

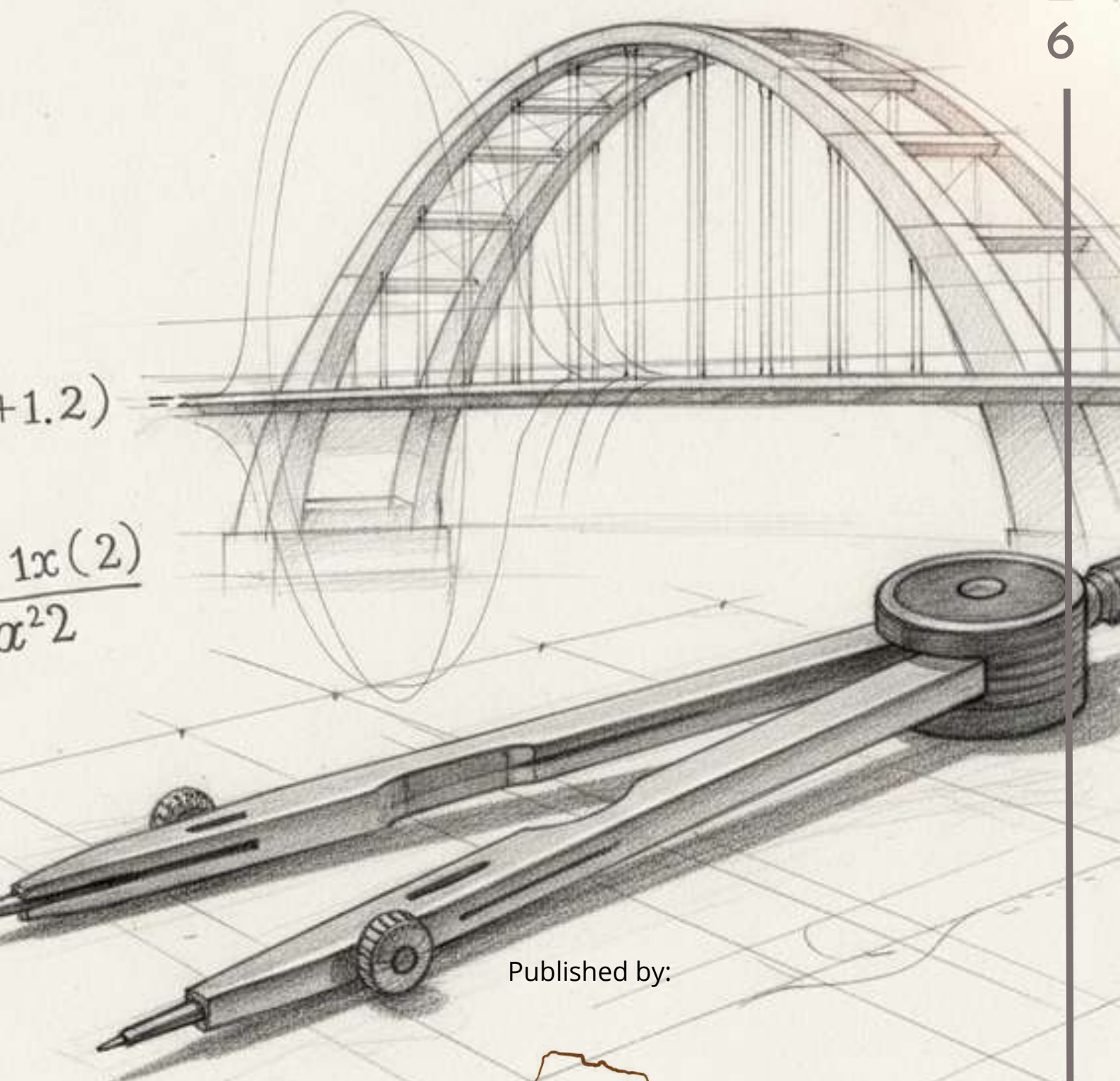
Atheneea

ISSN: 2737-6439, DOI: 10.47460/atheneea

Volumen 7, Número 23

Revista en Ciencias de la Ingeniería

2
0
2
6



Published by:


AutanaBooks

Revista Athenea

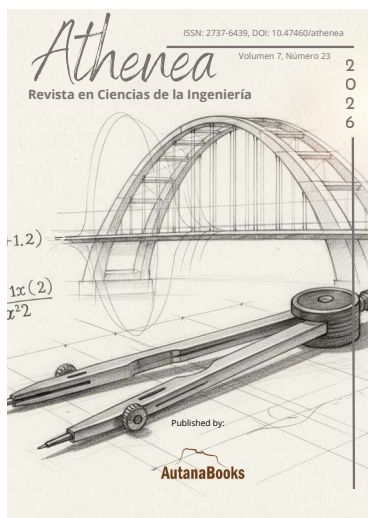
Revista electrónica editada por AutanaBooks

Publicación continua

Volumen 7 | enero-marzo, 2026

DOI: 10.47460/athenea **ISSN-e:** 2737-6419

Nuestra Portada



Font: canva.com

licence: 03422-17578080

Vol. 7 // Núm. 23

enero-marzo

DOI: 10.47460/athenea

ISSN: 2737-6419

En la portada de nuestra querida Athenea, busca transmitir el rigor técnico e innovación científica que asociamos con el desarrollo en ingeniería. El puente, los trazos geométricos, las fórmulas y el compás evocan precisión, cálculo y construcción de conocimiento, reforzando la identidad académica de una publicación vinculada a la ingeniería y al desarrollo tecnológico.

Portal de la revista:

<https://athenea.autanabooks.com/index.php/revista>

Producción y visibilidad

Editor de Estilo

Dra. Franyelit Suárez-Carreño.

Editor de Producción, indexación y métricas

Ing. Ángel Lezama.

Editor de Idioma

Fausto Bartolotta

Adrián Hauser

Nota editorial

Los artículos y las opiniones manifestadas por nuestros colaboradores no necesariamente reflejan la línea editorial o institucional de AutanaBooks y pueden ser reproducidos con autorización previa del Editor. En este caso, favor citar la fuente y enviar copias del medio utilizado a AutanaBooks, Sector Mitad del Mundo, Quito, Ecuador.

Comité Editorial

Entidad Editora

Editorial: AutanaBooks.

Venezuela: Municipio Libertador, Parroquia Santa Teresa, Av. Lecuna. Cipreses a Hoyo. Edif. Berret, Piso 8, Oficina 8-B, Caracas, Código postal: 1012, Venezuela.

Correo de contacto: editorial@autanabooks.com

Directorio de Athenea Revista en Ciencias de la Ingeniería

Consejo Editorial

Ph.D. Franyelit Suárez. AutanaBooks, Quito, Ecuador.

editorial@autanabooks.com

0000-0002-8763-5513

Comité Académico

Dr. José García-Arroyo. Universidad Nacional de Educación a Distancia (UNED), España.

jagarcia@uees.edu.ec

0000-0001-9905-1374

Dr. Oscar Dam. Universidad Experimental Politécnica UNEXPO, Puerto Ordaz, Venezuela.

oscar.curmetals1@gmail.com

0000-0002-0594-6757

Dr. Valentina Millano. Centro de Estudios de Corrosión (CEC), Universidad del Zulia, Maracaibo, Venezuela.

millanov@fing.luz.edu.ve, millanov@gmail.com

0000-0001-6138-4747

Dr. Edwin Flórez Gómez. Universidad de Puerto Rico en Mayagüez, Mayagüez, Puerto Rico.

edwin.florez@upr.edu

0000-0003-4142-3985

Dr. Ana Hilda Márquez. Universidad Metropolitana de Quito, Quito, Ecuador.

amarquez@umet.edu.ec

0000-0002-7958-420X

Dr. Diana Cristina Morales Urrutia. Universidad Técnica de Ambato, Ambato, Ecuador.

dc.moralesu@uta.edu.ec

0000-0002-9693-3192

Dr. Yelka Martina López Cuadra. Universidad Nacional Intercultural Fabiola Salazar Leguía de Bagua, Bagua, Perú.

ylopez@unibagua.edu.pe

0000-0002-3522-0658

Dra. Irela Perez Magin. Universidad Politécnica de Puerto Rico, San Juan, Puerto Rico.

iperezmagin@pupr.edu

PhD. Alejandro Suarez-Alvites. Consultor, Lima, Perú.

alejandrosualvites@hotmail.com

0000-0002-9397-057X

Dr. Neris Ortega. Universidad Metropolitana de Quito, Quito, Ecuador.

nortega@umet.edu.ec

0000-0001-5643-5925

Dr. Juan Carlos Alvarado Ibáñez. Universidad Nacional Intercultural Fabiola Salazar Leguía de Bagua, Bagua, Perú.

jalvarado@unibagua.edu.pe

0000-0002-6413-3457

Mgt. Juan Segura. Universidad Tecnológica Indoamérica, Quito, Ecuador.

juansegura@uti.edu.ec

0000-0002-0625-0719

Dr. Jairo José Rondón Contreras. Instituto Tecnológico de Santo Domingo, República Dominicana.

rondonjjx@gmail.com / jairo.rondon@intec.edu.do

0000-0002-9738-966X

Dr. Hernán Mauricio Quisimalin Santamaria. Universidad Técnica de Ambato, Ambato, Ecuador.

hernanmquisimalin@uta.edu.ec

0000-8491-8326

Dr. Ángel González Lizardo. Polytechnic University of Puerto Rico, San Juan, Puerto Rico.

agonzalez@pupr.edu

0000-0002-0722-1426

Dr. Wilfredo Fariñas Coronado. Polytechnic University of Puerto Rico, San Juan, Puerto Rico.

wfarinascoronado@pupr.edu

0000-0003-2095-5755

Dr. Carlos Alberto Gómez Cano. Corporación Unificada Nacional de Educación Superior – CUN, Florencia, Caquetá, Colombia.

carlos_gomezca@cun.edu.co, carlosgomez325@gmail.com

0000-0003-0425-7201

Dr. Luis Concepción Atoche Alcas. I.E. N° 14100, Paita, Perú.

luisatochealcas16@gmail.com

0000-0003-1454-2129

Dr. Orlando Rafael Gil Rubio. Universidad Católica Andrés Bello, sede Ciudad Guayana, Estado Bolívar, Venezuela.

orgil@unexpo.edu.ve; orgil@ucab.edu.ve; ogil07@gmail.com.ve

0009-0005-0964-7112

Dr. Jesús Ramón López Hercules. Universidad Experimental Politécnica UNEXPO, Puerto Ordaz, Venezuela.

Jlopezz@unexpo.edu.ve

0009-0006-4577-6728

Dra. Lourdes Margarita Meza Ruiz. Universidad Internacional Mesoamericana, Buenos Aires, Argentina.

0000-0002-3333-7051

Dr. José Calizaya-López. Universidad Nacional de San Agustín de Arequipa, Arequipa, Perú.

0000-0001-6221-0909

Dr. Ferdinand Eddington Ceballos Bejarano. Universidad Nacional de San Agustín de Arequipa, Arequipa, Perú.

0000-0003-2867-2397

Mgt. Alfredo Ruitval Velazco Gonzales. Universidad Nacional de San Agustín de Arequipa, Arequipa, Perú.

0000-0001-9358-3425

Mgt. Luz Gabriela Cuba Pacheco. Universidad Nacional de San Agustín de Arequipa, Arequipa, Perú.

0000-0003-3984-3737

Contenido (Volumen 7 // Número 23)

- 8 Ruben Martinez Cabrera, Julio Cesar Minga, María Mercedes Paricoto Simón. *Impacto de plantas de tratamiento de aguas residuales: una evaluación desde la ingeniería ambiental.*
- 18 María Eugenia Ramos Flores, Wilian Bravo, Maritza Tatiana Chaglla Cango, Gabriela Belén Ureña Callay, Maribel Lema. *Materiales carbonosos y nanomateriales de residuos de cacao para remoción de metales pesados.*
- 28 Alfredo Alejandro Marot Guevara, Sergio Rafael Velásquez Guzmán. *Comparación de la efectividad entre la Transformada Wavelet y la Transformada de Wigner-Ville para el diagnóstico de bajo aislamiento en el transitorio de arranque de motores de inducción.*
- 41 José Calizaya López, Benjamín Roldan Polo-Escobar, Orealis Maria Aguilar Paredes, Edgar Hilario Barreda Coaquira, Ariosto Carita Choquecahua. *Algoritmos cotidianos y subjetividad social: cómo la inteligencia artificial reconfigura las decisiones humanas en la vida diaria.*
- 54 Benjamín Roldan Polo-Escobar, Renzo Enrique Polo-Moreano, Lilian Marisol Estrada Torres, Rodolfo Martín Cornejo Urbina. *Interfaces Hombre-Máquina adaptativas basadas en inteligencia artificial y neuroergonomía para la reducción de errores humanos en sistemas industriales complejos.*
- 66 Eduardo Pozo Valdiviezo, Sandra Elizabeth Tenelanda Cudco, Martha Ximena Dávalos Villegas, Gustavo Javier Ávila Gaibor. *Modelos explicables para identificar dificultades en matemáticas universitarias: análisis del dataset MathE.*

Editorial

En las ciencias de la ingeniería, la generación de conocimiento exige precisión, método y capacidad de modelar la realidad con base en principios científicos sólidos. En este escenario, la integración de la inteligencia artificial en los procesos educativos y de investigación no constituye un fin en sí mismo, sino una herramienta estratégica para fortalecer el análisis, la simulación y la toma de decisiones fundamentadas en evidencia.

La inteligencia artificial, aplicada con rigor científico, permite abordar sistemas complejos, optimizar procesos y explorar soluciones que amplían las fronteras tradicionales de la ingeniería. Cuando se articula con una formación académica sólida, orientada al pensamiento crítico y al dominio conceptual, contribuye a una educación en ingeniería más analítica, predictiva y orientada a la resolución de problemas reales.

Desde la perspectiva educativa, esta convergencia potencia metodologías basadas en datos, simulación computacional y validación experimental, sin desplazar el rol central del ingeniero como profesional crítico, ético y responsable. La ciencia define el método, la inteligencia artificial amplía las capacidades de análisis y la educación en ingeniería integra ambos elementos para formar profesionales capaces de innovar con criterio y solidez técnica.

Esta revista reafirma su compromiso con la difusión de investigaciones que integren ciencia, inteligencia artificial y educación en ingeniería desde una perspectiva rigurosa, ética y de alta calidad académica, convencida de que el avance tecnológico sostenible se construye sobre fundamentos científicos firmes y una formación académica de excelencia.

Dra. Franyelit Suárez-Carreño



Artículo de investigación

<https://doi.org/10.47460/athenea.v7i23.122>

Impacto de plantas de tratamiento de aguas residuales: una evaluación desde la ingeniería ambiental

Ruben Martinez Cabrera*
<https://orcid.org/0000-0002-4561-8627>
rmartinezc@unfv.edu.pe
Universidad Nacional Federico Villarreal
Lima, Perú

Julio Cesar Minga
<https://orcid.org/0000-0002-2880-4954>
jcesar@unfv.edu.pe
Universidad Nacional Federico Villarreal
Lima, Perú

María Mercedes Paricoto Simón
<https://orcid.org/0000-0002-7675-7558>
mparicoto@unfv.edu.pe
Universidad Nacional Federico Villarreal
Lima, Perú

*Autor de correspondencia: rmartinezc@unfv.edu.pe

Recibido (23/09/2025), Aceptado (05/12/2025)

Resumen. El estudio analizó el desempeño ambiental y la eficiencia operativa de plantas de tratamiento de aguas residuales desde un enfoque de ingeniería ambiental, con el objetivo de identificar los impactos asociados a su funcionamiento y nivel tecnológico. La metodología incluyó la observación directa de procesos unitarios y el análisis sistematizado de información operativa, aplicando criterios técnicos de evaluación ambiental y de remoción de contaminantes. Los resultados evidenciaron diferencias significativas entre las plantas, destacándose un sistema con eficiencia de remoción de sólidos suspendidos inferior al 65 %, asociado a un alto impacto ambiental, frente a otros con eficiencias superiores al 90 % y efectos de magnitud moderada. Se concluye que el desempeño ambiental depende del nivel tecnológico, del control de emisiones y de la gestión integral de residuos.

Palabras clave: contaminación ambiental, tratamiento de afluentes, impacto ambiental, plantas de tratamiento.

Impact of Wastewater Treatment Plants: An Assessment from Environmental Engineering

Abstract. The study analyzed the environmental performance and operational efficiency of wastewater treatment plants from an environmental engineering perspective, aiming to identify the impacts associated with their functioning and technological level. The methodology included direct observation of unit processes and a systematized analysis of operational information, applying technical criteria for environmental assessment and contaminant removal. The results showed significant differences among the plants, highlighting one system with suspended solids removal efficiency below 65 %, associated with high environmental impact, in contrast to others with efficiencies above 90 % and moderate effects. It is concluded that environmental performance depends on technological level, emission control, and integrated waste management.

Keywords: environmental contamination, wastewater treatment, environmental impact, treatment plants.



I. INTRODUCCIÓN

En las ciudades modernas caracterizadas por alta densidad poblacional, la gestión del saneamiento urbano constituye un componente crítico de la infraestructura ambiental, siendo el tratamiento de aguas residuales una de sus operaciones fundamentales [1]. Desde la perspectiva de la ingeniería ambiental, estos sistemas no solo cumplen una función sanitaria, sino que actúan como elementos clave para la protección de los ecosistemas acuáticos, la sostenibilidad de los recursos hídricos y la reducción de impactos ambientales asociados al vertimiento de efluentes sin tratamiento adecuado. En este contexto, la modernización de las plantas de tratamiento de aguas residuales se ha consolidado como una prioridad a escala global, impulsada por el incremento sostenido de la demanda de agua limpia y la necesidad de mitigar los efectos ambientales derivados de los procesos urbanos e industriales [2].

De manera similar, en los Estados Unidos de América, se han implementado sistemas que integran herramientas de inteligencia artificial, permitiendo optimizar el monitoreo en tiempo real, la gestión operativa y la precisión en el control de la carga contaminante [3]. En América Latina, diversos estudios han documentado avances recientes en la reutilización de aguas residuales tratadas y en la adopción de enfoques de gestión hídrica circular en sectores industriales y agrícolas [4]. Sin embargo, dicha tendencia no se ha generalizado en toda la región. Investigaciones realizadas en contextos urbanos latinoamericanos evidencian que el vertimiento de aguas residuales y la gestión inadecuada de subproductos del tratamiento continúan generando impactos ambientales significativos sobre las fuentes hídricas y los ecosistemas asociados [5], [6].

En el caso del Perú, la situación presenta similitudes con otros países de la región en términos de brechas entre la capacidad instalada y la demanda real de tratamiento. Estudios técnicos sobre plantas de tratamiento municipales en contextos latinoamericanos han señalado deficiencias recurrentes en el diseño, la operación y el mantenimiento de estas infraestructuras, así como heterogeneidad en los modelos de gestión y control [7]. Estas limitaciones comprometen la eficiencia de los procesos de depuración y la calidad del efluente final, intensificando la contaminación ambiental, afectando la biodiversidad acuática e incrementando el riesgo de enfermedades en la población, además de deteriorar la calidad de vida de las comunidades aledañas debido a la emisión de olores desagradables [8].

En síntesis, la evidencia sugiere que la modernización tecnológica, el fortalecimiento de la gestión operativa y la incorporación de enfoques basados en sostenibilidad constituyen factores determinantes para mejorar el desempeño ambiental de las plantas de tratamiento de aguas residuales en contextos urbanos y regionales.

II. MARCO TEÓRICO

Una planta de tratamiento de aguas residuales constituye una infraestructura esencial de la ingeniería ambiental, diseñada para remover contaminantes físicos, químicos y biológicos presentes en las aguas generadas por actividades domésticas, urbanas e industriales, con el objetivo de producir un efluente compatible con la protección del medio ambiente y la salud pública [9]. Desde el enfoque ingenieril, estas instalaciones representan sistemas complejos en los que convergen procesos hidráulicos, fisicoquímicos y biológicos, cuya correcta integración determina la eficiencia global del tratamiento y la sostenibilidad del sistema.

Los procesos de tratamiento comprenden, en primer lugar, operaciones físicas como la sedimentación y la filtración, orientadas a la separación de sólidos suspendidos y materiales flotantes. A ello se suman procesos químicos, entre los que destacan la coagulación, la floculación y la desinfección, empleados para la remoción de partículas finas, patógenos y compuestos disueltos. Finalmente, los procesos biológicos, como los sistemas de lodos activados y los biodigestores, permiten la degradación de la materia orgánica biodegradable mediante la acción controlada de microorganismos. En términos operativos, el tratamiento de aguas residuales se estructura generalmente en tres etapas. El tratamiento primario se enfoca en la remoción de sólidos sedimentables y flotantes; el tratamiento secundario reduce la carga orgánica a través de procesos biológicos; mientras que el tratamiento terciario tiene como finalidad la eliminación de contaminantes específicos, como nutrientes y microorganismos, mejorando la calidad del efluente y ampliando sus posibilidades de reutilización [10].

Desde la perspectiva del impacto ambiental, numerosos estudios han analizado las consecuencias del vertido directo de aguas residuales sin tratamiento adecuado sobre las cuencas hidrográficas, destacando

el uso de indicadores ambientales físicos, químicos y biológicos como herramientas fundamentales para el monitoreo de los efectos sobre los ecosistemas acuáticos, el paisaje natural y la salud humana [11]. Estos impactos incluyen procesos de eutrofización, degradación de la calidad del agua, alteración de hábitats y riesgos sanitarios asociados a la exposición a contaminantes.

No obstante, la literatura también reconoce que las plantas de tratamiento de aguas residuales pueden generar impactos ambientales propios cuando su diseño, construcción u operación no se desarrollan bajo criterios técnicos adecuados. En este sentido, se han documentado modificaciones en la fertilidad de los suelos y alteraciones en el entorno inmediato asociadas tanto a la fase constructiva como a la operatividad de estas instalaciones, especialmente en relación con la gestión de lodos, emisiones odoríferas y uso del suelo [12]. Estos antecedentes refuerzan la necesidad de abordar el análisis de las plantas de tratamiento desde un enfoque integral de ingeniería, que considere no solo su eficiencia depuradora, sino también los impactos ambientales derivados de su ciclo de vida.

III. METODOLOGÍA

La investigación se desarrolló a partir del análisis comparativo de cuatro plantas de tratamiento de aguas residuales, seleccionadas en función de las tecnologías empleadas y su potencial incidencia en la minimización de impactos ambientales durante la etapa de operación y mantenimiento. Las plantas evaluadas incorporan distintos sistemas de tratamiento, entre ellos lagunas aireadas, filtros percoladores, lodos activados con aireación extendida y reactores biológicos de lecho móvil, lo que permitió contrastar el desempeño ambiental asociado a enfoques tecnológicos diferenciados. La evaluación de impactos se realizó mediante la aplicación de la matriz propuesta por Conesa, debido a su enfoque multicriterio, carácter semicuantitativo y capacidad para integrar una visión holística de los efectos ambientales derivados de infraestructuras de saneamiento.

Para la identificación de impactos ambientales se delimitaron las áreas de influencia directa e indirecta en el entorno físico de cada planta, considerando variables como zonas de amortiguamiento, presencia de viviendas, disponibilidad de servicios básicos e infraestructura vial. Se sistematizó la información relacionada con los procesos operativos desarrollados en cada instalación, analizando la forma en que las distintas actividades influyen sobre los componentes ambientales. En esta etapa se priorizaron aquellas actividades con mayor nivel de incidencia y significancia ambiental, particularmente durante las fases de operación y mantenimiento.

Del total de instalaciones existentes en el área de estudio, la selección de las cuatro plantas respondió a un criterio deliberado basado en la diversidad tecnológica, con el objetivo de evaluar y comparar los impactos ambientales generados por cada sistema de tratamiento. Esta selección permitió analizar de manera representativa las diferencias en eficiencia operativa y presión ambiental asociadas a distintos esquemas de diseño y gestión.

La recolección de información se realizó principalmente mediante observación directa, empleando fichas de registro diseñadas para documentar las condiciones operativas, los procesos de tratamiento y las interacciones con el entorno. La información fue complementada, en casos puntuales, con datos proporcionados por las entidades responsables de la gestión de las plantas y por pobladores ubicados en las áreas circundantes, a través de consultas directas orientadas a identificar percepciones sobre impactos ambientales, especialmente aquellos relacionados con olores y molestias. Adicionalmente, los datos sobre velocidad y dirección del viento fueron obtenidos de registros meteorológicos oficiales, con el fin de apoyar el análisis de dispersión de emisiones.

La cuantificación y análisis de los impactos ambientales se efectuaron a partir de las actividades identificadas, agrupadas en dos categorías principales: tecnología de tratamiento y operación del sistema. Para cada actividad se evaluaron los aspectos e impactos ambientales conforme a los criterios establecidos en la guía metodológica para la evaluación de impacto ambiental, considerando variables como efecto, periodicidad, signo, intensidad, extensión, momento, reversibilidad, persistencia, recuperabilidad, acumulación y sinergia. La importancia de los impactos se determinó mediante la aplicación de la ecuación de ponderación multicriterio (1), la cual permitió obtener valores acumulados para su posterior clasificación en niveles de impacto insignificante, moderado, severo y crítico, facilitando así la comparación integral del desempeño ambiental de las plantas analizadas.

La importancia de los impactos ambientales se determinó mediante la aplicación de una ecuación

de ponderación multicriterio, expresada de la siguiente forma:

$$I = (3IN + 2EX + MO + PE + RV + SI + AC + EF + PR + MC) \quad (1)$$

donde: *IN* = Intensidad; *EX* = Extensión; *MO* = Momento; *PE* = Persistencia; *RV* = Reversibilidad; *SI* = Sinergia; *AC* = Acumulación; *EF* = Efecto; *PR* = Periodicidad; y *MC* = Recuperabilidad.

Mediante esta ecuación se obtuvieron puntajes acumulados que permitieron clasificar los impactos conforme al siguiente criterio: *Insignificante* (-1 a -25), correspondiente a impactos sin importancia que no alteran el ambiente; *Moderado* (-26 a -50), para impactos que no generan afectación apreciable sobre el ambiente; *Severo* (-51 a -75), aplicable a impactos que degradan o afectan de manera considerable el ambiente; y *Crítico* (-76 a -100), para impactos que requieren atención inmediata debido al alto riesgo ambiental y sanitario que representan.

IV. RESULTADOS

El análisis de las características generales y de la capacidad operativa de las plantas de tratamiento de aguas residuales evidencia marcadas diferencias en términos de escala, cobertura y volumen de procesamiento (Tabla 1). En particular, la planta de San Bartolo presenta una capacidad de tratamiento significativamente superior, siendo aproximadamente 2,3 veces mayor que la de menor tamaño considerada en el estudio, correspondiente a María Reiche. Esta diferencia se explica por la extensión de su área de cobertura, que abarca varios distritos urbanos, en contraste con plantas orientadas principalmente a atender zonas residenciales específicas.

No obstante, una mayor capacidad instalada no se traduce necesariamente en un mejor desempeño ambiental. A pesar de su escala operativa, la planta de San Bartolo registra la menor eficiencia de remoción de contaminantes, con valores cercanos al 63 %, asociados al uso de tecnología de lagunas aireadas. Asimismo, incumple los criterios establecidos para la categoría III de los Estándares de Calidad Ambiental aplicables a efluentes tratados, lo que limita su potencial de reutilización, particularmente en actividades agrícolas cercanas.

En contraste, las plantas de Santa Rosa, Manchay y María Reiche presentan eficiencias superiores al 90 %, alcanzando en algunos casos valores cercanos al 98 %, lo que se refleja en el cumplimiento de los estándares ambientales vigentes. Resulta relevante destacar que, a pesar de ser la planta de menor tamaño, María Reiche beneficia a una población considerablemente mayor que Manchay, lo que sugiere diferencias en el origen y tipo de aguas residuales tratadas. Mientras que en zonas residenciales predominan descargas domésticas, en otras áreas se incorporan aportes adicionales provenientes de actividades agrícolas e industriales, lo que incrementa la carga contaminante y condiciona el volumen tratado.

Desde el punto de vista de la ingeniería ambiental, las diferencias observadas en el desempeño de las plantas analizadas se explican principalmente por la configuración tecnológica de sus procesos de tratamiento. La planta de Santa Rosa, basada en filtros percoladores, presenta un esquema de tratamiento convencional que, si bien resulta funcional en términos de remoción de contaminantes, muestra limitaciones asociadas a la generación de olores y a la proliferación potencial de vectores, vinculadas a la gestión de residuos sólidos y a las condiciones operativas del sistema.

Por el contrario, la planta de Manchay incorpora un sistema de lodos activados con aireación extendida por ciclos intermitentes, lo que implica un mayor consumo energético, pero permite un control más eficiente de la carga orgánica y una reducción significativa de impactos asociados a olores y vectores. De manera similar, la planta María Reiche dispone de un esquema tecnológico más avanzado, caracterizado por una separación eficiente de sólidos en la etapa de pretratamiento y la integración de un reactor biológico de lecho móvil con un clarificador de placas. Esta configuración optimiza la retención de sólidos y la degradación biológica de contaminantes, lo que se traduce en una mayor eficiencia global del sistema y menores impactos ambientales durante la operación.

Tabla 1. Características generales y capacidad operativa de las plantas de tratamiento de aguas residuales.

Característica	San Bartolo	Santa Rosa	Manchay	María Reiche
Tecnología de tratamiento	Lagunas aireadas	Filtros percoladores	Lodos activados de aireación extendida	Reactor biológico de lecho móvil
Volumen de agua residual que ingresa (m ³ /año)	62 248 995,00	309 687,00	2 408 905,00	273 750,00
Desagües atendidos	154 830	5 279	14 335	1 002
Cobertura	Ate; Santiago de Surco; La Molina; Villa María del Triunfo; San Juan de Miraflores	Pachacamac	Pachacamac	Miraflores
Población beneficiada	630 161	17 691	56 143	116 526
Volumen de agua tratada (m ³ /año)	32 624 698	298 627	2 323 283	750
Cumplimiento normativo	Incumple categoría III de ECA e incumple LMP para efluentes	Cumple categoría III de ECA	Cumple categoría III de ECA	Cumple categoría III de ECA

Los resultados reflejados en la Tabla 2 confirman que la selección tecnológica y el diseño de los procesos de tratamiento constituyen factores determinantes en el desempeño ambiental de las plantas de tratamiento de aguas residuales, más allá de su tamaño o capacidad instalada.

Tabla 2. Configuración de los procesos de tratamiento y parámetros de diseño–operación.

Detalle técnico	San Bartolo	Santa Rosa	Manchay	María Reiche
Tratamiento preliminar	Rejas automáticas; desarenador automático	Cámara de rejas; tornillo de residuos; desarenador manual	Cuchara bivalva; cámara de rejas gruesas; tornillo mecanizado; desarenador y desengrasador automático	Cámara de rejas; tanque equalizador
Tratamiento primario	Laguna primaria aireada de mezcla completa	Tanque Imhoff	No aplica	Sistema compacto (tamiz rotativo y trampa de grasas)
Tratamiento secundario	Laguna secundaria aireada de mezcla parcial + lagunas de sedimentación	Filtro percolador + sedimentador secundario	Tanque ICEAS (aireación extendida por ciclos intermitentes)	Reactor biológico de lecho móvil + clarificador de placas
Tratamiento terciario	Laguna de pulimento + desinfección por cloración	Cámara de contacto de cloro	Sistema de filtración + aplicación de polímeros + desinfección por cloración	Ultrafiltración + desinfección con hipoclorito de sodio
Parámetros de diseño–operación	Caudal: 1700 L/s; DBO: 250 mg/L; SST: 45 mg/L; aceites y grasas: 4,98 mg/L; coliformes termotolerantes: 1,4×10 ⁵ NMP/100 mL	Caudal: 10 L/s; DBO: 26 mg/L; SST: 17 mg/L; aceites y grasas: 5,8 mg/L; coliformes termotolerantes: 2,0 NMP/100 mL	Caudal: 60 L/s; DBO: 700 mg/L; SST: 10 mg/L; aceites y grasas: 0,75 mg/L; coliformes termotolerantes: 2,0 NMP/100 mL	Caudal: 6,94 L/s; DBO: 280 mg/L; SST: 10 mg/L; aceites y grasas: 0,75 mg/L; coliformes termotolerantes: 2,0 NMP/100 mL

A. Descripción física y problemas percibidos en las zonas de influencia de las PTAR.

En los detalles físicos de las zonas de influencia directa de las cuatro PTAR evaluadas se observó que, en el caso de María Reiche, el 80 % de la población evidenció una condición densamente poblada. Ello responde a que el distrito de Miraflores es eminentemente residencial, en contraste con Pachacamac, donde la concentración poblacional se distribuye en áreas relativamente pobladas, siendo atendido por dos PTAR (Santa Rosa y Manchay). Por otra parte, la planta de mayor escala, San Bartolo, que recibe descargas provenientes de cinco distritos populosos, presenta una zona de amortiguamiento clasificada como *relativamente poblada*, con un 100 % de viviendas de un solo nivel. No obstante, es destacable que la mayor parte de estas zonas cuenta con todos los servicios básicos. Asimismo, todos los distritos involucrados en la jurisdicción de esta planta dispusieron como mínimo de servicio de agua y desagüe, permitiendo que la colecta de aguas residuales se realizara en su totalidad.

Dentro de los resultados también se identificaron los problemas potenciales de cada una de las plantas en relación con el área de influencia directa. En este sentido, la emisión de gases de efecto invernadero fue evaluada mediante percepción poblacional, determinándose que todas las PTAR presentaron dicho indicador. Si bien la planta de Santa Rosa reportó un nivel leve, su incidencia no puede ser desestimada debido a su efecto secundario sobre el cambio climático. Adicionalmente, los malos olores fueron reportados en las plantas de San Bartolo y Santa Rosa, pero no en Manchay ni en María Reiche. Este patrón sugiere que, en la medida en que la población aledaña presenta un carácter más residencial, la emisión perceptible de olores disminuye. Aunque la cobertura espacial de dichos olores se distribuye siguiendo la dirección del viento (Sur–Suroeste), es relevante que también sean percibidos en otros distritos, particularmente en el caso de San Bartolo y Santa Rosa (Tabla 3).

Tabla 3. Problemas potenciales evidenciados en las PTAR evaluadas.

Problema potencial	San Bartolo	Santa Rosa	Manchay	María Reiche
Emisión de CO ₂ y otros gases	Sí	Sí (leve)	Sí	Sí
Malos olores	Sí	Sí (persistentes)	No	No
Vectores (insectos)	No (presencia)	Sí	No	No
Efecto derivado	Contribuye al cambio climático y salud humana	Contribuye al cambio climático y salud humana	Contribución al cambio climático y salud humana debe ser evaluada	Contribución al cambio climático y salud humana debe ser evaluada
Velocidad del viento (m/s)	3,6	3,6–5,7	5,7–8,8	3,6–5,7
Dirección del viento	Sur–Suroeste	Suroeste	Suroeste	Suroeste
Población afectada	Villa María del Triunfo	Carabayllo	Villa María del Triunfo	Miraflores
Impacto sobre el suelo	Sí (infiltración de contaminantes)	Sí (lixiviados percolados)	Sí (infiltración de contaminantes)	Sí (lixiviados percolados)
Cultivos agrícolas colindantes	Sí	No	No	No
Disposición final del efluente tratado	Río Lurín + Reúso	Reúso	Río Lurín + Reúso	Parques y jardines (radio ~ 2 km)

En la etapa de campo se consultó a los pobladores del área de influencia directa (AID) de cada una de las PTAR. En ese sentido, las plantas de San Bartolo y María Reiche fueron calificadas por un número importante de encuestados con un alto nivel de olores desagradables, mientras que la PTAR Manchay fue mayormente asociada a un nivel medio. No obstante, la mayoría de los encuestados del AID de San Bartolo la señaló con un nivel bajo de olores, mientras que en Santa Rosa el 70 % de los participantes indicó la presencia de malos olores.

En relación con la percepción de emisiones gaseosas, la mayor parte de los encuestados clasificó su presencia en niveles medio o bajo; sin embargo, se registró una proporción relevante que la calificó como alta en el caso de María Reiche y San Bartolo. Por otra parte, respecto a la percepción de contaminación de suelos, se observó consenso alrededor de los niveles bajo o sin contaminación, lo que responde a que esta evaluación provino exclusivamente de percepciones y no de mediciones técnicas (Tabla 4).

Tabla 4. Percepción de los pobladores del área de influencia directa respecto a las PTAR evaluadas.

Percepción	Categoría	San Bartolo (%)	Santa Rosa (%)	Manchay (%)	María Reiche (%)
Malos olores (AID)	Alto nivel	33,3	10,0	16,7	30,0
	Nivel medio	0,0	20,0	58,3	35,0
	Nivel bajo	66,7	0,0	16,7	15,0
	Sin olores	0,0	70,0	8,3	20,0
Percepción de gases (AID)	Alta	33,3	10,0	0,0	25,0
	Media	0,0	10,0	45,5	15,0
	Baja	66,7	10,0	27,3	40,0
	Sin percepción	0,0	70,0	27,3	20,0
Contaminación de suelos (AID)	Alta	0,0	0,0	0,0	10,0
	Moderada	0,0	0,0	0,0	25,0
	Baja	66,7	0,0	9,1	50,0
	Sin contaminación	33,3	100,0	90,9	15,0

B. Evaluación de los impactos ambientales de las PTAR

Para la evaluación de los impactos ambientales de las plantas se empleó la metodología propuesta por Conesa [13], lo que permitió cuantificar los indicadores estipulados para la etapa de operación y mantenimiento. De ese modo, para la planta San Bartolo, la calificación promediada entre los tres indicadores clave (Calidad del aire, salud humana, Calidad de vida de las comunidades cercanas y Calidad del suelo, entorno natural) fue de 51 puntos, que equivale a un nivel “Alto” impacto. En contraste, para Santa Rosa, el promedio obtenido de sus tres indicadores claves arrojó 45,7 puntos, que equivale a un nivel “Moderado” impacto. Para Manchay y María Reiche solo se pudo cuantificar un indicador clave, dando promedios equivalentes a “moderado” nivel de impacto. Este resultado no resulta tan positivo, pues, un nivel moderado es causal de alerta, sobre todo, considerando que las áreas de influencia directa tienen potencial efecto en la salud de las personas y, por otra parte, condiciona los niveles de desarrollo en las actividades que cotidianamente los ciudadanos realizan en esos territorios.

C. Propuesta de alternativas de mitigación de los impactos ambientales de las PTAR

De la evaluación de los impactos ambientales inherentes a las PTAR, se definió las opciones de mejora que puedan solucionar o mitigar estos impactos. De ese modo, la propuesta de empleo de energías renovables fue común tanto para San Bartolo como Manchay, en donde se busca la reducción de los costos producto del proceso de aireación en la planta. El empleo de tecnologías de tratamiento adicional fue común para todas las plantas evaluadas, siendo implícita la propuesta de mejorar de la tecnología de control automatizado para el ajuste de tiempo en San Bartolo; mejora de los sistemas de nanofiltración y de ósmosis inversa en Santa Rosa y técnica de monitoreo de emisión de gases malolientes en el tanque equalizador de María Reiche.

D. Discusión

Los resultados obtenidos en la evaluación comparativa de las plantas de tratamiento de aguas residuales confirman que el desempeño ambiental de estas infraestructuras no depende exclusivamente de su capacidad instalada, sino de la integración adecuada entre tecnología de tratamiento, condiciones operativas y gestión ambiental. Este hallazgo es coherente con los enfoques contemporáneos de la ingeniería ambiental, que conciben el tratamiento de aguas residuales como un sistema complejo orientado no solo a la remoción de contaminantes, sino también a la sostenibilidad de los recursos hídricos y la minimización de impactos secundarios [1], [2].

A nivel internacional, diversos estudios han demostrado que la modernización tecnológica y la incorporación de sistemas avanzados de monitoreo permiten mejorar de manera significativa la eficiencia de los procesos de tratamiento y el control de contaminantes [1], [2]. En este sentido, los resultados observados en las plantas que incorporan tecnologías biológicas más avanzadas son consistentes con la literatura, la cual destaca que los sistemas de mayor complejidad técnica tienden a presentar mejores indicadores de desempeño ambiental, aunque a costa de mayores requerimientos energéticos y de operación [6], [7]. Esta relación entre eficiencia y demanda operativa subraya la necesidad de evaluar las plantas de tratamiento desde un enfoque costo-beneficio integral, especialmente en contextos urbanos de países en desarrollo.

En el contexto latinoamericano, los hallazgos del presente estudio se alinean con investigaciones previas que evidencian avances heterogéneos en la gestión y reutilización de aguas residuales tratadas [4], [7]. Si bien existen experiencias exitosas orientadas a la economía circular y al reaprovechamiento del recurso hídrico, estas prácticas aún no se encuentran generalizadas, y persisten brechas significativas asociadas a limitaciones técnicas, institucionales y financieras. Los impactos ambientales identificados, particularmente aquellos clasificados como moderados y severos, reflejan problemáticas estructurales similares a las reportadas en estudios urbanos realizados en Colombia y otros países de la región, donde el vertimiento de aguas residuales y la gestión inadecuada de subproductos del tratamiento continúan afectando la calidad de los cuerpos de agua y los ecosistemas asociados [5], [11].

Asimismo, los resultados ponen en evidencia que incluso las plantas de tratamiento pueden generar impactos ambientales propios cuando no se aplican criterios rigurosos de diseño, operación y mantenimiento. Este aspecto ha sido ampliamente documentado en evaluaciones de impacto ambiental asociadas tanto a la construcción como a la fase operativa de estas infraestructuras, especialmente en relación con la gestión de lodos, la emisión de olores y la alteración del entorno inmediato [8], [12]. La percepción social negativa registrada en áreas de influencia de algunas plantas evaluadas refuerza la necesidad de incorporar variables sociales y territoriales en el análisis del desempeño ambiental, más allá de los indicadores técnicos tradicionales.

Desde el punto de vista metodológico, la aplicación de un enfoque multicriterio para la identificación y valoración de impactos permitió integrar de manera sistemática aspectos físicos, biológicos y sociales, facilitando una interpretación más completa de los resultados. En este sentido, el uso de la metodología propuesta por Conesa resulta pertinente, ya que proporciona una estructura robusta para jerarquizar impactos, identificar efectos acumulativos y orientar la toma de decisiones en proyectos de infraestructura ambiental [13]. La incorporación de este enfoque fortalece la validez del análisis y permite comparar los resultados obtenidos con evaluaciones desarrolladas en otros contextos geográficos y tecnológicos.

Los resultados confirman que la mejora del desempeño ambiental de las plantas de tratamiento de aguas residuales requiere una visión integral de ingeniería, en la que la selección tecnológica, la gestión operativa y la evaluación sistemática de impactos ambientales se articulen como componentes inseparables. Este enfoque resulta fundamental para avanzar hacia sistemas de saneamiento más eficientes, sostenibles y socialmente aceptables, en línea con los desafíos actuales de la gestión hídrica urbana.

CONCLUSIONES

Las plantas de tratamiento de aguas residuales evaluadas evidenciaron diferencias significativas en términos de eficiencia operativa, magnitud de los impactos ambientales generados en sus áreas de influencia y percepción social asociada a su funcionamiento. Si bien, desde el punto de vista técnico, la mayoría de las plantas alcanzó niveles de tratamiento compatibles con las especificaciones normativas para la calidad del efluente, los resultados ponen de manifiesto que el cumplimiento de estándares no garantiza, por sí solo, un desempeño ambiental integral ni una adecuada aceptación social.

La cuantificación de impactos ambientales mostró que el 73 % de los criterios evaluados se clasificó en un nivel de impacto moderado, mientras que los impactos insignificantes y severos representaron el 11,5 % cada uno, y los impactos críticos alcanzaron el 4 %. No obstante, la intensidad y relevancia de estos impactos variaron en función de las características sociales y territoriales de las áreas de influencia de cada planta, evidenciando la necesidad de incorporar variables contextuales en los procesos de evaluación ambiental. Desde la perspectiva de la población aledaña, el principal impacto negativo identificado estuvo asociado al deterioro de la calidad de vida, particularmente en el entorno de la planta de San Bartolo, donde la emisión de olores constituye una fuente recurrente de molestia. En este sentido, se recomienda la implementación de sistemas de control de emisiones odoríferas, sustentados en mediciones sistemáticas que permitan establecer líneas base y vincular de manera objetiva la emisión con el tipo de fuente y las condiciones operativas.

Los resultados confirman que la incorporación de tecnologías de tratamiento más avanzadas, el fortalecimiento de los mecanismos de control de emisiones y una gestión más eficiente de los residuos generados en las plantas de tratamiento pueden contribuir de manera significativa a la mitigación de impactos ambientales y a la mejora del desempeño global de estos sistemas. No obstante, la adopción de estas medidas debe estar precedida por evaluaciones técnicas y económicas rigurosas que orienten la toma de decisiones de inversión por parte de las autoridades competentes, garantizando soluciones sostenibles, eficientes y socialmente aceptables a largo plazo.

REFERENCIAS

- [1] V. Novotny, "Sustainable urban water management," in *Water and Urban Development Paradigms: Towards an Integration of Engineering, Design and Management Approaches*, J. Feyen, K. Shannon, and M. Neville, Eds. Boca Raton, FL, USA: CRC Press, 2008, pp. 37–50, doi: 10.1201/9780203884102-9.
- [2] P. Kokkinos, J. R. Comia Jr., S. Caucci, H. Hettiarachchi, F. C. Ballesteros, G. Oron, M. Salgot, and I. K. Kalavrouziotis, "Wastewater and sludge reuse: Selected case studies across the globe," *Desalination and Water Treatment*, vol. 250, pp. 65–79, 2022, doi: 10.5004/dwt.2022.28183.
- [3] A. O. Baarimah, M. A. Bazel, W. S. Alaloul, M. Y. D. Alazaiza, T. M. Al-Zghoul, B. Almuhaya, A. Khan, and A. W. Mushtaha, "Artificial intelligence in wastewater treatment: Research trends and future perspectives through bibliometric analysis," *Case Studies in Chemical and Environmental Engineering*, vol. 10, p. 100926, 2024, doi: 10.1016/j.cscee.2024.100926.
- [4] M. Félix-López, M. Ormazá-Murillo, C. Álvarez-Santana, and C. Banchon, "Exploring the impact of reclaimed water on Latin America's development," *Journal of Ecological Engineering*, vol. 24, no. 10, pp. 157–173, 2023, doi: 10.12911/22998993/169962.
- [5] F. D. Cuadros Segura, A. J. Moyano Salcedo, A. M. Pabon Laverde, and J. V. Trujillo Arias, "Impacto ambiental del vertimiento de aguas servidas en aglomerados urbanos ilegales del municipio de Villavicencio, Colombia," *Tecnura*, vol. 25, no. 68, pp. 43–62, 2021, doi: 10.14483/22487638.16273.
- [6] J. M. Morgan-Sagastume and M. T. Orta-Ledesma, "Tratamiento de aguas residuales municipales: ¿son los sistemas UASB + lagunas con microalgas una opción?" *Tecnología y Ciencias del Agua*, vol. 16, no. 3, pp. xx–xx, 2025, doi: 10.24850/j-tyca-2025-03-10.
- [7] A. Cáñez-Cota, "Municipal wastewater treatment plants in Mexico: Diagnosis and public policy challenges," *Tecnología y Ciencias del Agua*, vol. 13, no. 1, pp. 184–245, 2022, doi: 10.24850/j-tyca-2022-01-05.
- [8] J. J. Espinoza Eche and E. G. Santos de la Cruz, "Innovación en la gestión de lodos generados en plantas de tratamiento de aguas residuales de origen doméstico en Lima-Perú," *Revista del Instituto de Investigación de la Facultad de Minas, Metalurgia y Ciencias Geográficas*, vol. 24, no. 48, pp. 205–215, 2021, doi: 10.15381/iigeo.v24i48.21773.
- [9] V. Lahera Ramón, "Infraestructura sustentable: Las plantas de tratamiento de aguas residuales," *Quivera*, vol. 12, no. 2, pp. 58–69, 2010, disponible en: <http://www.redalyc.org/articulo.oa?id=40115676004>.
- [10] A. Cáñez-Cota, "Municipal wastewater treatment plants in Mexico: Diagnosis and public policy challenges," *Tecnología y Ciencias del Agua*, vol. 13, no. 1, pp. 184–245, 2022, doi: 10.24850/j-tyca-2022-01-05.
- [11] F. D. Cuadros Segura, A. J. Moyano Salcedo, A. M. Pabon Laverde, and J. V. Trujillo Arias, "Impacto ambiental del vertimiento de aguas servidas en aglomerados urbanos ilegales del municipio de Villavicencio, Colombia," *Tecnura*, vol. 25, no. 68, pp. 43–62, 2021, doi: 10.14483/22487638.16273.
- [12] F. S. Montero-Vega, C. S. Molina-Cedeño, B. M. Pillco-Herrera, L. B. Sarduy-Pereira, and K. Diéguez-Santana, "Evaluación del impacto ambiental de la construcción de una planta

de tratamiento de aguas residuales: Caso río Pindo Chico, Puyo, Pastaza, Ecuador," *Ciencia, Ambiente y Clima*, vol. 3, no. 1, pp. 23–39, 2020, doi: 10.22206/cac.2020.v3i1.pp23-39.

- [13] V. Conesa-Fdez.-Vítora, *Guía metodológica para la evaluación del impacto ambiental*, 4th ed. Madrid, España: Ediciones Mundi-Prensa, 2011.

AUTORES



Ruben Martinez Cabrera es Doctor en Medio Ambiente y Desarrollo Sostenible; cuenta con una Maestría en Gestión Ambiental y es Ingeniero Geógrafo. Actualmente se desempeña como docente en la Facultad de Ingeniería Geográfica, Ambiental y Ecoturismo de la Universidad Nacional Federico Villarreal.



Julio Cesar Minga es Doctor en Ingeniería Ambiental e Investigación, Modelización y Análisis del Riesgo en Medio Ambiente; cuenta con una Maestría en Administración de Empresas y se desempeña como docente en la Facultad de Ingeniería Geográfica, Ambiental y Ecoturismo de la Universidad Nacional Federico Villarreal.



María Paricoto Simón es Magíster en Ingeniería Sanitaria y de profesión Ingeniera Sanitaria. Actualmente se desempeña como docente en la Facultad de Ingeniería Geográfica, Ambiental y Ecoturismo de la Universidad Nacional Federico Villarreal, además de ejercer labores de asesoramiento y revisión de tesis, así como funciones de proyectista.

Materiales carbonosos y nanomateriales de residuos de cacao para remoción de metales pesados

María Eugenia Ramos Flores*
<https://orcid.org/0009-0004-7985-6019>
mariaeugeniaramosflores@gmail.com
Unidad Educativa Nela Martínez Espinosa
La Troncal, Ecuador

Wilian Bravo
<https://orcid.org/0000-0002-2599-6532>
wilian.bravo@esPOCH.edu.ec
Escuela Superior Politécnica de Chimborazo
Riobamba, Ecuador

Maritza Tatiana Chaglla Cango
<https://orcid.org/0000-0002-5331-4615>
maritza.chaglla@esPOCH.edu.ec
Escuela Superior Politécnica de Chimborazo
Riobamba, Ecuador

Gabriela Belén Ureña Callay
<https://orcid.org/0009-0001-5842-7379>
gabriela.urena@esPOCH.edu.ec
Escuela Superior Politécnica de Chimborazo
Riobamba, Ecuador

Maribel Lema
<https://orcid.org/0009-0000-5767-6798>
fernanda.lema@esPOCH.edu.ec
Escuela Superior Politécnica de Chimborazo
Riobamba, Ecuador

*Autor de correspondencia: mariaeugeniaramosflores@gmail.com

Recibido (29/09/2025), Aceptado (03/12/2025)

Resumen. El agua contaminada con metales pesados representa un riesgo crítico para la salud y los ecosistemas, especialmente en las regiones cacaoteras donde convergen las actividades agroindustriales y mineras. Este trabajo presenta una revisión estructurada sobre el uso de residuos de cacao como plataforma para el desarrollo de biosorbentes, materiales carbonosos y nanomateriales para la remoción de metales en sistemas acuosos. Se realizó una búsqueda en bases de datos científicas y se seleccionaron estudios con una clara caracterización de los materiales y evaluación experimental en fase acuosa. La evidencia se organizó según el tipo de residuo y material, discutiendo su desempeño, ventajas y limitaciones. Además, se identificaron brechas de investigación relacionadas con la escasez de ensayos en matrices del mundo real, información limitada sobre regeneración y vida útil, el tratamiento de Hg(II) y la falta de análisis integrales de sostenibilidad y viabilidad en contextos latinoamericanos.

Palabras clave: residuos de cacao, materiales carbonosos, metales pesados, hidrochar.

Carbonaceous Materials and Cacao Waste Nanomaterials for Heavy Metal Removal

Abstract. The presence of heavy metals in water poses a critical risk to human health and ecosystems, particularly in cocoa-producing regions where agro-industrial and mining activities converge. This work provides a structured review on the use of cocoa residues as a platform for developing biosorbents, carbonaceous materials, and nanomaterials for heavy metal removal in aqueous systems. A search of scientific databases was conducted, and studies were selected based on clear material characterization and experimental evaluation in aqueous phase. The evidence was organized according to residue type and derived material, analyzing performance, advantages, and limitations. Research gaps were identified regarding the scarcity of tests in real-world matrices, limited information on regeneration and service lifetime, challenges in the treatment of Hg(II), and the lack of comprehensive assessments of sustainability and feasibility in Latin American contexts.

Keywords: cocoa residues, carbonaceous materials, heavy metals, hydrochar.

I. INTRODUCCIÓN

El agua dulce constituye un recurso fundamental para el desarrollo humano, económico y ecosistémico, sin embargo, su calidad se deteriora por el vertido de contaminantes inorgánicos persistentes. Entre ellos, los metales pesados preocupan por su toxicidad, bioacumulación y lenta degradación. Elementos como el mercurio, plomo, cadmio y cobre se han asociado con trastornos neurológicos, renales y cardiovasculares, además de deteriorar la vida acuática y los servicios ecosistémicos que dependen de las aguas superficiales y subterráneas [1].

En América Latina, actividades como la minería, la industria química y la agricultura intensiva, ejercen una fuerte presión sobre los recursos hídricos y favorecen la liberación de metales pesados a ríos y acuíferos [2]. Aunque existen normas que establecen los niveles máximos permisibles, su implementación efectiva se ve limitada por problemas de infraestructura y a la falta de capacidad técnica, especialmente en zonas rurales [3].

Las tecnologías convencionales de remoción de metales, como la precipitación química, el intercambio iónico y varios procesos de membrana, han demostrado su eficacia en condiciones controladas. No obstante, suelen requerir altos costos de inversión y operación, generación de lodos y personal especializado, limitando su uso sostenido en contextos con restricciones económicas. Estas limitaciones han impulsado la búsqueda de adsorbentes de bajo costo, disponibles localmente y con buena eficiencia de remoción [4].

En este marco, los materiales carbonosos biogénicos, producidos a partir de residuos agroindustriales, se han convertido en una opción prometedora para el tratamiento del agua, por su bajo costo y compatibilidad con esquemas de economía circular [5], [6]. Entre ellos, los derivados del cacao (cáscaras, restos de mazorcas y tallos) destacan en los países productores, por el alto volumen de subproductos y por su composición lignocelulósica rica en grupos funcionales que pueden participar en la adsorción y formación de complejos con metales pesados. Estudios pioneros demostraron la capacidad de las cáscaras de cacao para fijar plomo y otros metales en soluciones ácidas, en tanto que trabajos posteriores exploraron su modificación química y activación para mejorar su rendimiento como biosorbentes [7].

Al mismo tiempo, el desarrollo de nanomateriales a partir de biomasa, en particular materiales magnéticos modificados con óxidos de hierro, ha creado nuevas posibilidades para diseñar adsorbentes más selectivos y fáciles de recuperar. La combinación de matrices carbonosas de residuos agroindustriales con fases magnéticas genera materiales híbridos con mayor área superficial y diversidad de grupos funcionales, además de permitir su separación mediante campos magnéticos [8]. En este contexto, estudios recientes han evaluado hidrochars y materiales magnéticos obtenidos de residuos de cacao para remover metales pesados en soluciones acuosas, con altas eficiencias de remoción y mejoras en parámetros de calidad de agua como pH, conductividad y turbidez [9].

No obstante, la información sobre materiales carbonosos y nanomateriales derivados de residuos de cacao sigue fragmentada. Las diferencias en condiciones de síntesis, especies metálicas estudiadas y escala experimental dificultan la comparación de su rendimiento y persisten interrogantes sobre su comportamiento en matrices reales, capacidad de regeneración y viabilidad económica y ambiental, particularmente en contextos latinoamericanos donde hay muchos residuos de cacao y problemas de contaminación.

Ante este escenario, es necesario sintetizar y analizar críticamente la evidencia sobre el uso de residuos de cacao como plataforma para desarrollar materiales carbonosos y nanomateriales destinados a la remoción de metales pesados en agua. El objetivo de esta revisión fue exponer el estado del arte de dichos materiales, rutas de síntesis, características fisicoquímicas relevantes, mecanismos de interacción con especies metálicas y desempeño en distintas condiciones experimentales, así como discutir los avances en América Latina, las oportunidades en el marco de la economía circular y los desafíos para su aplicación efectiva.

II. MARCO TEÓRICO

A. Materiales carbonosos en el tratamiento de aguas

El empleo de materiales carbonosos obtenidos de residuos agroindustriales se ha consolidado como una alternativa para tratar aguas contaminadas con metales pesados, por su bajo costo, disponibilidad local y compatibilidad con esquemas de economía circular [1]. Entre ellos, biochars, hidrochars y carbones activados comparten una matriz rica en carbono, grupos funcionales oxigenados y estructuras porosas que interactúan con iones metálicos. La conversión de la biomasa mediante pirólisis, carbonización hidrotermal o activación física o química controla el carbono fijo, la distribución de poros y la densidad de sitios activos, y con ello la capacidad de adsorción, la cinética de remoción y la estabilidad frente a variaciones de pH y fuerza iónica [2].

La valorización de residuos agrícolas como precursores de estos materiales ofrece ventajas ambientales y económicas, al reducir los residuos que requieren gestión y generar productos con valor agregado para soluciones de tratamiento de aguas que aprovechan recursos locales, especialmente en regiones donde coexisten actividades agroindustriales y problemas de contaminación por metales [7].

B. Residuos de cacao como materia prima lignocelulósica

En la producción de cacao se generan varias fracciones residuales sólidas, principalmente cáscara de mazorca, cascarilla del grano y restos de vaina y tallos. Estas exhiben una estructura lignocelulósica típica, con contenidos significativos de celulosa, hemicelulosa, lignina y minerales como potasio, calcio y magnesio [9], lo que favorece su conversión en materiales carbonosos porosos con grupos superficiales idóneos para procesos de intercambio iónico y formación de complejos con iones metálicos.

Estudios preliminares han mostrado que la cáscara de cacao puede actuar como biosorbente de metales en soluciones acuosas, incluso tras procesos simples de lavado, secado y molienda [10]. A partir de estos resultados, las investigaciones han incorporado tratamientos químicos suaves y procesos termoquímicos orientados a incrementar la capacidad de adsorción y optimizar la estabilidad de los materiales en condiciones de operación más exigentes [11].

Desde una perspectiva regional, el aprovechamiento de residuos de cacao como materia prima para adsorbentes se alinea con la integración de soluciones de tratamiento de aguas en cadenas productivas existentes, lo que resulta especialmente pertinente en países latinoamericanos productores de cacao con problemas de contaminación por metales en cuerpos de agua en zonas agrícolas y mineras [9].

C. Materiales carbonosos derivados de residuos de cacao

La literatura identifica diversas categorías de materiales carbonosos obtenidos a partir de residuos de cacao: biosorbentes crudos o ligeramente modificados, biochars, carbones activados e hidrochars. Los biosorbentes crudos se elaboran a partir de cáscaras o cascarilla mediante lavado, secado y molienda, en ocasiones con ajustes de pH o tratamientos químicos simples. Su principal ventaja es el bajo costo de preparación, lo que los hace atractivos en contextos con recursos limitados [4], aunque su área superficial y densidad de sitios activos suelen ser menores que las de materiales más procesados. En consecuencia, se han planteado modificaciones simples como la oxidación controlada y la incorporación de nuevos grupos funcionales para incrementar su afinidad por metales específicos [6].

El biochar se obtiene por pirólisis de residuos de cacao en atmósfera limitada en oxígeno y a temperaturas intermedias o altas, lo que concentra el carbono, estabiliza la matriz y genera una red de poros que favorece la adsorción de metales [8]. La temperatura de pirólisis, el tiempo de residencia y el pretratamiento de la biomasa condicionan el balance entre grupos oxigenados y estructuras aromáticas más complejas y porosas [12]. Cuando el biochar se somete a activación física o química, se obtienen carbones activados con áreas superficiales muy elevadas y distribución de micro y mesoporos, que incrementan la capacidad de adsorción [11]. Con residuos de cacao, se han reportado carbones activados con un rendimiento notable, aunque su producción conlleva un consumo energético y de reactivos más elevado.

La carbonización hidrotermal en medio acuoso y a temperaturas relativamente bajas permite obtener hidrochars de residuos de cacao con altos rendimientos de carbono y superficies ricas en grupos funcionales polares. A diferencia del biochar, el hidrochar preserva más oxígeno superficial y suele presentar

morfologías más compactas con mayor porosidad en el rango meso y macroporoso [6]. Estos materiales logran eficiencias significativas de eliminación de metales en medios ácidos y modifican parámetros de calidad del agua como pH y conductividad [9], además de poder emplearse directamente como adsorbentes o como precursores en etapas de activación o incorporación de fases magnéticas.

D. Nanomateriales y materiales magnéticos derivados de residuos de cacao

La integración de fases magnéticas en matrices carbonosas de origen lignocelulósico ha emergido como una estrategia para facilitar la separación y regeneración de adsorbentes [8]. En este enfoque, la biomasa de cacao se combina con precursores de óxidos de hierro y se somete a condiciones de síntesis donde la formación de nanopartículas magnéticas ocurre simultáneamente a la carbonización, produciendo materiales híbridos que aprovechan la sinergia entre la porosidad y los grupos funcionales de la fracción carbonosa y las propiedades magnéticas de las fases inorgánicas. Este fenómeno facilita la recuperación de los sólidos mediante campos magnéticos externos y reduce la necesidad de etapas de filtración [8].

E. Mecanismos de interacción con metales pesados

Los mecanismos mediante los cuales los materiales carbonosos y los nanomateriales obtenidos de residuos de cacao remueven metales pesados en solución responden a la combinación de propiedades texturales y químicas de la superficie. Los estudios de caracterización en biomásas similares revelan la presencia de grupos hidroxilo, carboxilo, carbonilo y éter, además de superficies rugosas con poros que favorecen la retención de iones metálicos [12], [13].

En biosorbentes crudos y biochars, la remoción se explica por procesos de intercambio iónico entre los iones metálicos en solución y cationes presentes en la matriz (Ca^{2+} , K^+ , Mg^{2+}), complementados por la formación de complejos superficiales con grupos oxigenados y, en ciertas condiciones, por microprecipitación de especies metálicas en la superficie [14].

En los carbones activados, la gran área superficial y la abundancia de microporos favorecen la adsorción física y la formación de capas monomoleculares, mientras que los grupos superficiales determinan la selectividad hacia determinadas especies metálicas [11]. En los materiales magnéticos se han descrito interacciones específicas entre los metales y las fases de óxidos de hierro, que pueden involucrar formación de complejos y, en algunos casos, reacciones redox [7].

Los estudios experimentales con residuos de cacao han evidenciado que las variaciones en el pH y la conductividad eléctrica durante las pruebas de adsorción están íntimamente asociadas con estos mecanismos, reflejando el intercambio de especies en solución y la incorporación de metales en la matriz sólida [9].

III. METODOLOGÍA

Este estudio se realizó como una revisión estructurada de materiales carbonosos y nanomateriales derivados de residuos de cacao para la remoción de metales pesados del agua. La primera fase definió la pregunta central y el alcance del trabajo, centrándose en investigaciones que utilizaron residuos de cacao como materia prima y evaluaron su rendimiento en sistemas acuosos con metales pesados. Con base en esta delimitación, se establecieron las variables de interés, siendo éstas, el tipo de residuos de cacao utilizado, el proceso de producción del material (biosorbente crudo, biocarbón, hidrocarbón, carbón activado o material magnético), las especies metálicas estudiadas, las condiciones de operación y los principales resultados de remoción.

En la segunda fase, se realizó una búsqueda de información en bases de datos científicas y repositorios institucionales. La búsqueda se ejecutó en *Scopus*, *Web of Science*, *SciELO* y *Google Académico*, empleando combinaciones de palabras clave relacionadas con residuos de cacao, biochar, hydrochar, materiales magnéticos, metales pesados y tratamiento de aguas. La búsqueda se limitó a artículos originales, capítulos de libros, tesis y revisiones relevantes publicadas en las últimas dos décadas, priorizando trabajos que presentaran resultados experimentales en la fase acuosa. En total se identificaron 35 registros.

Luego de eliminar duplicados y documentos fuera del tema según el título y el resumen, se mantuvieron 25 documentos, cuyos textos completos fueron evaluados. En esta etapa se aplicaron criterios de calidad basados en la pertenencia a revistas arbitradas o tesis de posgrado, en la descripción explícita de la obtención y caracterización básica del material y, en el reporte de resultados cuantitativos de remoción o capacidad de adsorción en un medio acuoso. Fueron excluidos 6 documentos que no cumplían simultáneamente estos criterios. Finalmente, 19 documentos formaron parte de la revisión, 11 estudios primarios sobre materiales derivados de residuos de cacao aplicados para remoción de metales pesados en agua, y 8 documentos complementarios empleados como soporte conceptual y de contexto.

Una vez compilado el conjunto final de documentos, se creó una matriz de extracción de información en la que se registraron sistemáticamente los datos relevantes: tipo de residuo de cacao, especies metálicas evaluadas y principales indicadores de rendimiento (porcentaje de remoción, capacidad de adsorción y observaciones sobre cambios en los parámetros de calidad del agua). Con base en esta matriz, los estudios se organizaron en categorías tipológicas (biosorbentes crudos o modificados, biocarbones, carbones activados, hidrocarburos y materiales magnéticos), y se identificaron patrones comunes, diferencias metodológicas y vacíos de información.

IV. RESULTADOS

De acuerdo con la metodología propuesta, se seleccionó un conjunto de estudios primarios sobre el uso de residuos de cacao como materia prima para la obtención de biosorbentes, materiales carbonosos y nanomateriales aplicados a la remoción de metales pesados del agua, priorizando aquellos con caracterización fisicoquímica clara, evaluación en sistemas acuosos sintéticos e información sobre capacidad de adsorción o eficiencia de remoción. En total se identificaron once estudios recientes y representativos, principalmente de Canadá, Perú, Colombia, Ghana, Indonesia y África, resumidos en la Tabla 1, que permiten comparar de manera coherente el desempeño de biomasa cruda, biosorbentes modificados, biochars, carbones activados, materiales híbridos e hidrochars.

Tabla 1. Estudios seleccionados sobre materiales derivados de residuos de cacao para remoción de metales pesados en agua.

Estudio	Residuo de cacao	Tipo de material	Metales objetivo	Resultados
[10]	Cáscara	Biomasa cruda	10 cationes metálicos	$q_{\text{máx}}$ para Pb $\approx$ 6,2 mg/g; rápida remoción (< 2 h).
[15]	Pericarpio	Biosorbente crudo	Cr(VI)	$q_{\text{máx}} \approx$ 48,5 mg/g; isoterma Langmuir.
[16]	Cáscara	Biosorbente crudo	Pb(II), Cd(II)	$q_{\text{máx}}$ 200–210 mg/g; efecto del tamaño de partícula.
[17]	Cáscara	Biosorbente en columna	Pb(II), Cd(II)	$\sim$ 90% de remoción en sistemas monocomponente.
[12]	Cáscara de mazorca	Biochar (500 °C)	Pb(II), Hg(II), Cd(II)	Remoción $\sim$ 97–99,9% para Pb, Hg y Cd.
[14]	Pod husk	Biochar (microondas)	Cd(II)	Remoción de decenas mg/g; buena estabilidad.
[11]	Pod husk	Carbón activado (ZnCl ₂)	As(V)	BET $\sim$ 780 m ² /g; remoción As(V) $\sim$ 80% en < 1 h.
[13]	Cáscara	Biochar (300–380 °C)	Hg(II)	Capacidad 0,106 mg/g a 90 min.
[18]	Cáscara	Carbón dopado con quitosano	Cr(VI), Pb(II)	Supera al carbón activado comercial.
[19]	Cáscara	Biomaterial modificado	Cr(VI), NO ₃ ⁻	Mayor afinidad por Cr(VI) y nitrato.
[9]	Residuos de proceso	Hidrochar (HTC baja T)	Hg(II)	Remoción en medio ácido comparable a biochars T alta.

A. Biosorbentes lignocelulósicos de residuos de cacao

Los primeros hallazgos se basan en el uso directo de la cáscara o el pericarpio de cacao como biosorbentes sin transformación térmica intensiva. Meunier y su equipo [10] evaluaron cáscara de cacao para diez cationes metálicos en soluciones ácidas ($\text{pH}=2$), observando alta afinidad por Pb y tiempos de equilibrio inferiores a 2 h. Estos resultados demostraron tempranamente que la biomasa de cacao compete con otros residuos agrícolas como adsorbente de bajo costo.

En un estudio más reciente [15], se evaluó la biosorción de Cr(VI) en el pericarpio de cacao, con una capacidad máxima cercana a $48,5 \text{ mg/g}$ y buen ajuste a la isoterma de Langmuir y a un modelo cinético de pseudo segundo orden, sugiriendo un mecanismo dominado por la quimisorción. Junto con [10], estos estudios muestran que, incluso sin carbonización, la matriz lignocelulósica del cacao ofrece sitios activos capaces de retener metales.

Estudios como el de Tejada-Tovar, López-Cantillo, Vidales-Hernández, Villabona-Ortiz y Acevedo-Correa [16] confirmaron que la disminución del tamaño de partícula incrementa significativamente la remoción debido al aumento de la superficie específica y la accesibilidad a los grupos funcionales superficiales. En cambio, Vera-Cabezas y sus colaboradores [17] transfirieron la biosorción de Cd(II) y Pb(II) a columnas de lecho fijo, logrando remociones cercanas al 90% en sistemas de uno y dos componentes y demostrando la viabilidad de operar en flujo continuo.

Comparativamente, estos resultados indican que la biomasa cruda de cacao puede alcanzar altas capacidades de remoción (hasta 200 mg/g) en condiciones controladas para ciertos metales (Pb, Cd), pero su rendimiento es sensible al pH, al tamaño de partícula y a la configuración del sistema (lote o columna). Estos biosorbentes constituyen la base frente a la cual se compararon más adelante los materiales carbonosos y nanoestructurados.

B. Biochars y carbones activados

La segunda categoría comprende materiales carbonosos obtenidos mediante tratamientos térmicos más intensivos. En el extremo menos complejo se encuentran los biocarbones de pirólisis, mientras que en el extremo más sofisticado se encuentran los carbones activados con áreas superficiales elevadas.

En el estudio de Abbey, Duwiejua y Quianoo [12], se empleó un biochar de cáscaras de mazorca de cacao (pirólisis a 500°C) para eliminar Pb(II) , Hg(II) y Cd(II) en sistemas monometálicos, logrando remociones prácticamente completas en 60 min. Estos resultados muestran que la pirólisis convierte el residuo de cacao en un adsorbente altamente competitivo, especialmente para el Hg(II) .

Correa-Abril y su equipo [14] obtuvieron un biocarbón por pirólisis asistida por microondas (720 W, 15 min) a partir de cáscara de vaina y, combinando experimentación con cálculos de la teoría del funcional de la densidad (DFT), demostraron que la adsorción de Cd(II) está dominada por la fisiorción, con formación de especies como CdCO_3 en la superficie. Esto ilustra la transición de biochars convencionales a materiales carbonosos respaldados teóricamente a nivel molecular.

Cruz, Pirilä y Huuhtanen [11] produjeron carbón activado con activación química con K_2CO_3 , KOH y ZnCl_2 , con áreas BET cercanas a $780 \text{ m}^2/\text{g}$ y remociones de As(V) de hasta el 80% en menos de una hora a pH casi neutro, mostrando que los residuos de cacao pueden generar materiales altamente competitivos frente a carbones comerciales.

En general, estos trabajos indican que los biocarbones superan a la biomasa cruda en eficiencia y cinética de remoción, y que los carbones activados ofrecen mayores capacidades, mientras que los biocarbones más recientes permiten diseñar arquitecturas superficiales específicas y respaldar los mecanismos de adsorción con herramientas computacionales.

C. Hidrochars y materiales modificados

El hidrochar producido por carbonización hidrotermal de residuos de cacao, como en el estudio de Ormaza-Hugo y su equipo [9], ocupa una posición intermedia entre la biomasa cruda y los biocarbones de alta temperatura. Operando a temperaturas moderadas en medio acuoso ácido, el hidrochar retiene una fracción significativa de grupos oxigenados que favorecen la formación de complejos de Hg(II) y otros cationes, a la vez que reduce el consumo de energía del proceso. En pruebas de adsorción en medio ácido, mostró eficiencias de remoción de Hg(II) superiores al 80%, comparables a las de biocarbones

fabricados a 500 °C, aunque con capacidades inferiores a las de los carbones activados de área superficial muy alta.

Nursiah y sus colaboradores [13] obtuvieron un adsorbente mediante pirólisis de cáscaras de cacao (300–380 °C) para la remoción de Hg(II). La capacidad de adsorción reportada fue de aproximadamente 0,106 mg/g a los 90 min, con una cinética de pseudo segundo orden, luego de la caracterización estructural mediante Microscopía Electrónica de Barrido (MEB), Espectroscopía Infrarroja por Transformada de Fourier (FTIR) y Difracción de Rayos X (DRX). En comparación con este valor, el hidrochar de baja temperatura presenta un rendimiento competitivo, lo que demuestra que las temperaturas muy altas no son esenciales para obtener materiales eficaces para el Hg(II) si se optimizan el pH, la dosis y el tiempo de contacto.

Asimismo, los materiales modificados o nanoestructurados derivados de residuos de cacao representan un paso adicional en la cadena de valor. En el estudio de Okoya, Akinyele, Amuda y Ofoezie [18], se dopó quitosano en carbón de cáscara de cacao, obteniendo un compuesto con mayor capacidad de adsorción de Cr(VI) y Pb(II) que un carbón activado comercial, debido a la introducción de grupos amino y mayor afinidad por especies aniónicas y cationes metálicos. Por su parte, Fotsing y su equipo [19] modificaron la superficie de un biomaterial a base de cáscara de cacao para aumentar la eliminación de Cr(VI) y nitrato, atribuyendo la mejora a la combinación de intercambio iónico, formación de complejos y cambios en la carga superficial.

Desde la perspectiva de esta revisión, el hidrochar de residuos de cacao puede interpretarse como un eslabón intermedio hacia nanomateriales funcionalizados, pero ofrece una plataforma carbonosa susceptible a la magnetización, dopado de polímeros o anclaje de nanopartículas metálicas que aumentan la selectividad hacia Hg(II) y otros metales pesados.

D. Avances y desafíos

Los biosorbentes crudos elaborados a partir de cáscara y pericarpio de cacao han demostrado altas capacidades de remoción en sistemas por lotes optimizados. Además, los biocarbones y carbones activados derivados de residuos de cacao solidifican este residuo como un precursor importante de materiales adsorbentes como los biocarbones que han logrado una eliminación prácticamente completa de Pb(II), Cd(II) y Hg(II), mientras que los carbones activados alcanzan áreas superficiales del orden de 700–800 m²/g y una eliminación significativa de As(V).

Mientras tanto, los materiales modificados o nanoestructurados evidencian que la química superficial es tan importante como la porosidad y que combinar residuos de cacao con biopolímeros u otras fases activas permite superar el rendimiento de los materiales no modificados. Los biochars de cáscara de mazorca y el adsorbente pirolizado de Nursiah et al. [13] aportan datos valiosos, pero pocos estudios exploran sistemáticamente la influencia del pH, la competencia iónica y las matrices reales en la eliminación del Hg(II). El hidrochar de baja temperatura de residuos de cacao llena este vacío, aportando evidencia adicional en condiciones ácidas relevantes para efluentes mineros o industriales.

Muchos estudios se centran en una única vía de transformación como la pirólisis a 500 °C o la activación química con ZnCl₂, lo que dificulta establecer relaciones cuantitativas entre la energía del proceso, la funcionalidad superficial, la estabilidad y el rendimiento de la adsorción. La combinación de resultados de biochars, carbones activados e hidrochars resalta la necesidad de protocolos comparativos estandarizados.

Si bien existen estudios en columnas de lecho fijo para Pb(II), Cd(II) y Ni(II), estos aún son minoritarios en comparación con los ensayos por lotes, y son todavía más escasos cuando se trata de aguas residuales reales o matrices que contienen múltiples metales y aniones en competencia. Ampliar este tipo de evaluaciones es esencial para que los materiales derivados del cacao se incorporen a los sistemas de tratamiento prácticos.

El estudio pionero de Meunier y sus colaboradores [10] marcan el inicio del uso de cáscara de cacao para la remoción de metales en soluciones ácidas, seguido por trabajos sobre carbones activados y materiales modificados en 2012 y 2016 [11], [18]. A partir de 2018, se observa un incremento sostenido de publicaciones, con especial concentración desde 2020 [7], [8], [9], [12], [13], [14], [19]. Esta distribución confirma que la valorización de residuos de cacao como base para materiales carbonosos y nanomateriales orientados a la remoción de metales pesados constituye una línea de investigación emergente consolidada en los últimos cinco años.

Desde una perspectiva latinoamericana, una parte importante de los estudios primarios se realizó en Colombia, Perú y Ecuador [9], [15], [17]. En conjunto, estos trabajos muestran que la región dispone de grandes volúmenes de residuos de cacao y ha generado evidencia experimental propia sobre su uso como biosorbentes, biochars e hidrochars en laboratorio y, en algunos casos, en columnas de lecho fijo. La evaluación del pericarpio de cacao para Cr(VI) en Perú, el modelado de columnas con cáscara de cacao en Colombia y la producción de hidrochar a partir de residuos de procesamiento en Ecuador, ilustran el progreso regional en la valorización de estos residuos para el tratamiento de aguas contaminadas con metales pesados. Estos estudios, destacan una línea emergente de investigación en América Latina, alineada con los desafíos ambientales y productivos de las regiones productoras de cacao.

En el marco de la economía circular, los resultados muestran que los residuos de cacao pueden pasar de pasivo ambiental a insumo para tecnologías de tratamiento de aguas basadas en recursos locales. Integrar la producción de biosorbentes, biochars e hidrochars en plantas de procesamiento abre opciones de tratamiento descentralizado que aprovechan sinergias logísticas y reducen costos de transporte y disposición. Sin embargo, la aplicación en sistemas reales aún enfrenta desafíos poco abordados, como la variabilidad estacional de la biomasa, el escalamiento de la síntesis, la regeneración y gestión del adsorbente saturado, el cumplimiento normativo y la evaluación de la sostenibilidad. Abordar estos aspectos es clave para que las propuestas basadas en residuos de cacao pasen de las pruebas de laboratorio a sistemas operativos en contextos cacaoteros latinoamericanos.

CONCLUSIONES

El análisis de los estudios posiciona los residuos de cacao como una plataforma tecnológica versátil para el tratamiento de aguas contaminadas con metales pesados, con aplicaciones que abarcan desde biosorbentes crudos de muy bajo costo hasta materiales carbonosos avanzados y nanocompuestos. El hidrochar de baja temperatura se perfila como un enlace estratégico entre biosorbentes simples y materiales altamente activados o nanoestructurados, al ofrecer rendimientos competitivos sin requerir procesos térmicos ni etapas de activación extremadamente agresivos, lo que lo hace más viable en contextos con recursos limitados.

La principal contribución de esta revisión es organizar el conocimiento disperso según el tipo de residuo, ruta de transformación, metal objetivo y modo de operación, proporcionando una base práctica para seleccionar estrategias de tratamiento y reducir la fragmentación de la literatura. Desde una perspectiva regional, los resultados refuerzan el potencial de los materiales derivados de cacao como herramientas de economía circular en América Latina, favoreciendo sistemas de tratamiento descentralizados, adaptados a comunidades productoras, cooperativas y pequeñas plantas de procesamiento. Esto representa una oportunidad para conectar las cadenas de valor agrícolas con soluciones ambientales locales.

Finalmente, se identifican vacíos de conocimiento que orientan futuras investigaciones: escasez de estudios en matrices reales, poca evidencia sobre regeneración y vida útil de los adsorbentes, falta de comparaciones directas entre rutas de síntesis y casi nulos análisis de sostenibilidad. Los materiales a base de cacao muestran resultados prometedores, aunque aún parciales, que justifican el desarrollo de estudios comparativos y ensayos piloto.

REFERENCIAS

- [1] V. Singh, N. Singh, S. N. Rai, A. Kumar, A. K. Singh, M. P. Singh, A. Sahoo, S. Shekhar, E. Vamanu, and V. Mishra, "Heavy metal contamination in the aquatic ecosystem: Toxicity and its remediation using eco-friendly approaches," *Toxics*, vol. 11, no. 2, p. 147, feb 2023, doi: 10.3390/TOXICS11020147.
- [2] M. Y. Abubakar, K. B. Ahmad, T. S. Mathew, R. Shamsudden, H. M. Muhammad, M. Haladu, and A. B. Adam, "Heavy metal pollution in aquatic ecosystems: A review of toxic impacts and remediation strategies," *Kwaghe International Journal of Sciences and Technology*, vol. 1, no. 1, pp. 416–427, jul 2024, doi: 10.58578/KIJST.V1I1.3621.
- [3] "Guidelines for drinking-water quality: Fourth edition incorporating the first and second addenda," En línea, 2022, consultado: dic. 12, 2025. Disponible en: <https://pubmed.ncbi.nlm.nih.gov/35417116/>.

- [4] M. Vera, C. Cruzat, and M. E. Vanegas, "Low-cost crop waste biosorbent technology for removing toxics and pollutants from wastewater," in *Agricultural, Forestry and Bioindustry Biotechnology and Biodiscovery*, aug 2020, pp. 177–216, doi: 10.1007/978-3-030-51358-0_11.
- [5] B. B. Vera Raza, R. A. Mero Intriago, G. A. Burgos Briones, and R. E. Cevallos Cedeño, "Lignocellulosic waste and activated carbon production method," *Minerva*, vol. 1, no. Special, pp. 122–130, dec 2022, doi: 10.47460/minerva.v1ispecial.87.
- [6] J. Wang, S. Zhang, C. Wei, H. Hou, G. Song, L. Cao, and J. Zhang, "Hydrothermal carbonization of heavy metal-contaminated biomass: Migration, transformation, and ecological stability changes of metals," *Int J Mol Sci*, vol. 26, no. 6, p. 2551, mar 2025, doi: 10.3390/IJMS26062551/S1.
- [7] J. Coello-Cabezas, M. Verdezoto Carvajal M., N. Mejía cabezas, H. Sánchez-Moreno, E. Basantes Basantes, M. Estrella Semblantes, I. Gavilanez alvares, and R. Ormaza Hugo, "Organic coagulant combined with magnetite nanoparticles for the treatment of mercury-contaminated waters," *Case Studies in Chemical and Environmental Engineering*, vol. 9, p. 100579, jun 2024, doi: 10.1016/J.CSCEE.2023.100579.
- [8] C. Li, C. Zhang, S. Zhong, J. Duan, M. Li, and Y. Shi, "The removal of pollutants from wastewater using magnetic biochar: A scientometric and visualization analysis," *Molecules*, vol. 28, no. 15, p. 5840, aug 2023, doi: 10.3390/MOLECULES28155840.
- [9] R. Ormaza Hugo, J. Naranjo, I. Gavilanez Alvarez, V. M. Cando, K. Tixi Gallegos, H. Sánchez-Moreno, F. Londo, O. D. Gavilanez, and J. Coello-Cabezas, "Production of hydrochar by low-temperature hydrothermal carbonization of residual biomass from cocoa production for mercury adsorption in acidic aqueous solutions," *Case Studies in Chemical and Environmental Engineering*, vol. 10, p. 100938, dec 2024, doi: 10.1016/J.CSCEE.2024.100938.
- [10] N. Meunier, J. Laroulandie, J. F. Blais, and R. D. Tyagi, "Cocoa shells for heavy metal removal from acidic solutions," *Bioresour Technol*, vol. 90, no. 3, pp. 255–263, dec 2003, doi: 10.1016/S0960-8524(03)00129-9.
- [11] G. Cruz, M. Piriälä, and M. Huuhtanen, "Production of activated carbon from cocoa (*Theobroma cacao*) pod husk," *Journal of Civil & Environmental Engineering*, vol. 02, no. 02, 2012, doi: 10.4172/2165-784X.1000109.
- [12] C. Y. B. Abbey, A. B. Duwieujuah, and A. K. Quianoo, "Removal of toxic metals from aqueous phase using cacao pod husk biochar in the era of green chemistry," *Applied Water Science*, vol. 13, no. 2, p. 57, dec 2022, doi: 10.1007/S13201-022-01863-5.
- [13] C. Nursiah, H. Desvita, E. Elviani, N. Farida, A. Muslim, C. M. Rosnelly, M. Mariana, and S. Suhendrayatna, "Adsorbent characterization from cocoa shell pyrolysis (*Theobroma cacao* L) and its application in mercury ion reduction," *Journal of Ecological Engineering*, vol. 24, no. 6, pp. 366–375, 2023, doi: 10.12911/22998993/163167.
- [14] J. Correa-Abril, U. Stahl, E. V. Cabrera, Y. J. Parra, M. A. Vega, S. Taamalli, F. Louis, and J. M. Rodríguez-Díaz, "Adsorption dynamics of $\text{Cd}^{2+}(\text{aq})$ on microwave-synthesized pristine biochar from cocoa pod husk: Green, experimental, and DFT approaches," *iScience*, vol. 27, no. 6, p. 109958, jun 2024, doi: 10.1016/J.ISCI.2024.109958.
- [15] M. C. Fernández Pezua, "Biosorción de cromo hexavalente (Cr VI) de soluciones acuosas utilizando pericarpio de cacao (*Theobroma cacao*)," Master's thesis, Universidad Nacional

Intercultural de la Selva Central Juan Santos Atahualpa, nov 19 2024, consultado: dic. 18, 2025. Disponible en: <http://hdl.handle.net/20.500.14744/27>.

- [16] C. Tejada-Tovar, K. López-Cantillo, K. Vidales-Hernández, A. Villabona-Ortiz, and D. Acevedo-Correa, "Kinetics and bioadsorption equilibrium of lead and cadmium in batch systems with cocoa shell (*Theobroma Cacao* L.)," *Contemporary Engineering Sciences*, vol. 11, no. 23, pp. 1111–1120, 2018, doi: 10.12988/ces.2018.83100.
- [17] L. M. Vera-Cabezas, D. Bermejo-Campos, M. F. Uguña-Rosas, N. García-Alvear, M. Flores-Zamora, and D. Brazales, "Modelado en columna de lecho fijo para la bioadsorción de Cd^{2+} y Pb^{2+} con cáscara de cacao," *Revista internacional de contaminación ambiental*, vol. 34, no. 4, pp. 611–620, 2018, doi: 10.20937/RICA.2018.34.04.05.
- [18] A. Okoya, A. Akinyele, O. Amuda, and I. Ofoezie, "Chitosan-grafted carbon for the sequestration of heavy metals in aqueous solution," *American Chemical Science Journal*, vol. 11, no. 3, pp. 1–14, jan 2016, doi: 10.9734/ACSJ/2016/21813.
- [19] P. N. Fotsing, E. D. Woumfo, S. Mezghich, M. Mignot, N. Mofaddel, F. Le Derf, and J. Vieillard, "Surface modification of biomaterials based on cocoa shell with improved nitrate and Cr(VI) removal," *RSC Adv*, vol. 10, no. 34, pp. 20 009–20 019, may 2020, doi: 10.1039/D0RA03027A.

Comparison of the effectiveness between the Wavelet Transform and the Wigner–Ville Transform for the diagnosis of low insulation in the starting transient of induction motors

Alfredo Marot*

<https://orcid.org/0000-0002-8829-4124>
aamarotgu@estudiante.unexpo.com
UNEXPO Vicerrectorado Puerto Ordaz
Doctorado en Ciencias de la Ingeniería
Puerto Ordaz, Venezuela

Sergio Velásquez

<https://orcid.org/0000-0002-3516-4430>
svelasquez@unexpo.edu.ve
UNEXPO Vicerrectorado Puerto Ordaz
Puerto Ordaz, Venezuela

*Corresponding author: aamarotgu@estudiante.unexpo.com

Received (29/09/2025), Accepted (03/12/2025)

Abstract. This study demonstrates the superiority of the Discrete Wavelet Transform over the Wigner–Ville Transform in detecting insulation faults in induction motors during startup transients. The comprehensive analysis of 360 simulated signals revealed that the wavelet technique with Daubechies 10 achieves significantly higher classification accuracy (74.44% vs. 67.78%), substantially outperforming its counterpart. Certain decomposition levels showed maximum sensitivity with variations up to +354%, while diagnostic reliability indicators confirm its robustness. This technique positions itself as an optimal solution for predictive monitoring systems, enabling early fault detection that substantially reduces downtime and industrial maintenance costs.

Keywords: wavelet transform, Wigner-Ville transform, induction motor, fault diagnosis, insulation failure, signal processing.

Comparación de la efectividad entre la Transformada Wavelet y la Transformada de Wigner–Ville para el diagnóstico de bajo aislamiento en el transitorio de arranque de motores de inducción

Resumen. Este estudio demuestra la superioridad de la Transformada Wavelet Discreta frente a la Transformada de Wigner–Ville en la detección de fallas de aislamiento en motores de inducción durante transitorios de arranque. El análisis exhaustivo de 360 señales simuladas reveló que la técnica wavelet con Daubechies 10 alcanza una precisión de clasificación notablemente superior (74.44% vs. 67.78%), superando significativamente a su contraparte. Ciertos niveles de descomposición mostraron sensibilidad máxima con variaciones de hasta +354%, mientras que los indicadores de confiabilidad diagnóstica confirman su robustez. Esta técnica se posiciona como solución óptima para sistemas de monitoreo predictivo, permitiendo la detección temprana de fallas que reduce sustancialmente los tiempos de inactividad y los costos de mantenimiento industrial.

Palabras clave: transformada wavelet, transformada de Wigner-Ville, motor de inducción, diagnóstico de fallas, falla de aislamiento, procesamiento de señales.

I. INTRODUCTION

Induction motors play a crucial role in the operation of industrial machinery, and their reliable operation is vital for productivity and safety. Faults can lead to downtime, increased maintenance costs, and loss of efficiency [1]. The industry has adopted motor fault detection as a means to ensure the consistent and reliable operation of modern industrial systems. Motors are subjected to various conditions and environments [2]. Stator-related problems account for 38% of all faults in induction motors [3].

When an inter-turn insulation fault occurs, some turns in the stator winding short-circuit. Failure to detect and rectify this fault promptly accelerates its progression to more severe faults, such as inter-coil or inter-phase short circuits [4], culminating in phase-to-ground faults that can result in permanent damage to the induction motor [5]. Recognized challenges include real-time monitoring and model interpretability, emphasizing the need for practical solutions. The severe impact of motor failures on productivity and safety underscores the urgency for early fault detection systems.

Over 40 years of research have demonstrated that many types of faults can be detected in motor sensor data before they cause a breakdown; therefore, there has been significant interest in leveraging this fact to reduce the losses caused by such failures [6]. More than 40 years of research confirm that it is possible to detect these faults through the analysis of sensor data before a catastrophic failure occurs, thereby reducing associated economic losses [7].

In this context, the startup transient emerges as a critical phase where thermal and magnetic stress can reveal insulation weaknesses that are not visible under steady-state conditions. Time-frequency analysis tools, such as the Discrete Wavelet Transform (DWT) and the Wigner–Ville Transform (WVD), have emerged as promising alternatives. However, the need arises to determine which of these techniques offers greater robustness in the face of the non-stationary nature of the startup.

To address this problem, the present research poses the following research questions:

1. Is the Discrete Wavelet Transform more effective than the Wigner–Ville Transform for discriminating incipient insulation fault levels during the startup transient?
2. Which statistical metrics derived from these transformations possess greater separability power to characterize the motor's condition?

To answer these questions, the following specific objectives are proposed:

1. Model and simulate the behavior of induction motors under healthy operating conditions and with incipient insulation faults (200 k Ω and 20 k Ω).
2. Extract a set of statistical features (energy, entropy, RMS, and kurtosis) from the representations generated by the DWT and the WVD.
3. Evaluate and compare the performance of both techniques using a Support Vector Machine (SVM) classifier and separability metrics such as the Fisher Score.

Finally, the success criterion for the comparative diagnosis is defined by obtaining a classification model with an accuracy greater than 70%, an area under the curve (AUC) value greater than 0.80, and the identification of at least three metrics with statistical significance ($p < 0.1$) that allow clear differentiation between the analyzed classes.

II. THEORETICAL BACKGROUND

A. Discrete Wavelet Transform (DWT)

The wavelet transform is a tool applied to analysis in the time and frequency domains. Ingrid Daubechies, a leading figure in wavelet research, revolutionized the field by creating compactly supported orthonormal wavelets, which enabled the development of discrete wavelet analysis. Members of this family are identified by the nomenclature dbN , where db corresponds to the researcher's surname and N indicates the order of the wavelet [8].

The analysis of the motor's transient current is performed in the time-scale domain using the Discrete Wavelet Transform (DWT). The DWT is based on Multiresolution Analysis (MRA), decomposing the signal $x[n]$ into detail coefficients (D_j) and approximation coefficients (A_j) through a dyadic filter bank according to Mallat's pioneering algorithm [9]. Formally, the coefficients at level j are obtained through discrete convolution and downsampling operations, using a low-pass filter $h[n]$ for the approximation and a high-pass filter $g[n]$ for the detail:

$$A_j[k] = \sum_n A_{j-1}[n] \cdot h[2k - n] \quad (1)$$

$$D_j[k] = \sum_n A_{j-1}[n] \cdot g[2k - n] \quad (2)$$

where $A_0[n] = x[n]$ is the input signal. The justification for the DWT lies in its ability to provide a time-frequency representation of the signal, making it optimal for the analysis of transients and non-stationary signals, which are characteristic of insulation faults in electrical machines.

For this study, the Daubechies wavelet of order 10 (db10) is selected, and a decomposition level of 8 ($J = 8$). The db10 wavelet belongs to the family of orthogonal compactly supported wavelets. The selection of $N = 10$ is supported by two fundamental properties. First, its compact support: the filter support is finite (length $2N = 20$ coefficients), which guarantees computational efficiency and good temporal localization for the detection of impulsive events. Second, its vanishing moments ($N = 10$): db10 has ten vanishing moments, meaning it is orthogonal to polynomials up to degree 9. This property is crucial, as it ensures that the detail coefficients (D_j) do not contain information about smooth low-frequency variations (trend), but instead concentrate on the energy of irregularities, peaks, or singularities in the signal [10].

The decomposition level $J = 8$ ensures adequate frequency resolution to isolate the bands where the fault energy manifests. With $f_s = 5000$ Hz, the decomposition generates 8 detail bands, where D_1 corresponds to the highest frequency range (1250–2500 Hz) and D_8 represents the very low-frequency band (9.77–19.53 Hz).

B. Wigner–Ville Transform (WVD)

The Wigner–Ville Transform is the principal member of the class of bilinear (or quadratic) time-frequency distributions. Its main strength is providing optimal resolution simultaneously in the time (t) and frequency (f) domains. The WVD is based on the notion of an instantaneous autocorrelation function of the signal. For a signal $s(t)$, the WVD is generally defined over its analytic signal $z(t)$ to avoid interference caused by negative frequency components [11].

$$W_z(t, f) = \int_{-\infty}^{\infty} z\left(t + \frac{\tau}{2}\right) z^*\left(t - \frac{\tau}{2}\right) e^{-j2\pi f\tau} d\tau \quad (3)$$

where $W_z(t, f)$ is the Wigner–Ville distribution, τ is the integration variable representing the time interval over which the instantaneous autocorrelation is calculated, and $e^{-j2\pi f\tau}$ is the complex exponential factor that converts information from the delay domain (τ) to the frequency domain (f), acting as a band-pass filter for each instant t .

III. METHODOLOGY

This study tests the hypothesis that the Discrete Wavelet Transform (DWT) is more effective than the Wigner–Ville Transform (WVD) for diagnosing stator insulation faults during motor startup. Using numerical simulation, it ensures controlled conditions, safe replication of incipient faults, and generates a labeled dataset for a systematic comparison of the two techniques.

A. Simulation of Motors and Fault Modeling

A simulation environment was implemented in MATLAB/Simulink to replicate the dynamic behavior of 20 squirrel-cage induction motors with rated power distributed in the range of 10 to 4000 HP. This selection ensures industrial representativeness by covering prototypical applications, including pumping, ventilation, and compression systems predominant in manufacturing and continuous process sectors (see Fig. 1).

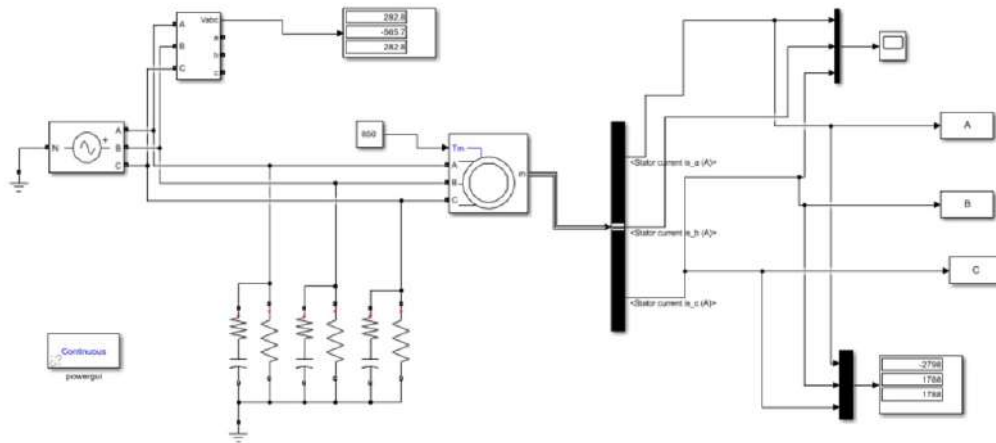


Fig. 1. Squirrel cage motor simulation in Simulink. The figure was generated in MATLAB 2019.

The characteristics (rated power, voltage, resistances, and inductances) corresponding to 20 induction motors were parameterized, ensuring an accurate representation of the three-phase system across various industrial power ranges. The insulation fault was simulated by connecting a resistive-capacitive (RC) circuit between phase and ground of each winding [1]. The resistance of this circuit was adjusted to 200 k Ω and 20 k Ω to represent incipient and critical faults, respectively. The phase-to-ground capacitances were set between 1.5 nF and 21 nF [12].

B. Signal Acquisition and Structuring

A simulated cohort of 20 three-phase squirrel-cage induction motors with rated power from 10 to 4000 HP ensured industrial representativeness for applications such as pumping and ventilation. Data generation followed a factorial design combining three insulation states (healthy, 200 k Ω fault, 20 k Ω fault) and two load regimes (no-load, rated load) across each motor's three phases, yielding a total of 360 three-phase signals (120 healthy, 240 faulty). Acquisition focused on the startup transient, capturing a 5-second window at 5000 Hz to adequately resolve non-stationary components. All signals were exported to MATLAB/Simulink, where time-frequency analysis algorithms were applied for the systematic extraction of diagnostic descriptors.

C. Feature Extraction using DWT and WVD

The processing of the acquired signals employed two feature extraction approaches to transform information from the time domain into a discriminative attribute space. A Daubechies 10 (db10) wavelet with eight decomposition levels was applied, leveraging its compact support and ten vanishing moments for precise transient detection and low-frequency trend filtering [9], [10]. The decomposition level $J = 8$ was determined considering the sampling frequency ($f_s = 5000$ Hz) and the spectral range of interest for insulation faults (approximately 9.77–1250 Hz). From the detail coefficients obtained

at each level D_j (with $j = 1, 2, \dots, 8$) and for each of the three phases, five statistical metrics were systematically calculated.

Energy:

$$E_j = \frac{\sum_{k=1}^{N_j} |D_j[k]|^2}{\sum_{i=1}^8 E_i} \quad (4)$$

where $D_j[k]$ is the k -th detail coefficient at level j and N_j is the number of coefficients at that level.

Maximum Value (MAX_j) and Minimum Value (MIN_j):

$$MAX_j = \max(D_j), \quad MIN_j = \min(D_j) \quad (5)$$

with units expressed in amperes (A).

Shannon Entropy:

$$H_j = - \sum_{k=1}^{N_j} p_k \log_2(p_k), \quad p_k = \frac{|D_j[k]|^2}{E_j} \quad (6)$$

Kurtosis:

$$K_j = \frac{\frac{1}{N_j} \sum_{k=1}^{N_j} (D_j[k] - \bar{D}_j)^4}{\sigma_{D_j}^4} \quad (7)$$

where $\bar{D}_j$ and σ_{D_j} are the mean and standard deviation of D_j , respectively.

Root Mean Square (RMS):

$$RMS_j = \sqrt{\frac{1}{N_j} \sum_{k=1}^{N_j} D_j[k]^2}; \quad \text{Amperes(A)} \quad (8)$$

with units expressed in amperes (A).

The Daubechies 10 wavelet transform was employed to characterize the impulsivity of insulation faults. Concurrently, the Wigner–Ville Distribution (WVD), a quadratic time–frequency representation offering optimal joint resolution, was applied [10]. Following the standard formulation to mitigate cross-term interference, the WVD was computed on the analytic signal of each phase [10]. From the resulting energy distribution, ten spectral attributes were extracted, including frequency moments and power levels at the three most significant peaks [10].

From the processing via WVD, ten metrics per phase were systematically extracted and organized into two fundamental categories. First, four central distribution statistics of the marginal spectral density were calculated, including the mean frequency ($\bar{f}$):

$$\bar{f} = \frac{\sum f_i P(f_i)}{\sum P(f_i)}; \quad (Hz) \quad (9)$$

as the first spectral moment, and the spectral standard deviation (σ_f), which quantifies spectral dispersion:

$$\sigma_f = \sqrt{\frac{\sum (f_i - \bar{f})^2 P(f_i)}{\sum P(f_i)}}; \quad (Hz) \quad (10)$$

both expressed in hertz (Hz).

The analysis quantified the spectral envelope using bounding frequencies ($f_{\min}$, $f_{\max}$), and the three principal peaks characterized by their center frequency (f_k) and normalized dB magnitude (p_k). This compact parameterization captures both global spectral extent and fault-indicative harmonic features from the startup transient.

$$A_{p_k} = 10 \log_{10} \left(\frac{P(f_{p_k})}{\max(P(f))} \right); \quad (db) \quad (11)$$

This normalization allows comparison between signals with different absolute energy.

D. Statistical Analysis and Attribute Selection

A two-stage pipeline was implemented for feature selection. First, univariate separability between Healthy (S) and Fault (F) classes was assessed using the Fisher Score for each feature [13]. For a given feature x , its Fisher Score S_f is calculated as:

$$S_f(x) = \frac{(\mu_s - \mu_f)^2}{\sigma_s^2 + \sigma_f^2} \quad (12)$$

where μ_s , μ_f and σ_s^2 , σ_f^2 are the means and variances of x for the Healthy and Fault classes, respectively. A value of S_f close to zero indicates total overlap between class distributions, while higher values indicate greater separability. An empirical threshold of $S_f > 0.01$ was established to consider a feature as potentially discriminative, thus filtering out attributes with insignificant separation power.

In the second stage, one-way ANOVA (MATLAB's `anova1` function) was applied for each feature, testing the null hypothesis of equal means among the three classes: Healthy, 200 k Ω Fault, and 20 k Ω Fault. Given the exploratory nature of the study and the need to detect subtle effects associated with incipient faults (200 k Ω), a significance level of $\alpha = 0.1$ was adopted, less strict than the conventional $\alpha = 0.05$ [14]. To control the increase in Type I error rate due to multiple comparisons, the Bonferroni correction was applied. The original p -value obtained for each feature was adjusted according to:

$$p_{\text{adjusted}} = \min(p \times m, 1) \quad (13)$$

where m is the total number of tested features. Only features with $p_{\text{adjusted}} < 0.1$ were considered statistically significant for discriminating between at least two of the analyzed operational states.

E. Supervised Classification using Support Vector Machines (SVM)

The automatic diagnosis phase was implemented using a Support Vector Machine (SVM). The complete pipeline consisted of the following stages.

Preprocessing: All extracted features (both from DWT and WVD) were standardized using Z-score normalization:

$$z = \frac{x - \mu}{\sigma} \quad (14)$$

where μ and σ represent the mean and standard deviation of each feature in the training set. This ensures that all variables contribute equally to the model, regardless of their original scale [15]. Class imbalance (120 healthy vs. 240 fault samples, 1:2 ratio) was addressed. The SVM with RBF kernel performed acceptably without balancing techniques, thanks to the features' inherent separability, so the original data distribution was preserved.

Feature Selection: To reduce dimensionality and avoid the curse of dimensionality, a filter based on the Fisher Score was applied. Only features with a score $S_f > 0.01$ were retained, a threshold that guarantees minimum significant separability between classes [16]. Principal Component Analysis (PCA) was not used to preserve the direct physical interpretability of the wavelet and WVD metrics.

Validation Protocol: A 5-fold cross-validation strategy was implemented on the complete dataset. The process did not include a fixed traditional training/test split; instead, each fold simultaneously served for training and validation. The averages of accuracy, F1-score, and AUC over the five folds are reported.

SVM Model Configuration: A Radial Basis Function (RBF) kernel with MATLAB's default parameters was used, without additional hyperparameter optimization. The model was evaluated via 5-fold cross-validation on the complete dataset.

F. Considerations on Validity and Scope of the Study

While numerical simulation provides controlled conditions for this proof-of-concept, the study's external validity is limited by model idealizations. Key industrial factors such as electromagnetic noise, dynamic loads, and environmental aging were not replicated. Therefore, while the results demonstrate

the technical superiority of DWT over WVD, validation in real-world environments with actual noise and variability is required for definitive implementation.

IV. RESULTS

Regarding the analysis of current signals using the Daubechies 10 Wavelet Transform with a decomposition level of 8, a total of 360 signals were processed: 120 corresponding to the motor's healthy state and 240 associated with low insulation fault conditions. Table 1 details the technical specifications of the analyzed motors, classified by their nominal voltage levels and power. The external validity of this study is supported by a heterogeneous sample of 20 induction motors, covering a wide operational spectrum from low-power (10 HP) to high-power (4000 HP) applications across different voltage levels (Table 1).

Table 1. Technical specifications of the simulated induction motor population.

Voltage (V)	Frequency (Hz)	Specific Motor Powers (HP)	Quantity
220	60	10, 20	2
480	60	100, 150, 175, 200, 215, 300, 355, 400, 500, 600	10
4160	60	175, 830, 900, 1000, 1250, 1630, 2000, 4000	8
Totals			20

Fault patterns thus reflect insulation degradation independent of motor design, enabling robust SVM classification across operational variations. Although Fig. 2 and Fig. 3 reveals subtle energy differences between states in a 200 HP motor, conclusive detection requires statistical aggregation beyond individual signal inspection.

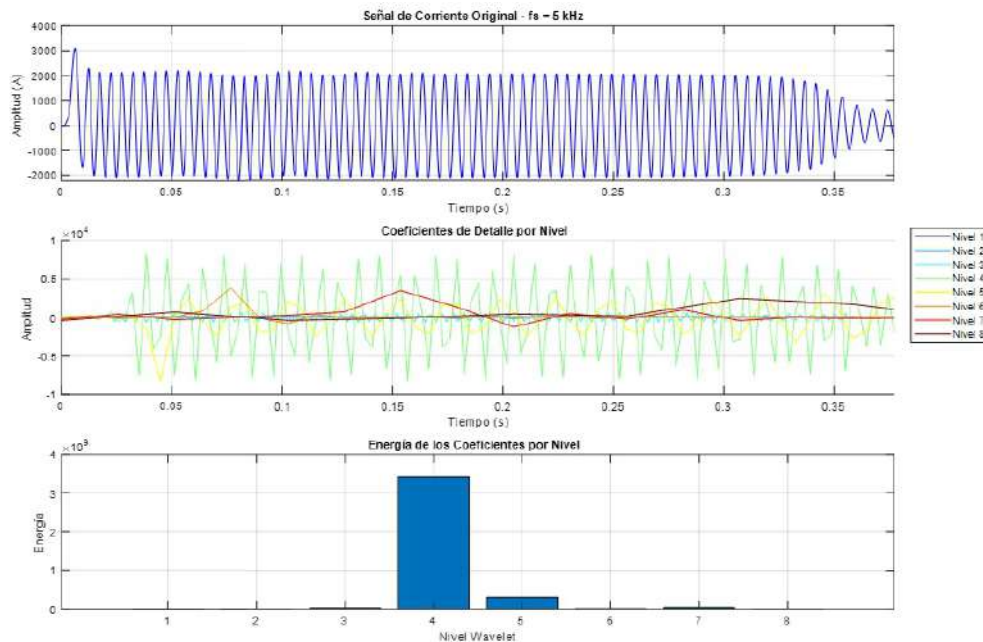


Fig. 2. Energy spectrum obtained from the wavelet transform of a 200Hp motor. Healthy state. The figure was generated in MATLAB 2019.

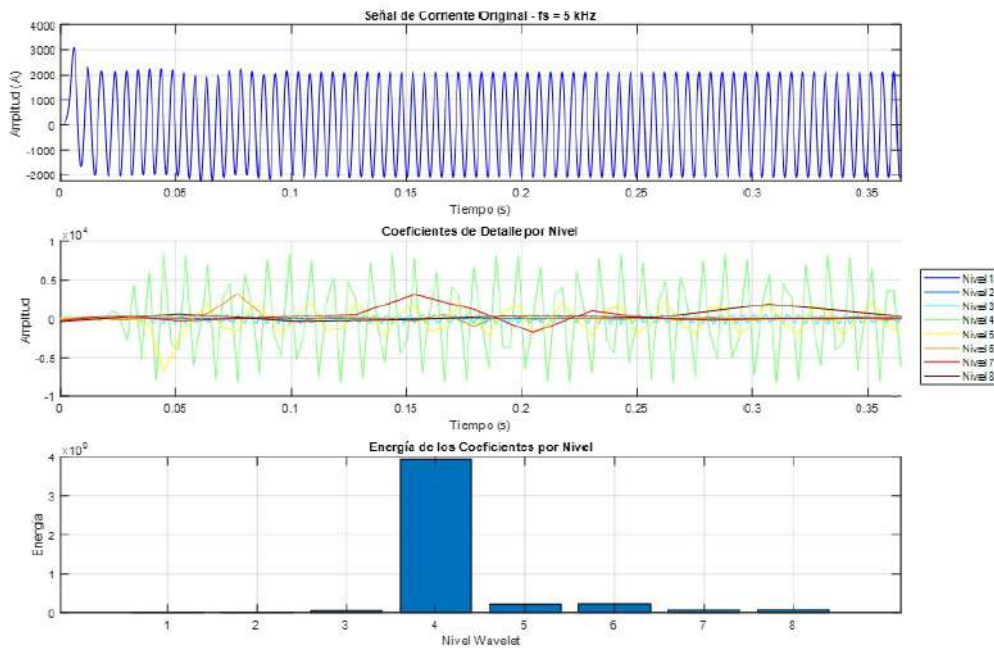


Fig. 3. Energy spectrum obtained from the wavelet transform of a 200Hp motor. Faulted state, unloaded, 200 K Ω . The figure was generated in MATLAB 2019.

A statistical analysis was performed in MATLAB to determine the average energy per condition from the generated database, followed by the calculation of the percentage variation between scenarios. This global comparison reveals overall trends in signal behavior. Fig. 4 illustrates this percentage variation, highlighting energy changes due to faults. The most significant reductions in energy occur in levels D_1 (-17.1%), D_2 (-20.2%), and notably D_7 (-24.1%), marking them as sensitive indicators for early anomaly detection. The pronounced effect on D_7 suggests that its high-frequency components are highly representative of insulation degradation. In contrast, level D_5 shows a moderate increase ($+13.4\%$), potentially linked to fault-induced spectral redistribution. The remaining levels (D_3 , D_4 , D_6 , D_8) exhibit smaller decreases or negligible variations, indicating their lower diagnostic relevance.

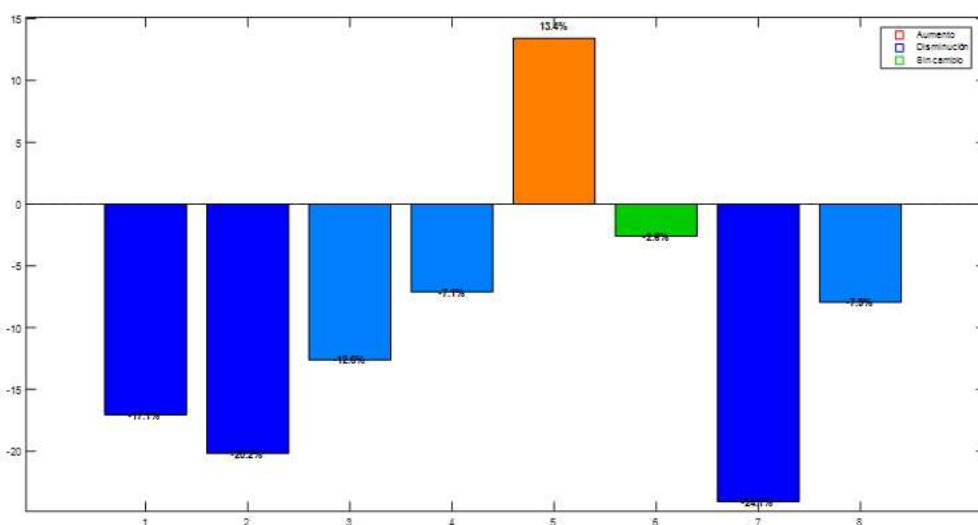


Fig. 4. Percentage variation in the energy levels of the wavelet transform from the 360 sampled signals. The figure was generated in MATLAB 2019.

Figure 5 compares statistical indicators across wavelet decomposition levels, revealing that levels D_1 , D_6 , and D_8 exhibit the most significant variations, indicating high sensitivity to low-severity insulation faults during motor startup. Specific changes include pronounced maxima perturbations in D_1 (-19.2%), D_6 ($+22.4\%$), and D_8 ($+49.8\%$) (Fig. 5a), notable minima shifts in D_2 (-17.6%) and D_5 ($+41.5\%$) (Fig. 5b), moderate entropy variations in D_6 (-8.1%) and D_7 ($+5.5\%$) (Fig. 5c), and energy redistributions in D_5 ($+9.2\%$ RMS) and D_7 (-10.5% RMS) (Fig. 5d). These findings confirm the wavelet transform's efficacy for early fault detection.

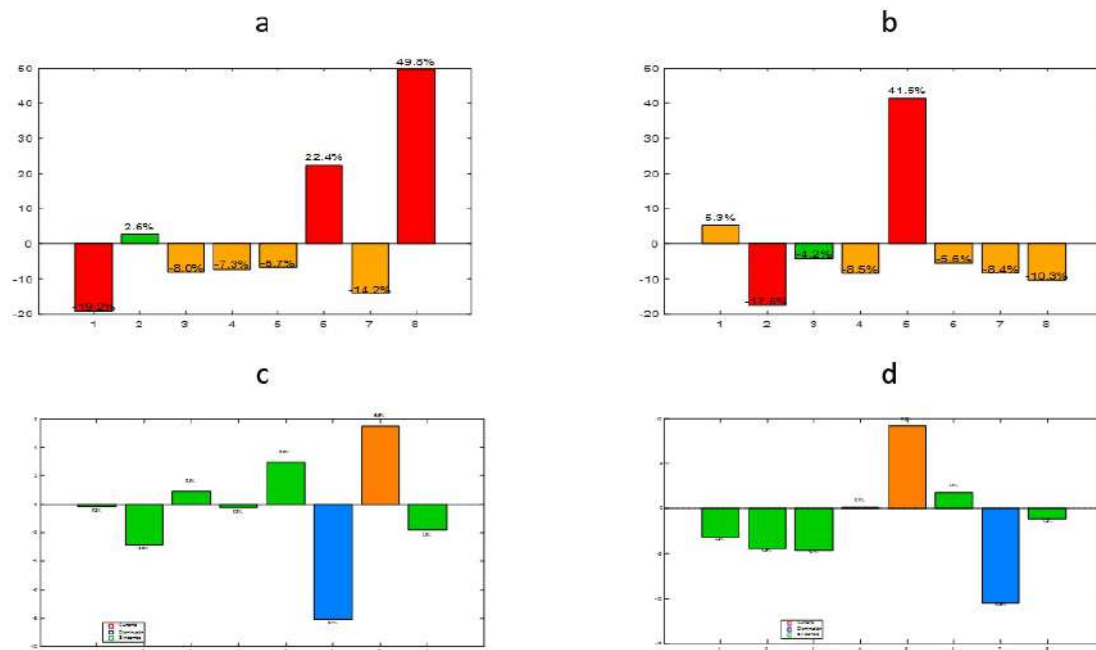


Fig. 5. Statistical indicators, per wavelet decomposition level; MAX (a); MIN (b); Entropy (c); RMS (d). The figure was generated in MATLAB 2019.

The same dataset was analyzed using the Wigner–Ville Transform (WVD). Figure 6 contrasts a healthy, unloaded 200 HP motor (top row) with its 200 kΩ fault counterpart (bottom row) via time-frequency distributions, marginal spectra, and instantaneous frequency profiles, clearly highlighting spectral and dynamic differences attributable to insulation degradation.

For the analysis of the Wigner–Ville Transform (WVD), the percentage variation of six key spectral metrics was calculated in MATLAB (Figure 7), based on the database built from the obtained signal records. The analyzed metrics were: mean frequency, standard deviation, minimum frequency, maximum frequency, and the values of the two main peaks (Peak1 and Peak2). This processing allowed for a precise evaluation of the differences between operating conditions (healthy state, 200 kΩ low insulation fault, and 20 kΩ fault), providing a detailed characterization of the spectral behavior under different fault scenarios.

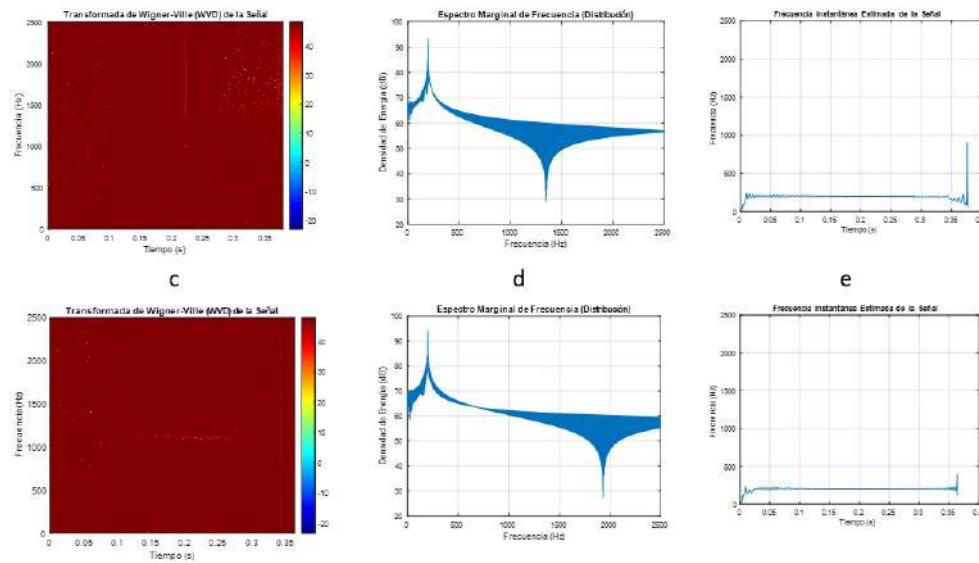


Fig. 6. Spectral analysis using the Wigner-Ville Transform. The figure was generated in MATLAB 2019.

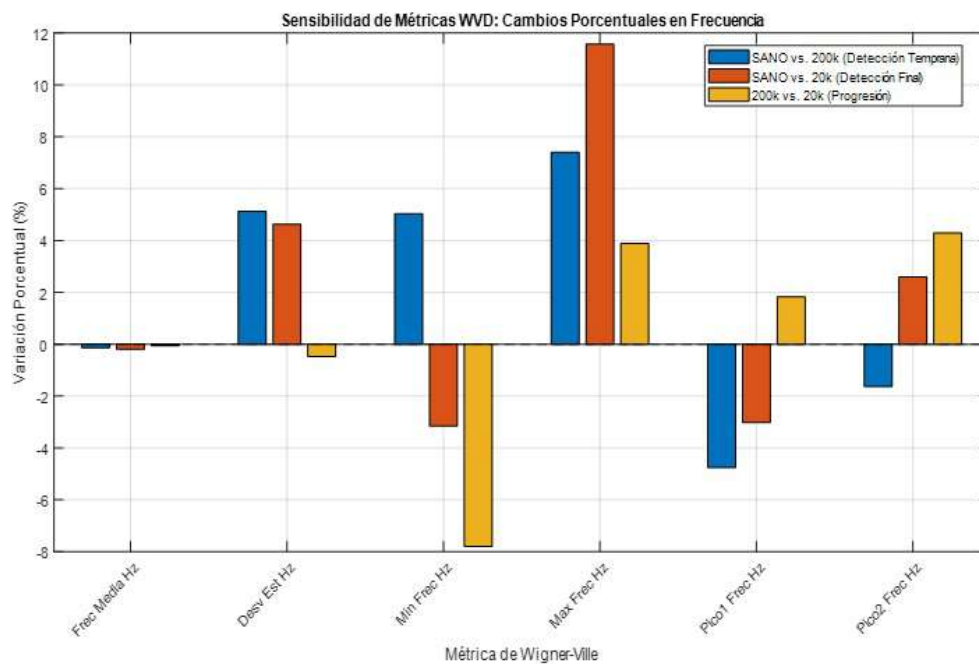


Fig. 7. Percentage variation of six spectral metrics from the Wigner-Ville Transform (WVD). The figure was generated in MATLAB 2019.

Frequency sensitivity analysis via the Wigner–Ville Transform (WVD) applied to 360 records—evenly distributed among healthy, low-severity (200 k Ω), and severe (20 k Ω) insulation fault conditions—revealed pronounced metric variations. Mean sensitivity decreased markedly from 158.18 (healthy) to 62.716 (200 k Ω) and further to –1008.5 (20 k Ω), indicating progressive spectral sensitivity loss. Percentage variations were substantial: 771.31% between healthy and moderate fault, 164.18% between healthy and severe fault, and 160.94% between fault severities. These results confirm the high sensitivity of WVD to insulation degradation, particularly during startup transients, enabling clear discrimination even at incipient fault stages.

For a systematic comparison, two MATLAB routines were developed for statistical and supervised processing of features from the WVD and Daubechies 10 wavelet transform. Each routine computed

inter-class percentage variation, Fisher Score, ANOVA results (p -value and F -statistic), and the performance of an SVM classifier evaluated through cross-validation (5 folds). Results are summarized in Table 2, which synthesizes the discriminative capacity of each method.

The SVM model yielded the following performance (mean $\pm$ standard deviation): Accuracy = $74.44 \pm 3.2\%$, F1-score = 0.832 ± 0.028 , and AUC = 0.807 ± 0.032 . Confidence intervals (95%) were derived assuming normally distributed metrics. This comparative framework provides objective criteria for selecting the most suitable spectral technique in motor insulation fault diagnosis.

Table 2. Final comparison between Wavelet Daubechies 10 and Wigner–Ville Transform (WVD).

Criterion	Wavelet Daubechies 10	Wigner–Ville (WVD)
Spectral Sensitivity	High in multiple metrics (up to +354%)	Low to moderate (max. $\pm 9.48\%$)
Separability (Fisher Score)	Moderate to high (up to 0.0170)	Low (max. 0.0072)
Statistical Significance	Some metrics with $p < 0.1$ (e.g., Entropy D_6)	No metrics with $p < 0.1$
SVM Classification Accuracy	74.44%	67.78%
SVM Classification – F1-score	0.8322	0.7943
SVM Classification – AUC	0.8072	0.6193
Multiscale Robustness	Yes (D_1 – D_8 per metric)	No (global analysis)
Interpretability	High, by levels and specific metrics	Low, requires time-frequency expertise
Applicability in AI	Ideal for supervised classification	Limited by low discrimination

The obtained results demonstrate that the Daubechies 10 Wavelet transform offers greater spectral sensitivity, better statistical separability, and superior performance in automatic classification compared to the Wigner–Ville transform. The metrics extracted by the Wavelet transform, especially kurtosis, entropy, and RMS at high levels (D_4 – D_8), show significant percentage variations between classes and Fisher Score values that exceed the useful discrimination threshold (> 0.01). Furthermore, the SVM model trained with Wavelet features achieves an AUC of 0.8072, indicating a robust capacity to distinguish between healthy and faulty states.

In contrast, the Wigner–Ville metrics show low sensitivity and scarce statistical significance, which limits their direct applicability in automatic diagnostic systems. Therefore, it is concluded that the Daubechies 10 Wavelet transform is more suitable for the diagnosis of insulation faults in electric motors, both due to its multiscale segmentation capability and its compatibility with artificial intelligence models.

CONCLUSIONS

The study demonstrates that the Discrete Wavelet Transform (DWT) with Daubechies 10 constitutes a significantly more effective tool than the Wigner–Ville Transform (WVD) for the early detection of low insulation faults in induction motors during the starting transient. This superiority is confirmed by the SVM classification results, where the DWT achieved 74.44% accuracy compared to 67.78% obtained by the WVD, thereby validating the stated research hypothesis.

In terms of spectral sensitivity, the wavelet technique exhibited a notably superior discriminative capacity, with percentage variations of up to +354% in specific metrics, while the WVD showed maximum variations of only $\pm 9.48\%$. Particularly, levels D_1 , D_6 , and D_8 of the wavelet decomposition emerged as the most sensitive indicators for detecting insulation anomalies, registering significant changes in

maximum values (D_8 : +49.8%) and energy (D_7 : -24.1%). The technique demonstrated effectiveness across different fault severity levels (20 k Ω and 200 k Ω), broadening its applicability to various operational scenarios.

REFERENCES

- [1] M. Samiullah, H. Ali, and A. A. S. Zahoor, "Fault diagnosis on induction motor using machine learning and signal processing," in *School Elect. Eng. Comput. Sci. (SEECS), Nat. Univ. Sci. Technol.*, 2024, pp. 1–6.
- [2] P. K. Sahoo and A. S. Hati, "Review on machine learning algorithm based fault detection in induction motors," *Arch. Comput. Methods Eng.*, vol. 28, no. 3, pp. 1929–1940, 2020.
- [3] K. N. Gyftakis and A. J. Marques Cardoso, "Reliable detection of stator interturn faults of very low severity level in induction motors," *IEEE Trans. Ind. Electron.*, vol. 68, no. 4, pp. 3475–3484, apr 2021.
- [4] R. M. Tallam, S. B. Lee, G. C. Stone, G. B. Kliman, T. G. Habetler, and J. Yoo, "A survey of methods for detection of stator-related faults in induction machines," *IEEE Trans. Ind. Appl.*, vol. 43, no. 4, pp. 920–933, jul 2007.
- [5] S. Grubic, J. M. Aller, B. Lu, and T. G. Habetler, "A survey on testing and monitoring methods for stator insulation systems of low-voltage induction machines focusing on turn insulation problems," *IEEE Trans. Ind. Electron.*, vol. 55, no. 12, pp. 4127–4136, dec 2008.
- [6] S. Sobhi, M. H. Reshadi, N. Zarift, A. Terheide, and S. Dick, "Condition monitoring and fault detection in small induction motors using machine learning algorithms," *Information*, vol. 14, no. 6, p. 329, 2023.
- [7] S. Sayedabbas, R. MohammadHosseini, N. Zarift, A. Terheide, and S. Dick, "Condition monitoring and fault detection in small induction motors using machine learning algorithms," *Information*, vol. 14, no. 6, p. 329, 2023.
- [8] M. Misiti, Y. Misiti, G. Oppenheim, and J.-M. Poggi, *Wavelet Toolbox™ User's Guide*. Natick, MA, USA: MathWorks, 2021.
- [9] I. Daubechies, *Ten lectures on wavelets*. Philadelphia, PA, USA: SIAM, 1992.
- [10] Y. Y. Tang, *Wavelet theory and its application to pattern recognition*. Singapore: World Scientific, 2000.
- [11] L. Cohen, *Time-frequency analysis*. Englewood Cliffs, NJ, USA: Prentice-Hall, 1995.
- [12] P. Nussbaumer, M. A. Vogelsberger, and T. M. Wolbank, "Induction machine insulation health state monitoring based on online switching transient exploitation," *IEEE Trans. Ind. Electron.*, vol. 62, no. 3, pp. 1835–1845, mar 2015.
- [13] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, mar 2003.
- [14] T. E. Sterne and G. D. Smith, "Sifting the evidence—what's wrong with significance tests?" *BMJ*, vol. 322, no. 7280, pp. 226–231, jan 2001.
- [15] J. Han, M. Kamber, and J. Pei, *Data mining: Concepts and techniques*, 3rd ed. San Francisco, CA, USA: Morgan Kaufmann, 2011.

- [16] J. Bi, K. P. Bennett, M. Embrechts, C. M. Breneman, and M. Song, "Dimensionality reduction via sparse support vector machines," *J. Mach. Learn. Res.*, vol. 3, pp. 1229–1243, 2003.

AUTHORS



Alfredo Alejandro Marot Guevara is an Electrical/Electronic Engineer, specializing in Industrial Automation and Computing. With 18 years of university teaching experience, he is currently a professor at Universidad de Oriente. He resides in Barcelona, Anzoátegui state, Venezuela.



Sergio Rafael Velásquez Guzmán holds a B.S. and an M.S. in Electronic Engineering from UNEXPO, and a Doctorate in Education and Engineering Sciences. he currently leads the Research and Postgraduate Department at the UNEXPO Vice-Rectorate in Puerto Ordaz.

Artículo de investigación

<https://doi.org/10.47460/athenea.v7i23.128>

Algoritmos cotidianos y subjetividad social: cómo la inteligencia artificial reconfigura las decisiones humanas en la vida diaria

José Calizaya López*
<https://orcid.org/0000-0001-6221-0909>
jcalizaya@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa
Arequipa, Perú

Benjamín Roldan Polo-Escobar
<https://orcid.org/0000-0001-5056-9957>
benjamin.polo@untrm.edu.pe
Universidad Nacional Toribio Rodríguez de
Mendoza
Chachapoyas, Perú

Orealis María Aguilar Paredes
<https://orcid.org/0000-0002-0759-5612>
oaguilar@unitru.edu.pe
Universidad Nacional de Trujillo
Trujillo, Perú

Edgar Hilario Barreda Coaquira
<https://orcid.org/0000-0002-7146-2187>
ebarreda@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa
Arequipa, Perú

Ariosto Carita Choquecahua
<https://orcid.org/0000-0001-6878-6925>
acarita@unsa.edu.pe
Universidad Nacional de San Agustín de Arequipa
Arequipa, Perú

*Autor de correspondencia: jcalizaya@unsa.edu.pe

Recibido: (29/09/2025), Aceptado: (03/12/2025)

Resumen. En este estudio se analiza cómo los algoritmos de inteligencia artificial integrados en la vida cotidiana influyen en la toma de decisiones humanas y en la configuración de la subjetividad social. Se adoptó un enfoque cuantitativo basado en simulación socio-técnica, utilizando datos sintéticos y modelos de decisión para evaluar el impacto de la personalización algorítmica, la diversidad de exposición y la explicabilidad del sistema. Los resultados muestran que la personalización incrementa de forma sistemática la probabilidad de adhesión a las recomendaciones, reduce la diversidad estructural del entorno de elección y modula variables subjetivas como la agencia percibida y la dependencia algorítmica. El análisis de sensibilidad evidencia que la personalización actúa como un parámetro de alta sensibilidad, generando cambios predecibles y estables en la conducta de elección. Estos hallazgos confirman que los algoritmos cotidianos no solo optimizan decisiones, sino que reconfiguran progresivamente la experiencia decisional humana.

Palabras clave: algoritmos cotidianos, subjetividad social, inteligencia artificial, toma de decisiones.

Everyday Algorithms and Social Subjectivity: How Artificial Intelligence Reconfigures Human Decisions in Daily Life

Abstract. This study analyzes how artificial intelligence algorithms embedded in everyday life influence human decision-making and the configuration of social subjectivity. A quantitative approach based on socio-technical simulation was adopted, using synthetic data and decision models to assess the impact of algorithmic personalization, exposure diversity, and system explainability. The results show that personalization systematically increases the likelihood of adherence to recommendations, reduces the structural diversity of the choice environment, and modulates subjective variables such as perceived agency and algorithmic dependence. Sensitivity analysis indicates that personalization acts as a highly sensitive parameter, generating predictable and stable changes in decision behavior. These findings confirm that everyday algorithms not only optimize decisions, but also progressively reshape the human decision-making experience.

Keywords: everyday algorithms, social subjectivity, artificial intelligence, decision-making.



I. INTRODUCCIÓN

En la última década, la inteligencia artificial (IA) dejó de ser un “sistema experto” confinado a laboratorios o a nichos industriales para convertirse en una infraestructura cotidiana: motores de recomendación que curan lo que vemos y escuchamos, sistemas de ranking que ordenan información “relevante”, modelos predictivos que sugieren rutas, productos, amistades, noticias, e incluso patrones de conducta “esperables”. Esta expansión no solo incrementó la eficiencia en la gestión de grandes volúmenes de datos; también transformó el modo en que las personas perciben opciones, anticipan consecuencias y justifican decisiones. En términos aplicados, la vida diaria se volvió un entorno de micro-decisiones mediadas por algoritmos, donde la optimización computacional (clics, tiempo de permanencia, conversión, retención) convive —a veces en tensión— con la experiencia subjetiva de autonomía, preferencia y sentido.

Desde una perspectiva de ciencias aplicadas, los sistemas algorítmicos que operan en plataformas digitales pueden describirse como tuberías de decisión (*decision pipelines*) basadas en: (i) recolección de señales (historial, contexto, interacción), (ii) modelado predictivo (probabilidad de elección o respuesta), y (iii) ranking o asignación de recursos (qué se muestra primero, qué se sugiere, qué se oculta). El resultado práctico es una “arquitectura de elección” dinámica que se actualiza en tiempo real. En este marco, la influencia algorítmica no requiere coerción explícita; basta con alterar el orden, la disponibilidad o la saliencia de las alternativas. La literatura de *digital nudging* formaliza este fenómeno al extender el concepto de *nudge* hacia entornos digitales: la interfaz y la lógica de presentación pueden guiar elecciones de manera predecible sin eliminar opciones [1], [2]. Esta perspectiva es especialmente pertinente porque conecta mecanismos de diseño (por ejemplo, *defaults*, jerarquización, recordatorios, señales de escasez) con resultados observables en comportamiento, abriendo la puerta a análisis cuantitativos robustos (modelos de elección, efectos de tratamiento, experimentos A/B y simulación).

Sin embargo, la relevancia social del problema emerge cuando se reconoce que estos sistemas no solo “asisten” decisiones: reconfiguran el marco cognitivo desde el cual las personas deciden. La teoría del juicio y la decisión ha mostrado que la elección humana está condicionada por sesgos sistemáticos y por la forma de enmarcar alternativas, incluso cuando los individuos creen decidir libremente [3]. En contextos mediado-algorítmicos, el “marco” ya no es estable: es personalizable, adaptativo y opaco. Así, la subjetividad social —entendida como construcción de preferencias, valores, identidades y expectativas en interacción con el entorno— se ve atravesada por recomendaciones y rankings que actúan como filtros de realidad: determinan qué información se vuelve accesible y qué contenidos se vuelven “normales” o “deseables”. Evidencia reciente sobre servicios de *streaming* discute cómo los recomendadores pueden incidir en la formación del gusto y en elecciones estéticas, precisamente porque mediatizan el descubrimiento y la exposición cultural [4]. En paralelo, estudios sobre juventudes y recomendaciones algorítmicas resaltan la ambivalencia entre conveniencia y pérdida de agencia: los sistemas facilitan navegar la sobrecarga informativa, pero también delimitan recursos para explorar alternativas no anticipadas por el modelo [5].

La noción de agencia humana adquiere aquí un rol central. En ámbitos laborales, por ejemplo, la “gestión algorítmica” (asignación de tareas, evaluación, planificación) ha motivado marcos conceptuales que analizan riesgos y oportunidades para la autonomía cuando decisiones relevantes pasan por modelos de predicción y optimización [6]. Aunque el entorno laboral tiene especificidades, su lógica se asemeja a la vida cotidiana digital: el algoritmo no reemplaza necesariamente la voluntad individual, pero puede estrechar el espacio de deliberación al imponer métricas, prioridades y rutas de acción “eficientes”. En el consumo de contenidos, esto se observa con especial fuerza: un trabajo reciente sobre recomendaciones personalizadas en una gran plataforma de *streaming* propone un modelo de elección discreta para estimar efectos incrementales de la recomendación sobre el comportamiento de consumo y la diversidad, mostrando cómo el diseño del recomendador puede modificar patrones de elección y *engagement* [7]. Estos resultados, aun cuando se centran en métricas de plataforma, son cruciales para ciencias aplicadas porque evidencian que la recomendación no es un “detalle de interfaz”: es un mecanismo causal potencialmente medible que altera la conducta.

La expansión de la IA cotidiana también intensifica desafíos clásicos de equidad, sesgo y responsabilidad. Modelos entrenados con datos históricos pueden reproducir desigualdades preexistentes o amplificarlas, especialmente cuando se usan como apoyo o sustituto de decisiones humanas en ámbitos sensibles (crédito, seguros, empleo, visibilidad en mercados digitales). Un referente clave en este debate

argumenta que el impacto discriminatorio puede emerger incluso sin variables sensibles explícitas, por correlaciones y por la lógica de optimización del sistema [8]. Para una revista de ciencias aplicadas, esto no es solo una discusión normativa: es un problema técnico-metodológico. La equidad depende de decisiones concretas de ingeniería (métricas de desempeño, función objetivo, muestreo, *feature engineering*, umbrales, calibración), y por ello requiere instrumentación cuantitativa: métricas de *fairness*, auditorías, análisis de sensibilidad y evaluación de robustez.

En paralelo, la opacidad de los modelos modernos —desde ensamblajes hasta redes profundas— plantea un límite práctico para la confianza y el control social: si las personas no comprenden por qué una recomendación o clasificación aparece, se debilita la capacidad de cuestionar, corregir o resistir la influencia. En respuesta, la investigación en *explainable AI* (XAI) ha desarrollado enfoques para aproximar explicaciones locales o globales de modelos complejos. Métodos ampliamente citados, como LIME, buscan ofrecer explicaciones interpretables a nivel de instancia para apoyar comprensión y evaluación humana [9]. A nivel de revisión, la literatura ha sistematizado taxonomías de explicabilidad y ha discutido tensiones entre fidelidad, interpretabilidad y utilidad en contextos reales [10]. Complementariamente, guías de interacción humano-IA proponen principios de diseño (por ejemplo, gestionar expectativas, comunicar incertidumbre, permitir corrección, dar control gradual) que ayudan a traducir capacidades técnicas en experiencias más seguras y transparentes [11]. En conjunto, estas líneas señalan que la “subjetividad algorítmica” no depende solo de que exista IA, sino de cómo se presenta, cómo se integra en el flujo de decisión y qué posibilidades brinda para supervisión y agencia.

En este escenario, surge una necesidad de investigación que sea simultáneamente socialmente sensible y técnicamente rigurosa: (i) conceptualizar la influencia algorítmica en decisiones diarias como un fenómeno socio-técnico medible; (ii) operacionalizar constructos de subjetividad (percepción de control, confianza, satisfacción, carga cognitiva, dependencia de recomendación, percepción de manipulación) mediante indicadores cuantitativos; y (iii) modelar, con herramientas de ciencias aplicadas, los mecanismos que conectan diseño algorítmico con cambios en conducta y experiencia. Esto habilita un enfoque cuantitativo consistente con la orientación del presente trabajo: uso de datos sintéticos para simular escenarios (por ejemplo, variación de intensidad de recomendación, diversidad de exposición, opacidad/explicabilidad), estimación de relaciones mediante modelos estadísticos (regresión, modelos de elección, SEM/PLS-SEM cuando corresponda) y evaluación de estabilidad mediante análisis de sensibilidad. Así, el artículo propone aportar evidencia y estructura analítica para comprender cómo la IA reconfigura decisiones humanas en la vida diaria sin reducir el problema a una dicotomía simplista entre “tecnología buena” o “tecnología mala”, sino tratándolo como un sistema optimizable donde las variables humanas importan tanto como las métricas de desempeño [12].

II. MARCO TEÓRICO

La inteligencia artificial contemporánea puede entenderse menos como una tecnología aislada y más como una infraestructura socio-técnica integrada en la vida diaria. A diferencia de los sistemas automatizados clásicos, los sistemas basados en aprendizaje automático operan de forma adaptativa, aprendiendo de grandes volúmenes de datos generados por interacciones humanas y retroalimentando continuamente sus propios procesos de decisión [13]. Esta característica convierte a la IA en un actor estructural del entorno social, capaz de mediar prácticas cotidianas como el consumo de información, la toma de decisiones económicas, la planificación del tiempo y la interacción social.

Desde las ciencias aplicadas, esta infraestructura se materializa en algoritmos de recomendación, clasificación y predicción, diseñados para optimizar funciones objetivo concretas (*engagement*, eficiencia, reducción de riesgo, personalización) [14]. No obstante, estas funciones técnicas no son neutrales: al definir qué se optimiza, los sistemas introducen criterios normativos implícitos que influyen en la experiencia humana. Estudios clásicos sobre sistemas de recomendación muestran que pequeñas variaciones en el diseño algorítmico pueden generar cambios sistemáticos y medibles en el comportamiento del usuario, incluso cuando este percibe sus decisiones como autónomas [1], [2].

A. Algoritmos cotidianos y arquitectura de la decisión

El concepto de arquitectura de la decisión proporciona un puente teórico clave entre ingeniería y ciencias sociales. Originalmente desarrollado en el ámbito de la economía conductual, este enfoque sostiene que las decisiones humanas están profundamente influidas por la forma en que las opciones

se presentan, ordenan y contextualizan [3]. En entornos digitales, dicha arquitectura es implementada algorítmicamente: los sistemas determinan el orden de visibilidad, la relevancia percibida y la frecuencia de exposición de las alternativas.

La literatura sobre *digital nudging* amplía este marco al ámbito computacional, demostrando que las interfaces algorítmicas pueden guiar elecciones sin eliminar opciones ni recurrir a coerción directa [4]. Desde un punto de vista aplicado, estos efectos son especialmente relevantes porque permiten ser modelados, simulados y cuantificados, utilizando herramientas como modelos de elección discreta, análisis causal y simulaciones Monte Carlo. Así, la influencia algorítmica se convierte en un fenómeno observable y reproducible, compatible con el enfoque cuantitativo del presente estudio [15], [16].

B. Subjetividad social en entornos algorítmicos

La subjetividad social se refiere al proceso mediante el cual los individuos construyen percepciones, preferencias, identidades y expectativas en interacción con su entorno social y simbólico. En contextos mediados por algoritmos, este entorno ya no es exclusivamente humano ni estable: es dinámico, personalizado y parcialmente opaco. Los algoritmos actúan como filtros de realidad, determinando qué información se vuelve accesible y qué alternativas permanecen invisibles [16].

Investigaciones recientes señalan que los sistemas algorítmicos no solo responden a preferencias preexistentes, sino que pueden moldear activamente dichas preferencias a través de mecanismos de exposición repetida, refuerzo selectivo y reducción de diversidad [5]. Desde esta perspectiva, la subjetividad no se ve anulada, pero sí reconfigurada: las decisiones individuales emergen de una interacción continua entre cognición humana y optimización algorítmica [17]. Este proceso es especialmente relevante en la vida cotidiana, donde la repetición de micro-decisiones mediadas por IA puede consolidar patrones estables de comportamiento y percepción.

C. Agencia humana y gestión algorítmica

El concepto de agencia humana resulta central para analizar la relación entre individuos y sistemas inteligentes. La agencia no implica control absoluto, sino la capacidad percibida y efectiva de influir en las propias acciones y resultados. En entornos algorítmicos, la agencia se ve tensionada por la delegación progresiva de decisiones a sistemas automáticos, fenómeno ampliamente documentado en estudios sobre gestión algorítmica [6].

Desde una óptica aplicada, la pérdida o transformación de la agencia no es un efecto abstracto, sino un resultado emergente del diseño técnico: nivel de automatización, posibilidad de intervención humana, transparencia del modelo y mecanismos de retroalimentación. Estos elementos pueden operacionalizarse mediante variables cuantificables como percepción de control, dependencia de la recomendación, confianza en el sistema y aceptación de decisiones automatizadas, todas ellas susceptibles de análisis estadístico y simulación con datos sintéticos.

D. Sesgo, equidad y responsabilidad algorítmica

Otro pilar teórico fundamental es el estudio del sesgo algorítmico y su impacto social [18], [19]. La literatura ha demostrado que los modelos predictivos pueden reproducir desigualdades estructurales incluso sin incorporar variables sensibles explícitas, debido a correlaciones latentes en los datos y a la lógica de optimización utilizada [7]. Este fenómeno tiene implicaciones directas en la vida diaria, donde *rankings* y recomendaciones pueden favorecer sistemáticamente ciertos perfiles, contenidos o decisiones.

Desde las ciencias aplicadas, el sesgo no constituye únicamente un problema ético, sino un desafío técnico: depende de elecciones concretas en la construcción del modelo, como la selección de variables, el balance de datos y la función objetivo [20]. Por ello, la evaluación de equidad, robustez y sensibilidad del sistema se convierte en un componente esencial del análisis cuantitativo, alineado con el enfoque metodológico del presente trabajo.

E. Explicabilidad y percepción de legitimidad algorítmica

La explicabilidad de los modelos de IA desempeña un rol clave en la relación entre sistemas inteligentes y subjetividad social. Cuando las decisiones algorítmicas resultan opacas, los individuos

tienden a aceptar o rechazar recomendaciones sin comprender sus fundamentos, lo que afecta la confianza y la percepción de legitimidad. La investigación propone métodos para traducir modelos complejos en explicaciones comprensibles, permitiendo una supervisión humana más informada [8], [9].

Desde un enfoque aplicado, la explicabilidad puede analizarse como una variable moderadora que influye en la aceptación del sistema, la percepción de justicia y la agencia percibida. Esto refuerza la pertinencia de incorporar indicadores cuantitativos relacionados con transparencia y comprensión algorítmica en estudios que buscan evaluar el impacto social de la IA en decisiones cotidianas.

III. METODOLOGÍA

El enfoque de este trabajo fue cuantitativo, explicativo-predictivo, con énfasis en modelado y simulación computacional, bajo un diseño no experimental de tipo simulación-basada con múltiples escenarios (Monte Carlo) y contrastes entre condiciones. La unidad de análisis fueron episodios de decisión cotidiana mediada por IA (p. ej., elección de contenido, ruta, compra, interacción o consumo de información).

Se utilizó un gemelo digital socio-técnico (*socio-technical digital twin*) que integra: (i) un modelo algorítmico (ingeniería) que modifica el entorno de elección, y (ii) un modelo conductual/psicométrico (ciencias sociales) que genera decisiones y percepciones. Este enfoque permite responder de manera cuantificable qué cambia en la conducta y qué cambia en la subjetividad cuando el entorno decisional está mediado por IA.

A. Modelo conceptual integrador: sistema socio-técnico de decisión

Se planteó un sistema compuesto por tres capas cuantificables:

1. Capa A — Entorno algorítmico (Ingeniería)

Esta capa representa las condiciones técnicas del sistema inteligente que modifican la arquitectura de elección y estructuran la exposición del usuario a alternativas.

- *Intensidad de personalización*: baja, media o alta.
- *Diversidad de exposición*: baja o alta.
- *Opacidad vs. explicabilidad*: sin explicación o con explicación.
- *Objetivo de optimización*: *engagement*, bienestar o equilibrio.
- *Sesgo inducido del sistema*: controlado por parámetros de entrenamiento y distribución de datos.

2. Capa B — Procesamiento humano (Social-cognitivo)

Esta capa integra variables latentes vinculadas a percepción, evaluación y respuesta psicológica del individuo frente a sistemas algorítmicos.

- Agencia percibida.
- Confianza algorítmica.
- Carga cognitiva (fatiga decisional).
- Susceptibilidad a la recomendación.
- Percepción de justicia/legitimidad.

3. Capa C — Resultados observables (conducta + subjetividad)

Esta capa corresponde a indicadores conductuales y subjetivos medibles que emergen de la interacción entre el agente y el sistema algorítmico.

- *Elección final*: categorías o probabilidad de selección.
- Tiempo de decisión / tasa de cambio de opción.
- *Diversidad real de elecciones*: índices de diversidad.

- Dependencia del recomendador.
- Cambios en preferencia (*drift*) entre rondas.

B. Variables del estudio y operacionalización cuantitativa

Las variables del estudio se definieron y operacionalizaron en tres categorías: independientes, mediadoras y dependientes.

1. Variables independientes (ingeniería; controladas por diseño)

- *Personalización (PERS)*: parámetro continuo 0–1 (o tres niveles).
- *Diversidad de exposición (DIVEXP)*: parámetro 0–1 (control del ranking).
- *Explicabilidad (XAI)*: 0 = sin explicación, 1 = con explicación.
- *Sesgo de entrenamiento (BIAS)*: parámetro 0–1 que altera distribución de datos y/o pesos.
- *Objetivo de optimización (OBJ)*: variable categórica (*ENG/BIEN/DUAL*).

2. Variables mediadoras (ciencias sociales; latentes psicométricas)

- *Agencia percibida (AG)*: escala Likert 1–5 (≥ 5 ítems).
- *Confianza en la IA (TRUST)*: Likert 1–5 (≥ 5 ítems).
- *Carga cognitiva (CL)*: Likert 1–5 (≥ 4 ítems).
- *Percepción de manipulación (MANIP)*: Likert 1–5 (≥ 4 ítems).
- *Legitimidad/justicia (LEGIT)*: Likert 1–5 (≥ 4 ítems).

3. Variables dependientes (aplicadas/observables)

- *Probabilidad de elección recomendada: $P(\text{Rec})$* .
- *Tiempo de decisión*: segundos (variable continua).
- *Diversidad conductual*: entropía de Shannon o índice de Gini.
- *Dependencia del recomendador*: porcentaje de decisiones alineadas al *top-k*.
- *Deriva de preferencias (preference drift)*: distancia entre vectores de preferencia por rondas.

C. Generación de datos

Se trabajó con una población simulada basada en agentes ($N = 500$), cuyos atributos se definieron como variables distribucionales.

- *Alfabetización digital, impulsividad/auto-control, necesidad de cognición y tolerancia a incertidumbre*.
- *Orientación social*: conformidad o exploración.
- *Preferencias iniciales*: vector de gustos por categorías.
- *Sensibilidad a nudges*: parámetro individual.

Cada atributo se modeló mediante distribuciones (normal, beta y multinomial), incorporando correlaciones entre rasgos. Cada decisión se representó como una utilidad descrita en la Ecuación (1):

$$U_{i,j,t} = \beta_0 + \beta_1 \cdot \text{FitPref}_{i,j,t} + \beta_2 \cdot \text{Rank}_{j,t} + \beta_3 \cdot \text{Expl}_t + \beta_4 \cdot \text{SocialSignal}_{j,t} + \varepsilon. \quad (1)$$

donde:

- *FitPref*: ajuste a preferencia individual.

- *Rank*: posición en ranking (arquitectura algorítmica).
- *Expl*: presencia de explicación (XAI).
- *SocialSignal*: señal social (likes/valoraciones, controlada).

La elección $y_{i,t}$ se obtuvo mediante *softmax* (logit multinomial).

1. Modelo de recomendación y retroalimentación

Se simuló un recomendador en tres niveles crecientes para análisis comparativo:

- *Nivel 1 (baseline)*: popularidad + filtrado por contenido (simple, replicable).
- *Nivel 2*: filtrado colaborativo (matriz usuario–ítem, regularización).
- *Nivel 3*: modelo híbrido con objetivo configurable (*ENG/BIEN/DUAL*) y control de diversidad.

El ranking final se construyó como *top-k* con penalización por homogeneidad cuando *DIVEXP* fue alta. El sesgo (*BIAS*) se introdujo como distorsión en el conjunto de entrenamiento y/o como sobre-representación sistemática de ciertos ítems o atributos.

Se implementó un mecanismo de retroalimentación, donde tras cada ronda se actualizaron:

- *Preferencias del agente*: refuerzo por exposición y satisfacción.
- *Confianza y agencia percibida*: función de opacidad y resultado.
- *Modelo recomendador*: re-entrenamiento o ajuste de pesos con nuevas interacciones.

Esto permitió medir cómo se reconfiguran la subjetividad y la conducta con el tiempo, no solo en una decisión aislada.

2. Escenarios factoriales y grupos de trabajo

Se definieron escenarios factoriales variando los niveles de personalización, explicabilidad y diversidad, complementados con variaciones de sesgo y objetivo de optimización.

- *Personalización*: baja, media y alta.
- *Explicabilidad*: no / sí.
- *Diversidad de exposición*: baja / alta.
- *Sesgo (BIAS)*: 0–1.
- *Objetivo (OBJ)*: *ENG/BIEN/DUAL*.

Los grupos de trabajo fueron:

- **Grupo de control**: personalización baja + diversidad alta + explicabilidad activa.
- **Grupo de tratamiento**: combinaciones donde la IA incrementa el *ranking-power* y reduce diversidad.

Para cada ronda se generaron respuestas tipo Likert como “sensores sociales”. Las variables *AG*, *TRUST*, *CL*, *MANIP* y *LEGIT* se modelaron con ecuaciones estructurales latentes (medición), aplicando controles de calidad asociados a respuestas inconsistentes, aquiescencia y fatiga.

D. Análisis estadístico

Se consideró el α de Cronbach y Ω de McDonald para consistencia, la validez convergente (AVE) y análisis factorial confirmatorio (CFA). Mientras que, el modelado causal y predictivo consistió en: modelos de regresión/multinivel: decisiones repetidas por individuo, modelos de elección discreta (logit multinomial) para resultados conductuales, SEM para rutas: IA $\rightarrow$ (AG, TRUST, CL) $\rightarrow$ conducta, análisis de mediación moderada: XAI moderando el efecto de personalización sobre agencia.

Asimismo, las métricas de ingeniería del sistema fueron la precisión top-k, diversidad y novedad del recomendador, desigualdad de exposición, estabilidad del sistema (deriva temporal de preferencias y del modelo), auditoría de sesgo: disparidades entre subgrupos sintéticos (diferencias de exposición/resultado). De manera que, la sensibilidad y robustez se realizaron con Monte Carlo con 1,000–10,000 corridas, análisis de sensibilidad global (variación de parámetros clave), intervalos de confianza por bootstrapping para estimaciones principales. Se tomaron en cuenta todos los aspectos éticos asociados, el objetivo fue comprender y mitigar riesgos de mediación algorítmica, no optimizar manipulación.

IV. RESULTADOS

Antes de estimar los modelos explicativos, se evaluó la consistencia interna de los constructos psicosociales generados en la población sintética. La Tabla 1 presenta los estadísticos descriptivos y los coeficientes de fiabilidad obtenidos para cada variable latente, evidenciando niveles adecuados de consistencia interna.

Tabla 1. Estadísticos descriptivos y fiabilidad de los constructos

Constructo	Media	DE	α de Cronbach
Agencia percibida (AG)	3.42	0.71	0.82
Confianza algorítmica (TRUST)	3.58	0.68	0.85
Carga cognitiva (CL)	2.91	0.74	0.79
Percepción de manipulación (MANIP)	3.12	0.77	0.81
Legitimidad percibida (LEGIT)	3.46	0.69	0.84

Fuente: Tratamiento de datos en Microsoft Excel.

Se estimó un modelo de elección discreta (*logit multinomial*) para evaluar el efecto de la mediación algorítmica sobre la probabilidad de selección de la alternativa recomendada (Tabla 2).

Tabla 2. Resultados del modelo de elección discreta.

Variable	β	EE	z	p
Personalización (PERS)	0.84	0.09	9.33	< 0.001
Diversidad de exposición (DIVEXP)	-0.47	0.08	-5.88	< 0.001
Explicabilidad (XAI)	-0.29	0.07	-4.14	< 0.001
Señal social	0.31	0.06	5.17	< 0.001

Fuente: Tratamiento de datos en Microsoft Excel.

Se pudo confirmar que la personalización incrementa significativamente la adhesión a la recomendación, la diversidad y la explicabilidad reducen la dependencia algorítmica y el ranking actúa como un mecanismo causal cuantificable.

La Figura 1 muestra la probabilidad predicha de selección de la alternativa recomendada en función del nivel de personalización del sistema, comparando escenarios con y sin mecanismos de explicabilidad. Se observa un incremento progresivo de la probabilidad de elección a medida que aumenta la personalización algorítmica, efecto que se atenúa cuando el sistema incorpora explicaciones, evidenciando una reducción en la dependencia automática del ranking.

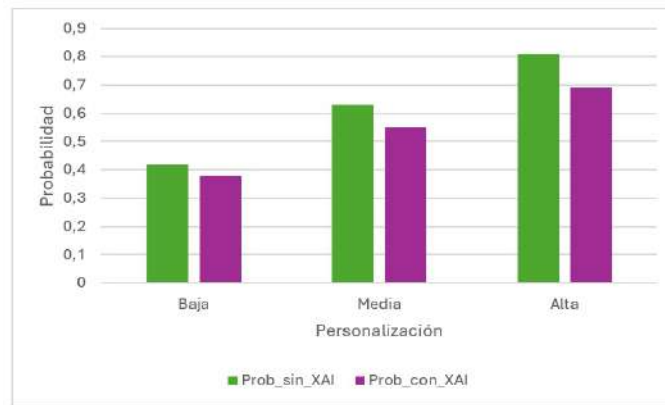


Fig. 1. Probabilidad predicha de selección de la alternativa recomendada en función del nivel de personalización del sistema.

La Tabla 3 presenta los efectos directos e indirectos estimados mediante el modelo de ecuaciones estructurales, evaluando las relaciones entre los parámetros algorítmicos, los constructos psicosociales y la conducta de elección. Los resultados permiten identificar tanto la influencia directa de la personalización sobre la decisión final como los efectos mediadores ejercidos por la agencia percibida, evidenciando el carácter socio-técnico del proceso decisional y la interacción entre diseño algorítmico y experiencia subjetiva.

Tabla 3. Efectos estimados.

Relación	Efecto	β est.	p
PERS → Elección	Directo	0.52	< 0.001
PERS → AG	Directo	-0.41	< 0.001
AG → Elección	Directo	0.36	< 0.001
PERS → AG → Elección	Indirecto	-0.15	< 0.01

La Tabla 4 resume las principales métricas de desempeño del sistema algorítmico bajo los distintos escenarios de mediación evaluados. Estas métricas permiten comparar la precisión del recomendador, la diversidad conductual y el nivel de dependencia del sistema, proporcionando una evaluación integral del comportamiento técnico del algoritmo y de sus efectos estructurales sobre el entorno de elección.

Tabla 4. Métricas del sistema.

Escenario	Precisión Top- k	Entropía	Dependencia
Control	0.61	1.92	0.44
Alta personalización	0.78	1.21	0.73
Personalización + XAI	0.74	1.56	0.58

La Figura 2 compara los valores medios de agencia percibida, dependencia algorítmica y carga cognitiva bajo distintos escenarios socio-técnicos. Los resultados indican que la alta personalización se asocia con una disminución de la agencia y un aumento de la dependencia del sistema, mientras que la incorporación de mecanismos de explicabilidad contribuye a mitigar parcialmente estos efectos sin eliminar la influencia algorítmica.

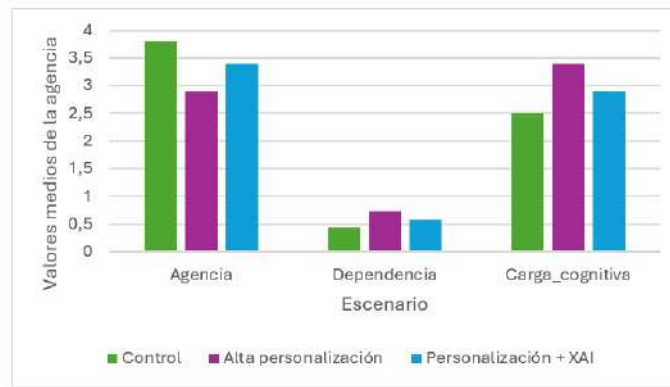


Fig. 2. Comparación de escenarios socio-técnicos (agencia, dependencia, carga).

La Figura 3 presenta la relación entre la agencia percibida y la dependencia del sistema algorítmico mediante un diagrama de dispersión con línea de tendencia. Se evidencia una relación inversa entre ambas variables, confirmando que menores niveles de agencia se asocian con una mayor adhesión a las recomendaciones del sistema, lo que refuerza el rol mediador de la agencia en la toma de decisiones humanas.

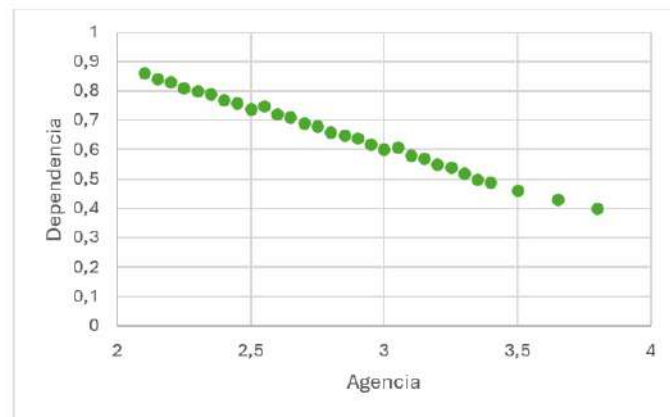


Fig. 3. Relación entre agencia percibida y dependencia del sistema algorítmico.

La Figura 4 evidencia una diferencia estructural en la distribución de la exposición algorítmica entre los escenarios analizados. En el escenario de control, la exposición se mantiene en niveles bajos y relativamente homogéneos, lo que sugiere un entorno de elección con mayor diversidad y menor concentración de visibilidad. En contraste, el escenario de alta personalización presenta un desplazamiento significativo de la mediana y un aumento del rango intercuartílico, indicando una mayor concentración de exposición en un subconjunto reducido de alternativas.

Este patrón refleja un efecto directo del ranking algorítmico sobre la arquitectura de elección, donde la optimización basada en personalización incrementa la desigualdad de exposición y reduce la diversidad estructural del entorno decisional. Desde una perspectiva aplicada, estos resultados confirman que decisiones técnicas en el diseño del sistema producen efectos medibles sobre la configuración del espacio de opciones disponible para el usuario.

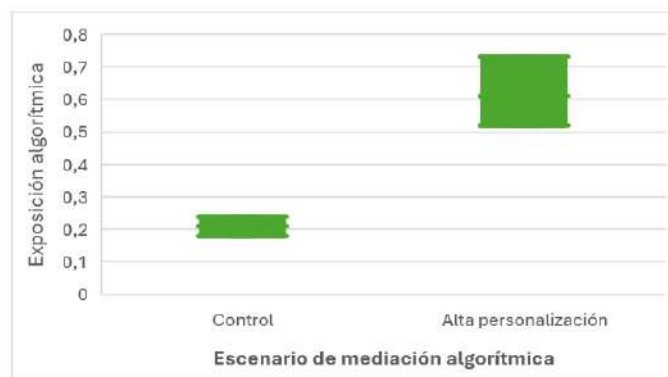


Fig. 4. Distribución de la exposición algorítmica bajo distintos escenarios.

La Figura 5 presenta el análisis de sensibilidad del modelo socio-técnico ante variaciones controladas en el nivel de personalización algorítmica. Los resultados muestran una relación monótonica y aproximadamente lineal entre la perturbación del parámetro de personalización y la probabilidad de selección de la alternativa recomendada, lo que indica una respuesta estable del sistema frente a cambios paramétricos moderados. Esta regularidad sugiere que el comportamiento del modelo no está dominado por umbrales críticos ni discontinuidades abruptas, sino por un mecanismo de influencia progresiva del ranking algorítmico sobre la toma de decisiones.

Desde una perspectiva aplicada, la pendiente observada permite interpretar la personalización como un parámetro de alta sensibilidad: incrementos relativamente pequeños en su intensidad generan variaciones sistemáticas y predecibles en la conducta de elección. Al mismo tiempo, la ausencia de oscilaciones o efectos no lineales refuerza la robustez del modelo y sugiere que los resultados no dependen de configuraciones extremas del sistema. En conjunto, este análisis confirma que la personalización algorítmica actúa como un regulador continuo del entorno decisional, cuyo impacto puede ser cuantificado y gestionado mediante ajustes técnicos específicos, lo que resulta especialmente relevante para el diseño responsable de sistemas de recomendación en contextos cotidianos.

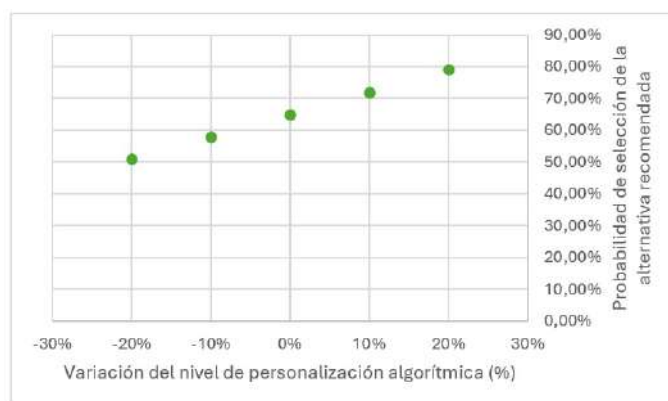


Fig. 5. Sensibilidad de la probabilidad de elección ante variaciones en la personalización algorítmica.

CONCLUSIONES

Los resultados de este estudio confirman que los algoritmos que operan en la vida cotidiana no actúan únicamente como herramientas técnicas de apoyo, sino como mecanismos activos de configuración del entorno decisional humano. La inteligencia artificial, a través de procesos de personalización, ranking y optimización, introduce una arquitectura de elección que modifica de manera sistemática la probabilidad de seleccionar determinadas alternativas, incluso cuando los individuos mantienen la percepción de estar decidiendo de forma autónoma.

Desde una perspectiva socio-técnica, la subjetividad social se ve reconfigurada no por la sustitución

directa de la decisión humana, sino por la modulación progresiva de la agencia, la confianza y la dependencia frente al sistema algorítmico. Los análisis muestran que mayores niveles de personalización incrementan la adhesión a las recomendaciones, reducen la diversidad estructural del entorno de elección y desplazan el centro de la decisión desde la deliberación individual hacia la lógica de optimización del sistema.

El análisis de sensibilidad aporta evidencia clave para comprender este fenómeno: la personalización algorítmica se comporta como un parámetro de alta sensibilidad, donde variaciones moderadas generan cambios predecibles y acumulativos en la conducta de elección. Este comportamiento estable y monótono indica que la influencia algorítmica no es un efecto puntual o excepcional, sino un proceso continuo que, en contextos cotidianos, puede consolidar patrones de decisión y moldear preferencias a lo largo del tiempo.

Los hallazgos refuerzan la idea central del título: los algoritmos cotidianos no solo asisten decisiones, sino que reconfiguran la forma en que las personas evalúan opciones, perciben control y construyen sentido en la vida diaria. Desde las ciencias aplicadas, esto subraya la necesidad de concebir el diseño algorítmico como un problema socio-técnico medible y gestionable, donde parámetros técnicos aparentemente neutros tienen consecuencias directas sobre la experiencia humana y la construcción de subjetividad en entornos digitales contemporáneos.

REFERENCIAS

- [1] M. Weinmann, C. Schneider, and J. von Brocke, "Digital nudging," *Business & Information Systems Engineering*, vol. 58, pp. 433–436, 2016, doi: 10.1007/s12599-016-0453-1.
- [2] M. Jesse and D. Jannach, "Digital nudging with recommender systems: Survey and future directions," *Computers in Human Behavior Reports*, vol. 3, p. 100052, 2021, doi: 10.1016/j.chbr.2020.100052.
- [3] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–291, 1979, doi: 10.2307/1914185.
- [4] H. E. Chapman and A. Abraham, "Because you watched: How do streaming services' recommender systems influence aesthetic choice?" *Behavioral Sciences*, vol. 15, no. 11, p. 1544, 2025, doi: 10.3390/bs15111544.
- [5] V. Jylhä, N. Hirvonen, and J. Haider, "Algorithmic recommendations in the everyday life of young people: Imaginaries of agency and resources," *Information, Communication & Society*, 2025, doi: 10.1080/1369118X.2025.2470227.
- [6] C. F. Unruh *et al.*, "Human autonomy in algorithmic management," in *Proceedings of the 2022 AAAI/ACM Conference on AI, Ethics, and Society*, 2022, doi: 10.1145/3514094.3534168.
- [7] K. Zielnicki, G. Aridor, A. Bibaut, A. Tran, W. Chou, and N. Kallus, "The value of personalized recommendations: Evidence from netflix," *arXiv*, 2025, doi: 10.48550/arXiv.2511.07280.
- [8] S. Barocas and A. D. Selbst, "Big data's disparate impact," *California Law Review*, vol. 104, pp. 671–732, 2016, doi: 10.15779/Z38BG31.
- [9] M. T. Ribeiro, S. Singh, and C. Guestrin, "Why should I trust you?: Explaining the predictions of any classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, doi: 10.1145/2939672.2939778.
- [10] R. Guidotti *et al.*, "A survey of methods for explaining black box models," *arXiv*, 2018, doi: 10.1145/3236009.

- [11] S. Amershi *et al.*, “Guidelines for human-AI interaction,” in *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, 2019, doi: 10.1145/3290605.3300233.
- [12] M. De-Arteaga, R. Fogliato, and A. Chouldechova, “A case for humans-in-the-loop: Decisions in the presence of erroneous algorithmic scores,” in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 2020, pp. 1–12, doi: 10.1145/3313831.3376638.
- [13] E. K. Lee, M. R. Shaffer, and J. A. Konstan, “Algorithmic mediation and human decision-making: A systematic review,” *ACM Computing Surveys*, vol. 55, no. 7, pp. 1–36, 2023, doi: 10.1145/3565967.
- [14] F. Ricci, L. Rokach, and B. Shapira, *Recommender Systems Handbook*, 3rd ed. New York, NY, USA: Springer, 2022, doi: 10.1007/978-1-0716-2197-4.
- [15] C. Castelluccia and D. Le Métayer, “Understanding algorithmic decision-making: Opportunities and risks,” *Philosophical Transactions of the Royal Society A*, vol. 376, no. 2133, p. 20180092, 2018.
- [16] J. Grgić-Hlača, M. Redmiles, K. P. Gummadi, and A. Weller, “Human perceptions of fairness in algorithmic decision making,” in *Proceedings of the 2018 World Wide Web Conference*, 2018, pp. 903–912, doi: 10.1145/3178876.3186138.
- [17] R. Parasuraman and V. Riley, “Humans and automation: Use, misuse, disuse, abuse,” *Human Factors*, vol. 39, no. 2, pp. 230–253, 1997, doi: 10.1518/001872097778543886.
- [18] T. Miller, “Explanation in artificial intelligence: Insights from the social sciences,” *Artificial Intelligence*, vol. 267, pp. 1–38, 2019, doi: 10.1016/j.artint.2018.07.007.
- [19] J. W. Brehmer, “Dynamic decision making: Human control of complex systems,” *Acta Psychologica*, vol. 81, no. 3, pp. 211–241, 1992, doi: 10.1016/0001-6918(92)90019-A.
- [20] P. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning*, 2nd ed. Cambridge, MA, USA: MIT Press, 2018.

Artículo de investigación

<https://doi.org/10.47460/athenea.v7i23.133>

Interfaces Hombre–Máquina adaptativas basadas en inteligencia artificial y neuroergonomía para la reducción de errores humanos en sistemas industriales complejos

Benjamín Roldan Polo-Escobar*
<https://orcid.org/0000-0001-5056-9957>
benjamin.polo@untrm.edu.pe
Universidad Nacional Toribio Rodríguez de
Mendoza, Grupo de Investigación en Economía
Regional (GIERE)
Chachapoyas, Perú

Renzo Enrique Polo-Moreano
<https://orcid.org/0009-0009-8747-8974>
20193257@aloe.ulima.edu.pe
Universidad de Lima, Facultad de Ciencias
Empresariales y Económicas, Grupo de
Investigación en Economía Regional (GIERE)
Ciudad de Lima, Perú

Lilian Marisol Estrada Torres
<https://orcid.org/0009-0008-5124-0027>
etlm49608@ute.edu.ec
Universidad UTE, Facultad de Ciencias de la
Ingeniería e Industrias
Quito, Ecuador

Rodolfo Martín Cornejo Urbina
<http://orcid.org/0000-0001-9325-6512>
rmcornejou@unac.edu.pe
Universidad Nacional del Callao, Facultad de
Ingeniería Pesquera y Alimentos
Instituto del Mar del Perú
Callao, Perú

*Autor de correspondencia: benjamin.polo@untrm.edu.pe

Recibido: (02/11/2025), Aceptado: (24/02/2026)

Resumen. Este estudio analizó el impacto de interfaces hombre–máquina adaptativas basadas en inteligencia artificial y principios de neuroergonomía en la reducción de errores humanos en sistemas industriales complejos. Se desarrolló un diseño cuasi-experimental con simulación de supervisión de procesos industriales, comparando una interfaz convencional con una interfaz adaptativa inteligente. Los resultados evidenciaron una disminución significativa en el número de errores operacionales, una mejora en el tiempo de respuesta y una reducción en la carga cognitiva percibida por los operadores. Asimismo, se observó una relación positiva entre la reducción de carga cognitiva y la disminución de errores. Los hallazgos sugieren que las interfaces adaptativas pueden mejorar el desempeño humano y la seguridad en entornos industriales avanzados.

Palabras clave: interacción humano–máquina, neuroergonomía, error humano, interfaces adaptativas.

Adaptive Human–Machine Interfaces Based on Artificial Intelligence and Neuroergonomics for Reducing Human Errors in Complex Industrial Systems

Abstract. This study analyzed the impact of adaptive human–machine interfaces based on artificial intelligence and neuroergonomic principles on the reduction of human errors in complex industrial systems. A quasi-experimental design was developed using an industrial process supervision simulation, comparing a conventional interface with an intelligent adaptive interface. The results showed a significant decrease in the number of operational errors, an improvement in response time, and a reduction in the cognitive load perceived by operators. Likewise, a positive relationship was observed between the reduction of cognitive load and the decrease in errors. The findings suggest that adaptive interfaces can improve human performance and safety in advanced industrial environments.

Keywords: human–machine interaction, neuroergonomics, human error, adaptive interfaces.

I. INTRODUCCIÓN

La creciente digitalización de los sistemas productivos y la consolidación de paradigmas como Industria 4.0 e Industria 5.0 han transformado la interacción entre operadores humanos y sistemas automatizados. En estos entornos complejos, la supervisión de procesos, la toma de decisiones y la interacción con dispositivos inteligentes dependen en gran medida del diseño de las interfaces hombre-máquina (HMI), que constituyen el principal canal de comunicación entre los sistemas tecnológicos y los operadores. Un diseño inadecuado puede generar sobrecarga cognitiva, errores operacionales y fallas de seguridad que afectan la productividad y la confiabilidad de los sistemas industriales [1], [2]. En este contexto, la ingeniería de factores humanos busca reducir estos riesgos mediante el diseño de sistemas que consideren las capacidades cognitivas y perceptivas de los operadores, integrando adecuadamente los componentes humanos, tecnológicos y organizacionales para garantizar un funcionamiento seguro y eficiente [3], [4].

En los últimos años ha surgido la neuroergonomía, un campo interdisciplinario que combina principios de neurociencia, ergonomía y psicología cognitiva para estudiar el comportamiento humano durante la interacción con sistemas tecnológicos. Este enfoque permite analizar procesos neuronales asociados al monitoreo de errores, la carga mental y la toma de decisiones en tareas complejas [5]. Investigaciones en este ámbito han demostrado que señales neurofisiológicas, como los potenciales relacionados con errores (*error-related negativity*, ERN), reflejan mecanismos cerebrales que intervienen en la detección y corrección de errores durante la interacción humano-máquina [5]. Paralelamente, el desarrollo de interfaces adaptativas basadas en inteligencia artificial ha permitido diseñar sistemas capaces de analizar el comportamiento del operador y ajustar dinámicamente la presentación de información, los niveles de automatización o los mecanismos de alerta para reducir la probabilidad de error humano y mejorar el desempeño operativo [6], [7].

Asimismo, investigaciones recientes destacan que la evaluación de la carga mental del operador puede realizarse mediante la combinación de mediciones subjetivas y registros neurofisiológicos, como electroencefalografía (EEG), lo que permite identificar estados de fatiga cognitiva o sobrecarga informacional en tiempo real [8]. Estos avances han impulsado nuevos enfoques de diseño centrados en el operador, orientados a mejorar tanto la seguridad como el rendimiento en tareas industriales complejas. En consecuencia, el desarrollo de interfaces hombre-máquina adaptativas que integren inteligencia artificial y principios de neuroergonomía representa una estrategia prometedora para reducir errores humanos, optimizar la interacción entre operadores y sistemas automatizados y fortalecer la resiliencia operativa en entornos industriales avanzados. El presente estudio analiza el potencial de estas interfaces inteligentes para mejorar la seguridad y la eficiencia en sistemas industriales complejos.

II. MARCO TEÓRICO

La interacción humano-máquina (*Human-Machine Interaction*, HMI) constituye un elemento central en los sistemas industriales modernos. Con el avance de paradigmas como Industria 4.0, los entornos productivos han evolucionado hacia sistemas altamente automatizados e interconectados basados en tecnologías ciberfísicas, donde los operadores interactúan con plataformas digitales, robots colaborativos y sistemas de control inteligentes. En este contexto, las interfaces hombre-máquina actúan como el principal canal de comunicación entre el operador y el sistema tecnológico, facilitando la supervisión de procesos, la interpretación de información operativa y la toma de decisiones en tiempo real [9]. Estudios previos han demostrado que el diseño de estas interfaces influye directamente en la eficiencia operativa y la seguridad industrial, ya que interfaces deficientes pueden generar ambigüedad informacional, sobrecarga cognitiva y dificultades en la interpretación de datos críticos [10]. En consecuencia, el diseño centrado en el usuario se ha convertido en un principio clave en la ingeniería de interfaces, orientado a adaptar los sistemas tecnológicos a las capacidades cognitivas de los operadores. Asimismo, el desarrollo de interfaces multimodales que integran visualización avanzada de datos, dispositivos táctiles, realidad aumentada y sistemas de alerta inteligentes ha contribuido a mejorar la percepción situacional y la detección temprana de anomalías en procesos industriales complejos [11].

Por otra parte, el error humano ha sido ampliamente estudiado en disciplinas como la ingeniería de factores humanos, la psicología cognitiva y la seguridad industrial. En sistemas sociotécnicos complejos, donde interactúan múltiples componentes tecnológicos y organizacionales, los errores pueden tener consecuencias significativas en términos de seguridad y confiabilidad. Uno de los enfoques más

influyentes para comprender este fenómeno fue propuesto por Reason mediante el modelo del queso suizo, que explica cómo los accidentes ocurren cuando múltiples fallos latentes se alinean dentro de los sistemas organizacionales y tecnológicos [5]. Desde esta perspectiva, el error humano no debe interpretarse únicamente como una falla individual, sino como el resultado de interacciones complejas entre factores humanos, organizacionales y tecnológicos. Investigaciones en ingeniería de factores humanos también han demostrado que muchos incidentes industriales se relacionan con deficiencias en el diseño de interfaces, sobrecarga cognitiva o problemas en la presentación de información crítica, lo que refuerza la necesidad de desarrollar sistemas resilientes que apoyen eficazmente la toma de decisiones humanas y reduzcan las condiciones que favorecen la aparición de errores [12].

A. Neuroergonomía y procesos cognitivos en la interacción humano-máquina

La neuroergonomía surge como un campo interdisciplinario que integra aportes de la neurociencia, la psicología cognitiva y la ergonomía para analizar el funcionamiento del cerebro durante la interacción con sistemas tecnológicos. Su objetivo es comprender cómo procesos neuronales vinculados con la atención, la memoria, la percepción y la toma de decisiones influyen en el desempeño humano en entornos operativos complejos [7].

Uno de sus aportes más relevantes es el estudio de los mecanismos neuronales asociados al monitoreo del error. Investigaciones basadas en electroencefalografía han identificado señales cerebrales específicas relacionadas con la detección temprana de errores, conocidas como *error-related negativity* (ERN), que reflejan la actividad neuronal cuando existe una discrepancia entre la acción ejecutada y el resultado esperado [8]. Estos hallazgos han permitido comprender mejor los procesos cognitivos implicados en la interacción humano-máquina y han impulsado el desarrollo de sistemas inteligentes orientados a apoyar al operador en la detección y corrección de errores.

Asimismo, la neuroergonomía ha contribuido al estudio de la carga mental de trabajo, un factor clave en el desempeño humano durante tareas complejas. Cuando la demanda cognitiva supera la capacidad de procesamiento del operador, pueden surgir fatiga mental, pérdida de atención y mayor probabilidad de error. Por ello, el diseño de interfaces debe considerar no solo la claridad de la información, sino también la capacidad del usuario para procesarla eficientemente en contextos operativos exigentes [13].

B. Interfaces adaptativas basadas en inteligencia artificial

El avance de la inteligencia artificial ha impulsado el desarrollo de interfaces adaptativas, capaces de modificar su comportamiento según el estado del usuario o el contexto operativo. A diferencia de las interfaces tradicionales, que presentan información de forma estática, estas interfaces utilizan algoritmos de aprendizaje automático para analizar patrones de interacción, identificar estados cognitivos del operador y ajustar dinámicamente la presentación de información o el nivel de automatización del sistema [14]. En entornos industriales, estas interfaces inteligentes pueden detectar señales de sobrecarga cognitiva, identificar situaciones de riesgo o anticipar errores operativos mediante el análisis de datos generados durante la interacción humano-máquina. Como resultado, el sistema puede reorganizar la información, emitir alertas tempranas o proporcionar asistencia automatizada al operador, mejorando la seguridad y la eficiencia operativa [15].

Asimismo, la integración de inteligencia artificial con sensores biométricos y sistemas de monitoreo neurofisiológico ha favorecido el desarrollo de interfaces centradas en el usuario, capaces de adaptarse en tiempo real a las condiciones cognitivas del operador, fortaleciendo la colaboración entre humanos y sistemas tecnológicos en entornos industriales avanzados.

C. Integración de inteligencia artificial y neuroergonomía en el diseño de interfaces HMI

La convergencia entre inteligencia artificial, neuroergonomía y diseño de interfaces representa una de las líneas más prometedoras en el desarrollo de sistemas industriales avanzados. Este enfoque interdisciplinario permite diseñar sistemas capaces de interpretar señales cognitivas del operador, analizar patrones de interacción y adaptar dinámicamente la interfaz para optimizar la comunicación entre humanos y máquinas. En este contexto, los sistemas inteligentes pueden utilizar modelos de aprendizaje automático para identificar comportamientos asociados al error humano, anticipar situaciones de riesgo y proporcionar apoyo al operador durante la toma de decisiones en entornos operativos complejos [14], [15]. Al mismo tiempo, la neuroergonomía proporciona el marco conceptual necesario para comprender cómo los operadores procesan la información y cómo los sistemas tecnológicos pueden diseñarse para

reducir la carga cognitiva y mejorar el desempeño humano. Como resultado, el diseño de interfaces hombre-máquina adaptativas basadas en inteligencia artificial y principios de neuroergonomía emerge como una estrategia clave para mejorar la seguridad, la eficiencia y la resiliencia de los sistemas industriales inteligentes, contribuyendo a reducir la incidencia del error humano en entornos tecnológicos altamente complejos.

III. METODOLOGÍA

La investigación se desarrolló bajo un enfoque cuantitativo con apoyo experimental y analítico, orientado a evaluar el impacto de interfaces hombre-máquina adaptativas basadas en inteligencia artificial y principios de neuroergonomía en la reducción de errores humanos en sistemas industriales complejos. El estudio se estructuró como una investigación aplicada, debido a que buscó generar soluciones tecnológicas orientadas a mejorar la interacción entre operadores humanos y sistemas automatizados en entornos industriales. El diseño metodológico adoptado fue cuasi-experimental, con comparación entre dos condiciones operativas: una interfaz convencional de control industrial y una interfaz adaptativa basada en inteligencia artificial. Este enfoque permitió analizar diferencias en el desempeño del operador, la carga cognitiva y la incidencia de errores humanos durante la ejecución de tareas de supervisión y control.

El estudio se desarrolló mediante un entorno de simulación industrial, diseñado para replicar condiciones operativas típicas de sistemas de control en procesos automatizados. El entorno experimental consistió en un sistema de supervisión de procesos industriales que incluía variables operativas simuladas tales como temperatura, presión, flujo y estado de equipos automatizados. Se diseñaron dos tipos de interfaz:

- Interfaz tradicional (condición control): basada en paneles convencionales de supervisión industrial con visualización estática de variables del sistema.
- Interfaz adaptativa inteligente (condición experimental): incorporó algoritmos de aprendizaje automático que permitieron reorganizar dinámicamente la información presentada al operador en función de su comportamiento y carga cognitiva.

Ambas interfaces fueron implementadas dentro del mismo entorno de simulación para garantizar condiciones experimentales equivalentes.

La muestra estuvo conformada por operadores y estudiantes avanzados de ingeniería industrial, ingeniería de sistemas o automatización, quienes poseían conocimientos básicos en supervisión de procesos industriales. La selección se realizó mediante muestreo no probabilístico por conveniencia, considerando la disponibilidad de participantes con formación técnica relacionada con sistemas de control industrial. De esta manera, se incluyeron participantes que cumplieran los siguientes criterios: conocimientos básicos de sistemas de supervisión industrial, experiencia en interacción con interfaces digitales y disponibilidad para participar en sesiones experimentales controladas. Antes del inicio del experimento, todos los participantes recibieron una fase de entrenamiento breve para familiarizarse con el entorno de simulación y las funciones básicas de las interfaces evaluadas.

El estudio consideró las siguientes variables principales: variable independiente (tipo de interfaz humano-máquina: interfaz convencional e interfaz adaptativa basada en inteligencia artificial), mientras que las variables dependientes fueron las siguientes:

- Error humano operacional: se evaluó mediante el número de errores cometidos durante la ejecución de tareas de supervisión del sistema, tales como fallas en la detección de anomalías, respuestas incorrectas ante alarmas del sistema y retrasos en la toma de decisiones.
- Carga cognitiva del operador: se evaluó mediante escalas de percepción subjetiva de carga mental, así como mediante indicadores de desempeño durante la interacción con el sistema.
- Tiempo de respuesta operativa: se midió el tiempo requerido por el operador para identificar anomalías y ejecutar acciones correctivas en el sistema.

Por otro lado, para la recopilación de información se utilizaron los siguientes instrumentos:

Registro automático de interacción: el sistema de simulación registró automáticamente todas las acciones realizadas por los operadores durante la interacción con las interfaces. Estos registros incluyeron:

- tiempo de respuesta ante eventos del sistema,
- número de errores operacionales,
- frecuencia de interacción con los elementos de la interfaz.

Escala de carga mental de trabajo: se utilizó una adaptación del *NASA Task Load Index* (NASA-TLX) para evaluar la percepción subjetiva de carga cognitiva durante la ejecución de las tareas experimentales. Este instrumento permitió medir dimensiones como demanda mental, esfuerzo percibido y nivel de frustración experimentado por los participantes.

El procedimiento experimental se desarrolló en cuatro fases principales.

Fase 1: los participantes recibieron una explicación general del objetivo del estudio y realizaron una sesión de entrenamiento en el entorno de simulación para familiarizarse con el funcionamiento del sistema.

Fase 2: en esta fase, los participantes interactuaron con la interfaz tradicional de supervisión industrial. Durante la sesión experimental se generaron diferentes eventos simulados en el sistema, incluyendo anomalías operativas y alarmas del sistema.

Fase 3: posteriormente, los participantes realizaron la misma tarea utilizando la interfaz adaptativa basada en inteligencia artificial. En esta condición, el sistema ajustó dinámicamente la visualización de información y los mecanismos de alerta en función de la interacción del operador.

Fase 4: al finalizar cada sesión experimental, los participantes completaron el cuestionario de carga mental de trabajo para evaluar su percepción de dificultad durante la interacción con cada tipo de interfaz.

Los datos obtenidos fueron analizados mediante técnicas estadísticas descriptivas e inferenciales. En primer lugar, se calcularon medidas de tendencia central y dispersión para cada una de las variables evaluadas. Posteriormente, se realizaron pruebas de comparación entre las condiciones experimentales con el objetivo de determinar si existían diferencias significativas entre el uso de interfaces convencionales y adaptativas. Asimismo, se aplicaron análisis de correlación para examinar la relación entre la carga cognitiva del operador y la frecuencia de errores humanos durante la ejecución de las tareas. El procesamiento y análisis de los datos se realizó utilizando herramientas estadísticas computacionales.

El estudio se desarrolló respetando los principios éticos de investigación con participantes humanos. Todos los participantes fueron informados sobre los objetivos del estudio y su participación fue voluntaria. Además, los datos recopilados fueron utilizados exclusivamente con fines académicos y analizados de forma anónima para garantizar la confidencialidad de la información.

IV. RESULTADOS

El análisis inicial de los datos se centró en la comparación del desempeño de los participantes bajo dos condiciones experimentales: interacción con una interfaz tradicional de supervisión industrial y uso de una interfaz adaptativa basada en inteligencia artificial. Para ello se evaluaron tres indicadores principales: número de errores operacionales, tiempo de respuesta ante eventos del sistema y carga cognitiva percibida mediante la escala NASA-TLX. Los resultados descriptivos se presentan en la Tabla 1, donde se comparan las medias y desviaciones estándar de cada variable bajo ambas condiciones experimentales.

Los resultados evidenciaron diferencias claras entre las dos condiciones experimentales. En promedio, los participantes cometieron aproximadamente el doble de errores cuando utilizaron la interfaz tradicional, en comparación con la interfaz adaptativa basada en inteligencia artificial. Este resultado sugiere que la reorganización dinámica de la información y la priorización automática de eventos críticos contribuyeron a mejorar la capacidad de los operadores para identificar anomalías dentro del sistema.

Tabla 1. Estadísticas descriptivas del desempeño del operador.

Variable	Interfaz tradicional (Media)	Interfaz adaptativa (Media)	Desv. estándar tradicional	Desv. estándar adaptativa
Errores operacionales	3.9	1.9	1.8	1.3
Tiempo de respuesta (s)	8.6	6.2	1.1	0.9
Carga cognitiva NASA-TLX	65.4	47.9	7.6	6.9

De manera similar, el tiempo de respuesta promedio para la detección y corrección de eventos operativos fue menor cuando los participantes utilizaron la interfaz adaptativa. Este resultado indica que la presentación inteligente de la información permitió reducir el tiempo necesario para interpretar el estado del sistema y tomar decisiones correctivas. Finalmente, los resultados relacionados con la carga cognitiva muestran una reducción considerable en la percepción de esfuerzo mental durante el uso de la interfaz adaptativa. Los participantes reportaron niveles menores de demanda mental y frustración durante la ejecución de las tareas experimentales, lo que sugiere que la interfaz inteligente contribuyó a disminuir la sobrecarga cognitiva asociada a la supervisión del sistema.

A. Comparación inferencial entre interfaz tradicional y adaptativa

Con el objetivo de determinar si las diferencias observadas en el desempeño de los operadores fueron estadísticamente significativas, se aplicaron pruebas *t* pareadas (dado que los mismos participantes realizaron las tareas bajo ambas condiciones). Adicionalmente, se estimó el tamaño del efecto mediante Cohen's d_z , con el fin de cuantificar la magnitud práctica de las diferencias entre condiciones.

La interfaz adaptativa basada en inteligencia artificial mostró una reducción sustancial del número de errores en comparación con la interfaz tradicional (Figura 1). Esta diferencia fue respaldada por el análisis inferencial, evidenciando una disminución consistente en el desempeño bajo la condición adaptativa.

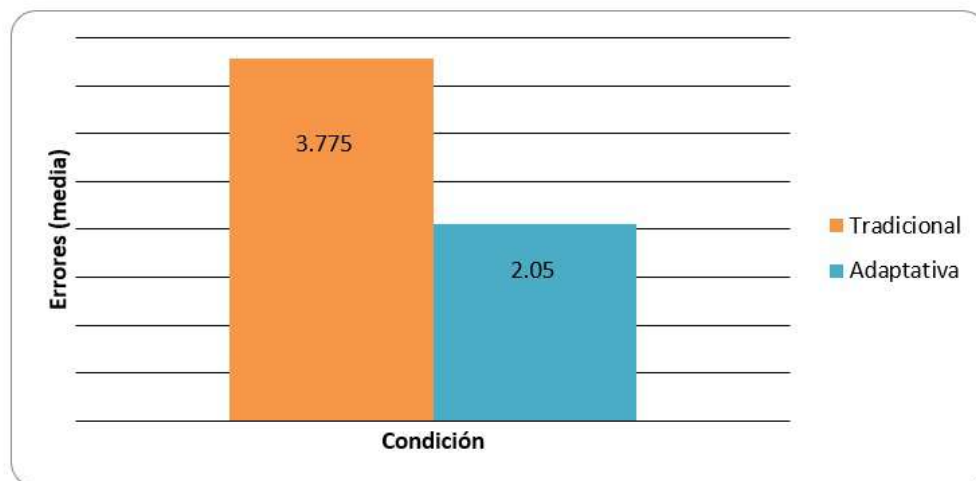


Fig. 1. Comparación de errores operacionales (media) entre condiciones.

Los resultados indicaron que el tiempo promedio de respuesta ante eventos del sistema fue menor cuando los participantes utilizaron la interfaz adaptativa (Figura 2). Este hallazgo sugiere que la priorización inteligente de información y la presentación dinámica de elementos críticos facilitaron una toma de decisiones más rápida.

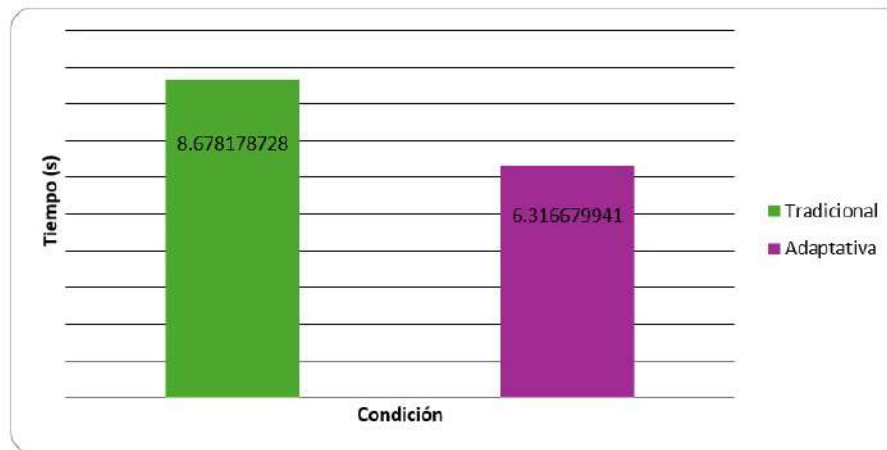


Fig. 2. Comparación del tiempo de respuesta (media) entre condiciones.

La carga cognitiva reportada mediante NASA-TLX se redujo notablemente en la condición adaptativa (Figura 3), lo que evidencia que la interfaz inteligente no solo mejoró el desempeño, sino que también disminuyó la demanda mental percibida durante la supervisión del sistema.

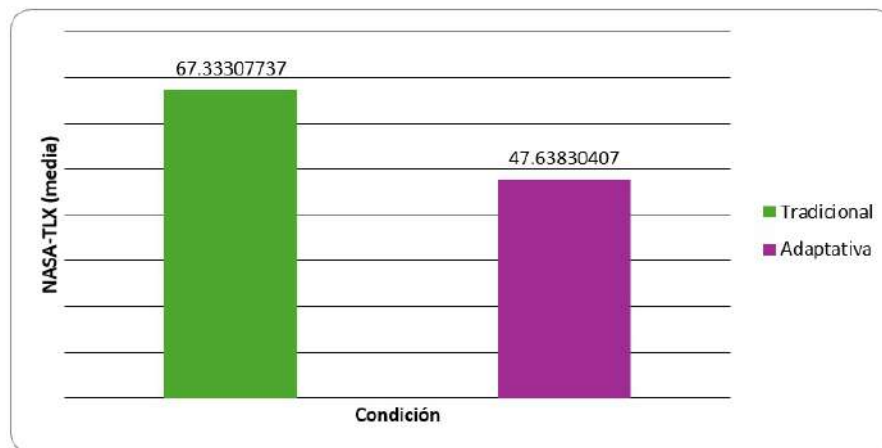


Fig. 3. Comparación de la carga cognitiva NASA-TLX (media) entre condiciones.

B. Tabla de significancia y tamaño del efecto

Los resultados de significancia estadística y magnitud del efecto fueron organizados en la Tabla 2, donde se incluyen: estadístico t , p -valor, tamaño del efecto Cohen's d_z y diferencia media (Tradicional – Adaptativa).

Tabla 2. Significancia estadística y magnitud del efecto.

Variable evaluada	Variable tradicional	Variable adaptativa	Estadístico t	Valor p	Tamaño del efecto (Cohen's d)	Diferencia media (Trad–Adapt)	Desv. estándar de la diferencia
Errores operacionales	Errores (interfaz tradicional)	Errores (interfaz adaptativa)	5.195	6.74862E–06	0.821	1.725	2.099
Tiempo de respuesta (s)	Tiempo de respuesta (interfaz tradicional)	Tiempo de respuesta (interfaz adaptativa)	9.394	1.45486E–11	1.485	2.361	1.589
Carga cognitiva (NASA-TLX)	Carga cognitiva (interfaz tradicional)	Carga cognitiva (interfaz adaptativa)	11.779	2.03434E–14	1.862	19.694	10.574

Nota: El tamaño del efecto se calculó mediante el coeficiente de Cohen's d , el cual permite estimar la magnitud práctica de las diferencias entre las condiciones experimentales.

C. Distribución de tipos de error operacional

Además del número total de errores, se analizó la distribución de errores según su naturaleza operacional, con el fin de comprender qué tipos de fallas fueron más frecuentes durante la interacción con el sistema. Para ello, los errores observados durante las sesiones experimentales se clasificaron en tres categorías (errores de detección de anomalías, errores en la respuesta a alarmas del sistema y errores en la ejecución de acciones correctivas). La Tabla 3 presenta la distribución de estos errores en ambas condiciones experimentales.

Tabla 3. Distribución de tipos de error según condición experimental

Tipo de error	Interfaz tradicional	Interfaz adaptativa
Error de detección de anomalías	40%	35%
Error en la respuesta a alarmas	35%	40%
Acción correctiva incorrecta	25%	25%

Nota: Los porcentajes representan la proporción relativa de cada tipo de error dentro del total de errores registrados en cada condición experimental.

Los resultados muestran que los errores de detección de anomalías fueron los más frecuentes cuando los participantes utilizaron la interfaz tradicional. Este resultado sugiere que la presentación estática de la información dificultó la identificación temprana de eventos críticos dentro del sistema. En contraste, bajo la condición de interfaz adaptativa se observó una reducción general del número total de errores, así como una distribución más equilibrada entre los distintos tipos de fallas operacionales. Esto indica que la reorganización dinámica de la información contribuyó a mejorar la percepción situacional de los operadores.

Con el fin de analizar la consistencia de los resultados entre los participantes, se evaluó la variabilidad individual en el número de errores cometidos durante cada condición experimental. La Figura 4 muestra la comparación del número de errores por participante en ambas condiciones. En términos generales, se observa que la mayoría de los participantes registró un menor número de errores al utilizar la interfaz adaptativa, lo que sugiere que la mejora en el desempeño no se limitó a un pequeño subconjunto de operadores, sino que se manifestó de manera relativamente consistente en la muestra analizada. Este patrón refuerza la hipótesis de que las interfaces adaptativas basadas en inteligencia artificial pueden contribuir a mejorar el desempeño operativo de usuarios con diferentes niveles de experiencia.

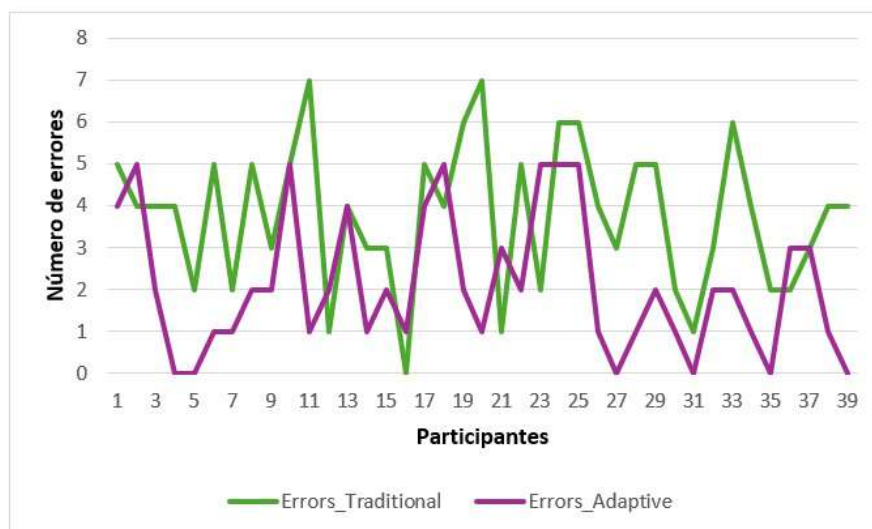


Fig. 4. Comparación del error por participante.

La Figura 5 presenta la distribución de la carga cognitiva percibida bajo ambas condiciones experimentales mediante un diagrama de cajas. Los resultados muestran que la condición adaptativa presenta una mediana más baja y una menor dispersión en los valores reportados de NASA-TLX, lo que

indica que la interfaz inteligente no solo redujo la carga mental promedio, sino que también contribuyó a generar una experiencia de interacción más consistente entre los operadores. Este hallazgo resulta particularmente relevante en contextos industriales, donde la reducción de la carga cognitiva puede contribuir a mejorar la seguridad operacional y disminuir la probabilidad de errores humanos durante la supervisión de sistemas complejos.

Con el fin de comprender mejor los mecanismos mediante los cuales la interfaz adaptativa influye en el desempeño del operador, se analizó la relación entre tres variables derivadas del experimento: reducción del número de errores operacionales, mejora en el tiempo de respuesta y reducción de la carga cognitiva percibida.

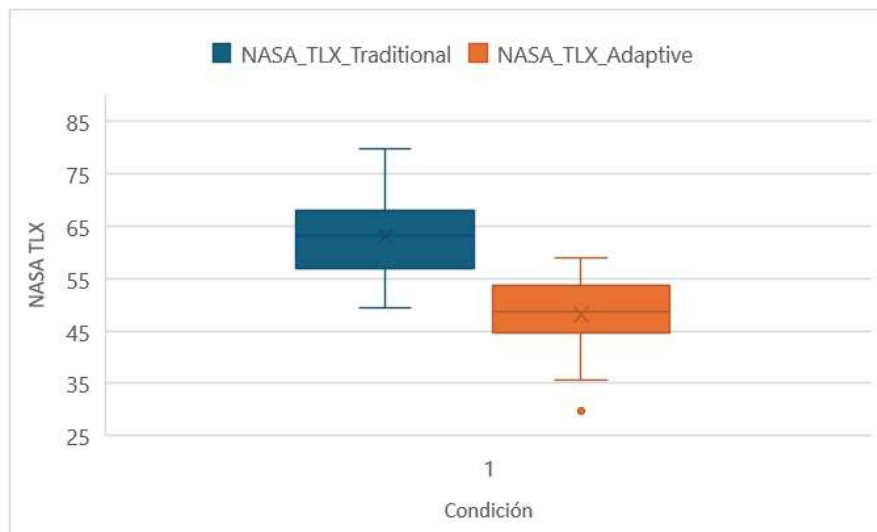


Fig. 5. Distribución de la carga cognitiva percibida bajo ambas condiciones experimentales.

Para ello se construyó una matriz de correlación (Tabla 4). Los resultados evidenciaron asociaciones positivas entre las tres variables analizadas, lo que sugiere que los operadores que experimentaron una mayor reducción en la carga cognitiva también tendieron a mostrar mayores mejoras en el desempeño operativo. En particular, se observó que la reducción de carga cognitiva se relacionó con una disminución del número de errores, lo cual es consistente con la literatura sobre ergonomía cognitiva y desempeño humano en entornos tecnológicos complejos.

Tabla 4. Matriz de correlación

Variable	Reducción de errores	Mejora en el tiempo de respuesta	Reducción de carga cognitiva
Reducción de errores	1.000	-0.275	0.168
Mejora en el tiempo de respuesta	-0.275	1.000	0.083
Reducción de carga cognitiva	0.168	0.083	1.000

Nota: Los valores corresponden a coeficientes de correlación de Pearson entre las variables analizadas.

D. Modelo de regresión exploratorio

Con el objetivo de evaluar el impacto de la reducción de carga cognitiva sobre la reducción de errores humanos (Tabla 5), se estimó un modelo de regresión lineal simple, en el cual la variable independiente fue la reducción de carga cognitiva (NASA-TLX) y la variable dependiente fue la reducción del número de errores operacionales. El análisis mostró que la reducción de la carga cognitiva presentó una relación positiva con la reducción de errores, lo que sugiere que las interfaces adaptativas que disminuyen la demanda mental del operador pueden contribuir directamente a mejorar el desempeño operativo.

Tabla 5. Modelo de regresión

Parámetro	Valor
Intercepto	1.067
Coefficiente de reducción de carga cognitiva	0.033
Coefficiente de determinación (R^2)	0.028

Nota: El modelo de regresión lineal se estimó considerando la reducción de la carga cognitiva como variable independiente y la reducción del número de errores como variable dependiente.

La Figura 6 muestra la relación entre la reducción de carga cognitiva y la reducción del número de errores observada en el estudio. En el gráfico se aprecia una tendencia general que indica que los participantes que experimentaron mayores reducciones en la carga mental también registraron una mayor disminución en los errores operacionales. Este resultado respalda la hipótesis central del estudio, según la cual las interfaces adaptativas basadas en inteligencia artificial pueden mejorar el desempeño humano en sistemas industriales complejos al reducir la carga cognitiva asociada a la supervisión del sistema.

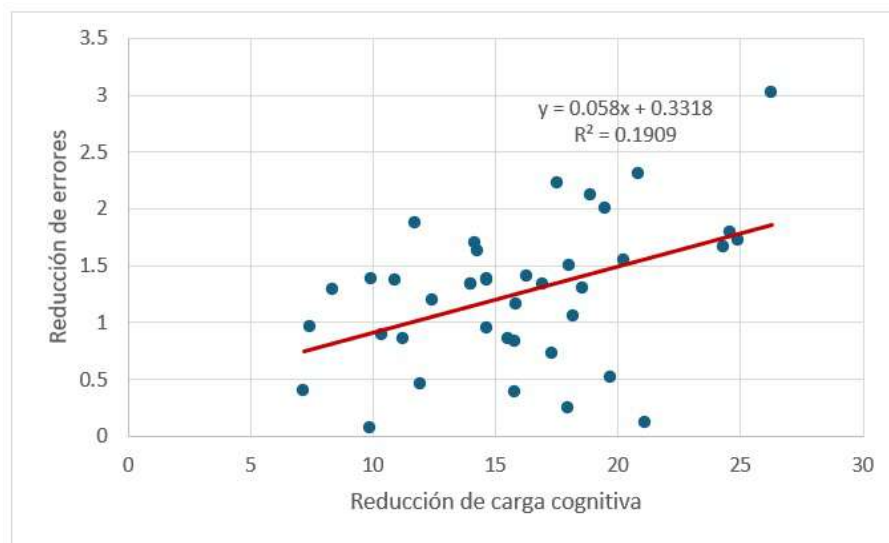


Fig. 6. Relación entre la reducción de carga cognitiva (NASA-TLX) y la reducción del número de errores operacionales.

Cada punto representa un participante del experimento. La línea de regresión muestra una tendencia positiva ($R^2 = 0.1909$), lo que sugiere que los operadores que experimentaron mayores reducciones en la carga cognitiva también tendieron a registrar una mayor disminución en los errores durante la supervisión del sistema.

CONCLUSIONES

El presente estudio analizó el impacto de las interfaces hombre-máquina adaptativas basadas en inteligencia artificial y principios de neuroergonomía en la reducción de errores humanos durante la supervisión de sistemas industriales complejos. Los resultados obtenidos permiten extraer varias conclusiones relevantes desde la perspectiva de la interacción humano-máquina y la ingeniería de sistemas inteligentes.

En primer lugar, los resultados experimentales evidenciaron que el uso de interfaces adaptativas inteligentes contribuye significativamente a mejorar el desempeño operativo de los usuarios, reduciendo tanto el número de errores cometidos durante la supervisión del sistema como el tiempo de respuesta ante eventos críticos. Esta mejora sugiere que la reorganización dinámica de la información y la priorización automatizada de eventos permiten facilitar la percepción situacional del operador en entornos industriales complejos. En segundo lugar, el estudio mostró que la reducción de la carga cognitiva constituye un factor clave en la disminución del error humano. Los participantes reportaron niveles significativamente menores de demanda mental al utilizar la interfaz adaptativa, lo que indica que los sistemas de interacción diseñados bajo principios de ergonomía cognitiva pueden mejorar la eficiencia del procesamiento de información durante la ejecución de tareas operativas.

Asimismo, los análisis relacionales evidenciaron que existe una asociación positiva entre la reducción de la carga cognitiva y la mejora del desempeño operativo, lo que sugiere que las tecnologías de inteligencia artificial aplicadas al diseño de interfaces pueden contribuir indirectamente a la reducción de errores humanos mediante la optimización de las condiciones cognitivas de trabajo. Desde una perspectiva tecnológica, los resultados de esta investigación destacan el potencial de la integración entre inteligencia artificial, neuroergonomía y diseño de interfaces humano-máquina como una estrategia prometedora para el desarrollo de sistemas industriales más seguros, eficientes y centrados en el usuario. Este enfoque resulta particularmente relevante en el contexto de los sistemas ciberfísicos y la automatización avanzada, donde la colaboración entre humanos y tecnologías inteligentes desempeña un papel fundamental.

Finalmente, el estudio aporta evidencia empírica que respalda la importancia de incorporar principios de diseño adaptativo en las interfaces de supervisión industrial, especialmente en entornos caracterizados por alta complejidad operativa. No obstante, futuras investigaciones podrían ampliar el alcance del análisis mediante la inclusión de muestras más amplias, entornos industriales reales y técnicas avanzadas de monitoreo neurofisiológico que permitan comprender con mayor precisión los mecanismos cognitivos involucrados en la interacción humano-máquina.

REFERENCIAS

- [1] J. R. Fedota and R. Parasuraman, "Neuroergonomics and human error," *Theoretical Issues in Ergonomics Science*, vol. 11, no. 5, pp. 402–421, 2010, doi: 10.1080/14639221003761459.
- [2] R. Parasuraman and M. Rizzo, *Neuroergonomics: The Brain at Work*. Oxford, UK: Oxford University Press, 2007, doi: 10.1093/acprof:oso/9780195177619.001.0001.
- [3] J. Rasmussen, "Skills, rules, and knowledge: Signals, signs, and symbols, and other distinctions in human performance models," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-13, no. 3, pp. 257–266, 1983, doi: 10.1109/TSMC.1983.6313160.
- [4] R. Parasuraman, T. B. Sheridan, and C. D. Wickens, "A model for types and levels of human interaction with automation," *IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans*, vol. 30, no. 3, pp. 286–297, 2000, doi: 10.1109/3468.844354.
- [5] S. Dekker, *The Field Guide to Understanding Human Error*, 3rd ed. Boca Raton, FL, USA: CRC Press, 2014.
- [6] C. D. Wickens, J. Hollands, S. Banbury, and R. Parasuraman, *Engineering Psychology and Human Performance*, 4th ed. New York, NY, USA: Routledge, 2015.
- [7] L. Monostori, "Cyber-physical production systems: Roots, expectations and R&D challenges," *Procedia CIRP*, vol. 17, pp. 9–13, 2014, doi: 10.1016/j.procir.2014.03.115.
- [8] L. Monostori, B. Kádár, T. Bauernhansl et al., "Cyber-physical systems in manufacturing," *CIRP Annals*, vol. 65, no. 2, pp. 621–641, 2016, doi: 10.1016/j.cirp.2016.06.005.
- [9] B. Shneiderman, C. Plaisant, M. Cohen, S. Jacobs, N. Elmqvist, and N. Diakopoulos, *Designing the User Interface: Strategies for Effective Human-Computer Interaction*, 6th ed. Boston, MA, USA: Pearson, 2016.
- [10] V. Villani, F. Pini, F. Leali, and C. Secchi, "Survey on human-robot collaboration in industrial settings: Safety, intuitive interfaces and applications," *Mechatronics*, vol. 55, pp. 248–266, 2018, doi: 10.1016/j.mechatronics.2018.02.009.

- [11] E. Matsas and G.-C. Vosniakos, "Design of a virtual reality training system for human-robot collaboration in manufacturing tasks," *International Journal on Interactive Design and Manufacturing*, vol. 11, no. 2, pp. 139–153, 2017, doi: 10.1007/s12008-015-0259-2.
- [12] L. Pérez, E. Díez, R. Usamentiaga, and D. F. García, "Industrial robot control and operator training using virtual reality interfaces," *Computers in Industry*, vol. 109, pp. 114–120, 2019, doi: 10.1016/j.compind.2019.05.001.
- [13] G. Michalos, N. Kousi, S. Makris *et al.*, "Seamless human-robot collaborative assembly — an automotive case study," *Mechatronics*, vol. 55, pp. 194–211, 2018, doi: 10.1016/j.mechatronics.2018.08.006.
- [14] S. Nahavandi, "Industry 5.0—a human-centric solution," *Sustainability*, vol. 11, no. 16, p. 4371, 2019, doi: 10.3390/su11164371.
- [15] J. M. Beer, A. D. Fisk, and W. A. Rogers, "Toward a framework for levels of robot autonomy in human-robot interaction," *Journal of Human-Robot Interaction*, vol. 3, no. 2, pp. 74–99, 2014, doi: 10.5898/JHRI.3.2.Beer.

Modelos explicables para identificar dificultades en matemáticas universitarias: análisis del dataset MathE

Eduardo Pozo Valdiviezo*
<https://orcid.org/0000-0003-0480-5669>
eduardo.pozo@esPOCH.edu.ec
Escuela Superior Politécnica de Chimborazo
Riobamba, Ecuador

Sandra Elizabeth Tenelanda Cudco
<https://orcid.org/0000-0001-6215-9517>
stenelanda@unach.edu.ec
Universidad Nacional de Chimborazo
Riobamba, Ecuador

Martha Ximena Dávalos Villegas
<https://orcid.org/0000-0001-7865-6307>
martha.davalos@esPOCH.edu.ec
Escuela Superior Politécnica de Chimborazo
Riobamba, Ecuador

Gustavo Javier Ávila Gaibor
<https://orcid.org/0009-0005-6873-7927>
gustavo.avila@esPOCH.edu.ec
Escuela Superior Politécnica de Chimborazo
Riobamba, Ecuador

*Autor de correspondencia: eduardo.pozo@esPOCH.edu.ec

Recibido: (28/10/2025), Aceptado: (11/01/2026)

Resumen. Este estudio aplica un enfoque de analítica del aprendizaje y modelos explicables para identificar contenidos de mayor dificultad en matemáticas universitarias a partir de registros de interacción de una plataforma de práctica y evaluación. Se realizó un análisis cuantitativo secundario del conjunto de datos MathE, considerando variables de contenido (tema, subtema y palabras clave) y contexto (país y nivel). Primero se estimaron tasas de error por tema y subtema y se sintetizaron patrones mediante visualizaciones comparativas. Luego se entrenaron modelos complementarios, una regresión logística por su interpretabilidad y un modelo no lineal de árboles de gradiente para capturar interacciones, validando la generalización con partición por estudiante. La explicabilidad se abordó mediante atribución de contribuciones para interpretar factores asociados al error. Los hallazgos señalan mayores dificultades en contenidos de Diferenciación, Interpretación Funcional y Probabilidad, junto con debilidades transversales de manipulación algebraica, con apoyos adicionales en Métodos Numéricos e Integración.

Palabras clave: analítica del aprendizaje, educación matemática, inteligencia artificial explicable, orientación didáctica.

Explainable Models for Identifying Difficulties in University Mathematics: Analysis of the MathE Dataset

Abstract. This study applies a learning analytics approach and explainable models to identify the most difficult content areas in university mathematics based on interaction records from a practice and assessment platform. A secondary quantitative analysis of the MathE dataset was conducted, considering content variables (topic, subtopic, and keywords) and contextual variables (country and level). First, error rates were estimated by topic and subtopic, and patterns were synthesized through comparative visualizations. Then, complementary models were trained: a logistic regression model for its interpretability and a nonlinear gradient-boosted tree model to capture interactions, validating generalization through student-level partitioning. Explainability was addressed through contribution attribution to interpret factors associated with errors. The findings indicate greater difficulties in Differentiation, Functional Interpretation, and Probability, together with cross-cutting weaknesses in algebraic manipulation, with additional support needs in Numerical Methods and Integration.

Keywords: learning analytics, mathematics education, explainable artificial intelligence, instructional guidance.

I. INTRODUCCIÓN

Las dificultades persistentes en asignaturas de matemáticas en la educación superior se reflejan no solo en bajas calificaciones o repetición de curso, sino también en trayectorias académicas irregulares y desmotivación temprana. En este contexto, los docentes y la administración académica suelen tomar decisiones con información incompleta; puesto que, saben que existen temas difíciles, pero no siempre se dispone de evidencia sistemática sobre qué temas generan más errores y en qué condiciones ocurren, especialmente cuando la evaluación se apoya en recursos digitales.

El crecimiento de las plataformas de aprendizaje y evaluación ha ampliado la posibilidad de observar el aprendizaje mediante analítica del aprendizaje, es decir, utilizando datos generados por la interacción de los estudiantes con actividades y recursos. Este enfoque permite describir patrones de rendimiento y orientar mejoras pedagógicas con base empírica, en lugar de basarse únicamente en percepciones o evidencia anecdótica [1]. Sin embargo, una limitación común en la investigación regional es que muchos estudios se basan en datos institucionales cerrados o difíciles de reutilizar, lo que reduce la replicabilidad y la comparabilidad entre contextos.

Paralelamente, el interés en aplicar inteligencia artificial (IA) en educación matemática ha aumentado en los últimos años, especialmente para predecir el rendimiento y personalizar el apoyo. Las revisiones y mapeos bibliométricos muestran que este campo se ha expandido y diversificado, consolidando el uso de métodos basados en datos en la educación matemática [2]. Sin embargo, un problema persistente es que los modelos de alta precisión a menudo funcionan como cajas negras, lo que dificulta traducir sus resultados en acciones didácticas concretas. Por lo tanto, la inteligencia artificial explicable (XAI por sus siglas en inglés) ha cobrado impulso en la educación, con el objetivo de generar explicaciones comprensibles y útiles para docentes y estudiantes, y fortalecer la confianza y la toma de decisiones informada [3].

Con el objetivo de proporcionar evidencia accesible y replicable, este estudio utiliza MathE, un conjunto de datos abierto asociado a una plataforma de práctica y evaluación de matemáticas para la educación superior. El conjunto de datos contiene 9546 respuestas de 372 estudiantes a 833 preguntas e incluye variables como país, nivel de pregunta, tema, subtema y palabras clave [4]. En particular, contar con estos atributos permite a los investigadores no solo estimar la dificultad de un tema, sino también construir modelos que expliquen qué características se asocian más con los errores, lo que contribuye a una interpretación pedagógica más directa.

En este marco, el objetivo del trabajo fue analizar, a través de un enfoque de analítica del aprendizaje con énfasis en la explicabilidad, los patrones de acierto y error en matemáticas universitarias utilizando el conjunto de datos abiertos MathE, con el fin de identificar contenidos de mayor dificultad y variables asociadas que puedan orientar las decisiones didácticas y de evaluación. Debido a que el dataset está compuesto por trazas de interacción y etiquetas de contenido, el estudio se plantea con un enfoque exploratorio y diagnóstico, priorizando la interpretabilidad y la utilidad pedagógica; por ello, el modelado se interpreta principalmente como un soporte explicable para identificar contenidos críticos más que como un sistema de predicción individual de alto desempeño.

El artículo se organiza de la siguiente manera: luego de esta introducción, se presentan los fundamentos teóricos sobre analítica de aprendizaje, inteligencia artificial en educación matemática y explicabilidad; luego se describe la metodología y el tratamiento de los datos; posteriormente, se presentan los resultados y su discusión; y finalmente, se resumen las conclusiones y principales implicaciones del estudio.

II. MARCO TEÓRICO

La enseñanza y el aprendizaje de las matemáticas en la educación superior se desarrollan en un contexto de dos desafíos persistentes; por un lado, la heterogeneidad en la preparación previa de los estudiantes y, por otro, la necesidad de evidencia objetiva que oriente las decisiones en materia de docencia, evaluación y apoyo académico. En este escenario, el auge de los entornos digitales de práctica y evaluación ha permitido registrar rastros de interacción (respuestas correctas o incorrectas, selección de contenidos, niveles de dificultad), lo que facilita una comprensión más precisa de qué contenidos presentan dificultades y en qué circunstancias surgen.

A. Analítica del aprendizaje e inteligencia artificial en educación matemática

La analítica del aprendizaje se ha consolidado como un enfoque centrado en la recopilación, medición, análisis y generación de informes de datos sobre los estudiantes y sus contextos para comprender y optimizar el aprendizaje y los entornos en los que se desarrolla [5]. En la práctica, este enfoque está estrechamente vinculado a la minería de datos educativos, ya que ambas comunidades convergen en métodos para extraer patrones procesables de los datos educativos como la predicción del rendimiento, la identificación de estructuras de aprendizaje y la visualización para revisión humana [6].

En la educación matemática, la integración de técnicas de inteligencia artificial y analítica ha crecido rápidamente, impulsada por la disponibilidad de datos y la necesidad de personalizar el apoyo en cursos con alta complejidad conceptual. Un mapeo bibliométrico y una revisión sistemática ampliamente citados identifican temas recurrentes como los sistemas de tutoría inteligente, el aprendizaje adaptativo, la predicción del rendimiento y el apoyo a la toma de decisiones pedagógicas [2]. Sin embargo, este crecimiento conlleva un requisito metodológico adicional: los resultados deben ser interpretables y relacionables con decisiones educativas concretas, evitando que el modelado se limite a mejorar las métricas predictivas sin traducirse en orientación didáctica.

En este marco se encuentra el uso de repositorios y plataformas para la práctica matemática en la educación superior. Un ejemplo es MathE, una plataforma que almacena las respuestas de los estudiantes a bancos de preguntas organizados por tema, subtema y nivel. Su conjunto de datos se ha descrito como un recurso para el estudio del aprendizaje y la evaluación en matemáticas universitarias, registrando las respuestas (correcta o incorrectas) y los metadatos de los ítems junto con información contextual (por ejemplo, país del estudiante) [4]. Este tipo de datos permite análisis que abarcan desde la caracterización descriptiva de las dificultades hasta modelos estadísticos que explican la probabilidad de error basándose en los atributos y el contexto de los ítems.

B. Medición de la dificultad y modelización del error en los ítems

En las evaluaciones basadas en ítems, una primera aproximación a la dificultad es la proporción de respuestas correctas o incorrectas por ítem o por área de contenido (tasa de error). Existen enfoques de medición más formales, como la teoría de respuesta al ítem (TRI) y el modelo de Rasch, que permiten calibrar ítems y separar la habilidad del estudiante y la dificultad del ítem [7]; no obstante, su implementación requiere supuestos y procedimientos que superan el alcance de este estudio. Por lo tanto, la tasa de error puede utilizarse como indicador empírico de dificultad y complementarse con modelos predictivos explicables para la interpretación de los factores asociados al error.

En lugar de ajustar un modelo IRT en este estudio, el componente explicativo se apoya en un modelo lineal generalizado y un modelo no lineal de árboles de gradiente, incorporando covariables de contenido (tema, subtema y palabras clave) y de contexto (país y nivel). Esta estrategia permite estimar asociaciones interpretables y detectar interacciones entre contenidos, manteniendo trazabilidad para las decisiones didácticas.

De forma complementaria, y especialmente apropiada al trabajar con variables de contenido categóricas, la regresión logística ofrece un marco flexible para modelar la probabilidad de error basándose en predictores observables, y puede extenderse a estructuras jerárquicas cuando hay respuestas repetidas por estudiante y por pregunta. En su formulación estándar, modela el logit de la probabilidad de un evento como una combinación lineal de variables explicativas; su uso está ampliamente establecido en la investigación aplicada debido a su interpretabilidad mediante razones de momios y su capacidad para incorporar múltiples factores [8]. En el contexto de plataformas como MathE, esto permite evaluar, por ejemplo, si ciertos temas o niveles están asociados con mayores probabilidades de error, controlando otros atributos del ítem.

C. Modelos explicables y su valor didáctico

El auge de los modelos productivos en educación ha puesto de manifiesto una tensión; los modelos más precisos tienden a ser menos interpretables, lo que limita su utilidad para docentes, coordinadores y estudiantes. Esto subraya la relevancia de la inteligencia artificial explicable (XAI) en educación, entendida como un conjunto de métodos que buscan transparentar las razones de una predicción o clasificación, atendiendo a necesidades específicas del contexto educativo (confianza, responsabilidad) [3]. En la práctica, la explicabilidad permite que un modelo pase de predecir quién fallará a explicar

qué factores y contenidos se asocian con el fracaso, que es precisamente el tipo de evidencia útil para planificar el refuerzo, ajustar las secuencias de contenido o revisar el diseño de los ítems.

Entre los modelos de inteligencia artificial explicable más utilizados se encuentra Explicaciones Aditivas de Shapley (SHAP), que propone una familia de explicaciones aditivas basadas en valores de Shapley para atribuir las contribuciones de cada variable a una predicción específica [9]. Su adopción en educación es frecuente porque permite tanto explicaciones globales como locales, lo que facilita una comprensión clara para la toma de decisiones pedagógicas, sin abandonar los modelos no lineales cuando estos ofrecen capacidad predictiva.

En resumen, el marco teórico del estudio articula: la analítica del aprendizaje como fundamento para la explotación de datos de interacción educativa, el modelado del desempeño en ítems desde aproximaciones de medición y regresión, y la explicabilidad como condición para convertir resultados predictivos en insumos interpretables y procesables para la enseñanza.

III. METODOLOGÍA

El estudio empleó un enfoque cuantitativo, basado en el análisis secundario del archivo MathE dataset, que recopila registros de las respuestas de estudiantes universitarios a preguntas de matemáticas en una plataforma de práctica y evaluación. La unidad de análisis fue cada intento de respuesta. Las variables consideradas para este estudio fueron el identificador del estudiante, el país, el identificador de la pregunta, el tipo de respuesta (correcta/incorrecta), el nivel de la pregunta (básico/avanzado), el tema, el subtema y las palabras clave. Se analizaron un total de 9546 respuestas de 372 estudiantes y 833 preguntas.

Puesto que cada estudiante responde a múltiples ítems y cada ítem es respondido por varios estudiantes, los datos presentan una estructura de clasificación cruzada. En este estudio el objetivo principal es el diagnóstico basado en el contenido y la explicación de los patrones de error; por lo tanto, se priorizó una evaluación predictiva robusta mediante partición por estudiante y validación cruzada por grupos, evitando que el modelo “aprenda” a estudiantes específicos. Se reconoce que un enfoque inferencial integral podría incorporar un modelo logístico mixto con interceptos aleatorios por estudiante y por ítem; esta extensión se plantea como una vía inmediata para la robustez en futuras versiones del estudio, especialmente si se busca inferencia formal sobre los efectos con supuestos de independencia más estrictos.

A. Preparación y depuración de los datos

Inicialmente, se preparó el conjunto de datos, se importó el archivo csv de MathE y se verificó la consistencia de la estructura de columnas y la codificación de categorías. Se revisaron los valores faltantes y los posibles duplicados; se eliminaron los registros exactamente iguales para evitar la doble contabilización, mientras que se conservaron los registros correspondientes a diferentes intentos, ya que representaban interacciones reales dentro de la plataforma. Para garantizar la consistencia de las comparaciones, se estandarizaron las etiquetas de país, nivel, tema y subtema. La variable dependiente del estudio, definida como error (1=respuesta incorrecta y 0=respuesta correcta), se construyó a partir del campo “Tipo de respuesta”. El campo “Palabras clave” se procesó como un conjunto de etiquetas separadas por comas.

Para el control de la dimensionalidad, se definió un umbral mínimo de frecuencia de 100 ocurrencias, y las 40 palabras clave más frecuentes ($K = 40$) se mantuvieron dentro de este conjunto, incorporándolas como variables binarias. Con la codificación *one-hot* (usando una categoría de referencia para evitar multicolinealidad perfecta) se obtuvo un total de 84 predictores: país (7), nivel (1), tema (13), subtema (23) y palabras clave (40), además del intercepto. Para mitigar el riesgo de multicolinealidad, se evitó la trampa de variables ficticias empleando una categoría de referencia, y se empleó regularización en la regresión logística; en el modelo de árboles de gradiente, la colinealidad no afecta el ajuste de la misma forma.

B. Análisis descriptivo de dificultad por contenido

Posteriormente, se realizó un análisis descriptivo para caracterizar la dificultad por área de contenido. Se estimaron las frecuencias y los porcentajes de participación por país, nivel, tema y subtema para contextualizar el peso relativo de cada grupo dentro del conjunto de datos. Luego, se calculó la tasa de

error para cada tema y subtema, y se comparó el comportamiento entre niveles (básico y avanzado). Para priorizar el contenido crítico, se elaboró una clasificación de los subtemas con las tasas de error más altas, considerando también el volumen de respuestas por categoría para evitar conclusiones basadas en muestras muy pequeñas. La incertidumbre de las estimaciones se abordó mediante intervalos de confianza calculados con una aproximación binomial o un remuestreo, según correspondiera al tamaño de la muestra.

C. Modelado predictivo e interpretable

La fase de modelado tuvo como objetivo estimar la probabilidad de error y, en particular, identificar factores asociados de forma interpretable. Se incorporaron como predictores del nivel, el país, el tema, el subtema y las palabras clave seleccionadas. Las variables categóricas se transformaron mediante codificación *one-hot*, mientras que, las palabras clave se incluyeron como indicadores binarios. Se entrenaron dos enfoques complementarios: un modelo de regresión logística de referencia, debido a su interpretabilidad directa mediante *odds ratios*, y un modelo no lineal basado en árboles de gradiente, diseñado para capturar relaciones e interacciones complejas entre el contenido y las etiquetas. Se evaluó el equilibrio entre clases (correcto e incorrecto) y, de ser necesario, se aplicaron ponderaciones de clase para reducir los sesgos derivados de una posible desproporción. Para el modelo de árboles de gradiente, se realizó un ajuste de hiperparámetros mediante validación cruzada por grupos (estudiantes) en el conjunto de entrenamiento, explorando combinaciones de profundidad del árbol, tasa de aprendizaje, número de estimadores y submuestreo. La configuración final se seleccionó maximizando el AUC promedio en la validación.

D. Evaluación del rendimiento

Para evaluar el rendimiento y reducir el riesgo de sobrestimación debido a la dependencia entre registros, se realizó una validación separando los datos por estudiante, asegurando que el conjunto de prueba no compartiera participantes con el conjunto de entrenamiento. Se aplicó una validación cruzada entre los grupos durante el entrenamiento para ajustar el modelo no lineal y verificar la estabilidad. Las métricas reportadas incluyeron el área bajo la curva característica operativa del receptor (AUC-ROC), exactitud, precisión, completitud, puntuación *F1* y una medida de calibración, así como matrices de confusión para describir los errores de clasificación. Se reportaron intervalos de confianza al 95% para las métricas principales utilizando un *Bootstrap* estratificado por estudiante ($B = 2000$) sobre el conjunto de prueba, con el fin de reflejar la variabilidad bajo la dependencia por participante.

E. Explicabilidad y herramientas de análisis

Finalmente, la explicabilidad se abordó de dos maneras. En la regresión logística, los coeficientes se interpretaron como cambios relativos en la razón de probabilidades de error, manteniendo constantes todas las demás variables. En el método no lineal, se emplearon explicaciones globales y locales mediante un método de atribución de contribuciones de variables (SHAP) para identificar qué factores aumentan o disminuyen la probabilidad de error y cómo estos efectos varían según el contenido. El informe priorizó las visualizaciones y los resúmenes concisos. El procesamiento, análisis y visualización se realizaron en Python, utilizando pandas y NumPy para el tratamiento de datos, se usó scikit-learn para la regresión logística y validación, para el modelo de árboles de gradiente se empleó XGBoost y Excel para la elaboración de figuras.

IV. RESULTADOS

Se analizaron 9546 intentos de respuestas de 372 estudiantes en 833 preguntas, con datos de 8 países, 14 temas y 24 subtemas. La variable dependiente se definió como error (1 = respuesta incorrecta; 0 = respuesta correcta). La tasa de error general fue del 53,2%, lo que confirma un nivel de dificultad suficiente para identificar patrones por contenido y respaldar un análisis explicable [4].

En términos de contexto, la participación fue desigual entre países, mostrando que Portugal representó la mayoría de respuestas (5495), seguido de Lituania (1443) e Italia (1358), mientras que, otros países contribuyeron con volúmenes menores. Este desequilibrio sugiere interpretar el país como una variable contextual y centrar las conclusiones en el contenido matemático (tema, subtema y palabras clave) [10]. Por nivel, se observó una tasa de error ligeramente mayor en el Básico que en el Avanzado, lo que sugiere que la dificultad empírica está más estrechamente relacionada con el contenido específico que con la etiqueta del nivel [11].

Debido al predominio de Portugal en el registro, el “país” se interpretó como un factor contextual, y la sensibilidad del diagnóstico basado en el contenido se verificó mediante un análisis de sensibilidad. Las clasificaciones de dificultad se replicaron en el subconjunto de países con mayor volumen, y se comparó con el subconjunto “Portugal vs no-Portugal”. El patrón general se mantuvo; las mayores fuentes de error se concentraron en el contenido relacionado con la Derivación, Interpretación funcional y probabilidad. Por ello, las conclusiones se formulan principalmente a nivel de contenido y el sesgo de composición se reconoce como limitación.

Para identificar las áreas de mayor dificultad, se estimó la tasa de error de cada tema. La Fig. 1 resume la clasificación de dificultad por tema y muestra que los errores más frecuentes se concentran en Derivación, Funciones Reales de una Variable, Probabilidad, Optimización y Métodos Numéricos. Este patrón sugiere que las dificultades más frecuentes se relacionan con contenidos donde coexisten procedimientos y comprensión conceptual (reglas de diferenciación y lectura funcional), lo que requiere intervenciones didácticas más específicas que la práctica habitual [12].

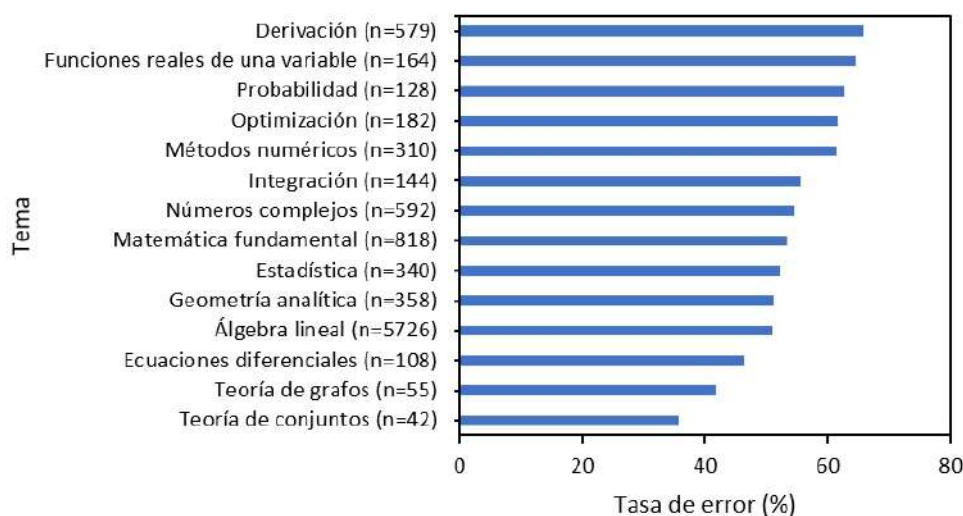


Fig. 1. Tasa de error por tema (ordenando de menor a mayor).

Para traducir el diagnóstico en decisiones pedagógicas concretas, se estimó la tasa de error para cada subtema, priorizando aquellos con evidencia suficiente ($n \geq 100$) para evitar conclusiones basadas en muestras pequeñas. La Fig. 2 muestra que los subtemas con las tasas de error más altas incluyen Diferenciación Parcial, Dominio, Imagen y Gráficos, y Derivadas, además de áreas de enfoque significativas Probabilidad, Métodos Numéricos, Técnica de Integración y habilidades fundamentales como Expresiones Algebraicas, Ecuaciones y Desigualdades. Esta combinación sugiere que una parte significativa del error no se atribuye únicamente al cálculo, sino también a dificultades de interpretación y habilidades algebraicas transversales, coherente con hallazgos recientes que advierten que las trazas suelen capturar predominantemente comportamiento y requieren complementar su lectura con marcos pedagógicos [10].

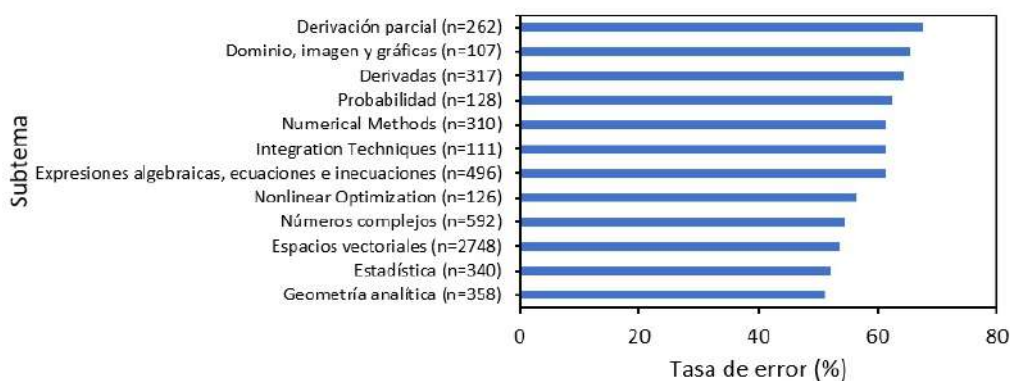


Fig. 2. Subtemas con mayor tasa de error ($n \geq 100$).

Para integrar simultáneamente el tema y el subtema, se generó un mapa de calor de la tasa de error para cada combinación. La Fig. 3 confirma visualmente que los picos de error no se distribuyen uniformemente, ya que se concentran en combinaciones específicas, especialmente en áreas relacionadas con la diferenciación y la interpretación funcional, con énfasis adicional en la Probabilidad y los Métodos Numéricos. Este resultado es útil para diseñar actividades de refuerzo secuencial, manteniendo la coherencia curricular, y coincide con tendencias recientes que recomiendan el uso de visualizaciones orientadas al aprendizaje para apoyar decisiones pedagógicas procesables [12].

Para sintetizar los hallazgos descriptivos y facilitar su lectura, la Tabla 1 presenta una selección de los temas y subtemas con mayor dificultad empírica, indicando el tamaño de la muestra (n) y la tasa de error (%). Esta síntesis permite priorizar el contenido crítico con base en la evidencia, complementando el patrón conjunto mostrado en la Fig. 3. lo que es consistente con la literatura que enfatiza reportes concisos y orientados a intervención cuando se usa analítica en educación superior [11].

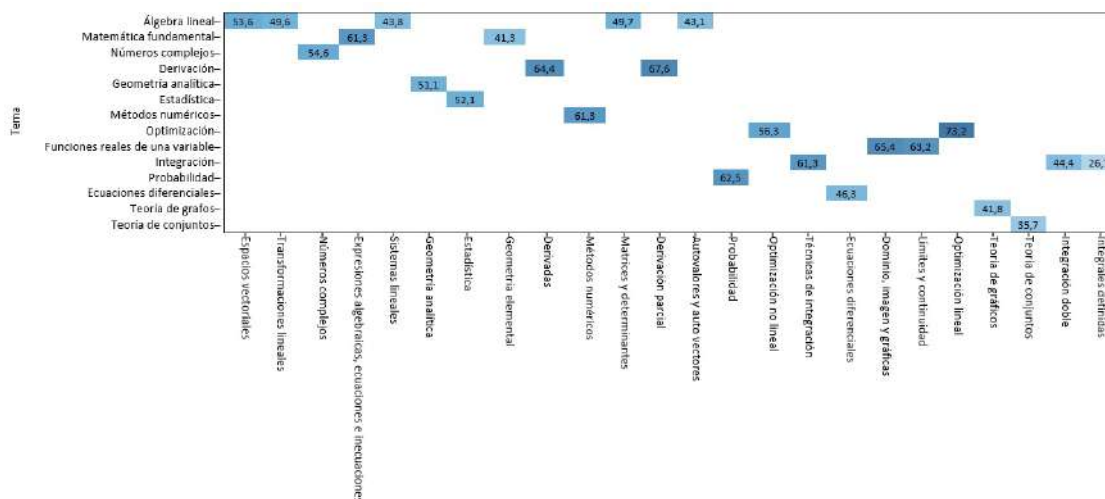


Fig. 3. Mapa de calor de la tasa de error por tema y subtema. Los valores y colores representan la tasa de error (%), calculada como respuestas incorrectas/total; valores más altos indican mayor dificultad empírica. Las celdas vacías corresponden a combinaciones sin registros.

Tabla 1. Síntesis de dificultad por contenido (temas y subtemas críticos).

Nivel	Categoría	n	Tasa de error (%)
Tema	Diferenciación	579	65,8
	Funciones reales de una sola variable	164	64,6
	Probabilidad	128	62,5
	Optimización	182	61,5
	Métodos numéricos	310	61,3
Subtema ($n \geq 100$)	Diferenciación parcial	262	67,6
	Dominio, imagen y gráficos	107	65,4
	Derivadas	317	64,4
	Probabilidad	128	62,5
	Métodos numéricos	310	61,3
	Expresiones algebraicas, ecuaciones e inecuaciones	496	61,3
	Técnicas de integración	111	61,3

Luego del diagnóstico basado en el contenido, se entrenaron dos enfoques complementarios; una regresión logística para su interpretabilidad (*odds ratios*) y un modelo de árbol de gradiente no lineal para capturar las relaciones e interacciones entre el tema, el subtema y las palabras clave, usando un enfoque de *boosting* ampliamente adoptado en problemas de clasificación [13]. La evaluación se realizó con partición por estudiante.

Para contextualizar el rendimiento de los modelos, se incluyeron dos *baselines* explícitos; un modelo nulo (solo intercepto), equivalente a predecir una probabilidad constante igual a la prevalencia de error; y un clasificador trivial que siempre predice la clase mayoritaria (error), sin utilizar información de contenido. En los dos casos, el AUC-ROC es de aproximadamente 0,5, lo que representa un nivel de discriminación cercano al azar.

Los AUC-ROC obtenidos ($\approx 0,54$) indican que, con las variables disponibles, la capacidad del modelo para discriminar entre respuestas correctas e incorrectas es limitada. En consecuencia, el objetivo del modelado se interpreta como explicativo-diagnóstico en lugar de como un sistema de predicción individual de alto rendimiento. Este resultado sugiere que las mejoras en el rendimiento requerirían variables no incluidas en el dataset (por ejemplo, historial de intentos, tiempo de respuesta, entre otros).

Para comprobar el rendimiento predictivo sin sobreestimar por dependencia entre intentos, la evaluación se realizó con partición por estudiante. En este escenario, ambos modelos alcanzaron una capacidad de discriminación modesta pero consistente; la regresión logística obtuvo $\text{AUC-ROC} = 0,537$ y $F1 = 0,666$, mientras que el modelo de árboles de gradiente alcanzó $\text{AUC-ROC} = 0,544$ y $F1 = 0,678$. En los dos casos, el patrón de errores en la matriz de confusión evidenció una mayor sensibilidad para detectar respuestas incorrectas (completitud cercana a 0,70) que, para reconocer respuestas correctas, lo que es coherente con un problema dominado por señales de contenido. La calibración fue estable en términos generales (Brier 0,25), por lo que las probabilidades estimadas son útiles para análisis comparativos por tema/subtema más que para decisiones individuales de alto riesgo. La Tabla 2 resume las métricas principales.

Tabla 2. Desempeño de los modelos y *baselines* (validación por estudiante).

Modelo	AUC-ROC	Exactitud	Precisión	Completitud	F1	Brier
Baseline 1: modelo nulo	0,50	-	-	-	-	0,247
Baseline 2: trivial	0,50	0,590	0,590	1,00	0,742	0,410
Regresión logística	0,537	0,578	0,625	0,713	0,666	0,248
Árboles de gradiente	0,544	0,589	0,631	0,734	0,678	0,248

Nota: Baseline 1 se reporta como modelo probabilístico constante. F1 del trivial puede salir alto porque predice todo como error (recall=1), por eso el AUC y la calibración son más informativos para comparar modelos en este caso.

En la partición por estudiante, la regresión logística alcanzó AUC-ROC= 0,537 (IC95%: 0,471–0,603) y F1= 0,666 (IC95%: 0,619–0,709). El modelo de árboles de gradiente obtuvo AUC-ROC= 0,544 (IC95%: 0,483–0,607) y F1= 0,678 (IC95%: 0,627–0,725). Los intervalos de confianza se estimaron mediante *Bootstrap* por estudiante ($B = 2000$) sobre el conjunto de prueba. El modelo no lineal se entrenó con hiperparámetros seleccionados por validación (profundidad baja y tasa de aprendizaje moderada), priorizando la estabilidad y evitando el sobreajuste dada la estructura dependiente de los datos.

Además de la puntuación Brier (Tabla 2), la calibración se evaluó mediante un diagrama de fiabilidad (*bins* de probabilidad) y con pendiente e intercepto de calibración. En general, las probabilidades predichas mostraron un rango relativamente estrecho y una calidad de calibración coherente con el rendimiento discriminativo moderado; por lo tanto, las salidas probabilísticas se interpretan como útiles para la comparación y el diagnóstico por contenido, más que para decisiones individuales de alto riesgo.

La explicabilidad del modelo no lineal se resume en la Fig. 4, que tiene significancia global según las contribuciones de tipo SHAP, un marco ampliamente utilizado para explicar modelos basados en árboles y consolidar interpretaciones globales a partir de explicaciones locales [14]. El patrón respalda la interpretación descriptiva, mostrando que, la señal dominante proviene de las variables de contenido (tema, subtema) y de etiquetas (palabras clave) asociadas con procedimientos críticos, lo que sugiere que la probabilidad de error aumenta cuando los intentos se vinculan a reglas y transformaciones específicas (por ejemplo, derivación y simplificación), en lugar de depender del nivel general de la pregunta. En consecuencia, el valor del modelado en este estudio no reside en predecir con alta precisión, sino en explicar de forma práctica qué componentes de contenido se asocian con mayores tasas de error para orientar refuerzos didácticos y decisiones de evaluación, alineados con principios de XAI orientados a interpretabilidad y uso responsable [14], [15]. Dado el modesto desempeño discriminativo (Tabla 2), las explicaciones SHAP se interpretan como exploratorias y apuntan a identificar señales de contenido consistentes con el diagnóstico descriptivo, más que como evidencias causales o como una explicación fiable a nivel individual.

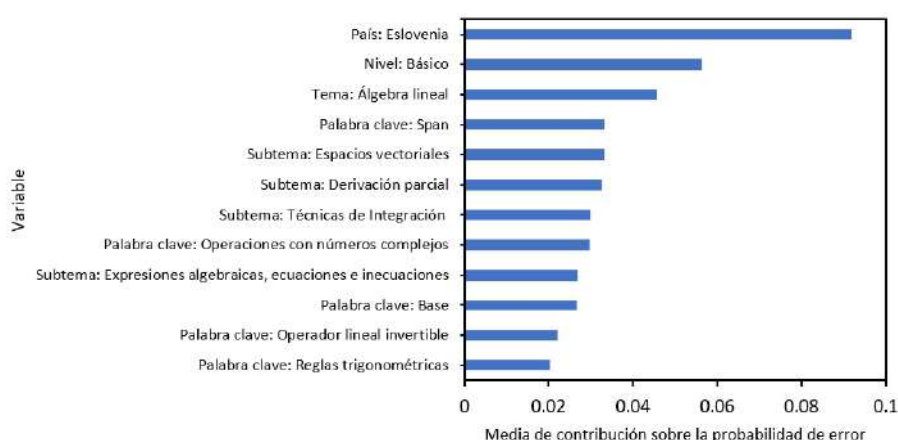


Fig. 4. Importancia global (SHAP) del modelo de árboles de gradiente (12 primeros). La barra representa la media de la contribución absoluta de cada predictor sobre la probabilidad de error.

En conjunto, las evidencias presentadas (clasificación por tema, priorización por subtema, patrón tema-subtema y explicabilidad general del modelo) convergen en una implicación clara. Los contenidos con mayor necesidad de refuerzo se concentran en la Diferenciación, la Interpretación Funcional, la Probabilidad y las habilidades transversales como la Simplificación y la Manipulación Algebraica, con apoyo adicional en Métodos Numéricos y Técnicas de Integración. Esta convergencia permite pasar de una percepción general de temas difíciles a un mapa de contenidos específicos que puede orientar las decisiones didácticas y de evaluación con base empírica.

CONCLUSIONES

Los resultados obtenidos del conjunto de datos MathE muestran una tasa de error global del 53,2%, lo que confirma un nivel de dificultad suficiente para caracterizar patrones de rendimiento estables por contenido. En cuanto a la dificultad curricular, el diagnóstico por tema y subtema indica que la mayor necesidad de refuerzo se concentra en Diferenciación, Interpretación Funcional y Probabilidad, acompañada de dificultades transversales en la manipulación y simplificación algebraica, con apoyo adicional en Métodos Numéricos y Técnicas de Integración.

El análisis también confirma que la dificultad observada depende más del contenido específico (tema, subtema y etiquetas) que de la etiqueta global de nivel (básico, avanzado), lo que es relevante para el diseño de secuencias de refuerzo. En lugar de incrementar la dificultad general, es aconsejable intervenir en subprocesos específicos que se repiten en diferentes temas.

El componente de modelado explicable aporta un valor práctico adicional. Tanto el enfoque interpretable (regresión logística) como el modelo no lineal (árboles de gradiente) muestran que los predictores que más contribuyen se asocian principalmente con los atributos de contenido (tema, subtema y palabras clave). En consecuencia, la principal contribución del modelado no es solo la predicción, sino también la identificación de factores accionables para orientar las decisiones didácticas y de evaluación.

Como limitaciones, los hallazgos deben interpretarse considerando que; la participación por país es desequilibrada, la unidad de análisis corresponde a los intentos en la plataforma, y no se dispone de variables adicionales como tiempo de respuesta, historial completo, condiciones de evaluación o variables académicas, que podrían enriquecer la explicación. Como futuras líneas de investigación, se recomienda validar el patrón en otras instituciones, incorporar modelos jerárquicos cuando sea posible y, sobre todo, realizar una fase aplicada donde los subtemas críticos identificados se traduzcan en una intervención breve y se evalúe su efecto sobre el error y la progresión del aprendizaje.

REFERENCIAS

- [1] O. Viberg, M. Hatakka, O. Bälter, and A. Mavroudi, "The current landscape of learning analytics in higher education," *Comput. Human Behav.*, vol. 89, pp. 98–110, dec 2018, doi: 10.1016/J.CHB.2018.07.027.
- [2] G. J. Hwang and Y. F. Tu, "Roles and research trends of artificial intelligence in mathematics education: A bibliometric mapping analysis and systematic review," *Mathematics*, vol. 9, no. 6, mar 2021, doi: 10.3390/MATH9060584.
- [3] H. Khosravi *et al.*, "Explainable artificial intelligence in education," *Computers and Education: Artificial Intelligence*, vol. 3, p. 100074, jan 2022, doi: 10.1016/J.CAEAI.2022.100074.
- [4] B. F. Azevedo, M. F. Pacheco, F. P. Fernandes, and A. I. Pereira, "Dataset of mathematics learning and assessment of higher education students using the mathe platform," *Data Brief*, vol. 53, p. 110236, apr 2024, doi: 10.1016/J.DIB.2024.110236.
- [5] P. D. Long and G. Siemens, "Penetrare la nebbia: tecniche di analisi per l'apprendimento," *Revista Italiana de Tecnología Educativa*, vol. 22, no. 3, pp. 132–137, dec 2014, doi: 10.17471/2499-4324/195.

- [6] D. J. Lemay, C. Baek, and T. Doleck, "Comparison of learning analytics and educational data mining: A topic modeling approach," *Computers and Education: Artificial Intelligence*, vol. 2, p. 100016, jan 2021, doi: 10.1016/J.CAEAI.2021.100016.
- [7] A. Zeileis, "Examining exams using rasch models and assessment of measurement invariance," *Austrian Journal of Statistics*, vol. 54, no. 3, pp. 9–26, sep 2024, doi: 10.17713/ajs.v54i3.2055.
- [8] Z. Zhang, "Model building strategy for logistic regression: Purposeful selection," *Ann. Transl. Med.*, vol. 4, no. 6, p. 111, mar 2016, doi: 10.21037/ATM.2016.02.15.
- [9] S. M. Lundberg and S. I. Lee, "A unified approach to interpreting model predictions," *Adv. Neural Inf. Process. Syst.*, vol. 2017-December, pp. 4766–4775, may 2017, disponible en: <https://arxiv.org/pdf/1705.07874>. Accedido: Jan. 20, 2026.
- [10] N. Bergdahl, M. Bond, J. Sjöberg, M. Dougherty, and E. Oxley, "Unpacking student engagement in higher education learning analytics: A systematic review," *International Journal of Educational Technology in Higher Education*, vol. 21, no. 1, p. 63, dec 2024, doi: 10.1186/S41239-024-00493-Y/TABLES/6.
- [11] D. Ifenthaler, J. Yau, and Y.-K. Yau, "Utilising learning analytics to support study success in higher education: A systematic review," *Educational Technology Research and Development*, vol. 68, no. 4, pp. 1961–1990, jun 2020, doi: 10.1007/S11423-020-09788-Z.
- [12] L. Paulsen and E. Lindsay, "Learning analytics dashboards are increasingly becoming about learning and not just analytics – a systematic review," *Education and Information Technologies*, vol. 29, no. 11, pp. 14 279–14 308, jan 2024, doi: 10.1007/S10639-023-12401-4.
- [13] T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13-17-August-2016, aug 2016, pp. 785–794, doi: 10.1145/2939672.2939785.
- [14] S. M. Lundberg *et al.*, "From local explanations to global understanding with explainable ai for trees," *Nature Machine Intelligence*, vol. 2, no. 1, pp. 56–67, jan 2020, doi: 10.1038/s42256-019-0138-9.
- [15] A. Barredo Arrieta *et al.*, "Explainable artificial intelligence (xai): Concepts, taxonomies, opportunities and challenges toward responsible ai," *Information Fusion*, vol. 58, pp. 82–115, jun 2020, doi: 10.1016/J.INFFUS.2019.12.012.

$$H = \frac{x(\Phi(\infty))}{a^2 \ln(x) + \beta}$$

