

IoT and Machine Learning Applications for Water Quality Monitoring in Rural Communities of Emerging Countries: Systematic Review

Mendoza Tinco, Tatiana ¹; Palacios Torres, Max Tony Enzo ¹; Huamán Aguirre, Arnold Anthony ¹

¹Universidad Tecnológica del Perú, Perú, 0831381@utp.edu.pe, U22235229@utp.edu.pe, c19532@utp.edu.pe

Abstract– *Monitoring water quality remains a persistent challenge in rural communities of emerging countries, where traditional methods are often costly, slow, and difficult to implement. In recent years, the integration of Internet of Things (IoT) technologies and Machine Learning (ML) techniques has emerged as an efficient alternative to improve the accuracy, continuity, and predictive capacity of water monitoring systems. This Systematic Literature Review aimed to identify and analyze the approaches, parameters, technological platforms, and models used in studies that combine IoT and ML for water quality monitoring in rural settings. A structured search was conducted exclusively in the Scopus database, resulting in the selection of 26 primary studies, which were analyzed using a PICOC-based extraction matrix. The results show that most investigations focus on rural environments and developing countries, with rivers being the most monitored water bodies and the Water Quality Index (WQI) the most frequently employed parameter. Additionally, fixed IoT nodes and ML models for predicting physicochemical variables were the most common technological solutions. Although these technologies show strong potential, the field still presents limitations, including a lack of standardized metrics and insufficient reporting of system costs and energy autonomy. Overall, the convergence of IoT and ML represents a promising pathway for strengthening water quality monitoring in rural communities, provided that future research improves methodological rigor and technical transparency.*

Keywords– *IoT, Machine Learning, rural communities, smart sensors, water monitoring.*

Aplicación de IoT y Machine Learning para el Monitoreo del Agua en Comunidades Rurales de Países Emergentes: Una revisión sistemática

Mendoza Tinco, Tatiana ¹; Palacios Torres, Max Tony Enzo ¹; Huamán Aguirre, Arnold Anthony ¹

¹Universidad Tecnológica del Perú, Perú, 0831381@utp.edu.pe, U22235229@utp.edu.pe, c19532@utp.edu.pe

Resumen— El monitoreo de la calidad del agua es un desafío persistente para las comunidades rurales de países emergentes, donde los métodos tradicionales suelen ser costosos, lentos y poco accesibles. En los últimos años, la integración de tecnologías como el Internet de las Cosas (IoT) y las técnicas de Machine Learning (ML) han emergido como una alternativa eficiente para mejorar la precisión, continuidad y capacidad predictiva del monitoreo hídrico. Esta Revisión Sistemática de Literatura tiene como objetivo identificar y analizar los enfoques, parámetros, plataformas tecnológicas y modelos utilizados en estudios que combinan IoT y ML para la vigilancia de la calidad del agua en contextos rurales; la búsqueda se realizó exclusivamente en Scopus, seleccionando 26 estudios primarios que fueron analizados mediante una matriz basada en el método PICOC. Los resultados muestran que la mayoría de las investigaciones se concentran en entornos rurales y en países en vías de desarrollo, siendo los ríos el cuerpo de agua más monitoreado y el índice de calidad del agua (WQI) el parámetro más empleado. Asimismo, prevalecen soluciones basadas en nodos IoT fijas y modelos ML orientados a la predicción de variables físico-químicas, aunque estas tecnologías muestran un alto potencial, persisten limitaciones relacionadas con la escasa estandarización de métricas, la falta de información sobre costos y autonomía energética. Se concluye que la convergencia IoT-ML constituye una herramienta prometedora para fortalecer el monitoreo del agua en comunidades rurales, siempre que futuros estudios mejoren la rigurosidad y la transparencia técnica.

Palabras clave— Comunidades rurales, IoT, Machine Learning, monitoreo hídrico, sensores inteligentes.

I. INTRODUCCIÓN

En los últimos años, el uso de tecnologías basadas en Internet de las Cosas (IoT) y técnicas de Machine Learning (ML) han adquirido especial relevancia en el monitoreo ambiental, particularmente en la vigilancia de la calidad del agua, por su capacidad para generar datos continuos, automatizar procesos analíticos y reducir los costos asociados al monitoreo tradicional [1], [2], [3]. La integración de sensores, conectividad inalámbrica y modelos predictivos ha permitido desarrollar sistemas capaces de detectar variaciones en parámetros físico-

químicos, optimizar la toma de decisiones y facilitar la gestión sostenible de los recursos hídricos, especialmente en entornos donde el acceso a servicios de laboratorio es limitado [4]. Como resultado, esta convergencia tecnológica ha impulsado un crecimiento significativo en los estudios orientados para el análisis, clasificación y predicción de la calidad del agua mediante enfoques híbridos IoT-ML [5].

A pesar de estos avances, la literatura existente presenta importantes desafíos, porque las investigaciones encontradas se desarrollan en contextos y condiciones técnicas muy heterogéneas, lo que dificulta establecer patrones comunes y evaluar la efectividad comparativa de las soluciones propuestas. Por otro lado, aunque existen revisiones previas centradas en IoT para monitoreo ambiental o en modelos ML aplicados a la calidad del agua, estas no abordan de manera integrada la relación entre el tipo de cuerpo de agua, los parámetros monitoreados, las arquitecturas IoT utilizadas y los modelos ML implementados. Esta falta de visión panorámica limita la posibilidad de comprender el estado actual del campo, identificar tendencias consolidadas y reconocer brechas metodológicas que requieren ser atendidas.

En este contexto, resulta necesario sintetizar y organizar la evidencia disponible para comprender cómo se están aplicando conjuntamente IoT y ML en el monitoreo de la calidad del agua [1], qué parámetros e indicadores se priorizan, qué cuerpos de agua se investigan con mayor frecuencia, qué modelos predictivos se emplean y cuáles son las limitaciones comunes en los estudios reportados [5]. Una Revisión Sistemática de Literatura permite cubrir este vacío, proporcionando un análisis estructurado y actualizado de las tendencias, desafíos y oportunidades futuras en el uso combinado de estas tecnologías.

El objetivo de esta revisión sistemática es identificar, analizar y clasificar los enfoques, parámetros, tecnologías y modelos utilizados en estudios que integran IoT y Machine Learning para el monitoreo de la calidad del agua, brindando una visión organizada del estado del arte y destacando sus principales avances, brechas y oportunidades de mejora.

El presente documento se organiza de la siguiente manera: La sección II presenta la metodología empleada para la ejecución de la RSL, detallando la búsqueda, selección y extracción de información. La sección III expone los resultados, organizados según las categorías derivadas del análisis de los estudios primarios. La sección IV discute los hallazgos en relación con la literatura existente, identificando convergencias,

divergencias y limitaciones del campo. Finalmente, la sección V presenta las conclusiones principales y las direcciones sugeridas para futuras investigaciones.

II. METODOLOGÍA

A. PICOC

La estrategia metodológica de esta RSL se formuló mediante la aplicación del modelo PICOC (Population, Intervention, Comparison, Outcomes, Context) [6]. Este marco, ampliamente reconocido, se empleó para estructurar de manera clara y precisa una interrogante de investigación bien delimitada para establecer de manera rigurosa los parámetros que determinarán la selección o exclusión de la literatura, garantizando así la coherencia y reproducibilidad del proceso de selección.

Población (P): La población de interés corresponde a comunidades rurales de países emergentes, donde la falta de acceso a un monitoreo confiable de la calidad del agua representa un reto fundamental para la salubridad colectiva y el progreso sustentable.

Intervención (I): La intervención corresponde a la implementación de sistemas distribuidos de monitoreo basados en el Internet de las Cosas y el aprendizaje automático. Estos sistemas integran sensores de bajo costo que miden: pH, turbidez, conductividad, temperatura, entre otros; integrando con los nodos de adquisición de datos (Arduino/ESP32), gateways para procesamiento local (Raspberry Pi) y redes de comunicación de bajo consumo (LoRa/LoRaWAN).

Comparación (C): La comparación establece el contraste entre las soluciones tecnológicas y los métodos tradicionales de monitoreo de calidad del agua, como el muestreo manual y los análisis de laboratorio. Si bien estos métodos ofrecen alta precisión analítica, presentan limitaciones significativas como los costos elevados, la baja frecuencia de muestreo y una logística compleja que los hace poco adaptables en zonas remotas.

Resultados (O): Se evaluará la capacidad de estas soluciones para mejorar métricas como: La precisión y frecuencia de muestreo, la eficiencia energética del sistema, la confiabilidad en la transmisión de datos, la reducción de costos operativos y la sostenibilidad general a largo plazo de los sistemas de monitoreo.

A partir de la desagregación de estos componentes, se formularon preguntas de investigación específicas, las cuales se pueden apreciar en la Tabla I, organizadas por cada elemento PICO junto con sus respectivas palabras clave.

Como resultado, estas preguntas parciales convergen en la pregunta PICO principal de la RSL: ¿Qué tan efectivas son las soluciones distribuidas basadas en IoT y Machine Learning, mediante sensores de bajo costo y redes de comunicación (LoRa/LoRaWAN, Wi-Fi mesh), para mejorar la precisión, frecuencia de muestreo y eficiencia energética en la vigilancia de las condiciones hídricas en comunidades rurales de países emergentes, en contraste con las técnicas clásicas de evaluación?

Finalmente, se desarrolló la ecuación de búsqueda con las palabras clave, producto de la aplicación de la metodología PICO, la cual se muestra a continuación: "Rural communities"

OR "Developing countries" OR "Emerging countries" OR "Remote areas" OR "Water quality monitoring" AND "Internet of Things" OR "IoT" OR "Machine learning" OR "Distributed monitoring" OR "Low-cost sensors" OR "Arduino" OR "ESP32" OR "Raspberry Pi" OR "LoRa" OR "LoRaWAN" OR "Wi-Fi mesh" AND "Traditional methods" OR "Laboratory analysis" OR "Manual sampling" OR "Conventional monitoring" AND "Accuracy" OR "Sampling frequency" OR "Energy efficiency" OR "Low-power consumption" OR "Cost reduction" OR "Sustainability" OR "Real-time monitoring"

TABLA I
 PREGUNTAS DE LOS COMPONENTES PICO Y PALABRAS CLAVE

Apariencia			
	Componente	Pregunta	Palabras clave
1	P (Población)	¿Cómo se definen las comunidades rurales de países emergentes que presentan problemas de acceso al agua?	Rural communities, developing countries, Emerging countries, remote areas, water quality monitoring
2	I (Intervención)	¿Qué tecnologías IoT y ML se han implementado en el monitoreo distribuido de agua en zonas rurales?	Internet of things, IOT, machine learning, distributed monitoring, Low-cost sensors, Arduino, ESP32, Raspberry Pi, LORA, LoraWAN, Wi-fi Mesh
3	C (Comparación)	¿Cómo se define un método tradicional de vigilancia de las condiciones hídricas y cuáles son sus características?	Traditional methods, laboratory analysis, manual sampling, conventional monitoring
4	O (Resultados)	¿Cómo contribuyen los sistemas IoT y ML a mejorar la precisión y la frecuencia de muestreo? ¿Qué técnicas de eficiencia energética se han aplicado para garantizar la sostenibilidad de los nodos de monitoreo?	Accuracy, sampling frequency, energy efficiency, Low-power consumption, cost reduction, sustainability, real-time monitoring

B. PRISMA

Para asegurar la transparencia y el rigor metodológico, el proceso de localización y filtrado de publicaciones se llevó a cabo siguiendo la metodología PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [7]. El proceso incluyó las fases de identificación, cribado, elegibilidad e inclusión, registradas en el diagrama de flujo correspondiente.

En la fase de identificación, la estrategia de búsqueda aplicada en Scopus permitió recuperar 106 artículos, posteriormente, se eliminaron duplicados y publicaciones que no

cumplan con filtros preliminares como: Antigüedad mayor a cinco años, documentos no clasificados como artículos científicos y textos en idiomas distintos al inglés.

En la fase de cribado, se seleccionaron 56 artículos para su revisión en texto completo. Sin embargo, 21 no pudieron ser recuperados, quedando un total de 35 artículos para el análisis detallado. Posterior a esta etapa se aplicaron los siguientes criterios de inclusión (CI):

- CI1. Implementación tecnológica orientada al agua: Estudios que presentan desarrollos basados en IoT, redes distribuidas (LoRa/LoRaWAN, Wi-Fi mesh) o Machine Learning aplicados al monitoreo de agua.
- CI2. Variables de interés alineadas al objeto de estudio: Inclusión de métricas clave (pH, turbidez, conductividad, oxígeno disuelto, temperatura, bacteriología) y resultados sobre precisión, frecuencia de muestreo o eficiencia energética.
- CI3. Contexto de aplicación relevante: Estudios desarrollados en zonas rurales, países emergentes, entornos con recursos limitados o aplicaciones similares (acuicultura, efluentes industriales).
- CI4. Evidencia empírica o experimental: Trabajos que incluyan pruebas de laboratorio, prototipos, simulaciones o validaciones de campo.
- CI5. Innovación y aporte al conocimiento: Artículos con propuestas metodológicas, arquitectónicas o algorítmicas innovadoras que fortalezcan el monitoreo distribuido de agua y permitan su contraste con métodos tradicionales.

Luego del análisis desarrollado, se determinó excluir 9 estudios: Cinco por no tratar el monitoreo de agua, dos por no incluir IoT/ML, uno por ausencia de comparación con métodos tradicionales y uno por contexto geográfico no pertinente. Para ello se utilizaron los siguientes criterios de exclusión (CE):

- CE1. Artículos cuyo enfoque no se relacione con la vigilancia de las condiciones hídricas (ej. Estudios de riego agrícola o gestión hídrica sin parámetros de calidad).
- CE2. Publicaciones que no utilicen IoT, Machine Learning o sensores electrónicos como parte de la intervención tecnológica (Ej. estudios únicamente con análisis estadístico en laboratorio).
- CE3. Estudios que no incluyen comparación directa o implícita con métodos tradicionales de monitoreo (muestreo manual, análisis de laboratorio).
- CE4. Investigaciones realizadas en contextos urbanos o países desarrollados, fuera del alcance de zonas rurales o países emergentes.

Finalmente, se concluyó que la muestra final será de 26 artículos los cuales constituirán un cuerpo de evidencia sólida y alineada con los objetivos de esta revisión sistemática. El procedimiento descrito previamente, se visualiza en la Figura 1.

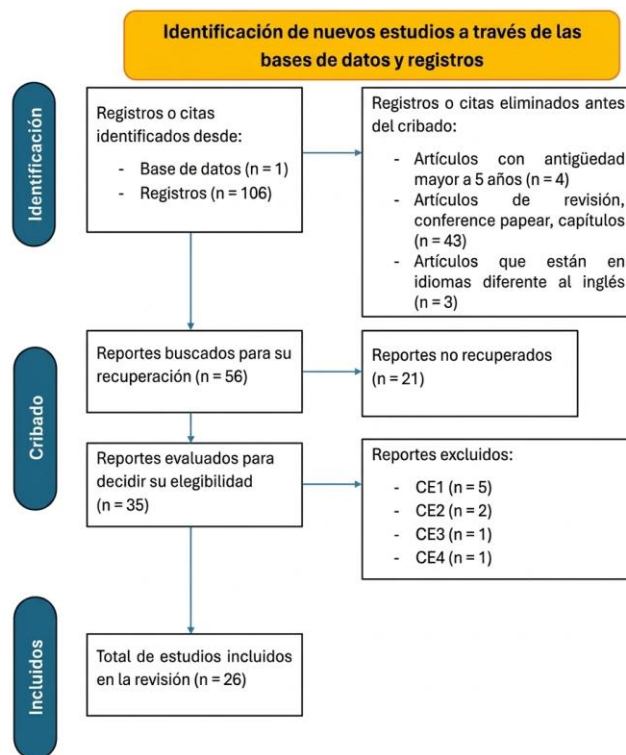


Fig. 1 Diagrama de flujo PRISMA aplicado a la identificación de nuevos estudios a través de las bases de datos y registros

III. RESULTADOS

La síntesis de los 26 artículos seleccionados se organiza en función de las 16 preguntas de extracción de información, las cuales se estructuran conforme a los componentes de la metodología PICO (Población, Intervención, Comparación y Resultado). Este esquema permite un análisis sistemático y coherente de los estudios, facilitando la identificación de patrones, vacíos y tendencias en el uso de tecnologías IoT y Machine Learning aplicadas al monitoreo de recursos hídricos en contextos rurales y de países en vías de desarrollo.

El bloque 1 (Preguntas 1–3) aborda el contexto y la aplicación (Población / Problema), analizando los entornos geográficos, los tipos de cuerpos de agua y las limitaciones del monitoreo tradicional que motivan la intervención tecnológica. El bloque 2 (Preguntas 4–8) se centra en la intervención tecnológica, evaluando los parámetros de calidad de agua medidos, los tipos de plataformas IoT utilizadas, los costos asociados, los sistemas energéticos y las estrategias de eficiencia implementadas. El bloque 3 (Preguntas 9–13) examina la dimensión analítica de Machine Learning (Intervención / Resultado), profundizando en los algoritmos aplicados, métricas de desempeño, niveles de error y propósito del aprendizaje automático. Finalmente, el bloque 4 (Preguntas 14–16) se orienta hacia la comparación y los resultados (Resultado / Comparación), identificando desafíos de escalabilidad, mecanismos de interpretabilidad de modelos y aportes científicos significativos reportados por cada estudio.

En conjunto, esta estructura de análisis permite una visión integral del estado del arte, articulando el contexto de la aplicación con los aspectos técnicos, energéticos y algorítmicos que determinan la efectividad y sostenibilidad de las soluciones tecnológicas para el monitoreo hídrico en entornos de alta vulnerabilidad.

A. ¿Cómo se definen las comunidades rurales de países emergentes que presentan problemas de acceso al agua?

Dentro del componente Población del método PICOC, se formularon tres preguntas orientadas a la extracción de información, con el propósito de caracterizar el contexto geográfico, socioeconómico y operativo en el que se desarrollan los estudios analizados. Los datos obtenidos a partir de estas preguntas permitieron identificar los escenarios de aplicación de las tecnologías de monitoreo hídrico basadas en IoT y Machine Learning, así como las principales limitaciones asociadas a los métodos tradicionales de vigilancia. Los resultados derivados de este análisis se muestran en el presente acápite, proporcionando una visión integral del entorno en el que se contextualiza la evidencia científica revisada.

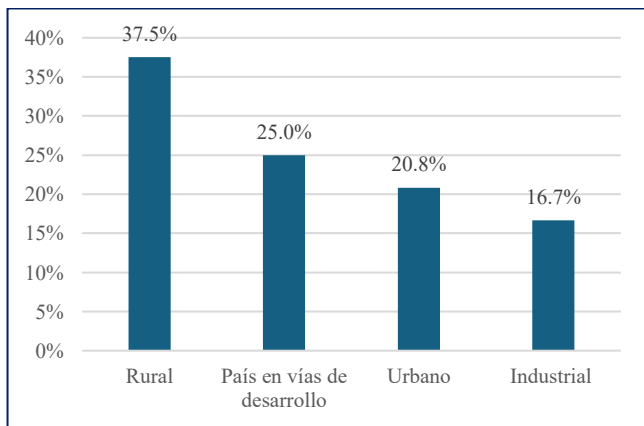


Fig. 2 Distribución de los estudios según el contexto geográfico y socioeconómico de aplicación

La figura 2 del contexto de aplicación de los estudios revela una distribución más concentrada en entornos con desafíos estructurales y limitaciones tecnológicas. El contexto rural constituye la categoría más representada, con 9 de los 24 estudios (37.5%), lo que evidencia un claro énfasis en escenarios donde las brechas de acceso al agua y la falta de infraestructura convierten al monitoreo inteligente en una necesidad prioritaria [8], [9], [10], [11], [12], [13], [14] [15] [16].

Le siguen los estudios desarrollados en países en vías de desarrollo, presentes en 6 artículos (25%), reforzando la orientación de la literatura hacia regiones con mayores restricciones económicas y operativas, donde las soluciones basadas en IoT y Machine Learning pueden generar impactos significativos [17], [18], [19], [20], [21], [22].

Los contextos urbanos aparecen en 5 estudios (20.8%), reflejando un interés creciente en la modernización de los sistemas de monitoreo hídrico en ciudades, aunque aún con menor frecuencia respecto a las zonas de menor desarrollo relativo [23], [24], [25], [26], [27]. Finalmente, los entornos

industriales, representados en 4 estudios (16.7%), constituyen la categoría menos frecuente, posiblemente debido a que estos escenarios suelen contar con infraestructuras más consolidadas y métodos tradicionales de control más estandarizados [28], [29], [30], [31].

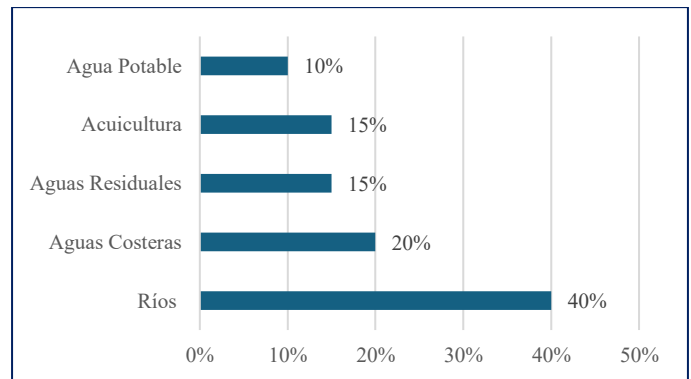


Fig. 3 Tipos de cuerpos de agua abordados en los estudios seleccionados

La figura 3 sobre los tipos de cuerpos de agua monitoreados evidencia una clara predominancia de estudios aplicados en ríos, con 8 de los 20 artículos analizados (40%), lo que confirma el interés prioritario de la literatura por los ecosistemas fluviales debido a su papel crítico en el suministro, transporte y gestión de la calidad del agua [10], [13], [14], [20], [24], [25], [28].

En segundo lugar, se encuentran las aguas costeras, presentes en 4 estudios (20%), reflejando la creciente preocupación por la contaminación marina, eventos de eutrofización y afectaciones a zonas litorales [21], [23], [27], [32]. Las categorías de acuicultura [8], [9] y aguas residuales [15], [18], [29] presentan cada una 3 estudios (15%), lo que indica una atención equilibrada hacia aplicaciones especializadas tanto en entornos productivos como en sistemas de tratamiento o descarga de efluentes.

Finalmente, el monitoreo de agua potable aparece en 2 investigaciones (10%), representando la menor proporción dentro del conjunto analizado, posiblemente debido a la existencia de normativas más estrictas que ya regulan el control de calidad en este tipo de sistemas [12], [33].

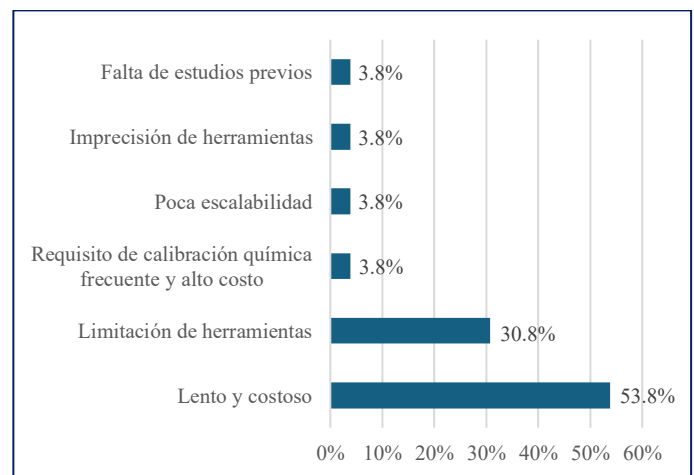


Fig. 4 Distribución de limitaciones reportadas en los estudios analizados

La figura 4 evidencia que la mayoría de los estudios (14 de 26; 53.8%) señalan como principal desafío el carácter lento y costoso de las técnicas convencionales, lo que limita su aplicabilidad a gran escala y en tiempo real [18][33].

Asimismo, 8 investigaciones (30.8%) destacan la limitación de las herramientas empleadas, especialmente en cuanto a cobertura espacial, capacidad de procesamiento y dependencia de infraestructura física [29][12][25][31][27].

En menor medida, se mencionan restricciones como la falta de estudios previos [22], la imprecisión de herramientas [23], la poca escalabilidad de los sistemas existentes [11], y la necesidad de calibración química frecuente [17], cada una reportada en un solo artículo (3.8%).

B. ¿Qué tecnologías IoT y ML se han implementado en el monitoreo distribuido de agua en zonas rurales?

Dentro del componente Intervención del método PICOC, se agruparon cuatro preguntas, orientadas a identificar las tecnologías, arquitecturas y estrategias de implementación aplicadas en el monitoreo distribuido de recursos hídricos. Este bloque analítico se centra en examinar los parámetros de calidad del agua monitoreados, los tipos de plataformas tecnológicas empleadas, los costos asociados, así como los sistemas de energía y mecanismos de eficiencia implementados para garantizar la sostenibilidad operativa de los nodos de monitoreo.

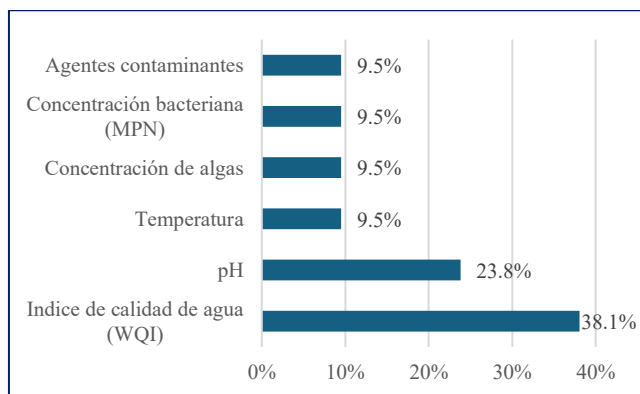


Fig. 5 Distribución de los parámetros de calidad del agua monitoreados en los estudios analizados

La figura 5 muestra una clara predominancia del uso del Índice de Calidad del Agua (WQI), presente en 8 de los 21 estudios (38.1%), lo que refleja la preferencia por indicadores sintéticos capaces de integrar múltiples variables y ofrecer una evaluación global del estado del recurso hídrico [8], [10], [12], [20], [22], [25], [29], [32]. En segundo lugar, el pH aparece en 5 investigaciones (23.8%), destacándose como un parámetro fundamental debido a su sensibilidad frente a cambios químicos y su relevancia en la caracterización básica de los cuerpos de agua [13], [14], [17], [18], [21].

Los parámetros agentes contaminantes [9], [11], concentración bacteriana (MPN) [24], [33], concentración de algas [27], [28], y temperatura [15], [16] registran cada uno 2 estudios (9.5%), lo que evidencia un interés distribuido, pero menos frecuente en variables específicas que suelen estar

asociadas a contextos particulares, como contaminación puntual, sanidad microbiológica o proliferación de algas.

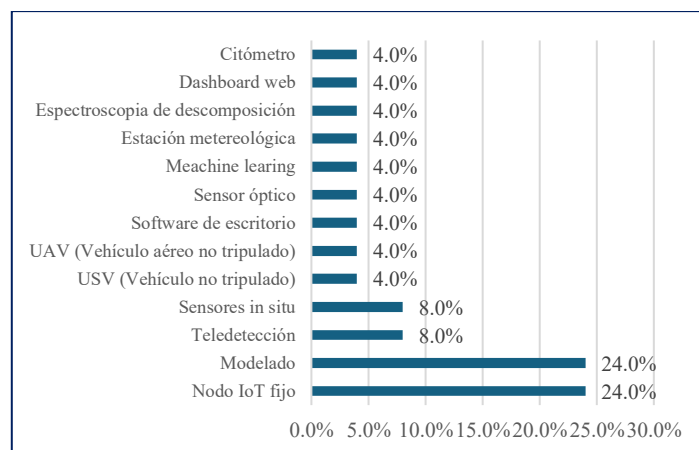


Fig. 6 Distribución de tecnologías y arquitecturas de monitoreo utilizadas en los estudios revisados

La figura 6 muestra una marcada preferencia por enfoques basados en modelado computacional [2], [23], [29] [10], [30][28] y nodos IoT fijos [22] [18], [21][14], cada uno presente en 6 de los 25 artículos (24%), lo que evidencia la relevancia de las soluciones híbridas que combinan simulación con monitoreo físico en campo. Las técnicas de teledetección [25] [11] y el uso de sensores in situ [15], [32] aparecen en 2 estudios respectivamente (8%), indicando un interés en métodos no invasivos y de observación remota. Por otro lado, se identifican tecnologías más específicas o de aplicación complementaria, como UAVs y USVs (1 estudio cada uno; 4%), así como citómetros, sensores ópticos, plataformas de software y dashboards web, que representan soluciones puntuales o de soporte en la gestión de datos.

El análisis de los resultados correspondientes a otras interrogantes relacionadas, revela una limitada transparencia en la documentación técnica y económica de las soluciones desarrolladas. En relación con el reporte de costos del sistema (P6), únicamente 5 de los 26 estudios (19.2%) proporcionan información explícita, clasificando sus propuestas como de “bajo costo” [13], [17], [33]. Este hallazgo coincide con la tendencia observada en la figura 4, donde la lentitud y el alto costo fueron identificados como limitaciones principales del monitoreo tradicional, evidenciando un esfuerzo por desarrollar soluciones más accesibles y eficientes. Sin embargo, la ausencia de datos en 21 artículos (80.8%) limita la posibilidad de establecer comparaciones precisas sobre la viabilidad económica entre distintas implementaciones.

Respecto a la autonomía energética, sólo 3 estudios (11.5%) describen los sistemas de energía empleados [13], [14]. 2 de ellos reportan el uso de paneles solares combinados con mecanismos de bajo consumo, mientras que uno menciona una autonomía basada en batería de hasta 119 horas (aproximadamente 5 días), lo que sugiere avances puntuales en sostenibilidad, pero sin una estandarización clara en la evaluación de la eficiencia energética.

C. ¿Cómo se define un método tradicional de vigilancia de las condiciones hídricas y cuáles son sus características?

Dentro del componente Comparación del método PICOC, se agrupan las preguntas orientadas a examinar los métodos o enfoques alternativos frente a los cuales se contrastan las soluciones basadas en IoT y Machine Learning. Este bloque tiene como objetivo identificar las referencias comparativas, ya sean métodos tradicionales de monitoreo, modelos estadísticos, o algoritmos de aprendizaje automático, que permiten evaluar el rendimiento, la precisión y la eficiencia de las propuestas analizadas. A través de este análisis, se busca comprender cómo las tecnologías emergentes superan las limitaciones de los enfoques convencionales, así como las métricas utilizadas para validar dichas mejoras. En conjunto, esta sección proporciona una visión crítica del posicionamiento relativo de las soluciones tecnológicas dentro del ecosistema de métodos existentes para el monitoreo y la gestión de la calidad del agua.

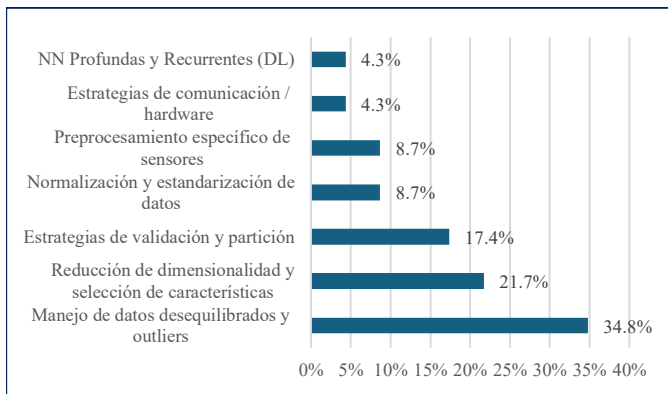


Fig. 7 Estrategias de eficiencia energética implementadas en los estudios revisados

La figura 7 muestra los artículos revisados sobre estrategias de eficiencia energética y gestión de datos para hardware evidencia una tendencia clara en las líneas de investigación actuales. La mayoría de los estudios priorizan enfoques orientados a la manipulación y optimización de los datos antes que la gestión directa sobre el hardware. Esto se refleja en el hecho de que la estrategia más recurrente es el manejo de datos desbalanceados y outliers, presente en un 34.8% de los artículos. Este tipo de técnica busca mejorar la calidad de la información procesada y, por extensión, el rendimiento de los modelos matemáticos utilizados [8], [10], [12].

La siguiente estrategia más frecuente es la reducción de dimensionalidad y selección de características, con un 21.7%. Esta práctica facilita la simplificación de los modelos computacionales y permite que los sistemas sean más eficientes en términos de procesamiento, aunque no necesariamente en consumo energético directo del hardware. También es destacable el uso de estrategias de validación y partición (17.4%), junto con la normalización y estandarización de datos, y el preprocesamiento específico de sensores (ambas con 8.7%). Estas acciones optimizan la preparación y la robustez de los datos, lo que repercute en la fiabilidad de los resultados [13], [14], [26], [32]. Por otro lado, el análisis revela que solo un 4.3% de los estudios implementan estrategias específicas sobre

eficiencia energética para hardware, como la utilización de protocolos de comunicación eficientes, modos de bajo consumo, o tecnologías avanzadas tipo LoRaWAN [14], [15], [26]. Esta proporción tan baja indica que el desarrollo de soluciones energéticas dentro de la revisión es aún incipiente y representa una oportunidad de mejora significativa para futuras investigaciones en el ámbito del Internet de las Cosas y sistemas embebidos.

D. ¿Cómo contribuyen los sistemas IoT y ML a mejorar la precisión y la frecuencia de muestreo? ¿Qué técnicas de eficiencia energética se han aplicado para garantizar la sostenibilidad de los nodos de monitoreo?

Dentro del componente Resultado del método PICOC, se agrupan las preguntas orientadas a identificar los principales hallazgos, beneficios y aportes científicos derivados de la implementación de tecnologías basadas en IoT y Machine Learning en el monitoreo hídrico. Este bloque tiene como finalidad analizar los niveles de precisión alcanzados, los tipos de métricas utilizadas, las estrategias de interpretación de modelos, y los desafíos de escalabilidad o sostenibilidad reportados en los estudios. Asimismo, permite evaluar cómo estas soluciones contribuyen a mejorar la frecuencia de muestreo, la eficiencia energética y la capacidad predictiva frente a los métodos tradicionales. En conjunto, esta sección busca sintetizar los impactos técnicos y científicos más relevantes, proporcionando evidencia empírica sobre la efectividad y madurez tecnológica de las herramientas inteligentes aplicadas al monitoreo de la calidad del agua.

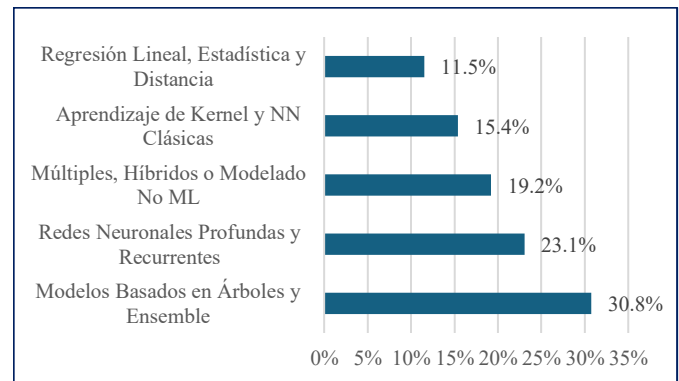


Fig. 8 Análisis de los algoritmos y arquitecturas reportado en los estudios revisados

La figura 8 permite identificar con precisión la distribución porcentual de los algoritmos y arquitecturas principales de Machine Learning y Deep Learning utilizada. Los datos se detallan a continuación, reflejando una clara tendencia en las preferencias metodológicas dentro del campo. La categoría más representativa es la de modelos basados en árboles y métodos ensemble, utilizada en el 30.8% de los estudios. Esta proporción evidencia su robustez para abordar problemas complejos y multidimensionales [8], [11], [12], [13], [15], [22], [26], [33]. Le siguen las redes neuronales profundas y recurrentes, presentes en el 23.1% de los artículos, lo que confirma su importancia en contextos donde el modelado de relaciones no lineales y

secuenciales es esencial [10], [14], [18], [20], [23], [25], [29], [30], [31].

Los enfoques múltiples, híbridos o modelado no basado en Machine Learning constituyen el 19.2% de los algoritmos empleados. Estos trabajos combinan diferentes técnicas para obtener resultados mejorados adaptándose a problemáticas específicas [9], [17], [19]. El aprendizaje de kernel y las redes neuronales clásicas suman el 15.4%, mostrando que siguen siendo valorados por su simplicidad y eficiencia en ciertos casos de uso. Por último, la regresión lineal, estadística y técnicas de distancia corresponden al 11.5%. Aunque su presencia es menor, estas técnicas tradicionales siguen siendo relevantes para escenarios específicos donde la interpretación y la simplicidad son prioritarias.

Esta distribución porcentual muestra que los algoritmos más utilizados son los modelos basados en árboles y ensemble, seguidos por redes neuronales profundas y técnicas híbridas, lo que evidencia una preferencia por enfoques capaces de manejar tanto grandes volúmenes de datos como relaciones complejas en las aplicaciones analizadas.

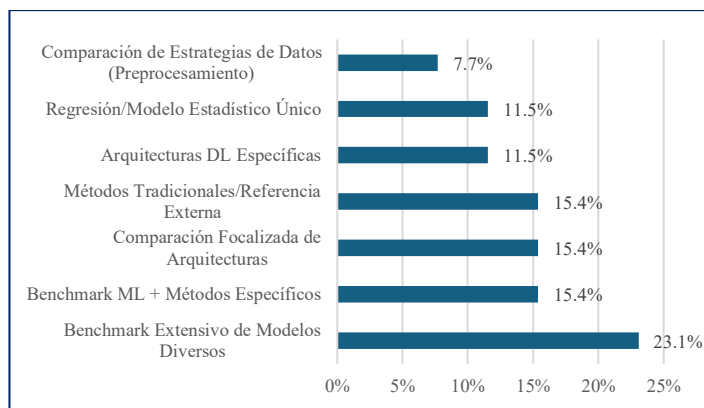


Fig. 9 Análisis de los enfoques de comparación empleados en los estudios revisados

La figura 9 revela la diversidad de enfoques utilizados para comparar el rendimiento de los modelos principales de Machine Learning y Deep Learning. La opción metodológica más frecuente es la comparación extensiva entre modelos diversos, presente en el 23.1% de los estudios, lo que refleja el esfuerzo de los autores por identificar la mejor alternativa entre múltiples algoritmos y arquitecturas. [19], [20], [30], [31]. Por otro lado, en el 15.4% de los artículos se realiza un benchmark entre métodos de Machine Learning y técnicas específicas, mientras que otro 15.4% emplea comparaciones focalizadas entre arquitecturas concretas. Un porcentaje igual (15.4%) corresponde al uso de comparaciones contra métodos tradicionales o referencias externas, permitiendo evaluar el valor agregado de los enfoques modernos respecto a técnicas ya consolidadas. [17], [24], [25], [28], [29]. Las comparaciones que utilizan arquitecturas de Deep Learning específicas y las que emplean únicamente regresión o modelos estadísticos independientes constituyen cada una el 11.5% de los casos. Finalmente, solo el 7.7% de los estudios compara estrategias relacionadas con el preprocesamiento de datos, enfocándose en cómo estos métodos impactan el rendimiento global de los

modelos probados [10], [18], [23], [25], [31]. Estos resultados reflejan que la comparación de rendimiento frente a una gran variedad de métodos (incluyendo SVR, redes neuronales estándar, métodos estadísticos y arquitecturas especializadas) es una práctica común, permitiendo una evaluación sólida y contextualizada de los nuevos modelos desarrollados para los problemas tratados.

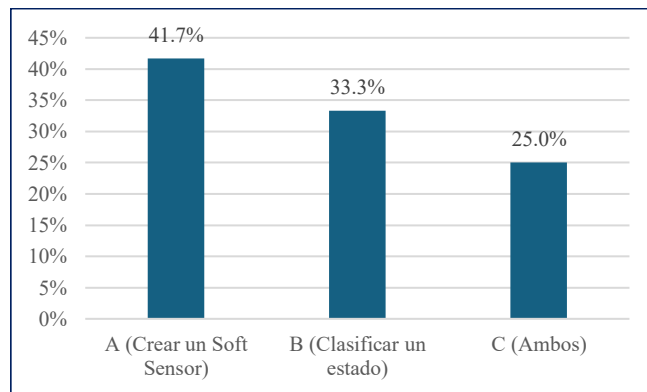


Fig. 10 Análisis de las aplicaciones de Machine Learning en los estudios revisados

La figura 10 permite identificar claramente los principales usos de Machine Learning en los estudios analizados. La mayor proporción corresponde a la creación de Soft Sensors, con un 41.7%. Esto evidencia el fuerte interés en desarrollar sistemas capaces de reemplazar sensores físicos mediante modelos computacionales, optimizando así el monitoreo y la medición en diferentes aplicaciones. [20], [23], [24], [25], [26], [33]. El segundo uso más frecuente es la clasificación de estados, como la distinción entre agua potable y no potable o la identificación de tipos de algas, que aparece en el 33.3% de los artículos.

Este enfoque confirma la relevancia de Machine Learning para resolver problemas de clasificación que requieren analizar patrones complejos en los datos. [22], [27], [28], [29], [30], [31]. Por último, el 25.0% de los estudios reporta la utilización de ambos enfoques, es decir, se emplea Machine Learning tanto para crear Soft Sensors como para clasificar estados en el mismo trabajo. Estos resultados reflejan que la utilización de Machine Learning en estos artículos se orienta principalmente hacia la generación de Soft Sensors, seguido por tareas de clasificación y, en menor proporción, implementaciones mixtas que integran ambos tipos de aplicación.

La figura 11 indica que existen varios desafíos específicos para la escalabilidad y la implementación a largo plazo de soluciones en grandes cuerpos de agua o a nivel nacional. El principal reto identificado es el manejo de la heterogeneidad y calidad de los datos, mencionado en el 37.5% de los casos. Esta dificultad se relaciona con la necesidad de integrar información proveniente de diversas fuentes, sensores y contextos, garantizando su validez y representatividad en aplicaciones a gran escala [15], [18], [20], [23], [24], [25], [33].

En segundo lugar, tres categorías comparten una incidencia del 18.8%: la validación y generalización externa de los modelos, la adaptabilidad regulatoria y contextual, y los desafíos

operacionales y de hardware (IoT/sensores). La validación externa implica asegurar que los modelos funcionen correctamente en distintos contextos y no sólo bajo condiciones específicas de laboratorio. La adaptabilidad regulatoria y contextual hace referencia a la necesidad de cumplir con normativas locales y adaptar las soluciones a distintos escenarios, mientras que los desafíos operacionales señalan dificultades técnicas y logísticas al desplegar sensores y sistemas IoT en ambientes reales y extensos. [12], [17], [19], [26], [28], [29], [31], [32].

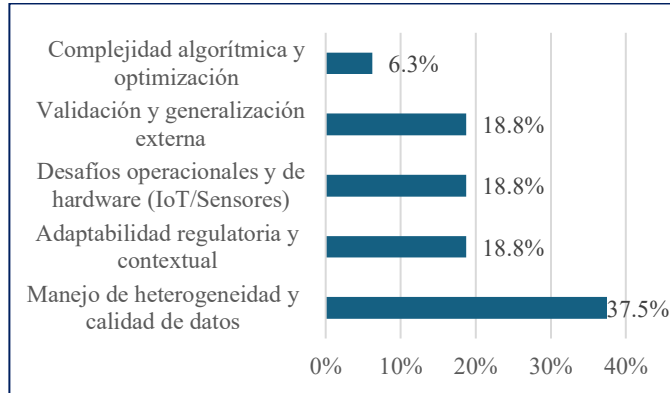


Fig. 11 Análisis de los desafíos de escalabilidad e implementación a largo plazo en los estudios revisados

Por último, la complejidad algorítmica y los retos de optimización fueron identificados en el 6.3% de los trabajos, subrayando la importancia de desarrollar algoritmos que sean tanto eficientes como escalables en infraestructura y costo operativo [14], [16], [21].

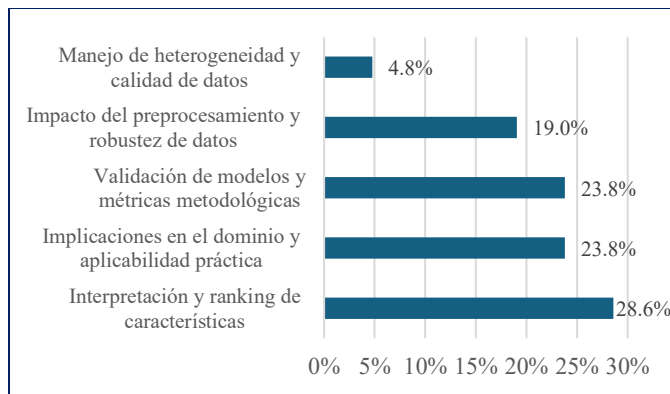


Fig. 12 Análisis de la proporción de insights sobre la interpretación del modelo en los estudios revelados

La figura 12 proporciona una visión clara sobre el tipo de insights ofrecidos en la interpretación de los modelos empleados. El aspecto más abordado es la interpretación y el ranking de características, presente en el 28.6% de los trabajos, lo que indica un énfasis en entender cuáles variables tienen mayor impacto en las predicciones y resultados generados por los modelos analizados [24], [25], [28], [30], [31].

En segundo lugar, se encuentran las implicaciones en el dominio y la aplicabilidad práctica, así como la validación de modelos y métricas metodológicas, ambas con el 23.8%. Esto

refleja el interés por justificar cómo los modelos se traducen en resultados relevantes para la práctica profesional y por validar la fiabilidad estadística y operativa de los enfoques aplicados. El impacto del pre-procesamiento y la robustez de los datos es considerado en el 19.0% de los estudios, evidenciando la preocupación por cómo el tratamiento de los datos puede fortalecer o debilitar la interpretación de los modelos [11], [15], [16], [22].

Por último, el manejo de la heterogeneidad y calidad de los datos se menciona en solo el 4.8% de los casos, mostrando que aunque es un aspecto importante, su discusión específica desde la perspectiva de interpretación es menos frecuente comparado con otros enfoques [10], [12], [19], [21], [29].

Estos resultados indican que los estudios tienden a proporcionar insights sobre la interpretación del modelo, principalmente a través del análisis de la relevancia de las características y la validación metodológica, promoviendo también su aplicación práctica en el dominio específico de interés.

IV. DISCUSIÓN

Los resultados de esta revisión sistemática permiten identificar patrones significativos sobre el uso combinado de IoT y Machine Learning en el monitoreo de la calidad del agua, especialmente en entornos rurales y en países emergentes. La alta realización de estudios en estos contextos sugiere que la adopción de tecnologías digitales se ha orientado hacia regiones donde los métodos tradicionales presentan limitaciones operativas, logísticas y económicas. Esta tendencia refleja una necesidad real de contar con sistemas de monitoreo continuos y de bajo costo que permitan compensar la falta de infraestructura especializada. Aunque una revisión previa señala mayor presencia de aplicaciones urbanas, la distribución encontrada en este estudio demuestra que el interés técnico actual se dirige hacia escenarios vulnerables donde el impacto de la digitalización puede ser más inmediato y significativo [34].

El predominio de los ríos como cuerpo de agua monitoreado confirma la importancia crítica de los ecosistemas fluviales tanto para el abastecimiento como para la vigilancia ambiental. Este hallazgo se alinea con tendencias generales descritas en la literatura, donde los cuerpos de agua superficial suelen priorizarse por su mayor susceptibilidad a descargas puntuales y difusas. Sin embargo, la baja representación de estudios centrados en agua potable evidencia una brecha que podría limitar la capacidad de estas tecnologías para contribuir a la seguridad hídrica directamente asociada a la salud pública. Esta diferencia contrasta parcialmente con informes recientes que reportan un creciente interés por la digitalización de redes de distribución urbana [35], lo que sugiere que el campo podría estar diversificándose hacia líneas aún poco reflejadas en nuestro conjunto de estudios.

La preferencia por el índice de calidad del agua (WQI) como parámetro evaluado muestra la necesidad de contar con métricas sintéticas que faciliten la integración con modelos ML y permitan interpretar de manera global el estado del recurso. No obstante, la limitada presencia de variables microbiológicas evidencia un desfase entre los estudios analizados y las

necesidades sanitarias de las comunidades rurales, donde la contaminación bacteriana suele ser uno de los problemas más críticos. Esta ausencia apunta a una oportunidad importante para el desarrollo de soluciones que integren bio-sensores, técnicas de muestreo inteligente o modelos predictivos orientados a la detección temprana de riesgos.

Desde el punto de vista tecnológico, el predominio de nodos IoT fijos y modelos basados en predicción computacional sugiere que la comunidad científica ha priorizado la estabilidad, el bajo costo y la escalabilidad básica por encima de la movilidad o la complejidad arquitectónica. Este comportamiento es coherente con las necesidades operativas de las zonas rurales, donde los sistemas deben funcionar con recursos energéticos limitados y mantenimiento reducido. Las revisiones previas señalan que las arquitecturas híbridas IoT-ML se están consolidando, lo cual coincide con la presencia creciente de soluciones que integran procesamiento local, conectividad de bajo consumo y análisis predictivo [34], [35]. Sin embargo, los resultados también muestran que la mayoría de los sistemas carece de reportes transparentes sobre costos y autonomía energética, lo que dificulta evaluar su viabilidad real para implementaciones a largo plazo.

Si bien los estudios analizados reportan altos niveles de precisión, correlación y exactitud, la heterogeneidad en las métricas empleadas continúa siendo un obstáculo para establecer comparaciones robustas. La falta de estandarización en criterios de validación, señalada también en revisiones previas, es evidente en los artículos incluidos en esta RSL y constituye una de las principales barreras para consolidar modelos reproducibles y extrapolables entre regiones.

V. CONCLUSIONES

La presente Revisión Sistemática responde al objetivo planteado, orientado a identificar y sintetizar la evidencia disponible sobre la aplicación de tecnologías basadas en IoT y Machine Learning para el monitoreo de la calidad del agua en diferentes contextos. El análisis realizado muestra que estas tecnologías se han consolidado como alternativas viables y potencialmente transformadoras frente a las limitaciones del monitoreo tradicional, especialmente en escenarios donde la infraestructura es limitada y los sistemas de vigilancia convencionales resultan costosos, discontinuos o poco accesibles.

Los hallazgos evidencian que la investigación reciente se concentra principalmente en entornos rurales y países en vías de desarrollo, lo que refleja la necesidad de soluciones tecnológicas adaptadas a realidades con restricciones operativas y mayores riesgos asociados al acceso al agua. Asimismo, la predominancia de estudios realizados en cuerpos de agua fluviales señala la importancia de los ecosistemas superficiales como fuentes críticas de abastecimiento y como zonas altamente vulnerables a variaciones en la contaminación.

La aparición reiterada del WQI como parámetro central muestra la preferencia por indicadores sintéticos que permitan integrar múltiples dimensiones de la calidad del agua, mientras que el uso frecuente de nodos IoT fijos y enfoques de modelado computacional confirma una tendencia hacia sistemas de

monitoreo continuo orientados a decisiones más oportunas y automatizadas. No obstante, la limitada transparencia en el reporte de costos, autonomía energética y métricas comparables revela vacíos persistentes que restringen la evaluación integral de la escalabilidad y sostenibilidad de estas soluciones.

El aporte principal de esta revisión reside en ofrecer una síntesis actualizada y estructurada de las líneas tecnológicas que están configurando el monitoreo hídrico inteligente, así como en evidenciar las brechas metodológicas y operativas que aún persisten en la literatura. Esta contribución resulta valiosa tanto para investigadores como para entidades responsables de la gestión del agua, pues proporciona una visión clara de las áreas donde la investigación ha avanzado y de aquellos aspectos que requieren mayor desarrollo para fortalecer la implementación real de estas tecnologías.

REFERENCIAS

- [1] S. Zyoud, "Exploring the promising role of internet of things in urban water systems: a comprehensive global analysis of insights, trends, and research priorities," *Discover Internet of Things*, vol. 5, no. 1, p. 31, Mar. 2025, doi: 10.1007/s43926-025-00129-1.
- [2] S. Sreejith et al., "IoT sensor-based systems in real-time monitoring of health and environment: a review," *Journal of the Korean Physical Society*, vol. 87, no. 1, pp. 1–27, Jul. 2025, doi: 10.1007/s40042-025-01378-7.
- [3] S. O. Ooko, L. Cheptegei, and S. M. Karume, "Application of machine learning for real-time water quality monitoring in developing countries: A review," *Sustainable Futures*, vol. 10, p. 100984, Dec. 2025, doi: 10.1016/j.sfr.2025.100984.
- [4] B. Ngwenya, T. Paepae, and P. N. Bokoro, "Monitoring ambient water quality using machine learning and IoT: A review and recommendations for advancing SDG indicator 6.3.2," *Journal of Water Process Engineering*, vol. 73, p. 107664, May 2025, doi: 10.1016/j.jwpe.2025.107664.
- [5] M. Flores-Iwasaki, G. A. Guadalupe, M. Pachas-Caycho, S. Chapa-Gonza, R. C. Mori-Zababurú, and J. C. Guerrero-Abad, "Internet of Things (IoT) Sensors for Water Quality Monitoring in Aquaculture Systems: A Systematic Review and Bibliometric Analysis," *AgriEngineering*, vol. 7, no. 3, p. 78, Mar. 2025, doi: 10.3390/agriengineering7030078.
- [6] E. E. Biney, C. Gyamfi, A. Y. Karikari, and D. Darko, "Reservoir ecological health assessment Methods: A systematic review," Feb. 01, 2025, Elsevier B.V. doi: 10.1016/j.ecolind.2025.113130.
- [7] M. J. Page et al., "The PRISMA 2020 statement: An updated guideline for reporting systematic reviews," Mar. 29, 2021, BMJ Publishing Group. doi: 10.1136/bmj.n71.
- [8] R. M. Frincu, "Artificial intelligence in water quality monitoring: A review of water quality assessment applications," *Water Quality Research Journal*, vol. 60, no. 1, pp. 164–176, 2025, doi: 10.2166/wqrj.2024.049.
- [9] Q. Ma, S. Li, H. Qi, X. Yang, and M. Liu, "Rapid Prediction and Inversion of Pond Aquaculture Water Quality Based on Hyperspectral Imaging by Unmanned Aerial Vehicles," *Water (Switzerland)*, vol. 17, no. 4, 2025, doi: 10.3390/w17040517.
- [10] S. Toumi et al., "Harnessing Deep Learning for Real-Time Water Quality Assessment: A Sustainable Solution," *Water (Switzerland)*, vol. 16, no. 23, 2024, doi: 10.3390/w16233380.
- [11] S. Zhang et al., "Evaluating agricultural non-point source pollution with high-resolution remote sensing technology and SWAT model: A case study in Ningxia Yellow River Irrigation District, China," *Ecol Indic*, vol. 166, 2024, doi: 10.1016/j.ecolind.2024.112578.
- [12] M. I. Qadir, R. Mumtaz, M. Manzoor, M. Saleem, M. A. Khan, and S. M. Charlesworth, "Unmanned surface vehicle for intelligent water quality assessment to promote sustainable human health," *Water Supply*, vol. 27, no. 7, pp. 2259–2270, 2024, doi: 10.2166/ws.2024.141.
- [13] P. Jayaraman, K. K. Nagarajan, P. Partheeban, and V. Krishnamurthy, "Critical review on water quality analysis using IoT and machine learning models," *International Journal of Information Management Data Insights*, vol. 4, no. 1, 2024, doi: 10.1016/j.jjimei.2023.100210.
- [14] Y. Li, Y. Fu, Z. Lang, and F. Cai, "A High-Frequency and Real-Time Ground Remote Sensing System for Obtaining Water Quality Based on a

- Micro Hyper-Spectrometer,” *Sensors*, vol. 24, no. 6, 2024, doi: 10.3390/s24061833.
- [15] H. Shyu, C. J. Castro, R. A. Bair, Q. Lu, and D. H. Yeh, “Development of a Soft Sensor Using Machine Learning Algorithms for Predicting the Water Quality of an Onsite Wastewater Treatment System,” *ACS Environmental Au*, vol. 3, no. 5, pp. 308–318, 2023, doi: 10.1021/acsenvironau.2c00072.
- [16] K. Belouz and S. Zereg, “Extreme learning machine for soil temperature prediction using only air temperature as input,” *Environ Monit Assess*, vol. 195, no. 8, 2023, doi: 10.1007/s10661-023-11566-2.
- [17] B. M. A. Zeta, S. U. Alam, G. M. A. E. U. Rahman, and K. I. U. Ahmed, “A Low-Cost pH Sensor for Real-Time Monitoring of Aquaculture Systems in a Multi-Layer Wireless Sensor Network,” *Sensors*, vol. 25, no. 9, 2025, doi: 10.3390/s25092824.
- [18] N. Chavhan et al., “APAH: An autonomous IoT driven real-time monitoring system for Industrial wastewater,” *Digital Chemical Engineering*, vol. 14, 2025, doi: 10.1016/j.dche.2025.100217.
- [19] E. Aslan, “DEVELOPMENT OF MALARIA DIAGNOSIS WITH CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES: A CNN-BASED SOFTWARE FOR ACCURATE CELL IMAGE ANALYSIS,” *Journal of Engineering and Technology for Industrial Applications*, vol. 11, no. 51, pp. 35–42, 2025, doi: 10.5935/jetia.v11i51.1392.
- [20] S. Chadalavada et al., “Gated-LNN: Gated Liquid Neural Networks for Accurate Water Quality Index Prediction and Classification,” *IEEE Access*, vol. 13, pp. 69500–69512, 2025, doi: 10.1109/ACCESS.2025.3561593.
- [21] J. Kustija, D. Fahrizal, M. S. B. Nasir, D. Setiawan, and I. Surya, “Design and development of coastal marine water quality monitoring based on IoT in achieving implementation of SDGs,” *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 3, pp. 1470–1484, 2024, doi: 10.11591/ijeecs.v36.i3.pp1470-1484.
- [22] N. R. Sivakumar and A. Z. Ibrahim, “Deep neural artificial intelligence for IoT based tele health data analytics,” *Computers, Materials and Continua*, vol. 70, no. 3, pp. 4467–4483, 2022, doi: 10.32604/cmc.2022.019041.
- [23] A. Elmotawakkil, N. Enneya, S. K. Bhagat, M. M. Ouda, and V. Kumar, “Advanced machine learning models for robust prediction of water quality index and classification,” *Journal of Hydroinformatics*, vol. 27, no. 2, pp. 299–319, 2025, doi: 10.2166/hydro.2025.290.
- [24] J. K. Pandya, S. S. Khandelwal, R. K. Kumar Tipu, and K. S. Pandya, “Advancing Water Quality Management: An Integrated Approach Using Ensemble Machine Learning and Real-Time Interactive Visualization,” *IEEE Access*, 2025, doi: 10.1109/ACCESS.2025.3573589.
- [25] Y. Zhou et al., “Spatial-temporal distribution of labeled set bias remote sensing estimation: An implication for supervised machine learning in water quality monitoring,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 131, 2024, doi: 10.1016/j.jag.2024.103959.
- [26] K. C. Kumar and B. P. Pande, “Air pollution prediction with machine learning: a case study of Indian cities,” *International Journal of Environmental Science and Technology*, vol. 20, no. 5, pp. 5333–5348, 2023, doi: 10.1007/s13762-022-04241-5.
- [27] J. Guo, Y. Ma, and H. W. J. Lee, “Real-time automated identification of algal bloom species for fisheries management in subtropical coastal waters,” *Journal of Hydro-Environment Research*, vol. 36, pp. 1–32, 2021, doi: 10.1016/j.jher.2021.03.002.
- [28] S. E. Breitmeyer et al., “Predicted Potential for Aquatic Exposure Effects of Per- and Polyfluorinated Alkyl Substances (PFAS) in Pennsylvania’s Statewide Network of Streams,” *Toxics*, vol. 12, no. 12, 2024, doi: 10.3390/toxics12120921.
- [29] G. Ozsert Yigit and C. Baransel, “Utilizing machine learning techniques for enhanced water quality monitoring,” *Water Quality Research Journal*, vol. 59, no. 4, pp. 187–204, 2024, doi: 10.2166/wqrj.2024.007.
- [30] D. A. Musleh et al., “Ensemble Learning Based Sustainable Approach to Carbonate Reservoirs Permeability Prediction,” *Sustainability (Switzerland)*, vol. 15, no. 19, 2023, doi: 10.3390/su151914403.
- [31] Y. Zheng, Q. Lu, A. Chen, Y. Liu, and X. Ren, “Rapid Classification and Quantification of Coal by Using Laser-Induced Breakdown Spectroscopy and Machine Learning,” *Applied Sciences (Switzerland)*, vol. 13, no. 14, 2023, doi: 10.3390/app13148158.
- [32] J. Xu et al., “An alternative to laboratory testing: Random forest-based water quality prediction framework for inland and nearshore water bodies,” *Water (Switzerland)*, vol. 13, no. 22, 2021, doi: 10.3390/w13223262.
- [33] A. Saavedra-Ruiz and P. J. Resto-Irizarry, “A Portable UV-LED/RGB Sensor for Real-Time Bacteriological Water Quality Monitoring Using ML-Based MPN Estimation,” *Biosensors (Basel)*, vol. 15, no. 5, 2025, doi: 10.3390/bios15050284.
- [34] C. Z. Zulkifli et al., “IoT-Based Water Monitoring Systems: A Systematic Review,” *Water (Basel)*, vol. 14, no. 22, p. 3621, Nov. 2022, doi: 10.3390/w14223621.
- [35] I. Essamlali, H. Nhaila, and M. El Khaili, “Advances in machine learning and IoT for water quality monitoring: A comprehensive review,” *Heliyon*, vol. 10, no. 6, p. e27920, Mar. 2024, doi: 10.1016/j.heliyon.2024.e27920.