos con menor dominio matemático son más susceptibles a experimentar tensión frente a la materia.

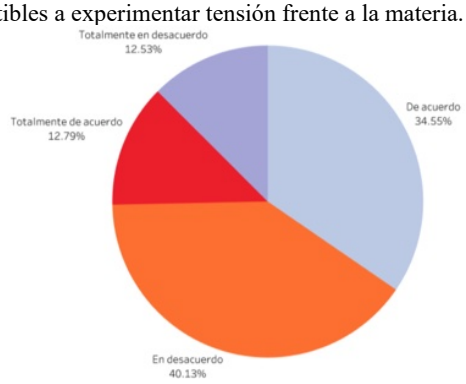


Fig. 3. Gráfico circular de la ansiedad de los estudiantes sobre las matemáticas

La Fig. 4 muestra cómo, conforme aumenta el nivel matemático, los estudiantes tienden cada vez con mayor frecuencia a estar en desacuerdo o totalmente en desacuerdo con esta afirmación, y esta tendencia se intensifica en los niveles superiores. Este patrón evidencia una "desacralización" de la materia entre los alumnos con mayor dominio y refleja cómo variables emocionales interactúan con el desempeño académico. En conjunto con los subconjuntos seleccionados, estas observaciones sugieren que los factores estudiantiles, especialmente aquellos ligados a la autoconfianza y al manejo de la ansiedad, pueden complementar la información de los métodos de selección de atributos para explicar diferencias en los niveles de matemáticas.

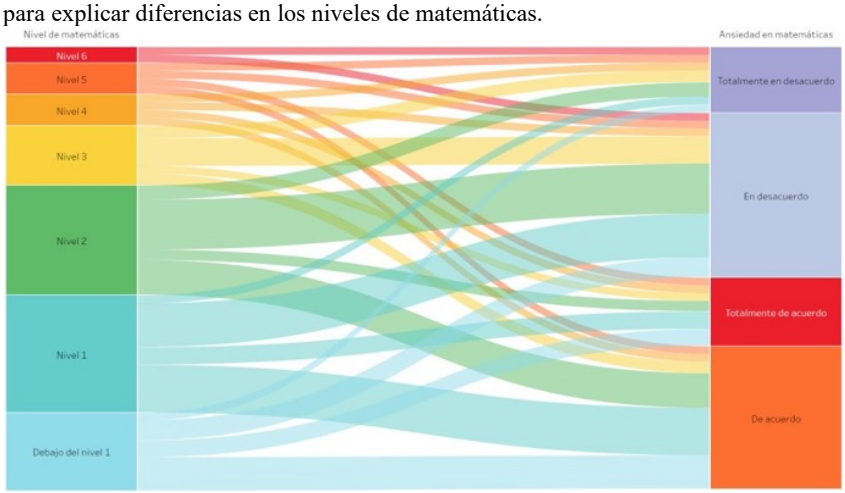


Fig. 4. Gráfico Sankey relacionando el nivel de matemáticas y la ansiedad

Para evaluar la relevancia de las variables seleccionadas, realicé una validación con un 70 % de datos para entrenamiento y 30 % para prueba, aplicando un muestreo previo para mejorar la representatividad y reducir el sesgo de categorías con pocos estudiantes en los niveles más altos de matemáticas. El objetivo no fue predecir perfectamente los puntajes, sino comprobar si las variables seleccionadas efectivamente capturan la información más relevante para clasificar el nivel de desempeño del estudiante. Para ello se utilizaron tres algoritmos: Naïve Bayes, que asume independencia entre variables; J48, un árbol de decisión que realiza particiones secuenciales; y RandomTree, que genera árboles con mayor variabilidad y puede capturar interacciones complejas.

Tabla 11. Precisión en porcentaje de clasificación de subsets con variables seleccionadas y preselección sobre el nivel de matemáticas.

	Familiar			**Escolar**		
Métodos	Naïve bayes	J48	Random Tree	Naïve bayes	J48	Random Tree
Con Selección	38.71	53.01	65.54	30.04	41.34	41.34
Sin Selección	37.74	54.32	64.77	27.33	41.58	41.35

	Estudiantil			**Conjuntos**		
Métodos	Naïve bayes	J48	Random Tree	Naïve bayes	J48	Random Tree
Con Selección	38.19	44.49	62.31	38.19	50.16	64.48
Sin Selección	40.04	54.67	64.39	39.11	53.51	64.39

Los resultados muestran que la selección de atributos no incrementa significativamente la exactitud, pero las mínimas diferencias entre modelos con y sin selección —con preselección de atributos— confirman que las variables más informativas sobre el nivel del estudiante están incluidas en nuestro subconjunto. Por ejemplo, en el contexto estudiantil, RandomTree alcanza 62.31 % con selección frente a 64.39 % sin selección, solo 2 % de diferencia, lo que indica que la reducción de 411 a 20 variables conserva la mayor parte de la información relevante. La menor exactitud observada con J48 (≈10 % menos) se explica porque este algoritmo realiza particiones secuenciales y puede no capturar patrones complejos que RandomTree sí detecta, especialmente interacciones entre variables, ver Tabla 11.

5 Conclusión

Este estudio demostró que la aplicación de múltiples métodos de selección de variables (Correlation, Gain Ratio, Relief-F, CFS y Wrapper) ofrece una visión más completa y precisa de los factores que influyen en el rendimiento matemático. La combinación de enfoques estadísticos y basados en clasificadores permitió reducir significativamente el número de atributos sin perder información esencial, al tiempo que reveló jerarquías e interacciones entre factores del estudiante, la familia y la escuela. De este modo, se construyó un protocolo sólido y transparente que aporta interpretabilidad y utilidad práctica para orientar decisiones en política educativa.

Si bien este enfoque aporta robustez, es importante reconocer ciertas restricciones. Métodos como CFS pueden subestimar interacciones complejas entre atributos, mientras que algoritmos como J48 tienden a privilegiar la interpretabilidad sobre la exhaustividad. Además, la dependencia de la calidad de los cuestionarios limita la

generalización de los hallazgos. A pesar de ello, la combinación de métodos fortalece la consistencia de los resultados y confirma que el rendimiento académico no puede entenderse desde un único factor, sino mediante la integración dinámica de contextos múltiples.

Finalmente, como línea de trabajo futuro, se está ampliando el análisis a las ediciones de PISA posteriores a 2012. Dado que cada ciclo introduce cambios en los cuestionarios y las variables, será necesario adaptar la metodología para mantener la consistencia comparativa. En una primera etapa, el objetivo es consolidar el método en el contexto mexicano; posteriormente, se plantea extender el estudio a otros países e incorporar una perspectiva interdisciplinaria mediante la colaboración con especialistas en educación, con el fin de interpretar de manera más integral los resultados y fortalecer la mejora de las políticas educativas a nivel nacional e internacional. Más allá del enfoque metodológico, la identificación de estos factores es esencial para comprender cómo los distintos contextos educativos y sociales influyen en el aprendizaje y para generar evidencia que sustente estrategias orientadas a mejorar la equidad y la calidad educativa. Esto garantiza la viabilidad de estudios longitudinales y la posibilidad de diseñar intervenciones focalizadas basadas en evidencia.

6 Referencias.

1. OECD: PISA: Programme for International Student Assessment. https://www.oecd.org/en/about/programmes/pisa.html, last accessed 2025/10/28
2. OECD: PISA Database 2012. https://webfs.oecd.org/pisa2022/index.html#section4, last accessed 2025/10/28
3. ANCONE Homepage. http://148.204.64.166/ANCONE, last accessed 2025/10/28
4. Ramírez Lemus, L.: Preparación académica y éxito en la educación secundaria en México (2023)
5. Singh, H., Kaur, M., Singh, B.: A hybrid feature weighting and selection-based strategy to classify high-dimensional and imbalanced medical data (2024)
6. OECD: PISA Data Analysis Manual: SPSS Second Edition, Ch. 6 – Plausible Values. https://www.oecd.org/content/dam/oecd/en/publications/reports/2009/03/pisa-data-analysis-manual-spss-second-edition_g1gh9e4d/9789264056275-en.pdf, last accessed 2025/10/28
7. Wu, M.: Plausible values in assessment data. Australian Council for Educational Research. https://www.rasch.org/rmt/rmt182c.htm, last accessed 2025/10/28
8. OECD: PISA Data Analysis Manual, p.139. https://www.oecd.org/content/dam/oecd/en/publications/reports/2009/03/pisa-data-analysis-manual-spss-second-edition_g1gh9e4d/9789264056275-en.pdf, last accessed 2025/10/28
9. WEKA Homepage, https://ml.cms.waikato.ac.nz/weka,last accessed 2025/10/28
10. WEKA Attribute Selection Package Documentation, https://weka.sourceforge.io/doc.stable-3-8/weka/attributeSelection/package-summary.html, accessed 2025/10/28.
11. Bhosale, D., Ade, R.: Feature Selection based Classification using Naive Bayes, J48 and Support Vector Machine. (2014)

Landscapes of memory and Sustainable Education in rural contexts: a STEAM learning project around hydraulic mills

Susana Gómez-Redondo, University of Valladolid,
susana.gomezr@uva.es, https://orcid.org/0000-0002-8285-2068

Raúl Maján Navalón, University of Valladolid, raul.majan@uva.es,
https://orcid.org/0009-0007-3354-0803

Lilian Johanna Obregón, University of Valladolid,
lilianjohanna.obregon@uva.es, https://orcid.org/0009-0000-5486-9383

1 Introduction

The present work aims to empower rural schools and natural and cultural heritage as the core of sustainable and environmentally friendly education. Framed within the theoretical approaches of inclusive schooling and education in the territory, it starts from the hypothesis that the school, especially the rural one, functions as a cohesive element open to the society in which it is embedded. From this perspective, the school institution not only provides exceptional opportunities for contextualized and situated learning but also acts as a constituent and sustaining factor of community and social interaction. In response to this aim, a project for sustainable and contextualized learning in the rural environment is presented, more specifically in the heart of depopulated Spain (province of Soria). The proposal combines educational innovation, attention to diversity, heritage memory, and the use of the immediate environment and social group as educational spaces and agents. In summary, it is a guide whose didactic design responds to the principles of an inclusive school included in the territory.

An inclusive school is one that, based on the pillars of inclusion, democratization, and interculturality, links its curricular practices with the territory [14, 19, 21, 22], with the aim of transforming itself and the territory. Open to the community and aware of its social function [5, 11], it is characterized by being a flexible institution, attentive to its physical and social environment, and whose objective of revaluing the territory implies contextualizing the educational process within it.

Rural territories present a multiplicity of unique economic activities and geographical, historical, cultural, landscape, and ecological assets, which give rise to lifestyles and local knowledge distinct from those of urban areas [7; 27, 28]. Such indicators are part of the identity markers of the school and can and should be incorporated into the school curriculum [3]. Rural schools are characterized by having their own identity [4]. In contrast to the greater neighbourhood dispersion of urban spaces, rural communities facilitate the participation and involvement of their members. This provides opportunities for contextualized learning [13], where students' daily experiences become one of their greatest motivations and educational sources.

The first sequence begins by addressing more technological and scientific skills and knowledge. The second block addresses the social area, to end with a third part that includes an artistic and creative dimension, especially in the plastic and literary fields, although it also considers other expressions. The guide concludes with a basic glossary that includes the most common terms in hydraulic milling.

3.1. Learning Sequence 1: *Three, two, one... devices in action*

The objective of this block is to understand what a mill is and how it works. For this, activities related to mills and hydraulic devices and their operation are proposed. Competencies and objectives closely related to a STEM (Science, Technology, Engineering, Mathematics) structure are considered as a framework to begin explaining what they consist of. Resources and activities are proposed for students that address competencies and stage objectives related to natural sciences, technology, and mathematics. It begins with an invitation to exploration, cooperative discovery, and techniques such as TPS (Think, Pair, Share), fostering intuition and inquiry without fear of making mistakes. Research tasks and group discussions are complemented with some basic diagrams and explanations, which include the investigation of mill parts and a series of inquiries that, among other intentions, seek to connect ancient devices with their current heritage (alternative energies, sustainability...).

Figure 1. Example of STEM Inquiry Activity from Sequence 1

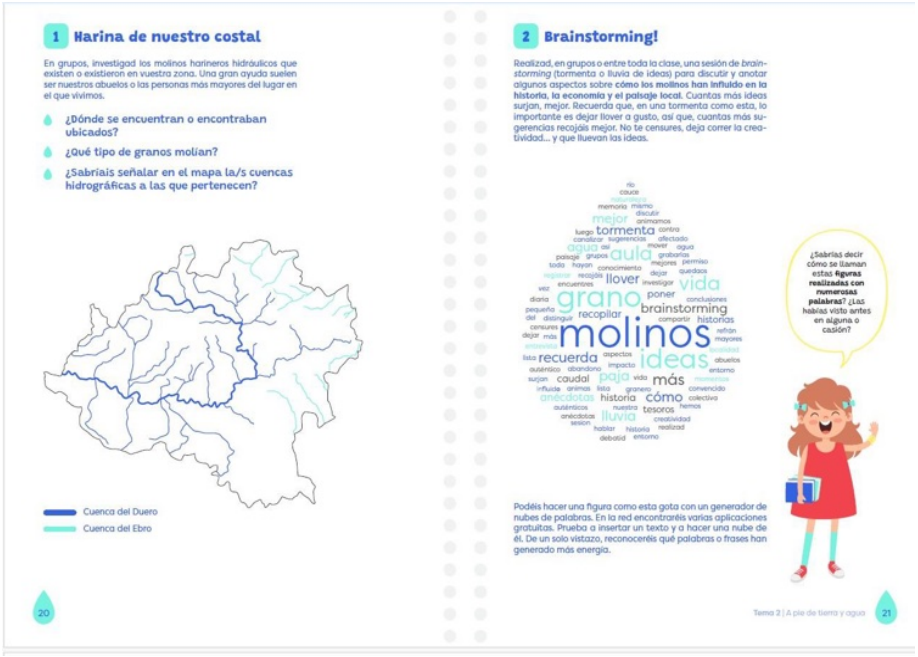


3.2. Learning Sequence 2: *On Land and Water*

This second sequence aims to explore local mills and traditions, proposing a journey from the local to the global that begins with the immediate environment and the students' context. The objective is to locate on a map their river basins, the socioeconomic context, the reason why mills or flour factories (or their remains or memory) are where they are, a bit of their history... The invitation is to explore and discover the mills of the area, the place names, the types of grain (what was black bread or smuggling all about...), legends, anecdotes, recipes... related to the landscape and the milling community.

The intention is to connect global context and history with local history and to talk with their neighbours, grandparents, elders... and other members of their community to learn, understand, and value the customs and hydraulic heritage of the area. Alongside this intergenerational dialogue, and with the aim of building bridges between the past, present, and future, they are encouraged to work on the Sustainable Development Goals (SDGs) of the 2030 Agenda through different methodological strategies around the ECO approach.

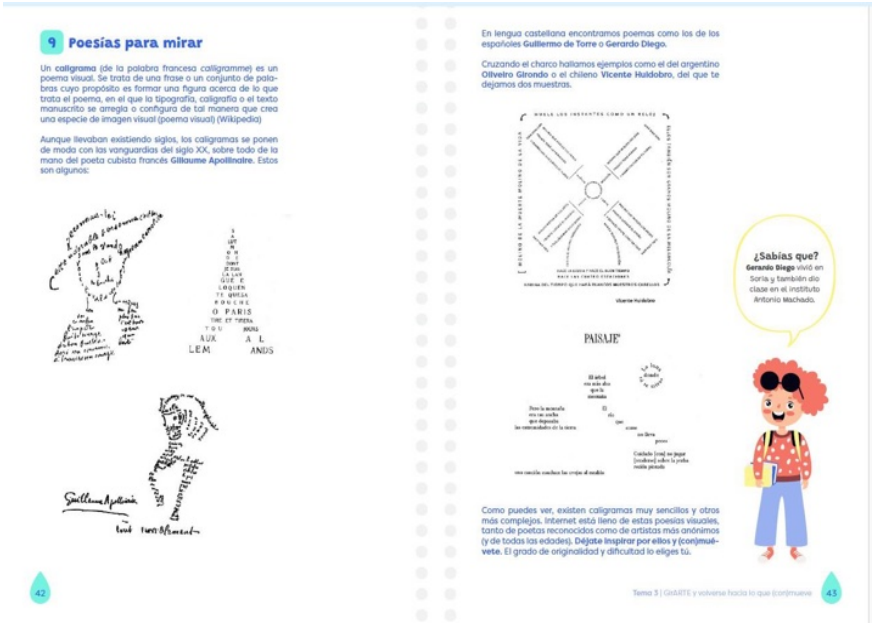
Figure 2. Example of local exploration from Sequence 2



3.3. Learning Sequence 3: *TurnArt... and look at what moves us*

The third sequence aims to understand and comprehend the importance of the mill in artistic imagination and its pictorial, literary, and musical representations, as well as to foster artistic creativity. As in the two previous blocks, it begins with exploration, inviting students to research some painters and works related to milling. This serves as an incentive to learn about aspects of their lives and works, in a process where information search skills lead to knowledge and creative abilities. Narrative, poetry, visual arts... share space with photographic and theatrical notes, as well as traditional proverbs and songs. Once again, discovery-based learning enhances the early stages of exploration, while creation is supported by the motivation and proposals of 20th-century avant-garde movements, such as the exquisite corpse or the creation of calligrams. As can be seen, the three projects seek to create bridges between the past and the present, each from its own perspective.

Figure 3. Example of creative and expressive activities from Sequence 3



3.4. Contribution to the SDGs

In addition to valuing and preserving a crucial element of cultural and technological heritage, such as flour mills, the proposal aligns with the Sustainable Development Goals (SDGs) established by the United Nations. Thus, through the three projects detailed in this guide, a direct contribution to several SDGs is promoted, particularly highlighting the following:

• SDG 4 (Quality Education): the projects promote inclusive and equitable quality education, providing lifelong learning opportunities for all by integrating local and global knowledge about flour mills within a STEAM framework.
• SDG 6 (Clean Water and Sanitation): by exploring the history and technology of water mills, students learn about sustainable water management, the importance of preserving our water resources, and practices that promote sanitation and access to clean water.
• SDG 7 (Affordable and Clean Energy): the relationship between ancient mills and current renewable energies highlights how traditional technologies can inspire sustainable energy solutions for the future.
• SDG 11 (Sustainable Cities and Communities): through intergenerational dialogue and exploration of the local environment, the project contributes to making cities and communities inclusive, safe, resilient, and sustainable, valuing cultural heritage as a pillar of sustainability.
• SDG 12 (Responsible Consumption and Production): understanding how historical milling relates to sustainable consumption patterns and the responsible production of food is promoted.
• SDG 13 (Climate Action): analyzing the environmental impact of mills and their relationship with renewable energies contributes to awareness-raising and action on climate change.
• SDG 15 (Life on Land): researching the environment of the mills fosters the protection, restoration, and promotion of the sustainable use of terrestrial ecosystems, managing forests sustainably, combating desertification, halting and reversing land degradation, and curbing biodiversity loss.

3.5. Key Competences

At the intersection of academic learning and its application in the real environment, contextualized education based on real situations emerges as a powerful tool for the holistic development of students. This approach is reflected and emphasized in Spanish educational legislation through Organic Law 3/2020, of December 29, which modifies Organic Law 2/2006, of May 3, on Education (LOMLOE), as well as in the specific context of Castilla y León through DECREE 26/2016, of July 21, which establishes the curriculum and regulates the implementation, evaluation, and development of Primary Education in the Community. These regulations highlight the importance of connecting curricular content with practical experiences and the cultural and natural context of students.

This perspective is particularly relevant for the fifth and sixth grades of Primary Education, where the aim is to foster not only knowledge but also understanding and appreciation of local and national heritage. In this context, the exploration of hydraulic flour mills offers a unique opportunity to address the key competencies defined by both national educational legislation and the autonomous legislation of Castilla y León. These ancient mills, beyond their heritage value, become cross-cutting axes for the development of critical, creative, and research skills, allowing students greater connection with their environment.

The table 1 summarizes how teaching materials focused on hydraulic flour mills contribute to the development of key competencies while addressing enriching and contextualized specific content, in accordance with DECREE 26/2016 of Castilla y León and Organic Law 3/2020 (LOMLOE), reflecting an evolution in the conception of education towards a more integral and connected with the students' environment.

Table 1. **Key competencies and tasks of the Learning Situations**

Key Competencies	Tasks	Sequence 1: *Three, two, one… ingenuity in action*	Sequence 2: *At the foot of the earth and water*	Sequence 3: TurnART … and look at what moves us
Competence in linguistic communication	The interpretation and production of written texts will be encouraged through the description and analysis of the parts of a mill and its operation, the elaboration of narratives and stories involving mills, and group discussion about their historical and cultural impact. Oral communication will be reinforced through presentations and exhibitions on the projects carried out.	Activities that stimulate description and debate, fostering oral and written expression.	Investigations and dialogues about local mills improve communication and presentation.	Analysis of works of art and creation of poems and narratives enhance linguistic skills.
Mathematical competence and basic competences in science and technology	Calculations will be made to understand the production of flour and the performance of the mills. The understanding of the mechanical operation will be achieved through the construction of mill models and experiments that demonstrate the use of hydraulic energy.	Experiments that demonstrate basic scientific principles through the construction of models.	Exploration of the history and technology of mills improves scientific understanding.	Reflection on energy and sustainability linked to mills and their technology.
Digital competence	Students will research mills using the internet and develop digital presentations of their projects. The use of digital tools for the design and simulation of mills will also be encouraged, as well as for the creation of educational videos that document the learning process.	Use the internet for research and project presentation.	Search for historical and current information about mills and their impact.	Use of digital tools for the creation of creative projects about art and mills.

Learning to learn	The development of research and study strategies will be encouraged through the realization of projects that require the search, analysis and synthesis of information about mills. Reflection on the learning process will be promoted through learning diaries and group discussions about the challenges encountered and the solutions applied.	Encourages exploration and discovery through simple experiments.	Promotes research and intergenerational dialogue, stimulating autonomous learning.	Motivates artistic research and critical analysis of works related to mills.
Social and civic competences	Teamwork will be essential in carrying out projects and experiments. Respect for local history and understanding of the socioeconomic impact of mills will be promoted through research and presentation of their history and their role in the community.	Promotes teamwork and respect for opinions during experiments.	Encourages collaboration and respect for local heritage through research.	Encourages cultural appreciation and creative expression, highlighting the community value of mills.
Sense of initiative and entrepreneurial spirit	Students will develop creative projects related to mills, fostering personal initiative and problem solving. This includes everything from creating functional models of mills to proposals for how the principles of mills can be applied in current sustainable solutions.	Stimulates creativity and problem solving in the construction of mill models.	Incentivizes initiatives for heritage conservation and environmental sustainability.	Promotes creativity and innovation through artistic projects inspired by mills.
Awareness and cultural expressions	The cultural heritage of windmills will be valued through activities that include the analysis of works of art and literature that represent them, as well as the creation of artistic expressions inspired by windmills and their impact on culture and the landscape.	Introduction to the knowledge of technological heritage and its historical impact.	Deepening the local history and cultural value of windmills through exploration.	Exploration of the representation of windmills in art and literature, enriching cultural appreciation.

3.6. Evaluation

Throughout the guide various intermediate or final products are proposed, which serve as evidence for their evaluation. Just as with issues of timing and sequencing, it is risky to establish rigid and decontextualized evaluation guidelines. The evaluation rubric, direct observation and its record, the field diary or group notebook… are some of the tools that teachers usually handle in their pedagogical practice.

The qualitative scales generated by these activities may also serve to refine and implement, where applicable, the quantitative or numerical ones. The evaluative framework can also include elements of self and peer evaluation, enriching and refining self-created rubrics that collect indicators and gradations to assist in grading.

4. Limitations of the work

This work presents a learning situation whose objective is to serve as a tool for teachers in rural schools. Its main limitation is that it does not present empirical results or a measurable impact analysis, so its empirical validity is limited. Future work can carry out fieldwork that addresses such analysis, either from quantitative, qualitative, or mixed perspectives, which can validate the design and implementation of this Learning Situation in the Primary classroom as a tool for meaningful and situated learning.

5. Conclusions

It is in accordance with the theoretical-practical budgets of the school included in the territory and the contextualized learning that the presented initiative is developed. Thus, while vindicating the rural architectural and socio-cultural heritage, sustainable energy and socioeconomic reflections are proposed, accompanied by STEAM contents and competencies from the last years of Primary Education. The physical and social framework is seen as a source of knowledge, education for sustainability, and a privileged context for meaningful and situated pedagogical processes. Active methodologies, cooperative and construction and discovery are based on interaction with the community, without dismissing other research formulas based on connectivity and information and communication technologies. This combination of resources from the collective imaginary and the post- digital era seeks to bridge the gap between the past, present, and future, contributing to the revaluation of the physical and social environment from the school curriculum. The acquisition of competencies based on the ECO model includes the pillars of knowledge, know-how, and being, fostering responsible reflection, self-esteem as an individual and member of a community, and, ultimately, sustainable education in the broadest sense of the word, rural landscape, and community memory.

References

1. Asociación de Ciudades Educadoras (AICE) *Carta Internacional de Ciudades Educadoras*. Barcelona (1990)
2. Benlloch C.. Una herramienta para el empoderamiento rural. Intervención en políticas públicas desde la perspectiva de "La ciudad educadora". *Rurali: revista interdisciplinar en estudios rurales*, *1*. Doi: https://doi.org/10.7203/Rurali.1.1.17613 (2021).
3. Boix Roser. "¿Qué queda de la escuela rural? Algunas reflexiones sobre la realidad pedagógica del aula multigrado." *Profesorado. Revista de Currículum y Formación de Profesorado*, vol. 15, no. 2, pp.13-23. https://www.redalyc.org/articulo.oa?id=5671912900 (2011)
4. Bustos, A. La escuela rural española ante un contexto en transformación. *Revista de Educación,* 350, pp. 449-461. **re35019-pdf.pdf** (2009).)
5. CAST. Universal Design for Learning. CAST Universal Design for Learning Guidelines version 2.0. Wakefield, MA: Author. Traducción al español version 2.0. (2013), modificado según la versión 2018 de las Pautas publicadas por CAST
6. Elboj Saso, C., Puigdellívol Aguadé, I., Soler Gallart, M. & Valls Carol, R. *Comunidades de aprendizaje. Transformar la educación.* Barcelona. Graó. (2018).
7. Escobar, A. *Territorios de diferencia: lugar, movimientos, vida, redes*. Popayán, Colombia, Universidad del Cauca (2014).
8. Figueras Bellot, P. Ciudades educadoras, una apuesta por la educación. *Participación educativa*, 22-27. http://hdl.handle.net/11162/91807 (2007).
9. García-Campos, M. D, Canabal, C. & Alba-Pastor, C. Executive functions in universal design for learning: moving towards inclusive education, International Journal of Inclusive Education, DOI: 10.1080/13603116.2018.1474955 (2018).
10. Imbernón, F. *Calidad de la enseñanza y formación del profesorado. Un cambio necesario.* Barcelona. Octaedro. (2014).
11. Jiménez Naranjo, Y., & Kreisel, M. Participación comunitaria en educación-reconfiguraciones de lo escolar y de la participación social. *Teoría De La Educación. Revista Interuniversitaria*, *30*(2), 223–246. https://doi.org/10.14201/teoredu302223246. (2018).
12. LOMLOE. Ley Orgánica 3/2020, de 29 de diciembre, por la que se modifica la Ley Orgánica 2/2006, de 3 de mayo, de Educación. *Boe*, *340*, 122868-122953. Permalink ELI: https://www.boe.es/eli/es/lo/2020/12/29/3/con (2020).
13. Miller, P. M. Contextual Learning May be a Better Teaching Model: a case for higher order learning and transfer. In *Allied Academies International Conference. Academy of Educational Leadership. Proceedings* (Vol. 11, No. 2, p. 19). Jordan Whitney Enterprises, Inc. (2006).
14. Monfort, C., & Sales, A. Vinculándose al territorio: el ApS y la participación comunitaria. Estudio de caso en una escuela rural. *IJNE:aInternational Journal of New Education*, *2*(1), 1-18. https://doi.org/10.24310/IJNE2.1.2019.5422 (2019).
15. Pastor, C. A. *El Diseño Universal para el Aprendizaje: Educación para todos y prácticas de enseñanza inclusivas*. Ediciones Morata. (2018).
16. Pastor, C. A. Diseño Universal para el Aprendizaje: un modelo teórico-práctico para una educación inclusiva de calidad. *Participación educativa*, *6*(9), 55-68. Diseño Universal para el Aprendizaje: un modelo teórico-práctico para una educación inclusiva de calidad- (2019).
17. Quitián Bernal, S. P. Ciudad educadora, lenguaje, sociedad y escuela. *Enunciación*, *26*(1), 12-13. DOI https://doi.org/10.14483/22486798.18668 (2021).
18. Ramos Fernández, L. F., & López Atrio, L. Ciudades educadoras y ciudades creativas, las nuevas fórmulas de la pedagogía social en el siglo XXI. *Revista de la SEECI*, (32), 160-175. DOI https://doi.org/10.15198/seeci.2013.32.160-175 (2013).

19. Sales A., Traver, J. A., & Moliner, O. Redefiniendo el territorio de la escuela: espacios educativos y currículum escolar para la transformación social. *Revista Fuentes, 21*(2), 177–188. DOI https://doi.org/10.12795/revistafuentes.2019.v21.i2.03 (2019).
20. Sterling, S *Sustainable Education – Re-visioning learning and change, Schumacher Briefing no6*. Schumacher Society/Green Books, Dartington. ISBN 1 870098 99 4 https://www.greenbooks.co.uk/sustainable-education (2001)
21. Traver, J. A., Moliner García, O., & Sales Ciges, A. Negociando el currículum: Aprendizaje-servicio en la escuela incluida. *Alteridad. Revista de Educación*, *14*(2), 195-206. DOI https://orcid.org/0000-0002-5948-1035 (2019).
22. Traver, J. A., Sales, M. A., Moliner, O., Sanahuja, A. & Benet, A. Hacia una escuela incluida en su territorio: Análisis de una práctica comunitaria. *Edetania. Estudios y propuestas socioeducativas*, (53), 99-116. DOI http://dx.doi.org/10.12795/revistafuentes.2019.v21.i2.03 (2018).
23. UNESCO Transforming our world: The 2030 agenda for sustainable development: United Nations. Available at: https://sdgs.un.org/2030agenda. (2015).
24. UNESCO Global education monitoring report. Available at: https://gem report-2016.unesco.org/en/chapter/education-and-sustainability-what-we-know-and what-we-need-to-do/ (2016).
25. Vygotsky, L. *Interaction between learning and development* (pp. 79-91). Linköping, Sweden: Linköpings universitet. (2011).
26. Vygotsky, L., & Cole, M. Lev Vygotsky: Learning and social constructivism. *Learning Theories for Early Years Practice. UK: SAGE Publications Inc*, 68-73. (2018).
27. Woods, M. *Rural geography: Processes, responses and experiences in rural restructuring*. SAGE Publications. DOI:10.4135/9781446216415 (2005).
28. Woods, M. *Rural*. Routledge. (2010).

Posters

The energy transition in Cuba: process characteristics

Miguel Castro Fernández[1][0000-0002-3983-469X], Miriam Vilaragut Llanes[2][0000-0002-5453-1136], Antonio Martínez García [3][0000-0001-7006-4454], Jesús C. Hernàndez [4][0000-0001-9117-1689], Ney R. Balderramo[5][0000-0001-8502-4332]

[1,2,3] Technological University of Havana José A. Echeverría, CUJAE. Havana Cuba
[4] University of Jaén, Jaén, Spain
[5] Technical University of Manabí
mcastro@electrica.cujae.edu.cu

Abstract. The current situation of the National Electricity System is quite complex; Cuba suffers daily blackouts, and every day it is learned that one or another thermoelectric plant lacks fuel to operate or has suffered a failure that has taken it off the grid. The same applies to the locations of power generators, while the process of changing the electricity matrix, to date, has been slow. The decision of the Cuban State and Government to change the electricity matrix based on a high integration of renewable energy sources (RES) is an established strategic axis, but the concept of energy transition encompasses more than the change and introduction of these technologies. It also requires the rest of society, and in particular, the business sector, to begin their journey toward fulfilling the Energy Transition Strategy for Cuba by 2030. The implementation of the energy transition in Cuban society, and in the business sector in particular, requires the development of a Roadmap that allows for the definition of objectives and goals for each company to jointly, with the electricity sector and the rest of Cuban society, achieve the objectives of the aforementioned strategy. If this is not achieved, the energy transition will ultimately be limited, at the very least, or applied to a portion of society, which will prevent the country from meeting its stated objectives or achieving the energy security required for the Cuban people to attain the status it deserves.

Keywords: energy transition, solar photovoltaic energy, road map

Preliminary Study on Convolutional Neural Network-Based Burn Diagnosis: Advancing Smart Health in Urban Environments

José de Jesús Álvarez Ramírez[1][0009-0004-2513-2701] Rocio Maciel Arellano[2][0000-0002-5548-2073].

[1] Universidad de Guadalajara, Centro Universitario de Ciencias Económico Administrativas, Doctorado en Tencologías de información, Periférico Norte Norte 799, Zapopan, Jalisco, México.

Abstract. Burn injuries remain a critical medical challenge where timely and accurate diagnosis strongly influences patient recovery. Traditional methods depend on Doctor expertise, often introducing subjectivity and variability. This paper presents a preliminary study on an automated burn classification system based on Convolutional Neural Networks (CNNs). The model was trained and validated on publicly available burn image datasets, using preprocessing and evaluation metrics such as accuracy, precision, recall, and F1-score. While preliminary results (accuracy ≈0.58) indicate that further optimization is needed and not clinical reliability reached, the system demonstrates the feasibility of CNNs for assisting clinical decision-making. Beyond the technical contribution, this research situates the model within the framework of smart and sustainable cities, emphasizing its role as a smart health service. By enabling deployment on cloud platforms and consumer-grade hardware, the solution becomes accessible for hospitals with limited resources, contributing to more equitable healthcare systems.

Keywords: Burn Classification, CNN, Deep Learning. Smart Health, Sustainable Cities.

Photovoltaic Potential in a City like Elche: Solar Plants and Agriculture.

D. López[1], S. Valero[2] [0000-0002-9222-9173], C. Senabre [3] [0000-0001-5749-7431], Francisco José Sánchez Sutil [4]

[1]Demetrio López Sánchez. Photovoltaic Engineer at Demesol Ingeniería FV, SLU. Co-chair of the "University Chair: Demesol of Photovoltaic Solar Energy at UMH"
tecnico@demesol.com

[2] Sergio Valero Verdú. Full Professor and PhD in Electrical Engineering, and co-chair of the "Demesol Chair University: Photovoltaic Solar Energy at UMH." Miguel Hernández University of Elche. Department of Mechanical and Energy Engineering. Avd. de la Universidad s/n. Innova-V Building, 03202 Elche.
svalero@umh.es

[3] Carolina Senabre Blanes. Full Professor and PhD in Mechanical Engineering, Miguel Hernández University of Elche. Department of Mechanical and Energy Engineering
csenabre@umh.es

[4] Francisco José Sánchez Sutil
fssutil@ujaen.es

Spain University of Jaen, Electrical Engineering Departament, Paraje de las Lagunillas s/n, Jaén, Spain C. Hernandez Full Professor.

Abstract. This article presents a study and analysis of the possibility of generating all the electrical energy we consume as a society through photovoltaic solar energy. A case study is presented, applied to the city of Elche and a portion of its municipal area, also known as “Camp d’Elx.” As a first step, it is necessary to determine the region's electrical consumption, as well as the potential for solar generation in the area. In this proposed example, no energy source other than solar has been considered; resources such as wind, hydroelectric, geothermal, or any other type of energy generation have not been taken into account. Only the incident solar irradiation in this region has been analyzed as the primary source of renewable energy. As will be detailed and explained later, generating the electrical energy consumed using photovoltaic solar power installations is initially feasible. However, while it is easy to question this type of analysis from a technical perspective, the article describes how such a solution would indeed be possible.

Keywords: Photovoltaic Solar Installations, Photovoltaic Self-Consumption, Agrivoltaics.

Hybrid solar lighting system with optical fiber and LED with continuous lighting control using PWM

Carlos Henriques Ventura do Rosário Oliveira[1] [0000-0002-6729-9209] , Marcio Zamboti Fortes[2] [0000-0003-4040-8126] and Francisco José Sánchez Sutil[3] [0000-0001-9508-1728]

[1] Department of Electrical Engineering, Fluminense Federal University, Rua Passo da Pátria, 156, Campus da Praia Vermelha, Niterói-RJ, 24210-240, Brazil
carloshenriques@id.uff.br

[2] Department of Electrical Engineering, Fluminense Federal University, Rua Passo da Pátria, 156, Campus da Praia Vermelha, Niterói-RJ, 24210-240, Brazil
mzamboti@id.uff.br

[3] Departamento de Engenieria Eléctrica, Universidad de Jaén, Paraje de las Lagunillas, s/n, Jaén, Spain
fssutil@ujaen.es

Abstract. This paper presents the development and integration of a direct solar lighting system with plastic optical fiber (POF) and an artificial lighting system with light-emitting diodes (LED), equipped with an electronic circuit for continuous lighting control. Emphasizing sustainability and energy efficiency, the utilization of direct solar lighting emerges as a crucial solution for indoor environments, especially considering that lighting accounts for approximately 19% of total energy consumption in buildings. Acknowledging that solar radiation is inherently variable due to weather conditions and cloud cover, the system proposes an innovative approach to compensate for these fluctuations through an adaptive artificial LED lighting component. Direct solar lighting solutions are vital for smart cities, as they combine advanced technology with sustainable practices, enhancing the energy efficiency of lighting systems and promoting healthier indoor environments. As the core proposed solution, a hybrid luminaire was meticulously designed and developed, integrating solar lighting via fiber optics and LED artificial lighting, controlled by a Pulse Width Modulation (PWM) system. This system is engineered to continuously adjust lighting intensity based on available solar radiation, thereby ensuring permanent ambient illumination and a smooth, imperceptible transition between the two lighting sources. The preliminary study conducted on the initial prototype confirmed the technical feasibility and viability of integrating solar and LED lighting. The proposed system's design and functionality pave the way for a more effective and energy-efficient lighting solution, with further comprehensive evaluations planned for specialized facilities like the UFF Lighting Engineering laboratories (LABLUX) to rigorously assess the system's performance under various simulated conditions.

Keywords: optical fiber, solar lighting system, solar luminaire, hybrid lighting system.

Aplicación de algoritmos genéticos para la optimización de la generación fotovoltaica en las redes de distribución de Puerto Rico mediante el software PowerFactory

Daniel Felipe Aragones Vargas[1 [0009−0004−2449−0363]],
Marcel J Castro-Sitiriche[2 [0000−0002−7650−4112]],
Fabio Andrade Reginfo[3 [0000−0002−8859−7336]], and
Jesús C. Hernández[4 [0000−0001−9117−1689]]

Department of Electrical and Computer Engineering, Recinto Universitario de Mayagüez, Mayagüez, PR 00682,USA
danielfelipe.aragones@upr.edu
marcel.castro@upr.edu
fabio.andrade@upr.edu
jcasa@ujaen.es

Abstract. El presente trabajo expone una estrategia para optimizar la generación distribuida fotovoltaica en redes de distribución eléctrica, con el propósito de reducir las perdidas de electricidad técnica, se empleó técnicas de optimización evolutivas inspiradas en biología como lo es el algoritmo genético (GA). La implementación se realizo en el software DigSILENT Power Factory, donde se integró información eléctrica con características geoespaciales a través del formato DGS, lo que permitió aumentar la precisión para el modelo de red. El presente estudio se desarrolló en alimentador de media tensión 6012-01, ubicado en un sector urbano residencial del municipio de Mayagüez-Puerto Rico, un área que posee una presencia significativa de sistemas fotovoltaicos en techos. Para lograr desarrollar el objetivo planteado del presente estudio se ajusto los niveles de generación de cada sistema fotovoltaicos con el fin de minimizar las pérdidas de electricidad, manteniendo siempre las condiciones operativas dentro de los límites de tensión aceptables y evitando sobre carga en transformadores y conductores. Los resultados obtenidos muestran que emplear algoritmos evolutivos funcionan para fortalecer la planificación y operación de las redes de distribución permitiendo una mayor integración de fuentes renovables de energía, particularmente en Puerto Rico.

Keywords: Generacion distribuida · Algoritmo Genetico · rooftop PV · GIS.

Implementation and Validation of an FPGA-Based System for Harmonic and Supraharmonic Analysis in Single-Phase Power Grids

Roger Challco[1,2][0000−0002−1379−2436], F. Sánchez-Sutil[1][0000−0001−9508−1728], José Carlos U. Peña[3][0000−0001−5958−6227], A. Cano-Ortega[1][0000−0002−6440−021X], and J.C. Hernández[1][0000−0001−9117−1689]

[1] University of Jaen, Electrical Engineering Departament, Paraje de las Lagunillas s/n, Jaén, Spain
[2] Universidad Nacional de Ingenieria, Av. Túpac Amaru 210 - Rímac, Lima Perú
[3] Universidade Estadual de Campinas, RCidade Universitária Zeferino Vaz, Campinas, Sao Paulo, Brazil

Abstract. The proliferation of power electronic switching devices has increased the presence of harmonics and supraharmonics in electrical networks, significantly affecting power quality in the 2 kHz to 150 kHz range. This work presents the design, development, and implementation of an FPGA-based prototype for real-time measurement of harmonics (50 Hz to 2 kHz) and supraharmonics (9 kHz to 150 kHz) in single-phase systems, strictly following the methodology established by the IEC 61000-4-7 standard. The system integrates a high-speed analog-to-digital module and an FPGA responsible for data acquisition, processing, and transfer, allowing simultaneous observation of voltage and current harmonics and supraharmonics. Experimental validation, conducted on a photovoltaic generator connected to an inverter, demonstrates a high correlation between the prototype's results and those of a commercial reference instrument, with maximum differences of 2.48 mA in current and 0.17 V in voltage for individual harmonics. The results demonstrate the feasibility of the proposed approach for advanced power quality monitoring and identify areas for improvement focused on optimizing supraharmonic processing directly on the FPGA.

Keywords: FPGA · Supraharmonics · Harmonics, Power Qualit · IEC 61000-4-7.

1 Introduction

The analysis of harmonics and supraharmonics in electrical networks has gained significant importance due to the proliferation of power electronic switching devices, which impact power quality in the 2 kHz to 150 kHz range [4,9,10]. Among the normative methods [1–3], IEC 61000-4-7 [1] stands as the main reference for the measurement of harmonics and interharmonics up to 9 kHz, using specific

Position Control of an 18 DoF Humanoid Robot Using a Saturated Control Law

Miguel Angel Ortega-Palacios[1], Fernando Reyes-Cortés[2]

Amparo Palomino-Merino[2],

Benemérita Universidad Autónoma de Puebla
[1]Facultad de Ciencias de la Computación, [2]Facultad de Ciencias de la Electrónica.
miguel.ortegap@alumno.buap.mx, fernando.reyes@correo.buap.mx,
amparo.palomino@correo.buap.mx

Abstract. Humanoid robots have been extensively studied by the scientific community to understand and replicate human motion during specific tasks. In recent years, numerous methods and strategies have been proposed to develop mathematical models that describe humanoid dynamics. Common approaches include analytical methods, geometric modeling, and the Denavit–Hartenberg convention. In this study, we propose a recursive method based on Euler's analytical mechanics and the use of the properties of the skew-symmetric matrix $S(k)$ to address key modeling challenges, such as forward kinematics, differential kinematics, linear velocity, angular velocity, kinetic energy, potential energy, and the Euler–Lagrange equations of an 18 degrees of freedom (DOF) humanoid robot. A controller based on a global saturated regulator is proposed to implement joint position control. Simulation results are presented using MATLAB/Simulink.

Keywords*:* Dynamics; forward kinematics; humanoid robot; recursive method; saturated controller.

NSN, AIoT en organizaciones civiles de base y redes comunitarias

Carlos Astroza [1] and Adolfo Mackay[1]

Centro de Investigación en TI y Aprendizaje, Antofagasta 1270524, Chile
castroza@citia.cl, amackay@citia.cl

Abstract. Este escrito analiza la evolución del concepto de *Red Social* (RRSS), argumentando que su significado original, fundamentado en la cohesión y la movilidad comunitaria, ha sido desvirtuado por la interpretación comercial de las plataformas masivas en línea.
A través de la visión antropológica que representa el trabajo de J. A. Barnes y la teoría de los lazos débiles de Mark Granovetter, se reposiciona el concepto hacia su esencia sociológica. Además, se propone el desarrollo de las *Redes Sociales de Barrio* (Neighborhood Social Nets o NSN) como un nuevo paradigma. Estas conceptualizaciones, asistidas por tecnologías como AIoT, no solo buscan la conectividad, sino que se centran en fortalecer lazos fuertes y débiles para proveer bienestar, cohesión y movilidad en comunidades civiles de base, como las agrupaciones que gestionan y cohabitan un territorio barrial.
Finalmente, el trabajo describe algunos aspectos del proyecto de I+D+i, referido a la inclusión de AIoT a la *Organización Civil de Base* (OCB), del *Centro de Investigación en TI y Aprendizaje* (CITIA). La iniciativa se enfoca en aplicar la Inteligencia Artificial e Internet de las Cosas (AIoT) a la conceptualización mundialmente conocida como agrupaciones o juntas de vecinos y/o comités vecinales o poblacionales, que es un tipo específico y posiblemente una de las OCB más funcionales de la sociedad civil y el planeta.

Urban energy modelling and data flow: review and application cases for decarbonization and resilience in Spanish cities

Pablo de Agustín-Camacho[1[0000-0001-9715-0803]], David García[2[0000-0002-1405-3053]], Emanuela Giancola[3[0000-0003-2450-1494]], Cecilia Pérez[4[0009-0002-3679-1097]], Carlos Alberto Rivera Gómez[5[0000-0001-5324-068X]], Maria Nuria Sánchez[3[0000-0002-0704-2989]], Silvia Soutullo[3[0000-0001-6420-2734]], María José Suárez[2[0000-0001-7468-831X]], and José Luis Vivancos[6[0000-0002-8524-5466]]

[1] TECNALIA, Basque Research and Technology Alliance (BRTA), Derio, Spain
[2] Universidad de Oviedo, Oviedo, Spain
[3] CIEMAT—Spanish Centre for Energy, Environmental and Technological Research, Madrid, Spain
[4] Institut de Recerca en Energia de Catalunya, Barcelona, Spain
[5] Escuela Técnica Superior de Arquitectura, Universidad de Sevilla, Sevilla, Spain
[6] Instituto de Ingeniería Energética, Universitat Politècnica de Valencia, Valencia, Spain

Abstract. The building sector is a primary contributor to urban energy consumption and associated carbon emissions, making the development of advanced urban energy models crucial for achieving city-level decarbonization and resilience goals. While current modeling approaches have shown promise, they face significant limitations, including data fragmentation, insufficient representation of complex urban interactions, and a critical trade-off between computational cost and predictive accuracy. This paper reviews and presents a standardized multi-domain framework for urban energy modelling and data flow to effectively integrate complex boundary conditions and advanced simulation techniques. The analysis points out the necessity of moving beyond traditional, siloed approaches to a unified workflow that integrates detailed data and diverse temporal and spatial resolution depending on the application case. It also shows how this framework has been applied to address different uses and urban energy challenges, especially on cities' climate change mitigation and/or adaptation.

Keywords: urban energy model, data flow, application cases, decarbonization, resilience.

Eficiencia del extractor de noticias fitosanitarias (FREM)

Olarte Astudillo Rodrigo[1 [0009−0005−3904−374X]], Martínez Luna Gilberto Lorenzo[1 [0000−0002−0105−1112]], Guzmán Arenas Adolfo[1 [0000−0002−8236−0469]], Muñiz Merino Manolo[2 [0009−0000−0339−8534]], Hernández Mata Sagrario[1 [0009−0004−8693−7431]], and Pascualito Jimenez Yadira Juliet[1 [0009−0003−3567−4670]]

[1] Centro de Investigación en Computación del Instituto Politécnico Nacional, Ciudad De México CDMX 07700, MEX https://www.cic.ipn.mx
[2] SENASICA, Ciudad de México 06100, México

Abstract. Los extractores web son herramientas útiles que permiten a las empresas obtener información sobre los últimos acontecimientos para la toma de decisiones. Estos extractores deben adaptarse a las necesidades de las empresas, ya que no pueden obtener toda la información disponible en Internet. Por lo tanto, deben desarrollarse continuamente para garantizar que su eficiencia sea adecuada para las empresas. En este artículo se calculó la eficiencia del extractor web FREM que utilizará SENASICA y su capacidad para clasificar las noticias recopiladas en función de un nivel de severidad. Para calcular la eficiencia, se compararán las noticias recopiladas por las herramientas con las noticias publicadas por SENASICA en el monitor de riesgos durante un rango de tiempo. Como resultado, se encontró que el extractor recopiló con éxito el 80% de las noticias publicadas en el monitor de riesgos. También se encontró que el 70% de las noticias publicadas en el monitor se clasifican como de alta severidad. Por lo tanto, el extractor es posible que envíe una alerta a SENASICA cada vez que recopila una noticia con un nivel de severidad alto.

Keywords: Procesamiento de lenguaje natural · Amenazas fitosanitarias · Extractores web. · Eficiencia de extracción

A machine-learning aid model for estimating lean manufacturing implementation in the Latin-American industry

Diego Rossit*[1], Jesús del Carmen Peralta-Abarca[2], Andrea Insfran-Rivarola[3], Pedro Moreno[2], Ana Pamela Arévalos[3], Daniel Rossit[1], and Sergio Nesmachnow[3]

[1] INMABB, Engineering Dep., Universidad Nacional del Sur-CONICET, Argentina
{diego.rossit,daniel.rossit}@uns.edu.ar
[2] Universidad Autónoma del Estado de Morelos, México
{carmen.peralta,pmoreno}@uaem.mx
[3] Methods Laboratory, Industrial Engineering Department, Universidad Nacional de Asunción, Paraguay aarevalos@fiuna.edu.py
andrea.insfran@ing.una.py
[4] Universidad de la República, Uruguay
sergion@fing.edu.uy

Abstract. Lean Manufacturing is recognized as a strategy to improve efficiency, reduce waste, and foster continuous improvement. In Latin America, adoption still varies significantly, and traditional maturity assessments often rely on qualitative surveys that are resource-intensive and difficult to replicate. This article proposes a predictive framework using supervised machine learning to categorize firms based on their level of Lean Manufacturing implementation. The dataset was created from closed-ended questions of the World Management Survey and refined with expert validation to ensure meaningful labeling of Lean adoption levels. Four ensemble-based classifiers, i.e., Random Forest, Gradient Boosting, XGBoost, and LightGBM, were trained and evaluated using accuracy, precision, recall, and F1-score. Results indicated that overall accuracy ranged from 0.878 (LightGBM) to 0.891 (Random Forest). The class-level F1-scores were consistently high, reaching 0.96 for the "Low" category and between 0.82 and 0.91 for intermediate groups. Results indicate that ensemble methods, particularly Random Forest and XGBoost, provide robust predictive performance while ensuring interpretability through feature analysis. The proposed framework provides a scalable, data-driven approach for assessing Lean maturity, supporting decision-making for managers and policymakers aiming to strengthen industrial competitiveness in Latin America.

Keywords: Smart industry · lean manufacturing · Latin-America · supervised learning · prediction models.

Agent-Based Model Simulation as a Decision Support Tool for Renewable Energy Communities Diffusion

Eduard Falcó Artieda[1], Roberto Villafafila-Robles[2*[0000-0003-4372-2575]], Marina Codina Escolar[2[0009-0002-4434-9741]], Alessandro Sebastiano Carrus[3*[0009-0003-0319-4023]], Ivan Blečić[3[0000-0002-4446-6705]] and Giuseppe Desogus[3[0000-0002-6989-3500]]

[1] Innovation department, Centre de Telecomunicacions i Tecnologies de la Informació (CTTI), Generalitat de Catalunya, Spain
eduard.falco@gencat.cat
[2] Centre d'Innovació Tecnològica en Convertidors Estàtics i Accionaments (CITCEA-UPC), Departament d'Enginyeria Elèctrica, ETS d'Enginyeria Industrial de Barcelona, Universitat Politècnica de Catalunya, Barcelona, Spain
roberto.villafafila@upc.edu; marina.codina@upc.edu
[3] Department of Civil and Environmental Engineering and Architecture, University of Cagliari, Cagliari, Italy
alessandros.carrus@unica.it; ivanblecic@unica.it; gdesogus@unica.it

* corresponding authors: roberto.villafafila@upc.edu; alessandros.carrus@unica.it

Abstract. The energy transition towards distributed generation highlights the role of Renewable Energy Communities (RECs) as complex socio-technical systems. To analyze the dynamics of these communities, a sophisticated Agent-Based Model (ABM) simulation has been developed on the NetLogo platform, designed to serve as a Decision Support System (DSS).

The simulator integrates interdependent modules: a physical-energy module, based on realistic hourly consumption and photovoltaic production profiles sensitive to climatic zones; an economic module for monetizing energy flows according to detailed market tariffs; and a financial module to evaluate investment sustainability (CAPEX, OPEX, Payback Period). The core of the model is the socio-technical module, where heterogeneous agents, with their own attributes and decision-making thresholds, evaluate perceived benefits to decide whether to join or leave the community on an annual basis.

This approach makes it possible to observe emergent phenomena such as adoption curves and the impact of different governance policies on benefit sharing. The model allows for in-silico experiments, including an analysis of the system's resilience to exogenous shocks like energy price crises. Future development will focus on a comparative analysis of different regulatory frameworks, specifically contrasting the Italian model with the Spanish-Catalan one, to evaluate the differential impact of policies on REC performance and growth.

Keywords: Renewable Energy Community; Local energy community; Energy transition; Social acceptance, Multi-Agent Systems.

OptIoT-ResourSense: Agriculture Optimization with IoT of Resources through Sensing

Enrique González-González [1[0000-0002-8025-2464]], Miguel Á. Sánchez-Herreros [1[0009-0001-2279-8477]], Alonso Garrido-Limiñana [1[0000-0002-6232-7012]], Ana del-Campo-Sánchez [1[0000-0002-1423-3572]], Pedro Antonio Plaza-Álvarez [1[0000-0003-1034-2576]] and J. Isidoro Martínez-Gutiérrez [1[0009-0008-0383-0241]]

[1] Universidad de Salamanca, Ávila 05003, Spain
egonzalezgonzalez@usal.es

Abstract. Efficient management of soil and water resources is a major challenge for sustainable agriculture under climate variability. This work presents OptIoT-ResourSense, a low-cost IoT sensor node based on a LoRaWAN-enabled platform, integrating soil moisture, electrical conductivity (EC), temperature, and a tipping-bucket rain gauge. The system operates with unlimited autonomy through photovoltaic energy harvesting and enables the connection of multiple EC probes, facilitating both spatial and potentially multi-depth monitoring. Preliminary field tests in irrigated potato crops (Rasueros, Ávila, Spain) confirmed stable LoRaWAN transmission, continuous outdoor operation, and a hardware cost substantially lower than most commercial solutions. Graphical results demonstrate the prototype's capacity to capture soil–water dynamics, nutrient mobilization, and the response of soil profiles to rainfall and sprinkler irrigation events. Beyond irrigation optimization, EC measurements provide insights into soil fertility and potential correlations with soil organic carbon. The modular RS485 interface also allows the integration of additional sensors (e.g., pH or flow meters), extending its applicability to broader agro-environmental monitoring. Overall, OptIoT-ResourSense offers an affordable, scalable, and sustainable solution for precision agriculture, bridging experimental monitoring with practical decision-making, and contributing to the achievement of Sustainable Development Goals (SDGs 6, 11, and 13).

Keywords: IoT, LoRaWAN, Smart Agriculture, Soil Moisture, Rainfall, Resource Optimization.

Energy-Saving Potential of Natural Ventilation in Temperate Climates: A Case Study of Spain

Carolina Aparicio-Fernández[1][0000-0003-0197-9452], José-Luis Vivancos[2][0000-0002-8524-5466]
Mar Cañada-Soriano[3][0000-0002-9435-2132], Ali Rafat Gigasari[4][0000-0002-5262-4574]

[1] Building Technology Research Centre, Universitat Politècnica de València, 46022, Valencia, Spain, caap@upv.edu.es
[2] Instituto de Ingeniería Energética, Universitat Politècnica de València, 46022 València, Spain, jvivanco@upv.es
[3] Applied Thermodynamics Department (DTRA), Universitat Politècnica de València, 46022, Valencia, Spain, macaaso@upv.es
[4] Universitat Politècnica de València, 46022, Valencia, Spain, arafgig@doctor.upv.es

Abstract. Natural ventilation is an effective strategy to reduce energy consumption and enhance thermal comfort in buildings, though its performance is strongly influenced by orientation, location, and climate. This study investigates the energy saving potential of natural ventilation in temperate climates using Spain as a case study. Different building configurations were simulated for four cities with distinct climatic conditions: Bilbao, Barcelona, Madrid, and Cáceres.

Results indicate that natural ventilation can reduce cooling demand by 20% to 60%, depending on orientation and location. The addition of fans further improved air change rates and lowered cooling loads.

These findings highlight the importance of accounting for climate variability and building orientation when designing natural ventilation systems. The study contributes practical insights for designers and planners, demonstrating that natural ventilation can be an effective strategy to lower energy consumption and improve thermal comfort in temperate regions.

Keywords: Thermal comfort, dynamic model, cooling demand.

“Implementación de Medidas para Mejorar la Eficiencia Energética y los Indicadores Claves de Rendimiento (ICRs) y de Medio ambiente (IIAs), en la Producción de Vapor en un Caldero Pirotubular de una nueva PyME - Pequeña y Mediana Empresa”

Félix Rustan Roca Subirana1 [1][0009-0001-1302-3311]
[1] Universidad Mayor de San Simón, Campus UMSS s/n, Sucre Street, Edificio Central, Cochabamba, Bolivia
rustanroca.s@fcyt.umss.edu.bo

Abstract

The PT09FC001 Universidad Mayor de San Simón (UMSS) project: “Method for relating key performance indicators (KPIs) with environmental impact indicators (EIIs) to improve energy efficiency in production” of the Universidad Mayor de San Simón as a “base project”, proposed a methodological proposal that combines energy and environmental indicators. The new 2025 management continuity thematic project is called "Implementation of Measures to Improve Energy Efficiency and Key Performance Indicators (KPIs) and Environmental Indicators (EPIs) in Steam Production in a Fire-Tube Boiler of a New Small or Medium Factory (SMF)."

The methodology consists of six indicators for a thermal production process: Lean Energy Indicator (LEI), Equality Indicator (Equality), Energy Usage Indicator (Eusage), Energy Saturation (Esat), Energy Availability (Eavail), and Opening Energy Indicator (Eopening), which identify opportunities for energy savings, production improvements, and investments in a company.

The base project includes a demonstration guide for transferring the methodology used in this new thematic continuity project to an SMF that operates with steam as an energy source instead of biomass as a fuel.

Steam operation will be simulated using the boiler's maximum thermal power as a supply to meet the demand of one of the SMF in the base project, in order to improve the ICRs and IIAs indicators. The boiler, with a steam production capacity of 150 kg/h, is located in a laboratory in the Department of Mechanical Engineering of the Faculty of Science and Technology of the UMSS.

Keywords: Boiler, Indicators, Efficiency

A System Dynamics Model for the Development of Innovation Districts

Work in Progress

Jorge A. Duran-Encalada, Universidad Popular Autónoma del Estado de Puebla
Carina A. Rapetti, La Salle – Ramón Llull University
Josep M. Piqué, La Salle – Ramón Llull University

Abstract

This work presents a System Dynamics (SD) model to examine the process of developing an Innovation District both structurally and empirically. Authors based the model on indicators of the four knowledge-based urban development, KBUD: Urban, Economic, Social, and Governance (Yigitcanlar, Adu-McVie, & Erol, 2020). The study associates the role of the Triple Helix actors, Government, Industry, and Academia (Etzkowitz & Leydesdorff, 2000) with the activation of those indicators along the different phases of evolution and development of an industrial district: inception, launching, growth, and maturity (Pique, Miralles, Teixeira, Gaspar, & Filho, 2019). Currently, after discussing the structural adequacy of the model, we are working on the empirical validation of the results using two real-world cases: the @22 Barcelona and Porto Digital districts.

Comparative Analysis of Power Demand Models for Electric Traction Systems in Colombian Public Transport

Harrynson Ramírez-Murillo[1], Juan Sebastián Martínez Donoso[1], Carlos Andrés Torres-Pinzón[2,3], Oscar Danilo Montoya[3], and Jesús C. Hernández[4]

[1] Faculty of Engineering, Universidad de La Salle, Bogotá D.C., Colombia haramirez@unisalle.edu.co

[2] Laboratorio de Investigación de Fuentes Alternativas de Energía (LIFAE), Facultad de Ingeniería, Universidad Distrital Fracisco José de Caldas, Bogotá D.C., 110121, Colombia caatorresp@udistrital.edu.co

[3] Grupo de Compatibilidad e Interferencia Electromagnética (GCEM), Facultad de Ingeniería, Universidad Distrital Fracisco José de Caldas, Bogotá D.C., 110121, Colombia odmontoyag@udistrital.edu.co

[4] Department of Electrical Engineering, University of Jaén, 23071 Jaén, Spain, jcasa@ujaen.es

Abstract This research presents a nonlinear control design for a Buck DC-DC converter that powers an unknown linear load. The proposed controller utilizes the backstepping approach, which ensures asymptotic stability in the sense of Lyapunov during closed-loop operation. The backstepping control law developed for the Buck converter is general and adaptable to a wide range of operating conditions. To address the uncertainty in the linear load, an immersion and invariance approach is employed, enhancing the robustness of the control strategy in the face of load variations. Numerical simulations, along with comparisons against the extended feedback linearization control method, demonstrate the superior performance of the proposed controller design, as evidenced by an evaluation of the integral time square error, integral absolute error, and integral time absolute error indices. All numerical simulations were conducted in the PLECs simulation tool of the MATLAB/Simulink environment.

Keywords: Electric vehicles · Energy consumption · Monte Carlo methods · Probability density function · Uncertainty.

Towards the Adaptability of Information Visualization for Smart City Technologies

Irma E. Romero-Arrioja[1*], Luis G. Montané-Jiménez[2] and Luis A. Rojano-Ruiz[3]

Facultad de Estadística e Informática, Universidad Veracruzana, Av. Xalapa S/N, Xalapa, 91020, Veracruz, México.

*Corresponding author(s). E-mail(s): iromero@uv.mx;
Contributing authors: lmontane@uv.mx;
zS24019583@estudiantes.uv.mx;

Abstract

Information visualization is fundamental to Smart City management, as it allows various stakeholders to understand the large volumes of data generated in these environments. Key factors involved in the visualization process in smart cities include urban component management, data collection, technological interoperability, information presentation, and urban collaboration. These factors are interrelated and present different levels of complexity and diversity, which should be considered to generate effective visualizations. The effectiveness of visualization comprehension depends not only on the functionalities of the visualization systems but also on user characteristics and the context in which these systems are utilized. This dependence poses a significant challenge, as these elements must be adequately considered to make visualizations effective and accessible to the various stakeholders involved. Through an exploration of literature review, we identify different frameworks, methodologies, models, and techniques to visualize information in Smart Cities using technologies and approaches such as machine learning techniques, IoT, Big Data, Fog Computing, Cloud Computing, Interactive dashboards and maps, virtual reality, blockchain, and 3D technology, these studies focus on the data to determine the visualization technique to use. We find that there is limited research on designing visualizations that consider user diversity and different contexts. We also emphasize the importance of incorporating adaptability into Smart City visualization tools, which combine both adaptable and adaptive elements. This enables the generation of visualizations tailored to users' characteristics and allows them to modify visualizations according to their needs and preferences. While some studies have explored user-centered design to tailor visualizations to users' preferences (adaptable), further research is needed

to integrate approaches, methodologies, and technologies that enable automatic personalization based on user characteristics and context (adaptive).

Keywords: smart cities; visualization; adaptability; adaptive systems

Desafíos metodológicos en la recolección de datos de movilidad urbana basados en encuestas con monitoreo personal

Cinthya Lady Butron Revilla[1[1,2]] and Jesus Santiago Ramos[2[1]]

[1] Universidad Pablo de Olavide Sevilla
[2] Universidad Nacional de San Agustin
cbutronr@unsa.edu.pe
jsanram@upo.es

Resumen.

Los desafíos metodológicos en la recolección de datos de movilidad urbana mediante encuestas con monitoreo personal constituyen una problemática crítica para el diseño de políticas de transporte basadas en evidencia, especialmente en contextos latinoamericanos marcados por desigualdades estructurales y heterogeneidad socioeconómica. Este estudio analiza las principales barreras metodológicas en la aplicación de encuestas con seguimiento personal en comunidades universitarias, a través de un estudio comparativo en dos universidades peruanas ubicadas en Arequipa y Lima. La metodología, de enfoque cualitativo, se basó en la realización de un focus group para interpretar la percepción de los participantes sobre el uso de dos aplicaciones móviles vinculadas a una encuesta de movilidad: el aplicativo especializado Smart City Arequipa y la plataforma de acceso libre Wikiloc. Los resultados evidencian una tensión metodológica entre el rigor científico y la accesibilidad social: mientras Smart City destaca por su control de datos y personalización contextual, presenta limitaciones en usabilidad y participación; Wikiloc, en cambio, muestra mayor aceptación social y colaboración comunitaria, aunque con menores garantías de control de calidad. El análisis con diagramas de Sankey reveló diferencias significativas en la percepción de los usuarios, subrayando la dicotomía entre herramientas académicas especializadas y plataformas abiertas. Se concluye que ninguna aplicación satisface plenamente los requerimientos metodológicos, pero la integración híbrida de ambos enfoques surge como una alternativa prometedora para equilibrar precisión científica y compromiso social en estudios de movilidad urbana longitudinales.

Palabras Clave: Tecnologias móviles, movilidad urbana, seguimiento personal.

The role of third-party logistic providers in Brazilian e-waste collection: A Case Study with Repense

Gabriela Rodrigues[1], Thiago Teixeira and Vanessa de Almeida Guimarães [0000-0001-7662-3499]

[1] COPPEAD Graduate School of Business, Federal University of Rio de Janeiro, Brazil
gabi.zuria@gmail.com
thiagoswt1@gmail.com
vanessa.guimaraes@coppead.ufrj.br

Abstract. Waste management is an important issue in circular and smart cities. This study examines the challenges and opportunities in Brazil's e-waste reverse logistics sector, focusing on the role of third-party logistics providers in supporting circular economy strategies. Through a case study of Repense, a Brazilian company specialized in e-waste management, the research explores how innovation trends and recent regulatory developments influence operational practices, stakeholder engagement, and environmental outcomes. Findings highlight persistent barriers, high logistical costs, strong bureaucracy and lack of interest of some stakeholders, alongside emerging opportunities driven by technological innovation, consumer education, and new policy instruments. The analysis underlines the importance of coordinated public-private efforts, collaboration with stakeholders and the integration of informal actors to advance scalable and sustainable e-waste management systems in Brazil.

Keywords: e-waste, circular and smart cities, third party logistics providers



CITIES

CIUDADES INTELIGENTES TOTALMENTE INTEGRALES, EFICIENTES Y SOSTENIBLES

ISBN 978-84-09-81656-9