

Inteligencia Artificial Explicable para la Predicción de la Calidad del Agua: Revisión del Estado del Arte

Yanqui-Cheneaux M. A.¹, Medina-Grados M. J.¹

¹ Departamento de Ingenierías de la Industria y el Ambiente, Universidad Católica San Pablo, Arequipa, Perú

Resumen:

La gestión actual de la calidad de los recursos hídricos recurre a modelos de aprendizaje automático con el propósito de predecir los parámetros críticos; no obstante, la intrínseca opacidad de los algoritmos limita la confianza en sus resultados y aplicabilidad en entornos operativos. La Inteligencia Artificial Explicable (XAI) emerge a consecuencia de las exigencias surgidas dentro de la academia y los marcos regulatorios. Así se devela un claro dominio de los modelos supervisados, especialmente los métodos de ensamble, como lo son *Random Forest* y *XGBoost*; estos han sido aplicados a múltiples conjuntos recurrentes de variables de entrada fisicoquímicas (e.g. pH, DO, COD). Los valores SHAP constituyen la herramienta de explicabilidad predominante, con la cual se cuantifica la importancia de las variables, sin profundizar en interpretaciones útiles para la gestión de recursos hídricos. La principal limitación del área de investigación es la falta de estandarización en conjuntos de datos y métricas de rendimiento, lo que impide contrastar el desempeño entre modelos. Se concluye así que el campo requiere avanzar hacia un enfoque de aplicación más profunda de la XAI.

Index words: Inteligencia Artificial Explicable (XAI), Calidad del Agua, Aprendizaje Automático, Predicción

1 Introducción

La Inteligencia Artificial Explicable (XAI, por sus siglas en inglés) son un conjunto de técnicas y métodos que permiten comprender y explicar las predicciones de modelos de aprendizaje automático, proporcionando interpretaciones inteligibles para el ser humano (Nallakuruppan et al., 2024). Dentro del contexto de la calidad del agua, dichos instrumentos ha emergido como una herramienta valiosa para el *forecasting*, estos facilitan la interpretación de los procedimientos y resultados de los modelos de aprendizaje automático usados en la predicción de indicadores de calidad del agua, tales como parámetros fisicoquímicos (p.e., clorofila-*a*, oxígeno disuelto, turbidez) (Park et al., 2022; Bolick et al., 2023). Mediante técnicas como LIME (*Local Interpretable Model-Agnostic Explanations*) o SHAP (*SHapley Additive exPlanations*), la XAI puede develar la contribución de cada parámetro imputado en la predicción, brindando transparencia (*white box nature*) a modelos opacos (*black box nature*) (Ramya et al., 2024). A efecto de estos, es posible obtener pronósticos precisos de la calidad del agua y entender las relaciones subyacentes entre las variables medidas y las condiciones del recurso hídrico (Kundu et al., 2025). Esto reviste gran importancia para la gestión del agua, dado que explicaciones provistas por modelos XAI permiten a especialistas y gestores confiar en dichas estimaciones y recabar información pertinente al comportamiento del sistema (Rudin, 2019; Nong et al., 2024).

Recientemente, el aprendizaje automático se ha consolidado como herramienta fundamental en el análisis de series temporales hidrológicas y la predicción de la calidad del agua con gran exactitud (Park et al., 2022; Savoy et al., 2023). Exploración de grandes volúmenes de datos medioambientales, apoyado en algoritmos de *machine learning*, han permitido mejorar las predicciones de disponibilidad y calidad del agua en diversos contextos (Savoy et al., 2023). Verbigracia, se han concebido modelos capaces de pronosticar la concentración de clorofila-*a* o el oxígeno disuelto en ríos y lagos, exhibiendo un desempeño preponderante frente a enfoques tradicionales (Park et al., 2022; Bolick et al., 2023). No obstante, la complejidad inherente de dichas técnicas los dota de un comportamiento opaco, lo cual obstaculiza la comprensión de cómo las entradas (e.g., temperatura, caudal, nutrientes) repercuten en las

salidas (Nong et al., 2024). La falta de interpretabilidad representa un obstáculo que infunde desconfianza en los resultados; análogamente limita la aplicabilidad práctica de los modelos en la toma de decisiones (Commission, 2024). En contextos hidrológicos críticos, se ha puesto de manifiesto la necesidad de dotar a los modelos de predicción de mecanismos explicativos, de forma que los responsables de la gestión de los recursos hídricos puedan entender las causas de un cambio en su calidad y prever acciones respecto a su control (Nong et al., 2024). En este sentido, las técnicas de XAI han cobrado relevancia: al interpretar modelos de *machine learning*, permiten identificar los factores determinantes de la contaminación o mejora de la calidad del agua, haciendo los pronósticos más transparentes y confiables (Ramya et al., 2024; Bolick et al., 2023).

Ahora también, aunque el uso de modelos para series temporales está ampliamente difundido, la literatura evidencia importantes limitaciones y brechas que motivan el presente estudio. En primer lugar, no existe un consenso claro sobre cuáles son las variables de entrada óptimas para predecir la calidad del agua: distintos trabajos emplean conjuntos de predictores muy diversos dependiendo del caso de estudio (Park et al., 2022; Bolick et al., 2023). Algunos estudios se basan solo en parámetros físico-químicos locales medidos *in-situ*, mientras que otros incorporan variables climáticas, hidrológicas o incluso datos de teledetección y factores socioeconómicos (Tian et al., 2024; Kundu et al., 2025), sin haber esclarecido aún los factores imprescindibles o más influyentes de forma general. En segundo lugar, la comparabilidad de los algoritmos de ML empleados es baja, dada la heterogeneidad en enfoques y métricas de evaluación utilizadas en cada investigación. Diferentes autores prueban algoritmos variados (p.e., *random forest*, redes neuronales, *boosting*) en distintas condiciones y con distintos indicadores de desempeño, dificultando la comparación directa de sus resultados (Kundu et al., 2025). Esto implica que no está esclarecido cuál modelo o tipo de modelo suele rendir mejor para el pronóstico de calidad del agua en determinados escenarios, ni bajo qué condiciones específicas. En tercer lugar, la integración de técnicas XAI en este ámbito ha sido limitada: solo en años recientes se encuentran unos cuantos estudios que incorporan XAI para interpretar mode-

los de calidad del agua (Park et al., 2022; Bolick et al., 2023; Ramya et al., 2024), y usualmente se emplea una única técnica interpretativa por estudio (p.e., SHAP o LIME) sin evaluarse múltiples enfoques de explicabilidad de manera integral (Nasir et al., 2022).

Como consecuencia, una falta de evidencia sistematizada continúa de gran manera sobre cómo aprovechar mejor la XAI para extraer conocimiento de los modelos de predicción para el recurso hídrico y sobre cuál es el estado global de esta relación entre XAI y pronóstico de calidad del agua. Por lo tanto, estos vacíos en el conocimiento justifican la necesidad de realizar una revisión exhaustiva de la literatura. Cabe añadir que durante los últimos años, el número de investigaciones que aplican ML en este campo ha crecido exponencialmente, puesto que este auge no solo responde a la necesidad científica, sino también a la presión regulatoria y ética por incorporar mecanismos de transparencia, trazabilidad y explicabilidad en los sistemas basados en inteligencia artificial. En este sentido, la entrada en vigor del *Artificial Intelligence Act* (AI Act) de la Unión Europea en 2024 constituye un hito normativo fundamental, al establecer exigencias y requerimientos específicos sobre sistemas de alto riesgo, incluyendo aquellos utilizados en la monitorización medioambiental y gestión de recursos naturales como es el recurso hídrico (Commission, 2024). Bajo este contexto, ha surgido el interés por técnicas de IA explicable (XAI) que permitan no solo predecir, sino también comprender y justificar las decisiones de los modelos, asegurando su aceptabilidad en contextos operacionales y regulatorios. Por ello es que el siguiente artículo busca responder de forma general: ¿Cuál es el estado del arte de la aplicación de técnicas de Inteligencia Artificial Explicable en los modelos de aprendizaje automático para la predicción de la calidad del agua? Dicho de otro modo, se busca detallar cómo y en qué medida la XAI se ha utilizado para mejorar la interpretabilidad y eficacia de los pronósticos de calidad del agua en la literatura existente. Para abordar la cuestión general, se plantean las siguientes preguntas específicas:

- ¿Qué variables de entrada se han utilizado en los modelos de predicción de la calidad del agua reportados en la literatura, y cuáles de ellas emergen como más influyentes en el desempeño de dichos modelos?
- ¿Qué algoritmos de aprendizaje automático se han aplicado para pronosticar la calidad del agua (e.g., árboles de decisión, ensambles, redes neuronales) y qué tan comparables son sus desempeños reportados entre distintos estudios y casos de uso?
- ¿Qué técnicas de XAI se han integrado con estos modelos de predicción de la calidad del agua, y de qué manera contribuyen a interpretar los resultados y a mejorar la transparencia o confiabilidad de los pronósticos?

En concordancia con estas preguntas, el objetivo general de la presente revisión sistemática es analizar el estado actual de la aplicación de XAI en la predicción de la calidad del agua, identificando patrones, avances y desafíos en el uso conjunto de aprendizaje automático e interpretabilidad para este dominio. Este objetivo general se centra en sintetizar el conocimiento disponible sobre qué se ha hecho hasta la fecha en la intersección de XAI y calidad del agua, y qué vacíos quedan por resolver en futuras investigaciones.

Para lograr lo anterior, se definen los siguientes objetivos específicos:

- Identificar y seleccionar las variables de entrada más signifi-

cativas para el modelado de la calidad del agua bajo diversas condiciones ambientales.

- Revisar y comparar el rendimiento de diferentes algoritmos de aprendizaje automático en la predicción de indicadores clave de la calidad del agua.
- Evaluar la precisión predictiva y la aplicabilidad operacional de los modelos, destacando su interpretabilidad, estructura e integración con herramientas de IA explicable (XAI).

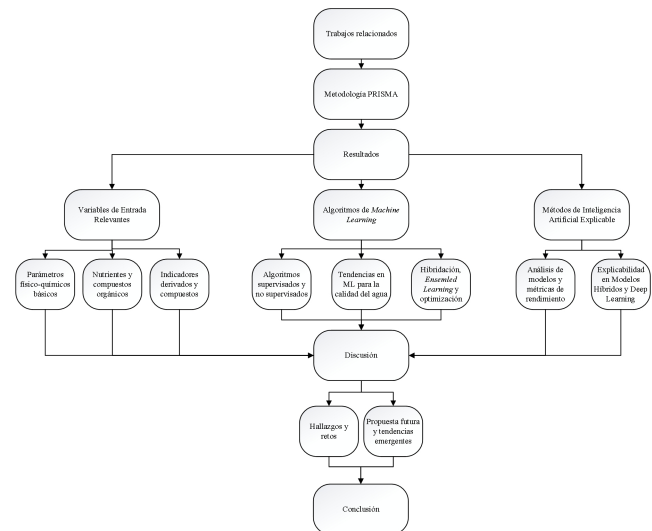


Figura 1: Trazado de la estructura del artículo

El mapa estructural presentado (Figura 1) organiza de forma jerárquica y temática los componentes principales de la revisión sistemática de literatura sobre modelos explicables de inteligencia artificial aplicados a la predicción de la calidad del agua. La estructura se inicia con los trabajos relacionados y continúa con la Metodología PRISMA, que detalla el proceso de selección de estudios. En la sección de Resultados, el análisis se divide en tres ejes centrales de las variables de entrada relevantes clasificadas, algoritmos de *machine learning* (agrupados por algoritmos supervisados y no supervisados, tendencias para la calidad del agua y técnicas de hibridación, *ensemble learning* y optimización) y métodos de inteligencia artificial explicable (XAI) (específicamente análisis de modelos y métricas de rendimiento y técnicas de XAI en modelos híbridos y de *deep learning*). La sección de Discusión se subdivide en el análisis de hallazgos y retos actuales, así como en propuestas futuras y patrones emergentes. Finalmente, se presentan las conclusiones, integrando los aportes clave de la revisión.

2 Metodología

La presente revisión del estado del arte se desarrolló siguiendo las directrices de la metodología PRISMA 2020 (Page et al., 2021). El protocolo inicial fue construido a fin de identificar, analizar y sintetizar la evidencia científica existente en la literatura respecto al uso de modelos XAI en la predicción de la calidad del agua. A continuación, se describen en detalle las fases implementadas.

2.1 Estrategia de búsqueda y fuentes de información

Se llevó a cabo la búsqueda en marzo de 2025 utilizando bases de datos indexadas *ScienceDirect* y *Scopus*, la delimitación tempo-

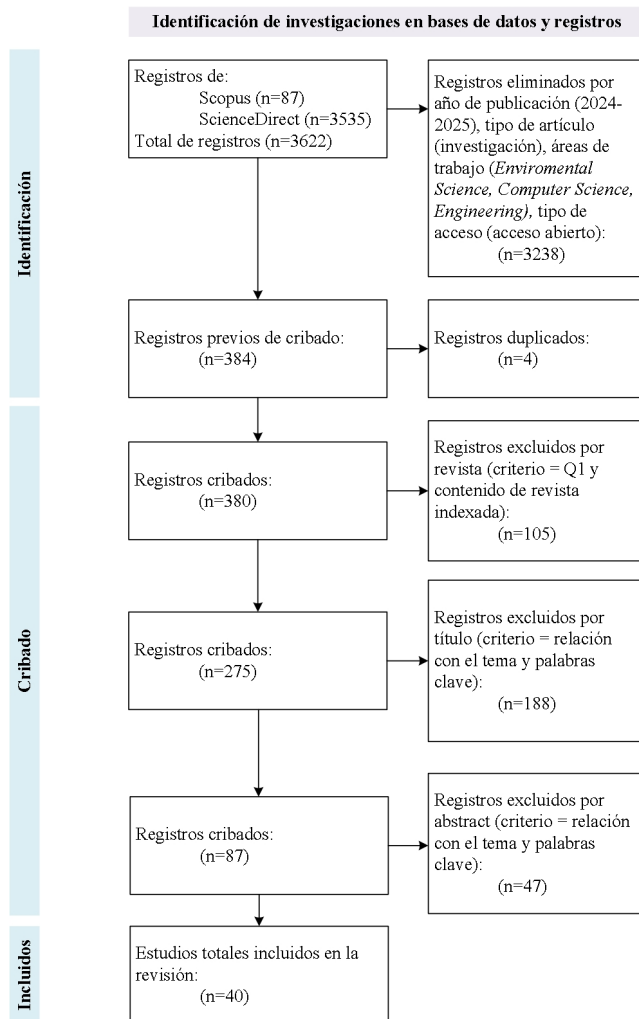


Figura 2: Desarrollo de Metodología PRISMA

ral se centró en los años 2024 y 2025; se adoptó dicho criterio de exclusión respondiendo a un criterio de actualidad científica y en concordancia con factores normativos emergentes, más aun el impacto del *Artificial Intelligence Act (AI Act)* de la Unión Europea, que entró en vigor durante 2024. Dicha disposición regula el desarrollo y uso de sistemas de inteligencia artificial en recursos de alto riesgo, incluyendo aquellos aplicados en monitoreo ambiental. Este acto ha generado una aceleración significativa en la publicación de investigaciones orientadas a la transparencia, explicabilidad y robustez de los modelos predictivos basados en aprendizaje automático (Commission, 2024). Se utilizó una composición *booleana* de palabras clave en inglés, estructuradas según los ejes de investigación:

- ‘machine learning’ AND ‘water quality’ AND ‘forecasting OR prediction’ AND ‘explainable AI’ OR ‘XAI’ OR interpretability

2.2 Criterios de inclusión y exclusión

Criterios de inclusión:

- Estudios publicados en revistas científicas indexadas entre 2024 y 2025.
- Uso explícito de técnicas de ML o DL para la predicción de variables relacionadas con calidad del agua.
- Implementación o análisis de técnicas de interpretabilidad (XAI), tales como *SHAP*, *LIME* o visualizaciones de importancia de variables.
- Estudios que incluyan evaluación cuantitativa del modelo (p.e., MAE, RMSE, R^2 , NSE).

Criterios de exclusión:

- Primero: Se excluyó del análisis el uso de artículos de investigación cuyo *journal* no se encuentra clasificado en Cuartil 1 (Q1) según *SCImago Journal Rank* (SJR), priorizando publicaciones con alto grado de relevancia en la academia y estén etiquetadas dentro de las áreas de ‘*Environmental Science*’, ‘*Computer Science*’ y ‘*Engineering*’.
- Segundo: Se descartó el uso de estudios cuyo título no incluía términos directamente relacionados con la temática tratada, tales como ‘*water quality*’, ‘*machine learning*’, ‘*explainable AI*’.
- Tercero: Se excluyó del estudio artículos cuyo resumen (*abstract*) no abordaba explícitamente el uso de técnicas de modelado predictivo en calidad del agua y se encontraban fuera de los objetivos del presente trabajo.

2.3 Proceso de selección

La depuración se ejecutó en cinco fases:

1. Identificación: Se obtuvieron inicialmente 380 artículos.
2. Filtrado por nombre de revista: Se eliminaron 105 artículos irrelevantes.
3. Filtrado por título: Se eliminaron 188 artículos que no cumplían con todos los criterios.
4. Filtrado por *abstract*: Se eliminaron 47 artículos que no cumplían con todos los criterios.
5. Inclusión final: Se seleccionaron los 40 artículos que cumplían con todos los criterios y que fueron leídos a profundidad.

Este proceso se documentó en el diagrama PRISMA (ver Figura 2).

2.4 Extracción y organización de datos

Para cada artículo se extrajeron sistemáticamente los siguientes campos:

- Modelo(s) de aprendizaje automático utilizados.
- Variables de entrada consideradas.
- Métricas de evaluación aplicadas.
- Técnicas XAI empleadas (*SHAP*, *LIME*).
- Región geográfica de aplicación.
- Tipo de agua (superficial, subterránea).
- Nivel de interpretabilidad reportado.

2.5 Plan de análisis de los estudios incluidos

Los 40 artículos seleccionados serán analizados en la sección de Resultados según un enfoque temático y comparativo alineado con los objetivos de esta revisión. Se establecerá una clasificación basada en:

- Tipo de variable predictiva (p.e., Water Quality Index (WQI), oxígeno disuelto (DO), nitrógeno total (TN), fósforo total (TP)).
- Algoritmo de aprendizaje automático empleado (*Random Forest*, *XGBoost*, *ANN*, *BiLSTM*).
- Técnica XAI utilizada (e.g. *SHAP* global, *SHAP* local, *LIME*, visualizaciones de importancia) e integración de herramientas explicables a actividades operacional.

3 Resultados

3.1 Análisis cuantitativo

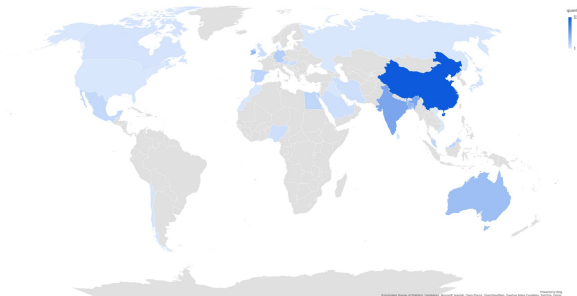


Figura 3: Distribución geográfica de la recurrencia de los autores

3.1.1 Distribución geográfica de la recurrencia de los autores

El mapa evidencia una alta concentración de autores residentes en China, India y Bangladesh, reflejando el fuerte impulso investigativo en regiones con alta presión hídrica y necesidad de herramientas predictivas robustas (Chen et al., 2024; Nong et al., 2024; Ramya et al., 2024). Europa también muestra presencia destacada, especialmente en Alemania, Irlanda y España, alineada con marcos regulatorios como el AI Act (Commission, 2024), que exige interpretabilidad en sistemas críticos como la gestión del agua (Maußner et al., 2025; Uddin et al., 2024; Pazo et al., 2024). La participación en América Latina y África es marginal, revelando una brecha investigativa que refuerza la necesidad de fomentar capacidades locales y cooperación científica internacional.

3.1.2 Principales revistas de publicación Se observa una fuerte inclinación hacia la aplicación de la inteligencia artificial y

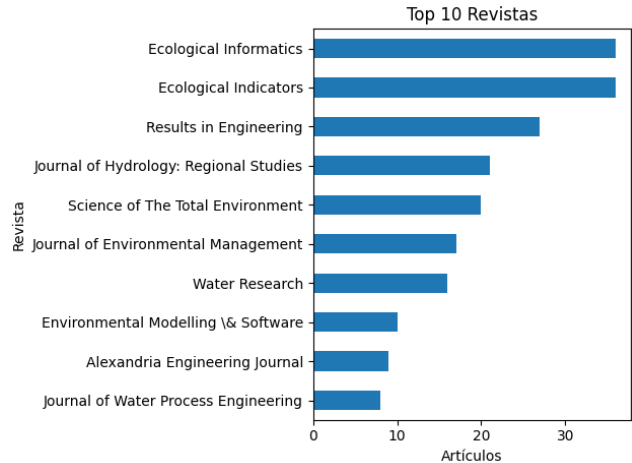


Figura 4: Frecuencia de aparición de artículos en revistas

el aprendizaje automático en el campo de la calidad del agua. Revistas como *Ecological Informatics* y *Ecological Indicators* lideran en la publicación de artículos que utilizan modelos avanzados y enfoques híbridos de aprendizaje automático para predecir y evaluar la calidad del agua. La importancia de la interpretabilidad de estos modelos se destaca en publicaciones de *Results in Engineering* y *Science of The Total Environment*, que exploran técnicas como *SHAP* y marcos de aprendizaje automático interpretables para entender el impacto de las regulaciones de descarga de flujo y la proliferación de cianobacterias. Además, la evaluación de riesgos para la salud y la gestión ambiental son temas recurrentes en el *Journal of Hydrology*, utilizando aprendizaje automático y métodos robustos para el monitoreo y la predicción. *Water Research* y *Alexandria Engineering Journal* también contribuyen significativamente con la evolución de modelos de calidad del agua basados en datos, la predicción de la demanda de agua y el uso de técnicas como Bi-LSTM y CFD. La presencia de *Environmental Modelling & Software* y *Journal of Water Process Engineering* también subraya la relevancia de la detección de anomalías y la clasificación de la calidad del agua a través de algoritmos de aprendizaje automático. En general, la gráfica refleja una tendencia clara hacia la adopción y el desarrollo de técnicas computacionales avanzadas para abordar los desafíos complejos en la predicción y gestión de la calidad del agua. (Kundu et al., 2025; Nallakaruppan et al., 2024; Nong et al., 2024; Commission, 2024).

3.1.3 Colaboración entre autores La red de coautoría revela dos núcleos colaborativos claramente diferenciados. El primer clúster, dominado por autores como Laith Alzubaydi y A.S. Albarhi, muestra una intensa colaboración interna, orientada principalmente a aplicaciones técnicas en redes inteligentes y evaluación predictiva del agua. El segundo grupo, representado por Mohammed A. Fadhel y Sayedali Mirjalili, se orienta a optimización de modelos e inteligencia computacional. Esta estructura sugiere una consolidación temática por afinidad metodológica, lo que respalda el enfoque dual observado en la literatura: predicción operativa y desarrollo metodológico, como en Makumbura et al. (2024) y Kundu et al. (2025).

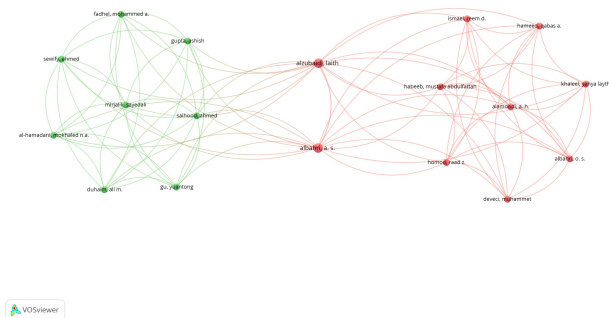


Figura 5: Red de colaboración entre autores

3.1.4 Red de co-ocurrencia de palabras clave La visualización revela una estructura temática bien definida en torno al nodo central ‘machine learning’, el cual actúa como eje articulador de múltiples líneas de investigación asociadas. Se observan cinco grandes conglomerados: (i) modelos supervisados y no supervisados, con términos como *random forest*, *regression* y *xgboost*; (ii) cambio climático y factores contextuales, como *influencing factors* y *climate change*; (iii) inteligencia artificial explicable, con menciones explícitas a *SHAP*, *explainable artificial intelligence*; (iv) índices de calidad del agua como *water quality index* y *principal component analysis*; y (v) técnicas de percepción remota y aprendizaje profundo, destacando términos como *deep learning*, *LSTM*, *convolutional neural network*, y de forma no anticipada, *remote sensing*. Este último término *remote sensing* constituye un hallazgo emergente no previsto en el diseño inicial de la revisión. Su fuerte conexión con términos como *chlorophyll-a*, *turbidity*, *landsat* y *Google Earth Engine* evidencia un enfoque en la medición indirecta de la calidad del agua mediante sensores satelitales y sistemas de observación remota. Esta línea temática se alinea con trabajos recientes que integran imágenes remotas como variable explicativa en modelos de ML/XAI, ampliando la aplicabilidad operacional más allá del monitoreo in-situ (Lu et al., 2024; Geng et al., 2024).

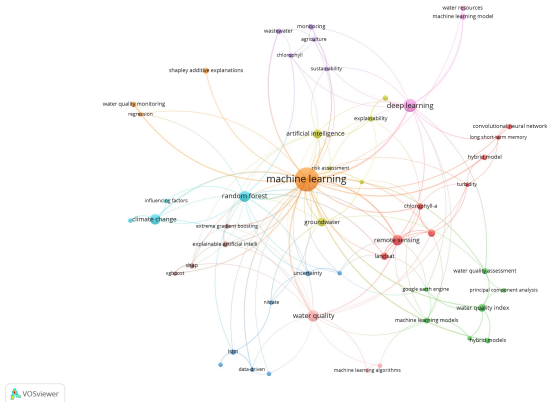


Figura 6: Red de co-ocurrencia de palabras clave

3.2 Análisis de contenido

3.2.1 Parámetros físico-químicos Los parámetros físico-químicos básicos dominan la construcción de modelos ML de calidad de agua: *pH* y *DO* son esenciales, seguidos por *EC*, temperatura (*T*) y *TDS*, reflejando prácticas estándar de monitoreo (Kundu et al., 2025; Makumbura et al., 2024; Wan et al., 2025; Roy et al., 2024; M. T. Islam et al., 2024). De manera más específica, el *pH* destaca como la variable más frecuente, presente en 26 de los 40 artículos analizados. Esto refleja su importancia fundamental como indicador general de las condiciones químicas del agua y su influencia en otros procesos (p.e., la toxicidad de contaminantes o la disponibilidad de nutrientes). De modo similar, el oxígeno disuelto (*DO*) aparece en 23 estudios, lo que subraya su rol para evaluar el estado trófico y la salud ecológica de cuerpos de agua (Kundu et al., 2025; Xu et al., 2024). Además, variables como la conductividad eléctrica (*EC*) (19 estudios) y la temperatura del agua (15 estudios) también muestran alta recurrencia, evidenciando que la salinidad/concentración iónica y el régimen térmico son consideraciones estándar en los modelos de predicción (Wan et al., 2025; Das, 2025; Kundu et al., 2025; Uddin et al., 2025; Xiao et al., 2025).

Cuadro 1

Parámetros físico-químicos básicos más frecuentes utilizados como variables de entrada (n=40).

| Variable | Frecuencia (n/40) | Artículos representativos |
|---------------------------------|-------------------|--|
| pH | 26 | (Kundu et al., 2025; Makumbura et al., 2024) |
| Oxígeno disuelto (DO) | 23 | (Kundu et al., 2025; Xu et al., 2024) |
| Conductividad eléctrica (EC) | 19 | (Wan et al., 2025; Das, 2025) |
| Temperatura del agua | 15 | (Uddin et al., 2025; Song et al., 2024) |
| Sólidos disueltos totales (TDS) | 11 | (Das, 2025; Eid et al., 2024) |

En menor medida, pero aún significativos, se encuentran los sólidos disueltos totales (*TDS*) (presentes en 11 estudios), parámetro relacionado con la mineralización del agua (Das, 2025; Eid et al., 2024). Cabe notar que otras variables físicas como la turbidez o sólidos suspendidos aparecen con frecuencia ligeramente inferior ($\approx 9/40$), mientras que ciertos iones mayoritarios (p.e., Cl^- , SO_4^{2-}) se incluyeron principalmente en estudios de agua subterránea o potable específicos. En conjunto, la prevalencia de estos cinco parámetros básicos concuerda con las prácticas estándar de monitoreo *in situ* y su uso amplio sugiere que son considerados predictores esenciales y fácilmente medibles para la calidad del agua en una variedad de contextos. Su omnipresencia indica un consenso implícito en la literatura de que las condiciones físico-químicas de base deben incorporarse en los modelos, proporcionando una línea base interpretativa común.

3.2.2 Nutrientes y compuestos orgánicos Los nutrientes y compuestos orgánicos (*COD*, NH_4^+ , *BOD*₅, fósforo *P* y nitrato NO_3^-) capturan cargas orgánicas y procesos de eutrofización (Ding et al., 2025; Alfwezan et al., 2024; Gupta et al., 2024; Irwan et al., 2025; Usang et al., 2025).

La demanda química de oxígeno (*DQO/COD*) encabeza la lista (19 estudios) como métrica integradora de la carga orgánica oxidable en el agua (Ding et al., 2025; Makumbura et al., 2024). Su amplio uso sugiere que muchos modelos buscan explicar y predecir la presencia de materia orgánica y contaminantes reductores, especialmente en contextos de calidad de agua superficial. El amonio (NH_4^+) aparece en segundo lugar (15 estudios), resaltando la atención a las formas reducidas de nitrógeno derivadas de descargas urbanas o procesos de descomposición (Alfwzan et al., 2024; Ding et al., 2025). Junto con el amonio, la demanda bioquímica de oxígeno a 5 días (*DBO₅*) está presente en 12 estudios, confirmando su rol como indicador tradicional de materia orgánica biodegradable (Alfwzan et al., 2024; Kundu et al., 2025). La inclusión concurrente de *DBO* y *DQO* en numerosos trabajos sugiere que se evalúa tanto la fracción fácilmente biodegradable como la carga orgánica total.

Cuadro 2
Nutrientes y compuestos orgánicos más frecuentes en los modelos revisados.

| Variable | Frecuencia (n/40) | Artículos representativos |
|--|-------------------|---|
| Demanda química de oxígeno (<i>COD</i>) | 19 | (Ding et al., 2025; Makumbura et al., 2024) |
| Amonio (NH_4^+) | 15 | (Alfwzan et al., 2024; Ding et al., 2025) |
| Demanda bioquímica de oxígeno (<i>BOD₅</i>) | 12 | (Alfwzan et al., 2024; Kundu et al., 2025) |
| Fósforo total/disuelto | 11 | (Ding et al., 2025; Song et al., 2024) |
| Nitrato (NO_3^-) | 10 | (Kushwaha et al., 2024; Das, 2025) |

Por otra parte, los nutrientes inorgánicos relacionados a la eutrofización tienen una presencia algo menor pero notable: el fósforo *P* (ya sea fósforo total u orto-fosfato) aparece en 11 estudios, mientras que el nitrato (NO_3^-) figura en 10 (Ding et al., 2025; Song et al., 2024; Kushwaha et al., 2024; Das, 2025). Estos nutrientes suelen ser cruciales para modelar florecimientos algales y calidad ecológica, por lo que su presencia en alrededor de una cuarta parte de los modelos sugiere que una proporción importante de investigaciones aborda escenarios de eutrofización. No obstante, la relativamente menor frecuencia de nitratos/fósforo en comparación con pH o DO indica que no todos los estudios incorporan el componente nutrientes, posiblemente porque algunos se enfocan más en contaminación orgánica o parámetros físico-químicos básicos. De hecho, solo unos pocos trabajos (p.e., (Song et al., 2024)) se centran explícitamente en pronosticar algas (clorofila-*a*) o nutrientes, reflejando una brecha en la cobertura temática: la intersección de modelos de predicción y eventos de eutrofización aún está emergiendo. En conjunto, estos hallazgos evidencian un énfasis predominante en variables de calidad relacionadas con materia orgánica y nutrientes más comunes, alineado con los problemas de contaminación orgánica y eutrofización que son frecuentes en la literatura reciente. Sin embargo, la diversidad en la selección de variables (distintos tipos de nitrógeno, fósforo, *DBO* vs *DQO*) refuerza la observación de que no existe un consenso claro sobre cuáles son las variables predictoras óptimas para calidad del agua en general. Cada estudio tiende a elegir los parámetros según su problemática local, lo que dificulta la comparación directa entre casos y deja abierta la pregunta de cuáles variables son más informativas de

forma universal.

3.2.3 Indicadores derivados o compuestos Los indicadores derivados o compuestos (precipitación, SAR, radiación solar, viento y temperatura del aire) aparecen esporádicamente; su baja frecuencia sugiere una oportunidad para integrar forzantes climáticos y métricas sintéticas en futuros modelos (Nong et al., 2024; Das, 2025; Ramya et al., 2024; Lloyd et al., 2024; M. T. Islam et al., 2024; Xue et al., 2024).

Este grupo engloba tanto índices sintéticos calculados a partir de variables primarias, como variables externas (climáticas, hidrológicas o de teledetección). En general, estas variables muestran frecuencias de uso mucho menores en comparación con los parámetros *in situ* tradicionales, evidenciando que solo una fracción reducida de estudios ha explorado integrar factores adicionales más allá de las mediciones directas de agua. Variables climáticas como la precipitación (incluida en 4 estudios) y la temperatura del aire (2 estudios) aparecen únicamente en trabajos que intentan vincular condiciones meteorológicas con la calidad del agua (Nong et al., 2024; Maußner et al., 2025). Su presencia sugiere que algunos modelos recientes han comenzado a incorporar la influencia de lluvias, estacionalidad térmica y otros forzantes atmosféricos, aunque esto todavía no es una práctica común en 2024–2025.

De modo similar, factores hidroclimáticos derivados, como la radiación solar (3 estudios) o la velocidad del viento (2 estudios), se han considerado esporádicamente para capturar efectos de evaporación, insolación o mezcla en cuerpos de agua (Nong et al., 2024; Lloyd et al., 2024). La relativamente baja frecuencia de estas variables externas corrobora que la mayoría de investigaciones se han circunscrito a datos de calidad medidos localmente, reservando la inclusión de datos climáticos/hidrológicos solo para casos específicos (p.e., pronóstico en lagos con fuertes fluctuaciones estacionales). Por otro lado, dentro de este grupo emerge la relación de adsorción de sodio (*SAR*), índice compuesto relevante en contextos de riego y salinidad del suelo, incluido en 4 estudios (principalmente evaluaciones de calidad de agua para uso agrícola) (Das, 2025; M. T. Islam et al., 2024). La presencia de *SAR* como uno de los indicadores derivados más frecuentes indica un interés particular por parte de algunos autores en evaluar la idoneidad del agua para riego y los impactos de la salinidad, integrando consideraciones agronómicas en los modelos explicables. Junto a *SAR*, unos pocos estudios (≈ 1 -2) han utilizado otros índices de teledetección o calidad integrales (p.e., índices espectrales como el *NDWI* (Normalized Difference Water Index) para detectar agua/turbidez, o el *FAI* (Floating Algae Index) para estimar floraciones algales pero cada uno de estos aparece solo de forma puntual en la revisión (usados en un único estudio cada uno) (Tian et al., 2024).

En conjunto, la escasa recurrencia de los indicadores compuestos refleja que la aplicación de modelos de aprendizaje aún se ha centrado mayoritariamente en variables convencionales, y la integración de fuentes de datos avanzadas (clima, sensores remotos, índices globales) está en sus primeras etapas. Esta situación deja ver una brecha y oportunidad para futuros trabajos: incorporar de manera más sistemática variables derivadas podría mejorar el poder predictivo y la interpretabilidad de los modelos, permitiendo capturar aspectos holísticos de la calidad del agua (como la influencia climática o la dinámica de algas) que actualmente están

Cuadro 3

Indicadores derivados o compuestos y variables externas más frecuentes.

| Variable | Frecuencia (n/40) | Artículos representativos |
|--------------------------------------|-------------------|---|
| Precipitación | 4 | (Nong et al., 2024; Maußner et al., 2025) |
| Relación de adsorción de sodio (SAR) | 4 | (Das, 2025; M. T. Islam et al., 2024) |
| Radiación solar | 3 | (Nong et al., 2024; Lloyd et al., 2024) |
| Velocidad del viento | 2 | (Nong et al., 2024; Lloyd et al., 2024) |
| Temperatura del aire | 2 | (Nong et al., 2024; Maußner et al., 2025) |

subrepresentados. Conforme aumenta la disponibilidad de datos satelitales y pronósticos meteorológicos, es previsible que próximos estudios exploren la inclusión de tales indicadores, avanzando hacia enfoques más integrales.

3.2.4 Indicadores de compuestos metálicos y minerales disueltos Un grupo de variables de entrada para poder cuantificar la calidad del agua se basa en indicadores de compuestos metálicos o minerales disueltos en el agua (cadmio *Cd*, cobre *Cu*, zinc *Zn*, magnesio *Mg*, manganeso *Mn*, aluminio *Al*, cromo *Cr*). Ello sugiere la existencia de una línea de evaluación centrada en la toxicidad por metales pesados o en la presencia de minerales disueltos como parte del análisis de calidad del agua, especialmente en contextos de agua potable, subterránea o cuerpos receptores cercanos a fuentes industriales y mineras; además abre la posibilidad de incorporar análisis de metales como capa explicativa adicional en futuros modelos aplicados a la predicción y gestión integral de la calidad del agua (Ding et al., 2025; Kushwaha et al., 2024; Usang et al., 2025; Xiao et al., 2025; Sajib et al., 2024; Sajib et al., 2025; Eid et al., 2024).

Variables metálicas o minerales disueltas constituye una dimensión crítica de análisis, orientada principalmente hacia la evaluación de la toxicidad inorgánica y la presencia de elementos trazas, particularmente en contextos de abastecimiento humano, aguas subterráneas o zonas con influencia industrial y minera. A diferencia de los parámetros fisicoquímicos convencionales (p.e., *pH*, *DO*, *EC*) que reflejan condiciones generales del cuerpo de agua, los metales y minerales permiten detectar contaminantes específicos, muchos de los cuales poseen efectos acumulativos, carcinogénicos o disruptivos sobre la salud humana y la biota acuática. La inclusión de elementos como cadmio (*Cd*), cromo (*Cr*), plomo (*Pb*) o arsénico (*As*) en un subconjunto de estudios (1 a 5 apariciones cada uno) revela que algunos autores han abordado explícitamente la necesidad de evaluar riesgos tóxicos mediante modelos explicables, sobre todo en escenarios de contaminación localizada o regulaciones normativas estrictas (Ding et al., 2025; Usang et al., 2025; Sajib et al., 2024; Sajib et al., 2025). A su vez, minerales esenciales como magnesio (*Mg*), manganeso (*Mn*) o hierro (*Fe*) también se han considerado, ya sea por su presencia natural o como indicadores de procesos geoquímicos que afectan la potabilidad y corrosividad del agua (Kushwaha et al., 2024; Xiao et al., 2025; Eid et al., 2024). Si bien el magnesio alcanza una frecuencia moderada (8 estudios), la mayoría de los demás elementos aparecen en menos del

20 % de los trabajos revisados, lo que sugiere que su incorporación aún no forma parte de un estándar metodológico en los modelos de calidad del agua. Esta baja recurrencia puede deberse a la limitada disponibilidad de datos para estos elementos, la necesidad de métodos analíticos costosos o a la falta de integración sistemática con bases de datos de calidad del agua.

Al igual que ocurre con los indicadores climáticos o derivados por teledetección, la baja frecuencia de uso de variables metálicas disueltas apunta a una brecha clara en el diseño de modelos predictivos actuales. No obstante, su presencia en múltiples estudios recientes evidencia una tendencia incipiente hacia el desarrollo de modelos explicables capaces de detectar, cuantificar y atribuir riesgos específicos relacionados con metales pesados y minerales. En este sentido, la incorporación futura de estas variables como capas explicativas complementarias representa una oportunidad clave para ampliar el alcance predictivo y el valor normativo de los sistemas *XAI*, especialmente en regiones afectadas por actividades extractivas, suelos contaminados o deterioro de fuentes de agua potable. Conforme se disponga de bases de datos más integrales y sensores específicos, se espera que estos modelos puedan evolucionar hacia una gestión más precisa y preventiva de contaminantes inorgánicos.

Cuadro 4

Frecuencia de inclusión de variables metálicas y minerales disueltos en los modelos predictivos de calidad del agua (n=40).

| Elemento /IÓN | Frecuencia (n/40) | Artículos representativos |
|-------------------------|-------------------|---|
| Cobre (Cu) | 5 | (Ding et al., 2025; Usang et al., 2025; Sajib et al., 2025; Eid et al., 2024; Sajib et al., 2024) |
| Zinc (Zn) | 3 | (Ding et al., 2025; Sajib et al., 2025; Xiao et al., 2025) |
| Magnesio (Mg) | 8 | (Kushwaha et al., 2024; Eid et al., 2024; Das, 2025; Song et al., 2024) |
| Manganeso (Mn) | 3 | (Sajib et al., 2025; Eid et al., 2024; Xiao et al., 2025) |
| Aluminio (Al) | 1 | (Usang et al., 2025) |
| Cromo (Cr) | 3 | (Ding et al., 2025; Usang et al., 2025; Sajib et al., 2025) |
| Cadmio (Cd) | 3 | (Castillo et al., 2024; Sajib et al., 2025; Sajib et al., 2024) |
| Hierro (Fe) | 4 | (Xiao et al., 2025; Sajib et al., 2025; Eid et al., 2024; Sajib et al., 2024) |
| Plomo (Pb) | 3 | (Ding et al., 2025; Castillo et al., 2024; Sajib et al., 2024) |
| Arsénico (As) | 1 | (Usang et al., 2025) |
| Vanadio (V) | 1 | (Usang et al., 2025) |

Cuadro 5

Variables de entrada utilizadas en los estudios analizados.

| Referencia | Año | Nombre del Artículo | Variables medidas |
|----------------------------|------|--|--|
| (Alfwzan et al., 2024) | 2024 | Water quality assessment using Bi-LSTM and computational fluid dynamics (CFD) techniques | <i>BOD, COD, TC, FC, DO, pH, WT, AMM, TKN</i> |
| (Ding et al., 2025) | 2025 | Using multiple machine learning algorithms to optimize the water quality index model and their applicability | <i>DO, BOD₅, COD_{Cr}, COD_{Mn}, TN, TP, NH₃-N, Pb, Cu, Zn</i> |
| (Kundu et al., 2025) | 2025 | Unveiling the hidden connections: Using explainable artificial intelligence to assess water quality criteria in nine giant rivers | <i>Temperatura, pH, DO, BOD, FC, NO₃-NO₄, EC</i> |
| (Gupta et al., 2024) | 2024 | Time Series Analysis and Forecasting of Water Quality Parameters along Yamuna River in Delhi | <i>pH, COD, BOD, DO</i> |
| (Uddin et al., 2025) | 2025 | The role of optimizers in developing data-driven model for predicting lake water quality incorporating advanced water quality model | <i>Temperatura, Turbidez, TON, MRP, pH, DO, DIN</i> |
| (Wan et al., 2025) | 2025 | Temporal and spatial feature extraction with graph neural networks for river water quality prediction | <i>NH₃-N, DO, COD_{Mn}, TN, Turbidez, EC, pH, PCP, Temperatura</i> |
| (Kushwaha et al., 2024) | 2024 | Stacked hybridization to enhance the performance of artificial neural networks (ANN) for prediction of water quality index in the Bagh river basin, India | <i>NO₃, SO₄, Ca, Mg, K</i> |
| (Das, 2025) | 2025 | Spatiotemporal evaluation and impact of superficial factors on surface water quality for drinking using innovative techniques in Mahanadi River Basin, Odisha, India | <i>Temperatura, pH, EC, TDS, TSS, DO, COD, TH, Na, Ca⁺, Mg⁺, Cl⁻, SO₄, NO₃, F⁻, SAR, BOD</i> |
| (Xiao et al., 2025) | 2025 | Spatiotemporal dynamics and optimization of water quality assessment in the Nantong section of the Yangtze River Basin: A WQImin approach | <i>Temperatura, pH, EC, Turbidez, DO, COD_{Mn}, BOD₅, SO₄, TN, NH₄-N, NO₃-N, TP, Cl⁻, F⁻, Fe, FC</i> |
| (Irwan et al., 2025) | 2025 | River water quality monitoring using machine learning with multiple possible in-situ scenarios | <i>DO, BOD, COD, SS, pH, NH₃-NL, SS</i> |
| (Castillo et al., 2024) | 2024 | Improving river water quality prediction with hybrid machine learning and temporal analysis | <i>Cd, Cr, BOD₅, DO, FC, FI, FOG, Hg, NH₃, NO₃, Pb, pH, TSS, SULF, TDS, Temperatura, Zn</i> |
| (Usang et al., 2025) | 2025 | Integrating principal component analysis, fuzzy inference systems, and advanced neural networks for enhanced estuarine water quality assessment | <i>pH, EC, TDS, Turbidez, TH, DO, TA, Cl, SO₄, PO₄, NO₃, BOD, COD, O&G, Cu, Cr, As, V, Al</i> |
| (M. T. Islam et al., 2024) | 2024 | Regional irrigation water quality index for the Old Brahmaputra River, Bangladesh: A multivariate and GIS-based spatiotemporal assessment | <i>EC, TDS, Ca, PO₄, SO₄, NH₄, NO, CO₃, HCO₃, SAR</i> |
| (Song et al., 2024) | 2024 | Quantitative prediction of water quality in Dongjiang Lake watershed based on LUCC | <i>chl-a, Temperatura, TP</i> |
| (Roy et al., 2024) | 2024 | Principal component analysis incorporated water quality index modeling for Dhaka-based rivers | <i>Temperatura, pH, EC, TA, Cl, TS, TDS, TSS, DO, BOD, COD, Turbidez</i> |

Cuadro 5 – continuación de la página anterior

| Referencia | Año | Nombre del Artículo | Variables medidas |
|----------------------------|------|---|---|
| (Xue et al., 2024) | 2024 | Predictive modeling of nitrogen and phosphorus concentrations in rivers using a machine learning framework: A case study in an urban-rural transitional area in Wenzhou China | Land use/cover (Construcción & bosques), GS, Agua, Wetland, DEM, Soil _{TP} , Soil _{TN} , Soil ADH, CEC, CLF, Sand, DIA, GDP |
| (Lloyd et al., 2024) | 2024 | Predicting recreational water quality and public health safety in urban estuaries using Bayesian Networks | Enterococcus, EC, HR, SE, SD, CC, WS, WD, SL, UV, ubicación, hora de muestra |
| (Han et al., 2025) | 2025 | Predicting and investigating water quality index by robust machine learning methods | pH, EC, TDS, CO ₃ , HCO ₃ , Cl, F, NO ₃ , SO ₄ , Na, K, Ca, Mg, TH, SAR, RSC |
| (Sajib et al., 2025) | 2025 | Novel groundwater quality index (GWQI) model: A reliable approach for the assessment of groundwater | Ph, EC, TDS, Zn, Fe, Mn, Cr, Cd, Cu |
| (Rigueira et al., 2025) | 2025 | Multivariate functional data analysis and machine learning methods for anomaly detection in water quality sensor data | NH ₄ , EC, DO, pH, Turbidez, Temperatura |
| (Lausch et al., 2025) | 2025 | Monitoring and modelling landscape structure, land use intensity and landscape change as drivers of water quality using remote sensing | Turbidez, chl-a, HAB, SOA, SIA |
| (Anand et al., 2024) | 2024 | Machine learning approach for water quality predictions based on multispectral satellite imagery | DO, TDS, pH, TSS |
| (M. S. Islam et al., 2024) | 2024 | Long-term trend prediction of surface water quality of two main river basins of China using Machine Learning Method | DO, COD _{Mn} , NH ₃ -N |
| (Chen et al., 2024) | 2024 | A coupled model to improve river water quality prediction towards addressing non-stationarity and data limitation | TP, TN, NH ₃ -N, COD _{Mn} |
| (Nong et al., 2024) | 2024 | A novel coupling interpretable machine learning framework for water quality prediction and environmental effect understanding in different flow discharge regulations of hydro-projects | COD _{Mn} , DO, pH, FD, SR, Temperatura, PCP, SP, WS |
| (Eid et al., 2024) | 2024 | An advanced approach for drinking water quality indexing and health risk assessment supported by machine learning modelling in Siwa Oasis, Egypt | TDS, Ph, EC, Na ⁺ , K ⁺ , Mg ₂ ⁺ , Ca ₂ ⁺ , Mn, Fe, Cl, SO ₄ , HCO ₃ , NO ₃ , TH |
| (Ehteram et al., 2024) | 2024 | An advanced deep learning model for predicting water quality index | DO, BOD, COD, SS, NH ₃ -NL, pH |
| (Yousefi et al., 2024) | 2024 | An innovative combination of extra trees within adaboost for accurate prediction of agricultural water quality indices | K ⁺ , HCO ₃ ⁻ , F ⁻ , Cl ⁻ , SO ₄ , Ca ₂ ⁺ , Mg ₂ ⁺ , Na ⁺ , EC, SAR, PS |
| (Xu et al., 2024) | 2024 | Assessing and predicting water quality index with key water parameters by machine learning models in coastal cities, China | Temperatura, pH, DO, EC, Turbidez, COM, AN, TP, TN |
| (Sajib et al., 2024) | 2024 | Assessing water quality of an ecologically critical urban canal incorporating machine learning approaches | Temperatura, DO, EC, TDS, Pb, Cd, Cu, Fe, Turbidez, Na ⁺ , K ⁺ , Ca ₂ ⁺ , Mg ₂ ⁺ , Cl |

Cuadro 5 – continuación de la página anterior

| Referencia | Año | Nombre del Artículo | Variables medidas |
|--------------------------|------|--|---|
| (Ramya et al., 2024) | 2024 | Comprehensive analysis of multiple classifiers for enhanced river water quality monitoring with explainable AI | DO, Temperatura, pH , EC , $NO_2 + NO_3$, $Si-O_2$, TA , TH , F , B , Mg , N , K , Cl , SO_4 , HCO_3 , BOD |
| (Uddin et al., 2024) | 2024 | Data-driven evolution of water quality models: An in-depth investigation of innovative outlier detection approaches-A case study of Irish Water Quality Index (<i>IEWQI</i>) model | pH , DOX , SAL , BOD_5 , Temperatura, $TRAN$, TON , MRP , DIN |
| (Karpagam et al., 2024) | 2024 | Deep learning-based water quality index classification using stacked ensemble variational mode decomposition | ph , TH , $TDSS$, SO_4^{2-} , EC , OC , Potabilidad |
| (Kandasamy et al., 2025) | 2025 | Enhanced remote sensing and deep learning aided water quality detection in the Ganges River, India supporting monitoring of aquatic environments | $a_{pH}(\lambda)$, $a_{dg}(\lambda)$, $a(\lambda)$, $b_{bp}(\lambda)$, $b_{dg}(\lambda)$, $b_b(\lambda)$, $Lu(O^-)/Ed(O^-)$, $r_{rs}(\lambda)$, $Lw(0+)/Ed(O^+)$, $R_{rs}(\lambda)$, $Eu(O^-)/Ed(O^-)$, $a_w(\lambda)$ |
| (Tsegay et al., 2025) | 2025 | Enhancing long-term water quality modeling by addressing base demand, demand patterns, and temperature uncertainty using unsupervised machine learning techniques | Demanda de agua, Chlorine Concentration |
| (Pazo et al., 2024) | 202 | Enhancing water quality prediction for fluctuating missing data scenarios: A dynamic Bayesian network-based processing system to monitor cyanobacteria proliferation | Chl-a, Nivel de agua, EC , pH , NH_4 , Temperatura |
| (Maußner et al., 2025) | 2025 | Explainable artificial intelligence for reliable water demand forecasting to increase trust in predictions | Demanda de agua de los últimos 7 días, Mes, Día de la semana |
| (Tian et al., 2024) | 2024 | Exploring the effects of climate change and urban policies on lake water quality using remote sensing and explainable artificial intelligence | FAI , TWI , ABI , $NDWI$ |
| (Liu et al., 2025) | 2025 | FNN-BiLSTM-Attention-DA: A hybrid fuzzy neural network and BiLSTM with multi-sensor information fusion for water quality monitoring and warning | DO , I_{Mn} , NH_3-N , TP , TN |
| (Makumbura et al., 2024) | 2024 | Advancing water quality assessment and prediction using machine learning models, coupled with explainable artificial intelligence (XAI) techniques like shapley additive explanations (SHAP) for interpreting the black-box nature | pH , EC , DO , COD , BOD , Cl^- , NO_3^- , PO_4 |

3.3 Machine Learning para Predicción de Calidad del Agua

El aprendizaje automático (ML) en la predicción de calidad del agua consiste en aplicar algoritmos computacionales para analizar series temporales hidrológicas y predecir parámetros fisicoquímicos clave tales como pH, oxígeno disuelto (DO) y turbidez (Park et al., 2022; Bolick et al., 2023), asimismo estos aprenden de patrones encontrados en datos hidrológicos y ambientales para estimar indicadores de calidad.

Se emplean desde modelos de series temporales clásicas (p.e., ARIMA, regresiones lineales) hasta redes neuronales profundas (LSTM, CNN, FNN) y ensamblados (Random Forest, XGBoost), capaces de capturar dependencias temporales y no lineales en parámetros fisicoquímicos (pH, oxígeno disuelto, nitratos, turbidez, demanda bioquímica) (Gupta et al., 2024; Alfwzan et al., 2024).

Estos modelos procesan datos multivariados provenientes de redes de sensores in situ (p.e., monitores fluviales, boyas) y de sensores remotos (p.e., imágenes satelitales multiespectrales como Sentinel-2 o Landsat-8) (Anand et al., 2024; Kandasamy et al., 2025). Con ellos es posible tanto predecir valores continuos de calidad (p.e., concentración de contaminantes) como calcular índices compuestos (Water Quality Index, WQI) que resumen múltiples variables. Esta aproximación permite modelar relaciones complejas no lineales en datos ambientales, superando limitaciones de métodos tradicionales (Savoy et al., 2023), y en consecuencia, ofrece mecanismos avanzados para integrar series temporales hidroquímicas y espaciales en la vigilancia y pronóstico de la calidad hídrica (Gupta et al., 2024; Anand et al., 2024).

No obstante, los modelos más precisos suelen ser ‘cajas negras’ difíciles de interpretar. Existe una tensión inherente entre precisión predictiva y explicabilidad: mientras las redes profundas o ensamblados maximizan la exactitud, su estructura compleja impide entender cómo influyen las variables en las predicciones. Ante esto, la Inteligencia Artificial Explicable (XAI) juega un rol crítico en gestión ambiental e institucional, pues proporciona mecanismos para desentrañar los modelos opacos. SHAP (Shapley Additive exPlanations) atribuyen a cada variable predictora una contribución aditiva a la salida del modelo, ofreciendo explicaciones tanto globales (importancia promedio) como locales (impacto en cada predicción) (Kundu et al., 2025). La técnica LIME (Local Interpretable Model-Agnostic Explanations) genera aproximaciones locales lineales de cualquier modelo, facilitando la comprensión de casos individuales. De modo similar, los PDP (Partial Dependence Plots) e ICE (Individual Conditional Expectation) grafican los efectos marginales promedio versus individuales de una variable, y los ALE (Accumulated Local Effects) corrigen sesgos debidos a colinealidad entre predictores. Estas herramientas permiten validar científicamente la coherencia del modelo (p.e., la influencia esperable de la turbidez o la temperatura) y proveen ‘transparencia caja blanca’ en el análisis de calidad hídrica (Kundu et al., 2025).

Se ha enfatizado que, en contextos de alto riesgo, conviene privilegiar modelos interpretables o enriquecidos con XAI antes que depender de sofisticadas cajas negras sin explicaciones (Rudin, 2019; Commission, 2024). Estudios recientes aplican estas técnicas: Makumbura et al. (2024) interpretan modelos de ML de calidad hídrica usando SHAP (Makumbura et al., 2024), y Ramya et al. (2024) incorporan explicabilidad en varios clasificadores de calidad del agua (Ramya et al., 2024). En suma, la XAI es esencial para que los modelos predictivos sean transparentes, confiables y útiles para tomadores de decisión ambientales.

Este interés por la trazabilidad y la explicabilidad está fuertemente impulsado por el marco regulatorio. La Unión Europea publicó en 2024 el *Artificial Intelligence Act*, que exige transparencia, explicabilidad y trazabilidad en sistemas de IA considerados de alto riesgo, incluyendo aplicaciones en monitoreo medioambiental y gestión de recursos naturales (Maußner et al., 2025; Commission, 2024). De este modo, cualquier modelo de predicción de calidad del agua usado en entornos institucionales deberá permitir auditar su funcionamiento interno. Maurer et al. (2025) destacan que estas regulaciones elevan la confianza en pronósticos hídricos al requerir explicaciones comprensibles para expertos y reguladores (Maußner et al., 2025). En consecuencia, la trazabilidad de las decisiones (p.e., registrar qué variables influyeron en cada pronóstico) se vuelve un requisito indispensable para la adopción institucional de modelos basados en ML.

Finalmente, la revisión del estado del arte revela problemas metodológicos aún abiertos. Existe gran heterogeneidad de entradas (qué variables incluyen) y de métricas de evaluación entre los estudios de calidad del agua. Cada autor suele probar conjuntos disímiles de predictores, tales como índices compuestos o variables químicas específicas, bajo diferentes condiciones, complicando la comparación directa de resultados (Chen et al., 2024; Park et al., 2022). Asimismo, se emplea una diversidad de criterios cuantitativos (coeficientes de determinación, MAE, RMSE, exactitud de clasificación), de modo que no hay consenso sobre cuál métrica usar como estándar. Esta variabilidad dificulta establecer qué métodos funcionan mejor en general. Como ejemplo, Gupta et al. (2024) comparan modelos ARIMA y LSTM mediante errores de pronóstico para el río Yamuna (Gupta et al., 2024), mientras otros estudios comparan clasificadores usando precisión o puntuaciones F1 (Park et al., 2022). Para avanzar hacia conclusiones robustas, la comunidad subraya la necesidad de protocolos estandarizados sobre entradas y métricas, que permitan evaluar consistentemente los modelos de predicción de la calidad del agua.

3.3.1 Modelos supervisados de Machine Learning En los estudios hidrológicos recientes es común el uso de modelos de aprendizaje supervisado para predecir indicadores de calidad del agua. Estos modelos de datos etiquetados aprovechan datos históricos ambientales e hidrológicos para ajustar relaciones complejas entre variables de entrada (precipitación, caudales, parámetros fisicoquímicos) y salidas de calidad.

Como muestra de ello, Song et al. desarrollaron un modelo cuantitativo basado en SVM para predecir la calidad del agua en función de los cambios de uso y cobertura del suelo, integrando datos espaciales y técnicas de ML para la gestión ambiental de cuencas (Song et al., 2024). Verbigracia, Ding et al. (2025) optimizan un modelo de Índice de Calidad del Agua (WQI) empleando múltiples algoritmos (CatBoost, SVM, Regresión Lineal, XGBoost, LightGBM) (Ding et al., 2025). De igual modo, Makumbura et al. (2024) destacan que ensambles de bosques aleatorios y métodos de boosting junto a redes neuronales pueden mejorar la evaluación predictiva de la calidad hídrica (Makumbura et al., 2024). En general, los estudios revisados muestran que los enfoques supervisados son capaces de incorporar numerosas variables predictoras y captar su interacción no lineal, adecuándose bien al carácter complejo de los sistemas acuáticos (Xu et al., 2024; Sajib et al., 2024).

Dentro de las técnicas más utilizadas por los autores dentro de la literatura se encuentran:

Cuadro 6
Algoritmos de machine learning supervisados en los modelos predictivos de calidad del agua.

| Algoritmo | Frecuencia (n/40) | Artículos representativos |
|------------------------------------|-------------------|--|
| Random Forest (RF) | 22 | (Anand et al., 2024; Chen et al., 2024; Das, 2025; Eid et al., 2024; Han et al., 2025; M. S. Islam et al., 2024; Kundu et al., 2025; Kushwaha et al., 2024; Lausch et al., 2025; Makumbura et al., 2024; Maußner et al., 2025; Nong et al., 2024; Ramya et al., 2024; Rigueira et al., 2025; Sajib et al., 2024; Sajib et al., 2025; Song et al., 2024; Uddin et al., 2024; Uddin et al., 2025; Xu et al., 2024; Xue et al., 2024; Yousefi et al., 2024) |
| Support Vector Machine (SVM) | 20 | (Anand et al., 2024; Ding et al., 2025; Eid et al., 2024; Han et al., 2025; Irwan et al., 2025; M. S. Islam et al., 2024; Kundu et al., 2025; Kushwaha et al., 2024; Lausch et al., 2025; Makumbura et al., 2024; Nong et al., 2024; Ramya et al., 2024; Sajib et al., 2024; Sajib et al., 2025; Song et al., 2024; Uddin et al., 2024; Uddin et al., 2025; Xu et al., 2024; Yousefi et al., 2024; Castillo et al., 2024) |
| XGBoost | 13 | (Anand et al., 2024; Ding et al., 2025; Eid et al., 2024; Kundu et al., 2025; Makumbura et al., 2024; Maußner et al., 2025; Nong et al., 2024; Ramya et al., 2024; Sajib et al., 2024; Sajib et al., 2025; Tian et al., 2024; Uddin et al., 2025; Xu et al., 2024) |
| LSTM/BiLSTM | 8 | (Alfwzan et al., 2024; Chen et al., 2024; Ehteram et al., 2024; Han et al., 2025; M. S. Islam et al., 2024; Karpagam et al., 2024; Liu et al., 2025; Wan et al., 2025) |
| Convolutional Neural Network (CNN) | 4 | (Ehteram et al., 2024; Kandasamy et al., 2025; Karpagam et al., 2024; Uddin et al., 2025) |

- **Random Forest (RF):** Ensamble de árboles de decisión que combina múltiples predictores mediante *bagging*. Destaca por manejar datos no lineales y cuantificar importancia de variables (Kundu et al., 2025; Makumbura et al., 2024). Presente en 22 de los 40 estudios analizados.
- **XGBoost:** Algoritmo de *boosting* que optimiza funciones de pérdida mediante descenso de gradiente. Sobresale en precisión para predicción de índices de calidad del agua (WQI) (Ding et al., 2025; Uddin et al., 2025). Utilizado en 13 estudios.
- **Redes Neuronales (LSTM/BiLSTM):** Arquitecturas recurrentes que capturan dependencias temporales en secuencias largas. Ideales para modelar series hidrológicas con memoria de estados anteriores (Alfwzan et al., 2024; Chen et al., 2024). Implementadas en 8 trabajos.
- **SVM (Máquinas de Vectores de Soporte):** Clasificadores que definen hiperplanos óptimos en espacios de alta dimensionalidad. Eficaces con datasets pequeños o ruidosos (Gupta et al., 2024; Song et al., 2024). Empleados en 20 estudios.
- **Redes Neuronales Convolucionales (CNN):** Especializadas en procesar datos espaciales (p.e., imágenes satelitales) para predecir variables como clorofila-a (Anand et al., 2024; Lausch et al., 2025). Usadas en 4 investigaciones.

La preferencia por técnicas de aprendizaje supervisado en la predicción de la calidad del agua se fundamenta en la naturaleza de la tarea: la mayoría de los problemas se definen como tareas de regresión o clasificación donde existen datos históricos con entradas (parámetros fisicoquímicos, meteorológicos) y salidas (índices de calidad, concentraciones de contaminantes) conocidas. En la literatura analizada, la práctica totalidad de los trabajos emplean modelos supervisados, destacando los ensambles como Random Forest (RF), XGBoost, AdaBoost y LightGBM, junto con las máquinas de vector soporte (SVM) y una variedad de redes neuronales (ANN, LSTM, GRU, CNN) (Ramya et al., 2024; Uddin et al., 2025). Ramya et al. (2024) ofrecen un análisis comparativo detallado sobre el rendimiento de múltiples clasificadores supervisados para el monitoreo de la calidad fluvial, mientras que Uddin et al. (2025) aplican regresiones avanzadas y redes neuronales recurrentes para predecir un índice compuesto de calidad del agua en lagos (Ramya et al., 2024; Uddin et al., 2025). La flexibilidad y alta capacidad predictiva de estos algoritmos explican su dominio, permitiendo además el aprovechamiento de grandes volúmenes de datos medidos en campo, como series temporales de sensores, datos geoespaciales e imágenes satelitales (Ding et al., 2025; Kundu et al., 2025).

Dentro de los modelos supervisados, la elección del algoritmo está intrínsecamente ligada a la naturaleza de los datos. Para capturar las complejas interacciones no lineales entre predictores, los métodos basados en árboles como Random Forest y Gradient Boosting son especialmente eficaces, ya que integran automáticamente estas relaciones sin necesidad de predefinir funciones matemáticas (Kundu et al., 2025; Xu et al., 2024). Por otro lado, cuando los datos exhiben dependencias temporales, como es común en las series de tiempo de calidad del agua, las redes neuronales recurrentes (LSTM, GRU) y las redes neuronales convolucionales (CNN) demuestran una capacidad superior. Uddin et al. (2025) emplearon arquitecturas LSTM y CNN para extraer características secuenciales y temporales de los datos de calidad de lagos, logrando una precisión que superó a los métodos lineales tradicionales (Uddin et al., 2025). De manera similar, para modelar dependencias espaciales como la calidad en un punto del río afecta a otro aguas abajo,

arquitecturas más avanzadas como las redes neuronales de grafos (GNN) han comenzado a utilizarse para capturar la estructura de la red fluvial (Wan et al., 2025). Este nivel de especialización permite a los modelos no solo ajustar tendencias globales, sino también detectar patrones locales complejos, como picos de contaminación estacionales o efectos no lineales, que los métodos estadísticos clásicos como ARIMA no logran identificar fácilmente (Makumbura et al., 2024; Uddin et al., 2025).

Las variables de salida (objetivos de predicción) más frecuentes en la literatura son tanto los índices compuestos de calidad del agua (WQI) como los parámetros físico-químicos individuales. El WQI y sus variantes (p.e., RMS-WQI, WQI_{min}) son populares porque resumen la condición general de un cuerpo de agua en un único valor, facilitando su interpretación y comunicación (Xu et al., 2024; Sajib et al., 2024). Paralelamente, muchos estudios se centran en predecir variables clave con un alto impacto ecológico o regulatorio, tales como el Oxígeno Disuelto (DO), la turbidez, los sólidos disueltos totales (TDS), el pH, el amonio (NH_4^+ o NH_3-N), nitratos, fosfatos y clorofila. Bolick et al. (2023) compararon diversos algoritmos para la predicción específica de DO en un arroyo urbano, mientras que Wan et al. (2025) abordaron la predicción de múltiples parámetros, incluyendo NH_3-N y fósforo total, en diversas estaciones de monitoreo (Bolick et al., 2023; Wan et al., 2025). La modelización de amoníaco-nitrógeno (NH_3-N) (Chen et al., 2024) y diversos metales pesados también son objetivos comunes, todos gestionados eficazmente mediante algoritmos supervisados.

Si bien la precisión de estos modelos es una ventaja fundamental, su naturaleza de ‘caja negra’ puede ser un obstáculo para su implementación en contextos regulatorios y de toma de decisiones críticas. Por ello, ha surgido un fuerte interés en el campo de la Inteligencia Artificial Explicable (XAI) para dotar de transparencia a los modelos de ML. Un enfoque destacado es SHAP (*SHapley Additive exPlanations*), una técnica basada en la teoría de juegos que permite cuantificar la contribución de cada variable de entrada a la predicción final del modelo. Trabajos como el de Kundu et al. (2025) utilizan SHAP no solo para evaluar la importancia global de cada parámetro en la predicción del WQI, sino también para interpretar decisiones a nivel local (explicaciones para una única predicción) (Kundu et al., 2025). Mediante visualizaciones como los *summary plots* y los *dependence plots*, los investigadores pueden desentrañar el impacto direccional (positivo o negativo) de variables como la turbidez o el TDS y detectar interacciones no lineales complejas. Este enfoque en la explicabilidad es crucial para fomentar la confianza, facilitar la validación de modelos y asegurar la trazabilidad de las decisiones algorítmicas, un paso indispensable para la adopción de estas tecnologías en la gestión ambiental del mundo real (Kundu et al., 2025; Rudin, 2019).

3.3.2 Modelos no supervisados de Machine Learning Los algoritmos no supervisados operan sin recurrir a datos etiquetados de salida y se emplean frecuentemente para el análisis exploratorio, reducción de dimensionalidad, agrupamiento y detección de valores atípicos en calidad del agua. Deriva de estas metodologías el análisis de componentes principales (PCA); una técnica estadística lineal que transforma variables posiblemente correlacionadas en un conjunto de componentes ortogonales que capturan la mayor parte de la varianza. En varios estudios de calidad del agua se ha aplicado PCA como paso de preprocesamiento: reduce la cantidad de indicadores redundantes de calidad del agua

(pH, oxígeno disuelto, nutrientes) en unos pocos ejes principales (Roy et al., 2024; Eid et al., 2024). Al agrupar información altamente correlacionada (e.g. múltiples parámetros físico-químicos de un río), PCA mejora la estabilidad del modelo y simplifica la predicción sin perder información crítica (Roy et al., 2024; Tsegay et al., 2025). Nath Roy et al. (2024) incorporan PCA al cálculo del WQI para reducir la redundancia de variables y hacer más objetivo el índice aplicado a los ríos de Daga (Roy et al., 2024). De modo similar, Usang et al. (2025) usan PCA para reducir la dimensionalidad previa a inferencia difusa en sistemas estuarinos, reforzando la precisión de la evaluación (Usang et al., 2025).

Cuadro 7

Frecuencia de técnicas de preprocesamiento y análisis no supervisado.

| Algoritmo | Frecuencia (n/40) | Artículos representativos |
|---|-------------------|---|
| Análisis de Componentes Principales (PCA) | 5 | (Eid et al., 2024; Roy et al., 2024; M. T. Islam et al., 2024; Tsegay et al., 2025; Usang et al., 2025) |
| Detección de valores atípicos (Outlier detection) | 3 | (Rigueira et al., 2025; Tsegay et al., 2025; Uddin et al., 2024) |
| Agrupamiento K-means (K-means) | 1 | (Tsegay et al., 2025) |
| Mezclas gaussianas (GMM) | 1 | (Wan et al., 2025) |
| Bosque de aislamiento (Isolation Forest, IF) | 1 | (Uddin et al., 2025) |
| Estimación de densidad de kernel (KDE) | 1 | (Uddin et al., 2025) |

Además de PCA, se emplean técnicas de agrupamiento (*clustering*) para identificar patrones similares en los datos. El *K-means* es el ejemplo más citado: particiona las muestras en *k* grupos basados en la cercanía a centroides. Tsegay et al. (2025) usan *K-means* para identificar patrones de comportamiento en la demanda de agua en redes de distribución, agrupando perfiles horarios de consumo similares (Tsegay et al., 2025). El agrupamiento permite segmentar días o estaciones según condiciones hidrometeorológicas, sin requerir etiquetas de calidad, facilitando así la caracterización de escenarios de calidad del agua. En estudios de uso multispectral y de cuencas, también se menciona el análisis de clúster jerárquico o de densidad (e.g. DBSCAN) para entender la heterogeneidad espacial, aunque con menor frecuencia. En algunos casos se exploran modelos de mezclas gaussianas (GMM), que asumen que los datos provienen de una combinación de distribuciones normales; este enfoque probabilístico permite agrupar dinámicas hidrológicas complejas según distribuciones de parámetros clave (Tsegay et al., 2025). Chen et al. (2024) destacan en sus palabras clave el uso de ‘*cluster analysis*’ y ‘*Gaussian mixture model*’ para modelar la demanda hídrica y la calidad del agua (Chen et al., 2024), lo que sugiere el interés en este tipo de modelado no supervisado.

Otra familia importante son las técnicas de detección de anomalías (*outlier detection*). Estas identifican mediciones atípicas o extremos en parámetros de calidad del agua (p.e., valores inusualmente altos de contaminantes o lecturas erróneas de sensores). Aunque en los artículos revisados no se aplican directamente algoritmos como *Isolation Forest* o estimación de densidad por núcleos

(KDE), si se encuentran estudios específicos enfocados en detectar anomalías ambientales. Rigueira et al. (2025) desarrollan un sistema basado en análisis funcional multivariante para detectar anomalías en datos de sensores de calidad del agua (etiquetado binario normal vs. anómalo) (Rigueira et al., 2025). Estas técnicas suelen usarse en preprocesamiento de datos hidricos para filtrar mediciones erráticas o resaltar eventos de contaminación repentina. En general, métodos no supervisados como *isolation forest*, *LOF* o *KDE* permitirían estimar automáticamente umbrales de excepción basados en la densidad de datos, apoyando la limpieza de datos antes de la modelación supervisada.

En cuanto a los roles en el modelado de calidad del agua, los algoritmos no supervisados se emplean principalmente en preprocesamiento y exploración de datos. PCA y técnicas similares se usan para reducción de dimensionalidad, simplificando conjuntos de variables ambientales antes de la predicción supervisada (Roy et al., 2024; Eid et al., 2024). El *clustering* facilita el descubrimiento de regímenes de calidad o grupos de condiciones hidrometeorológicas homogéneas sin conocer de antemano la respuesta. Se puede ilustrar este método segmentando zonas de lluvia o sequía según su perfil de contaminación. La detección de anomalías sirve para filtrado de datos: permite identificar lecturas atípicas de sensores o emergencias contaminantes para luego tratarlas o señalarlas. Estos métodos no supervisados aportan una capa de análisis exploratorio ('pre-modelo') que complementa los modelos predictivos de calidad (supervisados), mejorando la calidad y relevancia de los datos de entrada.

Las variables más comúnmente abordadas con estos enfoques incluyen los parámetros fisicoquímicos clásicos (pH, turbidez, oxígeno disuelto, nitratos, fósforo, demanda bioquímica de oxígeno), así como datos ambientales asociados (temperatura, caudal o demanda hídrica, índices GIS de uso de suelo). Tsegay et al. (2025) aplican PCA sobre variables de 'demanda base, temperatura ambiental y patrones horarios' de consumo (Tsegay et al., 2025). De igual forma, otros trabajos agrupan indicadores compuestos como el *WQI* o el *RIWQI* regional (como Touhidul et al. 2024) integrando múltiples parámetros en grupos representativos (M. T. Islam et al., 2024; Roy et al., 2024). Cualquier conjunto de variables ambientales multivariantes puede beneficiarse de métodos no supervisados para resumir su estructura intrínseca.

Finalmente, frente a los modelos supervisados, las técnicas no supervisadas ofrecen ventajas claras en los contextos revisados: no requieren datos de salida etiquetados ni índices definidos, por lo que son útiles cuando dichos datos son escasos o inciertos. También ayudan a evitar sobreajuste al trabajar sobre la estructura interna de los datos, y pueden revelar relaciones latentes que un modelo supervisado pasaría por alto. Al reducir dimensionalidad o agrupar datos de entrada, estas técnicas mejoran la eficiencia computacional y la robustez de los modelos posteriores. En varios estudios se observa que integrar PCA o *clustering* previo eleva la fidelidad de simulaciones multianuales y la estabilidad frente a variabilidad estacional (Tsegay et al., 2025; Roy et al., 2024).

3.4 Tendencias en ML para Calidad del Agua

3.4.1 Modelos híbridos Se combinan diversos algoritmos para aprovechar ventajas complementarias. Un ejemplo avanzado es la arquitectura *FNN-BiLSTM Attention*, que integra redes neuronales difusas (*FNN*) para manejar incertidumbre con *BiLSTM* y un mecanismo de atención para enfocar las entradas más relevantes

(Liu et al., 2025). Liu y Cheng (2025) demostraron que este modelo híbrido supera en precisión a enfoques estándar al predecir el *WQI* bajo datos escasos (Liu et al., 2025). De manera similar, se emplean diseños *stacked* donde se apilan redes neuronales y *SVM*: se entrenan varias *ANN* y luego una *SVM* actúa en la capa de salida para combinar las salidas, uniendo la capacidad no lineal de las *ANN* con la robustez discriminativa de las *SVM* (Kushwaha et al., 2024).

3.4.2 Técnicas de optimización algorítmica Incluyen variantes avanzadas de algoritmos base. Métodos de *boosting* adaptativo (como *AdaBoost*) potencian iterativamente modelos débiles, corrigiendo errores sucesivos y afinando la combinación final. Asimismo, se integran mecanismos de atención en redes profundas (*RNN* o transformadores) para ponderar dinámicamente las entradas más relevantes en cada paso de la predicción. Estas optimizaciones permiten entrenar modelos más robustos y precisos con datos heterogéneos y ruidosos, mejorando el desempeño en problemas de calidad del agua (Liu et al., 2025).

3.4.3 Técnicas de Remote Sensing Un hallazgo de creciente relevancia en el campo de la predicción de la calidad del agua es la integración de datos de teledetección (*remote sensing*) con modelos de aprendizaje automático (Song et al., 2024; Lausch et al., 2025; Tian et al., 2024). Este enfoque, no anticipado, ha demostrado ser fundamental para la medición indirecta y el monitoreo a gran escala de la calidad del agua, superando las limitaciones del muestreo *in situ* (Anand et al., 2024; Kandasamy et al., 2025).

La teledetección permite el uso de imágenes satelitales multiespectrales, como las proporcionadas por Sentinel-2 y Landsat-8, para capturar bandas sensibles a parámetros ópticamente activos del agua, incluyendo turbidez, clorofila-a y sólidos suspendidos (Anand et al., 2024; Kandasamy et al., 2025). Esta capacidad ofrece una cobertura espacial amplia y una alta frecuencia temporal, facilitando la evaluación en regiones de difícil acceso y la detección de cambios longitudinales y temporales en la calidad del agua (Anand et al., 2024; Kandasamy et al., 2025).

- **XGBoost + Teledetección:** Se ha utilizado para predecir el índice trófico (*TSI*) en lagos urbanos, cuantificando el impacto de políticas urbanas y el cambio climático mediante el análisis *SHAP* para la explicabilidad (Tian et al., 2024). Este modelo integra variables climáticas y de cobertura del suelo extraídas de teledetección.
- **Random Forest (RF) + Teledetección:** Modelos como el desarrollado por Lausch et al. utilizan imágenes satelitales (*Landsat8/9*) para predecir la concentración de clorofila-a, identificando la estructura del paisaje, la intensidad del uso del suelo y los cambios del paisaje como impulsores clave de la calidad del agua (Lausch et al., 2025).
- **SVM + LUCC (Land Use and Land Cover Change):** Song et al. desarrollaron un modelo cuantitativo basado en *SVM* para predecir la calidad del agua en función de los cambios de uso y cobertura del suelo, integrando datos espaciales y técnicas de ML para la gestión ambiental de cuencas (Song et al., 2024).

Cuadro 8

Modelos utilizados en los estudios analizados.

| Referencia | Año | Nombre del Artículo | Modelos utilizados | Variables de salida |
|-------------------------|------|--|---|--|
| (Alfwzan et al., 2024) | 2024 | Water quality assessment using Bi-LSTM and computational fluid dynamics (CFD) techniques | CFD-Bi-LSTM | Water Quality Index (WQI) |
| (Ding et al., 2025) | 2025 | Using multiple machine learning algorithms to optimize the water quality index model and their applicability | CatBoost, SVM, LR, XGB, LightGBM | Clases estadísticas de WQI |
| (Kundu et al., 2025) | 2025 | Unveiling the hidden connections: Using explainable artificial intelligence to assess water quality criteria in nine giant rivers [*] | SVR, XGB, RF, AB, GB | Valores de estimación SHAP |
| (Gupta et al., 2024) | 2024 | Time Series Analysis and Forecasting of Water Quality Parameters along Yamuna River in Delhi | ARIMA | Auto-regression (AR), differencing (I), moving average (MA) con análisis en PowerBI |
| (Uddin et al., 2025) | 2025 | The role of optimizers in developing data-driven model for predicting lake water quality incorporating advanced water quality model | LR, DT, RF, GBR, XGB, CatBoost, AdaBoost, HBR, GPR, SVR, DNN, CNN | RMSE Water Quality Index (WQI) |
| (Wan et al., 2025) | 2025 | Temporal and spatial feature extraction with graph neural networks for river water quality prediction | GAT, GraphSAGE, MPNN, LSTM, GRU | Valores de calidad del agua en múltiples estaciones de monitoreo de una red fluvial. |
| (Kushwaha et al., 2024) | 2024 | Stacked hybridization to enhance the performance of artificial neural networks (ANN) for prediction of water quality index in the Bagh river basin, India | ANN, RSS, RF, AR, M5P | Water Quality Index (WQI) |
| (Das, 2025) | 2025 | Spatiotemporal evaluation and impact of superficial factors on surface water quality for drinking using innovative techniques in Mahanadi River Basin, Odisha, India | MEREC, ARAS, RF | Turbidez, TDS, DO, pH |
| (Xiao et al., 2025) | 2025 | Spatiotemporal dynamics and optimization of water quality assessment in the Nantong section of the Yangtze River Basin: A WQI _{min} approach | ANOVA, SMK, Sen's slope | Minimal Water Quality Index WQI_{min} |
| (Irwan et al., 2025) | 2025 | River water quality monitoring using machine learning with multiple possible in-situ scenarios | NN, GPR, Ensemble, SVM | Clases estadísticas de WQI |
| (Castillo et al., 2024) | 2024 | Improving river water quality prediction with hybrid machine learning and temporal analysis | ANFIS, ANN, SVM | Water Quality Index bajo valores continuos de parámetros de calidad |
| (Usang et al., 2025) | 2025 | Integrating principal component analysis, fuzzy inference systems, and advanced neural networks for enhanced estuarine water quality assessment [*] | PCA, FIS | Índice compuesto de calidad del agua estuarina |

Cuadro 8 – continuación de la página anterior

| Referencia | Año | Nombre del Artículo | Modelos utilizados | Variables de salida |
|----------------------------|------|---|---|--|
| (M. T. Islam et al., 2024) | 2024 | Regional irrigation water quality index for the Old Brahmaputra River, Bangladesh: A multivariate and GIS-based spatiotemporal assessment | PCA, GIS, IDW | Regional Irrigation Water Quality Index (RIWQI) |
| (Song et al., 2024) | 2024 | Quantitative prediction of water quality in Dongjiang Lake watershed based on LUCC | SVM, MLE, RF | Uso de tierra usando Cellular Automata-Markov chain model (CA-Markov) |
| (Roy et al., 2024) | 2024 | Principal component analysis incorporated water quality index modeling for Dhaka-based rivers | PCA | Water Quality Index (WQI) |
| (Xue et al., 2024) | 2024 | Predictive modeling of nitrogen and phosphorus concentrations in rivers using a machine learning framework: A case study in an urban-rural transitional area in Wenzhou China [*] | RF, SHAP, RSS | Concentraciones de nitrógeno y fósforo en ríos |
| (Lloyd et al., 2024) | 2024 | Predicting recreational water quality and public health safety in urban estuaries using Bayesian Networks | BNs | Calidad recreacional del agua (enterococos) y riesgo para la salud pública (probabilidad de enfermedad gastrointestinal) |
| (Han et al., 2025) | 2025 | Predicting and investigating water quality index by robust machine learning methods | LSTM, SVM, DT, RF | Water Quality Index (WQI) |
| (Sajib et al., 2025) | 2025 | Novel groundwater quality index (GWQI) model: A reliable approach for the assessment of groundwater | RF, XGB, SVM, DT, GB, GPR, HBR, LR, DNN | GWQI (GroundWater Quality Index) |
| (Rigueira et al., 2025) | 2025 | Multivariate functional data analysis and machine learning methods for anomaly detection in water quality sensor data | MMSA, RF | Detección de anomalías en los datos sensoriales de calidad del agua (etiquetado binario: normal vs anómalo) |
| (Lausch et al., 2025) | 2025 | Monitoring and modelling landscape structure, land use intensity and landscape change as drivers of water quality using remote sensing | GLM, SVM, DL, DT, GBT, RF | Chlorophyll-a como indicador de calidad del agua. |
| (Anand et al., 2024) | 2024 | Machine learning approach for water quality predictions based on multispectral satellite imageries | LR, KNN, SVM, DT, RF | DO, BOD, COD, TDS, TSS, pH |
| (M. S. Islam et al., 2024) | 2024 | Long-term trend prediction of surface water quality of two main river basins of China using Machine Learning Method | LSTM, GRU, SVR, RFR, ARIMA | COD_{Mn} , un indicador de contaminación orgánica en agua superficial. |
| (Chen et al., 2024) | 2024 | A coupled model to improve river water quality prediction towards addressing non-stationarity and data limitation | WT, BiLSTM | NH_3-N , parámetro de calidad del agua |

Cuadro 8 – continuación de la página anterior

| Referencia | Año | Nombre del Artículo | Modelos utilizados | Variables de salida |
|--------------------------|------|---|-------------------------------|---|
| (Nong et al., 2024) | 2024 | A novel coupling interpretable machine learning framework for water quality prediction and environmental effect understanding in different flow discharge regulations of hydro-projects | RF, XGB, SVR, GBRT | Concentración de amonio NH_4^+-N |
| (Eid et al., 2024) | 2024 | An advanced approach for drinking water quality indexing and health risk assessment supported by machine learning modelling in Siwa Oasis, Egypt | RF, SVM, GBM, XGB, PCA | Groundwater Quality Index (GWQI) |
| (Ehteram et al., 2024) | 2024 | An advanced deep learning model for predicting water quality index | MLP, LSTM, GRU, CNN, CNN-M5T | Water Quality Index (WQI) |
| (Yousefi et al., 2024) | 2024 | An innovative combination of extra trees within adaboost for accurate prediction of agricultural water quality indices | SVR, KNN, RF, GB, ET-AdaBoost | AWQIs: SAR, RSC, MH, PI, Porcentaje de sodio, KR, Wilcox Classification |
| (Xu et al., 2024) | 2024 | Assessing and predicting water quality index with key water parameters by machine learning models in coastal cities, China | MLR, SVR, RF, XGB, LightGBM | Water Quality Index (WQI) |
| (Sajib et al., 2024) | 2024 | Assessing water quality of an ecologically critical urban canal incorporating machine learning approaches | RF, SVM, KNN, GBDT, XGB | Water Quality Index (WQI) |
| (Ramya et al., 2024) | 2024 | Comprehensive analysis of multiple classifiers for enhanced river water quality monitoring with explainable AI [*] | DT, RF, SVM, KNN, NB, LR, XGB | Water Quality Index (WQI), dividida en niveles cualitativos (excelente, buena, media, pobre, muy pobre) |
| (Uddin et al., 2024) | 2024 | Data-driven evolution of water quality models: An in-depth investigation of innovative outlier detection approaches-A case study of Irish Water Quality Index (IEWQI) model | RF, SVM, GB, ANN, KNN | IEWQI (Irish Environmental Water Quality Index) |
| (Karpagam et al., 2024) | 2024 | Deep learning-based water quality index classification using stacked ensemble variational mode decomposition | CNN, LSTM, BiLSTM, GRU, VMD | Water Quality Index (WQI) dividida en niveles excelente, buena, media, pobre, muy pobre |
| (Kandasamy et al., 2025) | 2025 | Enhanced remote sensing and deep learning aided water quality detection in the Ganges River, India supporting monitoring of aquatic environments | CNN | Water Quality Index (WQI) etiquetada por niveles: buena, moderada, pobre u otras categorías ordinales. |

Cuadro 8 – continuación de la página anterior

| Referencia | Año | Nombre del Artículo | Modelos utilizados | Variables de salida |
|--------------------------|------|--|--------------------------------|--|
| (Tsegay et al., 2025) | 2025 | Enhancing long-term water quality modeling by addressing base demand, demand patterns, and temperature uncertainty using unsupervised machine learning techniques | PCA, K-means clustering, t-SNE | Mejorar la modelación del comportamiento a largo plazo de la calidad del agua. |
| (Pazo et al., 2024) | 202 | Enhancing water quality prediction for fluctuating missing data scenarios: A dynamic Bayesian network-based processing system to monitor cyanobacteria proliferation | DBN | Nivel de proliferación de cianobacterias, indicador de la calidad del agua y de riesgo ecológico en cuerpos de agua dulce. |
| (Maußner et al., 2025) | 2025 | Explainable artificial intelligence for reliable water demand forecasting to increase trust in predictions [*] | RF, XGB, LightGBM, CatBoost | Demanda de agua en términos de volumen diario (litros/día) |
| (Tian et al., 2024) | 2024 | Exploring the effects of climate change and urban policies on lake water quality using remote sensing and explainable artificial intelligence [*] | XGB | Trophic State Index (TSI) que cuantifica el estado trófico de cuerpos de agua |
| (Liu et al., 2025) | 2025 | FNN-BiLSTM-Attention-DA: A hybrid fuzzy neural network and BiLSTM with multi-sensor information fusion for water quality monitoring and warning | FNN-BiLSTM-Attention-DA | Water Quality Index (WQI) |
| (Makumbura et al., 2024) | 2024 | Advancing water quality assessment and prediction using machine learning models, coupled with explainable artificial intelligence (XAI) techniques like shapley additive explanations (SHAP) for interpreting the black-box nature [*] | RF, GBM, XGBoost, SVM | Water Quality Index (WQI) |

[*] Denota estudios que, además de la predicción, incluyen un componente explícito de Inteligencia Artificial Explicable (XAI) para la interpretación de los resultados del modelo.

- **Redes Neuronales Convolucionales (CNN) + Teledetección:** Modelos *CNN* han sido diseñados específicamente para la clasificación de la calidad del agua a partir de imágenes satelitales, aprendiendo patrones espaciales en imágenes multiespectrales para inferir la clase de calidad del agua (Kandasamy et al., 2025). Este enfoque es escalable y automatizable para el monitoreo de grandes ecosistemas acuáticos.

La integración de la teledetección con el aprendizaje automático amplía significativamente la aplicabilidad operacional de los modelos de predicción de la calidad del agua más allá del monitoreo *in situ*, proporcionando una supervisión continua y proactiva del recurso hídrico (Tian et al., 2024; Anand et al., 2024).

3.4.4 Inteligencia Artificial Explicable La Inteligencia Artificial Explicable (XAI, por sus siglas en inglés) ha cobrado creciente relevancia en la predicción de la calidad del agua debido a la opacidad inherente de muchos modelos de *Machine Learning*, particularmente los de tipo ensamblado y profundo como *Random Forest* (RF), *eXtreme Gradient Boosting* (XGBoost), y redes neuronales. La necesidad de interpretabilidad no solo es técnica sino también regulatoria y ética (Commission, 2024), ya que los sistemas de apoyo a decisiones ambientales requieren una trazabilidad clara de sus inferencias (Ramya et al., 2024; Park et al., 2022; Nallakaruppan et al., 2024). En este contexto, las técnicas XAI permiten descomponer las predicciones de modelos complejos en componentes comprensibles por humanos, atribuyendo cuantitativamente la importancia de cada variable de entrada en el resultado final (Sajib et al., 2025; Ding et al., 2025; Irwan et al., 2025).

Entre las técnicas más utilizadas se encuentra *SHAP* (*SHapley Additive exPlanations*), que proporciona explicaciones consistentes basadas en teoría de juegos. *SHAP* asigna una contribución ponderada a cada característica para una predicción específica, facilitando tanto la interpretación global del modelo como el análisis individual de casos. En los estudios analizados, *SHAP* ha sido aplicado mayoritariamente sobre modelos *RF* y *XGBoost* para cuantificar la influencia de parámetros como *pH*, temperatura, conductividad, turbidez, nutrientes disueltos (nitratos, fosfatos) y metales pesados sobre índices de calidad del agua como *WQI* o *IWQI*. Park et al. (2022) y Ramya et al. (2024) utilizaron *SHAP* para identificar las variables más determinantes en la predicción de la calidad del agua superficial y subterránea, revelando que el contenido de nitratos y el *pH* eran recurrentemente las más influyentes (Park et al., 2022; Ramya et al., 2024). Asimismo, Sajib et al. y Ding et al. integraron análisis *SHAP* a sus modelos de regresión para afinar la sensibilidad del sistema ante fluctuaciones de entrada (Sajib et al., 2025; Ding et al., 2025).

Otra técnica relevante es *LIME* (*Local Interpretable Model-agnostic Explanations*), que genera explicaciones locales al aproximar el modelo complejo mediante uno más simple en la vecindad de una predicción puntual. Aunque su presencia es menos frecuente que *SHAP* en los artículos analizados, ha sido utilizada por Irwan et al. como método complementario para contrastar la validez de las interpretaciones generadas por *SHAP*, especialmente en entornos con alta colinealidad entre variables (Irwan et al., 2025).

Más allá del análisis de importancia de variables, otras aplicaciones de XAI en los estudios revisados incluyen: detección de sesgos algorítmicos, análisis de robustez ante perturbaciones de datos, identificación de estructuras latentes en modelos no lineales y visualización de reglas emergentes en redes neuronales profundas.

Nallakaruppan et al., utilizaron técnicas XAI para evaluar la coherencia lógica de las predicciones de un modelo híbrido *CNN-GRU* aplicado a datos temporales de estaciones hidrometeorológicas, lo cual permitió filtrar predicciones espurias derivadas de errores en la señal de entrada (Nallakaruppan et al., 2024).

Cabe destacar que la aplicabilidad de XAI está estrechamente ligada al tipo de modelo base. Modelos lineales o árboles de decisión poco profundos ofrecen por sí mismos interpretabilidad inherente, mientras que XAI es esencial en modelos tipo ‘caja negra’. No obstante, incluso en casos donde el modelo sea parcialmente interpretable, las técnicas XAI añaden un nivel adicional de validación cruzada explicativa, lo que las convierte en una herramienta indispensable para la implementación operativa de sistemas de predicción de calidad del agua en contextos reales.

3.5 Herramientas de Análisis de Rendimiento y Explicabilidad en Modelos de Calidad del Agua

3.5.1 Análisis de Métricas de Rendimiento

- **CNN-M5T: Superioridad en la Captura de Patrones Espacio-Temporales**

El modelo CNN-M5T (Convolutional Neural Network con M5T) ha demostrado un rendimiento robusto en la predicción del Índice de Calidad del Agua (WQI). Específicamente, el modelo ‘CNN-Cos RNN-M5T’ logró un Error Absoluto Medio (MAE) de 0.79, una Eficiencia Nash-Sutcliffe (NSE) de 0.99 y una Eficiencia de Reducción (RE) del 94 % (Ehteram et al., 2024). Este desempeño superó significativamente a modelos comparativos como CNN, CRNN y LSTM (Ehteram et al., 2024).

La ventaja arquitectónica del CNN-M5T radica en su estructura convolucional, que incorpora múltiples bloques M5 y capas de dropout, optimizadas para series temporales multivariadas (Ehteram et al., 2024). Este diseño le permite capturar patrones espacio-temporales complejos y dependencias no lineales inherentes a los datos hidrológicos, lo que se traduce en una mayor precisión predictiva. La evaluación del modelo se realizó mediante validación cruzada *K - fold* (con *k*=10), una elección metodológica que asegura la capacidad de generalización del modelo y reduce el sesgo de validación al utilizar sistemáticamente todas las particiones del conjunto de datos para entrenamiento y prueba (Ehteram et al., 2024). Este enfoque de validación riguroso refuerza la fiabilidad de su rendimiento superior reportado.

- **XGBoost y Random Forest (RF): Equilibrio entre Precisión y Eficiencia**

XGBoost y Random Forest (RF) son modelos de alto rendimiento que se destacan consistentemente en diversas tareas de predicción de la calidad del agua. Estos modelos son particularmente efectivos en el manejo de interacciones complejas entre variables y en la consecución de altos valores de R^2 . En el caso de Random Forest (RF), para la predicción de las concentraciones de nitrógeno total (TN) y fósforo total (TP) en áreas de transición urbano-rural, el modelo alcanzó un R^2 de entrenamiento de 0.85 para TN y 0.87 para TP (Xue et al., 2024). Aunque el R^2 de prueba para TN disminuyó a 0.62, estos valores aún indican una sólida capacidad predictiva, especialmente en los datos de entrenamiento.

Por su parte, XGBoost ha demostrado una alta precisión en la

predicción del Índice de Calidad del Agua (WQI). Si bien los valores específicos de R^2 para XGBoost por sí solo no se detallan explícitamente en los materiales de investigación para una comparación directa con la predicción de N/P de RF, su mención constante como uno de los modelos de mejor rendimiento en varios estudios (Ding et al., 2025; Xu et al., 2024; Ramya et al., 2024; Makumbura et al., 2024) subraya su robustez y eficiencia.

El rendimiento consistentemente alto de métodos de ensamble como XGBoost y Random Forest, a pesar de su naturaleza de ‘caja negra’, subraya un enfoque predominante en la precisión predictiva dentro del campo. Sin embargo, la mención explícita de su carácter ‘menos explicable’ (Ding et al., 2025) resalta la tensión inherente entre el rendimiento y la interpretabilidad, lo que a su vez enfatiza la importancia de las técnicas de XAI. Estos modelos, al combinar las predicciones de múltiples modelos más simples, suelen lograr una mayor precisión y robustez en comparación con los modelos individuales. La adopción generalizada y el éxito de estos métodos de ensamble sugieren que, para muchas aplicaciones prácticas, la consecución de un alto poder predictivo es un objetivo primordial. No obstante, esto a menudo implica una menor transparencia inmediata, lo que hace necesaria la integración de técnicas de Inteligencia Artificial Explicable (XAI) para que sus decisiones sean comprensibles y confiables para los usuarios finales y los organismos reguladores.

3.5.2 Explicabilidad (XAI) por Tipo de Modelo Para los modelos de ensamble basados en árboles de decisión, como Random Forest (RF) y XGBoost, SHAP (SHapley Additive exPlanations) es una técnica de XAI ampliamente adoptada. SHAP cuantifica el impacto individual de los predictores en la salida del modelo utilizando la teoría de juegos, lo que proporciona una medida robusta de la importancia de las características (Makumbura et al., 2024; Ramya et al., 2024). Esto permite una comprensión global de la importancia de las características y su impacto direccional.

En el contexto de la predicción del Índice de Calidad del Agua (WQI) utilizando XGBoost, SHAP identificó parámetros clave de calidad del agua y sus contribuciones precisas. Verbigracia, la Demanda Química de Oxígeno (COD) tuvo un valor SHAP de 6.92, la Demanda Bioquímica de Oxígeno (BOD) 4.45, el pH 2.29, el Fosfato (PO_4^{3-}) 2.16, y el Oxígeno Disuelto (DO) 2.01, lo que indica su influencia relativa en el WQI (Makumbura et al., 2024).

Por otro lado, LIME (Local Interpretable Model-agnostic Explanations) genera explicaciones locales para predicciones individuales ajustando un modelo sustituto interpretable alrededor de esa instancia específica (Ramya et al., 2024). Este enfoque agnóstico al modelo es crucial para comprender por qué se realizó una predicción particular para un punto de datos específico, lo que ayuda a validar la coherencia del modelo con la intuición experta y a detectar comportamientos anómalos (Ramya et al., 2024).

Si bien las técnicas de XAI como SHAP son muy efectivas para los modelos basados en árboles, su aplicación y cuantificación detallada para arquitecturas híbridas complejas y de deep learning presentan desafíos continuos. Para el modelo FNN-BiLSTM-Attention-DA (Liu et al., 2025), que combina Redes Neuronales Difusas (FNN), LSTMs Bidireccionales (BiLSTM), Mecanismos de Atención y Adaptación de Dominio para el monitoreo de la calidad del agua, los materiales de investigación proporcionados no contienen detalles específicos sobre cómo se aplicaron SHAP u otras

técnicas de XAI para explicar sus mecanismos internos o cuantificar contribuciones como la mitigación de ruido (Liu et al., 2025).

Esta ausencia destaca una brecha actual en la interpretabilidad integral de algunos modelos de deep learning híbridos altamente complejos y multicomponente. Aunque estos modelos sobresalen en precisión y adaptabilidad (p.e., al manejar la escasez de datos mediante la Adaptación de Dominio (Liu et al., 2025)), su funcionamiento interno sigue siendo menos transparente en comparación con modelos más sencillos integrados con XAI.

3.6 Técnicas de Modelado para la Predicción de Calidad del Agua: Complejidad vs. Eficiencia

3.6.1 Técnicas Complejas (Alta Complejidad Computacional)

■ CFD-BiLSTM: Integración Físico-Hidrodinámica para Ríos Urbanos

El modelo CFD-BiLSTM representa un enfoque avanzado que combina la dinámica de fluidos computacional (CFD) con redes neuronales recurrentes bidireccionales de memoria a largo plazo (BiLSTM). Esta hibridación única permite la integración de simulaciones hidrodinámicas con la capacidad de las redes BiLSTM para capturar patrones temporales complejos en los datos. El objetivo principal de este modelo es predecir la calidad del agua en entornos dinámicos como el río Yamuna, buscando reducir el Error Cuadrático Medio (MSE) en comparación con modelos físicos puros (Alfwzan et al., 2024). Una contribución distintiva de CFD-BiLSTM es su capacidad para integrar la física de fluidos con el aprendizaje de patrones temporales, lo que resulta en una reducción del Error Absoluto Porcentual Medio (MAPE) a 15.08 % en el río Yamuna (Alfwzan et al., 2024). Esta combinación es particularmente valiosa para modelar la calidad del agua en ríos urbanos, donde los procesos hidrodinámicos y las variaciones temporales de los contaminantes son altamente interconectados y no lineales. Al incorporar las simulaciones de CFD, el modelo puede representar con mayor precisión el transporte y la dispersión de contaminantes, mientras que las capas BiLSTM manejan las dependencias temporales de los datos de calidad del agua. Este enfoque aborda una limitación fundamental de los modelos predictivos tradicionales: la dificultad de capturar tanto los procesos físicos subyacentes como las dinámicas temporales observadas en los datos. La integración de CFD con BiLSTM permite que el modelo no solo aprenda de los datos históricos, sino que también incorpore el conocimiento físico del comportamiento del flujo y la mezcla de contaminantes en el río. Esto es crucial para la predicción de la calidad del agua en sistemas fluviales interconectados, donde la propagación de contaminantes está fuertemente influenciada por la hidrodinámica. La capacidad de este modelo para reducir el MAPE a 15.08 % en el río Yamuna demuestra su eficacia en la integración de la física de fluidos con patrones temporales, lo que lo hace particularmente adecuado para entornos fluviales urbanos complejos.

■ Hibridación Apilada (ANN-SVM): Optimización con Datos Limitados

La hibridación apilada, ejemplificada por el ensamble ANN-SVM (Kushwaha et al., 2024), representa un enfoque avanzado para mejorar la precisión en la predicción del Índice de

Calidad del Agua (WQI). Este modelo combina múltiples versiones optimizadas de Redes Neuronales Artificiales (ANN) con otras técnicas de aprendizaje automático como *Support Vector Machine (SVM)*, *Random Subspace (RSS)*, *Random Forest (RF)*, *Additive Regression (AR)* y *M5 pruned (M5P)* como modelos base (Kushwaha et al., 2024). El objetivo es superar las limitaciones de los modelos individuales y lograr una mayor precisión predictiva, especialmente en escenarios con datos limitados.

Aunque los materiales de investigación no especifican un valor de R^2 de 0.94 para este ensamble ni detallan RSS y *Additive Regression* como técnicas de optimización para ANN-SVM, sí se destaca que SVM obtuvo el 'MEJOR RESULTADO' entre los modelos base (Kushwaha et al., 2024). La arquitectura de hibridación apilada permite que el modelo aprenda de las fortalezas de cada componente, lo que es particularmente beneficioso cuando se trabaja con conjuntos de datos escasos, como los del río Bagh en India, donde se aplicó este enfoque (Kushwaha et al., 2024).

■ **FNN-BiLSTM-Attention-DA: Adaptación Cross-Domain y Robustez ante Datos Escasos**

El modelo *FNN-BiLSTM-Attention-DA* es una arquitectura híbrida neuro-difusa avanzada diseñada para el monitoreo y la advertencia de la calidad del agua, especialmente en entornos con escasez de datos o heterogeneidad. Este modelo integra una Red Neuronal Difusa (FNN) para manejar la incertidumbre y la lógica difusa en los datos ambientales, una red *BiLSTM* para capturar dependencias temporales bidireccionales, un Mecanismo de Atención para priorizar características de entrada relevantes, y Adaptación de Dominio (DA) para transferir conocimiento entre dominios con datos limitados (Liu et al., 2025).

La combinación híbrida de estos componentes permite que el modelo *FNN-BiLSTM-Attention-DA* supere las limitaciones individuales de los modelos tradicionales. La FNN proporciona robustez frente al ruido y la variabilidad difusa de los parámetros ambientales, El Mecanismo de Atención mejora la capacidad del modelo para enfocarse en las entradas con mayor peso causal en la predicción del WQI. La Adaptación de Dominio permite trasladar las capacidades predictivas a nuevos cuerpos de agua sin necesidad de un reentrenamiento desde cero, lo que es invaluable en entornos con escaso monitoreo directo (Liu et al., 2025).

Aunque los materiales de investigación no proporcionan detalles específicos sobre cómo se aplicaron técnicas de *XAI* como *SHAP* para explicar los mecanismos internos de este modelo o cuantificar la mitigación de ruido (Liu et al., 2025), su arquitectura multicomponente está diseñada para mejorar la precisión, estabilidad y capacidad de adaptación entre dominios. La ausencia de detalles explícitos de *XAI* en modelos híbridos tan complejos resalta una brecha en la investigación actual. A medida que los modelos se vuelven más sofisticados, la interpretabilidad se vuelve más difícil pero también más crítica para la confianza y la adopción generalizada, especialmente en la toma de decisiones ambientales de alto riesgo.

3.6.2 Técnicas Sencillas (Alto Rendimiento y Eficiencia)

■ **ARIMA: Modelado Clásico de Series Temporales**

El modelo *ARIMA (AutoRegressive Integrated Moving Average)*

es un enfoque clásico de series temporales utilizado para la predicción de parámetros de calidad del agua, como el oxígeno disuelto (DO) en el río Yamuna (Gupta et al., 2024). Este modelo es competitivo en escenarios con patrones estables, con un Error Absoluto Medio (MAE) de 0.8 para series estables de DO (Gupta et al., 2024).

Sin embargo, *ARIMA* presenta limitaciones significativas al no manejar adecuadamente la no linealidad ni las fluctuaciones irregulares presentes en los datos de calidad del agua. Esto restringe su utilidad para predicciones finas o de largo plazo en sistemas hídricos complejos. Aunque el modelo *ARIMA* mostró cierta capacidad para capturar la tendencia estacional y la persistencia de series temporales, su precisión fue inferior en comparación con modelos basados en redes neuronales recurrentes, como *LSTM* (M. S. Islam et al., 2024). En la predicción del índice de permanganato, *LSTM* superó a *SARIMA* (una variante estacional de *ARIMA*) con un MAPE del 7 % frente al 9.1 % en el río Yangtsé, y del 14 % frente al 13.3 % en el río Amarillo (M. S. Islam et al., 2024). La aplicabilidad de *ARIMA* se limita a escenarios donde los patrones son estables, no hay multivariantes ni entradas externas, y la interpretabilidad se prioriza sobre la precisión (M. S. Islam et al., 2024). La incapacidad de *ARIMA* para manejar la no linealidad y las fluctuaciones irregulares en los datos de calidad del agua es una limitación crucial, ya que los fenómenos ambientales rara vez son puramente lineales o estacionarios.

■ **RF + SHAP: Precisión y Explicabilidad en Monitoreo Rutinario**

La combinación de *Random Forest (RF)* con *SHAP* (SHapley Additive exPlanations) es un enfoque que equilibra la precisión predictiva con una interpretabilidad intuitiva, lo que lo hace óptimo para el monitoreo rutinario de la calidad del agua con recursos limitados. Este modelo se ha utilizado para predecir concentraciones de nitrógeno (TN) y fósforo (TP) en ríos, logrando un R^2 de entrenamiento de 0.85 para TN y 0.87 para TP (Xue et al., 2024). Aunque los materiales de investigación no proporcionan valores *SHAP* específicos superiores a 3.0 para los predictores clave, se destaca que el análisis *SHAP* fue fundamental para identificar la importancia de las variables (Xue et al., 2024).

El uso de *SHAP* en conjunto con RF permite no solo obtener predicciones precisas, sino también comprender qué factores influyen más en esas predicciones. En cuencas urbanas, el análisis *SHAP* puede identificar el uso de suelo agrícola como un predictor clave de las concentraciones de nitrógeno y fósforo (Xue et al., 2024). Esta capacidad de explicar las relaciones causales es vital para la toma de decisiones informadas y la formulación de políticas ambientales.

La combinación de RF y *SHAP* es particularmente valiosa para sistemas de monitoreo continuo y rutinario, donde la transparencia operativa es tan importante como la precisión. *SHAP* proporciona una visión global y local de la importancia de cada atributo de entrada en las predicciones del modelo, lo que permite identificar cómo cada parámetro del agua contribuye positiva o negativamente a la clasificación final de la calidad del agua (Ramya et al., 2024). Esta capacidad de trazabilidad es fundamental para la confianza en los sistemas de monitoreo y para la validación de los modelos en entornos regulados.

4 Discusión

El análisis de los estudios incluidos en esta revisión permite responder a las preguntas de investigación planteadas, revelando un campo dinámico pero con importantes desafíos metodológicos. Los hallazgos confirman patrones claros en la selección de variables de entrada y algoritmos, una adopción creciente pero a menudo superficial de técnicas de XAI, y una serie de brechas que limitan la aplicabilidad y comparabilidad de los modelos actuales.

4.1 Hallazgos y retos

En respuesta a la primera pregunta de investigación sobre las variables de entrada, se observa un dominio abrumador de los parámetros fisicoquímicos y de nutrientes medidos *in situ*. Indicadores como el *pH*, Oxígeno Disuelto (*DO*), Demanda Química de Oxígeno (*COD*) y Amonio ($\text{NH}_3\text{--N}$) son la base de la mayoría de los modelos predictivos. Esta preferencia refleja las prácticas de monitoreo convencionales y un consenso implícito sobre su rol como predictores fundamentales de la calidad del agua. Sin embargo, esta dependencia de mediciones puntuales contrasta con un hallazgo emergente y de gran relevancia: la integración de datos de teledetección (*remote sensing*). Estudios como los de Kandasamy et al., Anand et al., Lausch et al. y Tian et al. utilizan imágenes satelitales (Sentinel-2, Landsat 8/9) para derivar indicadores ópticamente activos como la clorofila-*a* o la turbidez. Este enfoque supera las limitaciones espaciales del muestreo tradicional y permite una monitorización a gran escala, demostrando el potencial de las variables derivadas de sensores remotos para enriquecer los modelos (Kandasamy et al., 2025; Anand et al., 2024; Lausch et al., 2025; Tian et al., 2024).

En cuanto a los algoritmos, la revisión muestra una clara preferencia por los modelos de ensamblaje, con Random Forest (RF) y XGBoost como los más recurrentes, presentes en 22 y 13 de los 40 estudios, respectivamente. Su popularidad se debe a su alto rendimiento y su capacidad para capturar interacciones complejas sin una configuración extensa. (Makumbura et al., 2024; Ramya et al., 2024). Modelos de *Deep Learning* como *LSTM* y *BiLSTM* se emplean de forma más específica para series temporales con dependencias a largo plazo. Esta especialización contrasta con modelos de alta complejidad como el *CFD - BiLSTM*, que integra la física de fluidos con el aprendizaje de patrones temporales, o los modelos híbridos apilados. Si bien estos últimos ofrecen una mayor fidelidad física o predictiva, su complejidad computacional y dificultad de implementación limitan su uso generalizado frente a la eficiencia y el sólido rendimiento de RF y XGBoost en la mayoría de los escenarios (Kushwaha et al., 2024; Alfwzan et al., 2024).

La integración de XAI, aunque es un objetivo central de esta revisión, muestra una aplicación incipiente y metodológicamente homogénea. La técnica dominante es SHAP (SHapley Additive exPlanations), utilizada en estudios como los de Kundu et al., Makumbura et al. y Tian et al. para una función principal: cuantificar la importancia global de las variables de entrada. Por ejemplo, Makumbura et al. utilizan SHAP para identificar que el *COD* y el *BOD* son los predictores más influyentes en su modelo de *WQI*. De manera similar, Tian et al. emplean SHAP para demostrar cómo la cobertura urbana y el *NDVI* impactan el estado trófico de los lagos, proporcionando explicaciones globales y locales (Kundu et al., 2025; Makumbura et al., 2024; Tian et al., 2024).

Pese a su poder, la aplicación de SHAP se limita en la mayoría de

los casos a la generación de gráficos de dependencia o importancia, lo que representa una explicabilidad superficial. Rara vez se profundiza en cómo estas explicaciones pueden guiar acciones de manejo concretas, lo que se alinea con la advertencia de Rudin (2019) sobre el riesgo de enfocarse en explicar ‘cajas negras’ en lugar de construir modelos inherentemente interpretables. La aplicación de LIME es aún menos frecuente, apareciendo como un método complementario para validar explicaciones locales en contextos de alta colinealidad. Esta dependencia de una única técnica de XAI y su uso limitado para la clasificación de características subraya una brecha crítica: la transición de la interpretabilidad (entender qué hace el modelo) a la accionabilidad (usar ese entendimiento para tomar mejores decisiones) aún no se ha consolidado (Rudin, 2019).

La discusión de los hallazgos revela una serie de limitaciones transversales en los estudios revisados. La más importante es la falta de estandarización en la selección de variables y métricas de evaluación. Los estudios utilizan conjuntos de datos heterogéneos y métricas de rendimiento diversas (R^2 , MAE, RMSE), lo que impide una comparación directa y rigurosa del desempeño entre diferentes algoritmos y enfoques. Por ejemplo, mientras que se comparan ARIMA y LSTM con errores de pronóstico, otros evalúan clasificadores con métricas de precisión (Gupta et al., 2024).

Un segundo reto es la calidad y disponibilidad de los datos. La no estacionariedad, la escasez de datos y los valores faltantes son problemas recurrentes que afectan la robustez de los modelos. Se directamente este tipo de desafíos mediante el uso de transformadas wavelet y aprendizaje por transferencia, demostrando que el preprocesamiento avanzado es clave para mejorar la precisión en contextos de datos limitados. La mayoría de los estudios, sin embargo, no implementan soluciones tan sofisticadas, lo que cuestiona la generalización de sus resultados (Chen et al., 2024).

Finalmente, la revisión evidencia una tensión entre complejidad y eficiencia. Por un lado, modelos de alta complejidad como los de Alfwzan et al. o Liu et al. ofrecen un rendimiento superior en escenarios específicos pero son difíciles de implementar y explicar. Por otro lado, enfoques más sencillos como el de Gupta et al. con ARIMA son altamente interpretables pero limitados en su capacidad para manejar la no linealidad inherente a los sistemas hídricos. La combinación de RF+SHAP, como se ve en Xue et al. y Makumbura et al., parece ofrecer el mejor equilibrio actual entre rendimiento predictivo y explicabilidad práctica, aunque, como se mencionó, esta explicabilidad necesita ser profundizada para tener un verdadero impacto operativo (Alfwzan et al., 2024; Liu et al., 2025; Gupta et al., 2024; Xue et al., 2024; Makumbura et al., 2024).

4.2 Propuesta futura y tendencias emergentes

Las brechas identificadas en la literatura actual abren múltiples y prometedoras vías para futuras investigaciones, las cuales podrían incrementar significativamente el impacto y la fiabilidad de la IA explicable en la gestión de la calidad del agua. Las tendencias emergentes, como la integración de datos de teledetección, el desarrollo de modelos híbridos y la aplicación de técnicas XAI más sofisticadas, son clave para superar los retos actuales.

Una de las tendencias más claras es el desarrollo de modelos híbridos que fusionan distintas arquitecturas para capturar la complejidad multidimensional de los sistemas hídricos. En lugar de depender de un único algoritmo, estos enfoques aprovechan ventajas complementarias. Por ejemplo, la combinación de Redes Neuronales Convolucionales (CNN) para extraer características espas-

ciales y Redes Neuronales Recurrentes (LSTM) para dependencias temporales ya ha demostrado su superioridad en la predicción de índices como el WQI. De forma aún más avanzada, modelos como FNN-BiLSTM-Attention-DA integran lógica difusa (FNN) para manejar la incertidumbre, BiLSTM para capturar patrones temporales complejos y mecanismos de atención para priorizar las entradas más relevantes. Este tipo de arquitectura no solo mejora la precisión predictiva, sino que también aborda la escasez de datos mediante la adaptación de dominio, permitiendo transferir conocimiento entre diferentes cuencas (Ehteram et al., 2024; Liu et al., 2025).

Simultáneamente, la integración de datos de teledetección (*remote sensing*) se consolida como un campo de expansión crucial. Trabajos como los de Tian et al., Lausch et al. y Anand et al. han demostrado que las imágenes satelitales (e.g., Sentinel-2, Landsat 8/9) son una fuente de datos invaluable para el monitoreo a gran escala. Estos sensores permiten derivar variables ópticamente activas (clorofila-a, turbidez, etc.) y analizar la influencia de la cobertura del suelo (LUCC), superando las limitaciones espaciales y de costo del muestreo *in situ*. La combinación de XGBoost y datos de teledetección, explicada mediante SHAP, ha permitido cuantificar el impacto de políticas urbanas y del cambio climático en la calidad del agua de lagos. El potencial futuro reside en la fusión sistemática de estos datos multimodales: integrar datos de sensores remotos, mediciones *in situ* y variables meteorológicas en un único marco híbrido y explicable. Aquí, la XAI será indispensable para desentrañar cómo el modelo pondera y prioriza cada fuente de información (Tian et al., 2024; Lausch et al., 2025; Anand et al., 2024; Song et al., 2024).

La investigación futura debe trascender la simple clasificación de la importancia de las características, que es el uso predominante de SHAP en la actualidad. Es necesario explorar técnicas de XAI que ofrezcan explicaciones más ricas y operativas:

- **Explicaciones Contrafactuales:** Futuros modelos podrían responder a preguntas como: ‘¿En qué medida tendría que reducirse la concentración de fosfatos para que el WQI mejorara en un 10 %?’ o ‘¿Qué patrón de caudal habría evitado la proliferación de cianobacterias observada?’. Este tipo de explicación es directamente aplicable para la planificación y gestión de recursos hídricos.
- **Descubrimiento de Conocimiento Causal:** En lugar de solo explicar el modelo, se puede utilizar XAI para extraer conocimiento científico. Por ejemplo, identificar umbrales críticos de contaminantes, descubrir interacciones sinérgicas no documentadas previamente (como las que Makumbura et al. comenzaron a explorar) o generar hipótesis causales que puedan ser validadas experimentalmente (Makumbura et al., 2024).
- **Adopción de Modelos Inherentemente Interpretables:** Siguiendo la recomendación de Rudin (2019), se debe investigar más a fondo el uso de modelos como árboles de decisión generalizados, regresión con restricciones o redes bayesianas. El trabajo de Lloyd et al. con Redes Bayesianas para predecir riesgos para la salud pública en estuarios es un excelente ejemplo de cómo estos modelos pueden ofrecer interpretabilidad causal por diseño (Lloyd et al., 2024).

Para que el campo madure y sus soluciones sean adoptadas a gran escala, es imperativo abordar la falta de estandarización. Se

necesita el desarrollo de conjuntos de datos de referencia (*benchmark datasets*) y protocolos de evaluación estandarizados. Esto permitiría una comparación justa y reproducible entre diferentes modelos, superando la heterogeneidad actual en métricas y variables de entrada que dificulta la generalización de resultados.

Además, es fundamental mover la validación de los modelos ‘fuera del laboratorio’. Se requieren más estudios que implementen y evalúen sus soluciones XAI en entornos operativos, en colaboración con agencias ambientales o empresas de servicios públicos. La validación con datos de alta resolución temporal y espacial, así como en escenarios de eventos extremos (inundaciones, sequías), es crucial para demostrar el valor real de estas herramientas en la toma de decisiones en tiempo real. Finalmente, el éxito a largo plazo de estas tecnologías dependerá de una colaboración interdisciplinaria estrecha entre científicos de datos, hidrólogos, ecólogos y responsables políticos. Esta sinergia es esencial para asegurar que los modelos sean no solo precisos, sino también relevantes, fiables y éticos. La discusión debe abordar explícitamente cuestiones de gobernanza de datos y responsabilidad algorítmica, especialmente a la luz de marcos regulatorios como el *AI Act* de la Unión Europea, para garantizar una gestión del agua que sea tecnológicamente avanzada, transparente y socialmente responsable (Commission, 2024).

5 Conclusión

La presente revisión de la literatura ha permitido analizar y sintetizar el estado del arte de la Inteligencia Artificial Explicable (XAI) aplicada a la predicción de la calidad del agua, respondiendo a las preguntas de investigación planteadas. Los hallazgos confirman que el campo está experimentando un crecimiento acelerado, impulsado tanto por la necesidad de una gestión hídrica más precisa como por marcos regulatorios emergentes como el *AI Act* de la Unión Europea, que exigen transparencia y fiabilidad en sistemas de alto riesgo.

Se ha evidenciado un claro dominio de los modelos de aprendizaje supervisado, con una marcada preferencia por algoritmos de ensamblaje como *Random Forest* y *XGBoost*, y modelos de soporte vectorial (*SVM*), debido a su robustez y alto rendimiento en el manejo de datos no lineales. Estos modelos se nutren predominantemente de parámetros fisicoquímicos y de nutrientes medidos *in situ*, como *pH*, oxígeno disuelto (*DO*) y demanda química de oxígeno (*COD*), lo que refleja las prácticas de monitoreo convencionales. No obstante, una tendencia emergente de gran relevancia es la integración de datos de teledetección (*remote sensing*), que permite superar las limitaciones espaciales del muestreo tradicional y ofrece una monitorización a gran escala, como demuestran trabajos que combinan imágenes satelitales con técnicas de ML.

En cuanto a la explicabilidad, la técnica XAI predominante es *SHAP*, utilizada principalmente para la clasificación de la importancia de las variables, seguida a distancia por *LIME*. Si bien este es un paso crucial hacia la transparencia, la revisión ha revelado que su aplicación es a menudo superficial, sin profundizar en cómo las explicaciones pueden traducirse en acciones de gestión concretas. Esto subraya una brecha fundamental entre la interpretabilidad del modelo y la accionabilidad de sus resultados, un reto que el campo debe abordar para maximizar su impacto práctico, tal como lo advierte la literatura sobre la necesidad de modelos inherentemente interpretables.

Las limitaciones identificadas son significativas. La falta de es-

tandarización en los conjuntos de datos y en las métricas de evaluación dificulta la comparabilidad entre los estudios. La calidad y disponibilidad de los datos, marcada por la no estacionariedad y los valores faltantes, sigue siendo un obstáculo para la robustez de los modelos, aunque ya existen enfoques avanzados para mitigar este problema.

De cara al futuro, esta revisión identifica varias direcciones de investigación prometedoras. Es fundamental avanzar hacia la integración sistemática de datos multimodales, combinando mediciones *in situ*, variables climáticas y datos de teledetección en marcos de modelos híbridos y explicables. Asimismo, es necesario explorar técnicas de XAI más allá de la importancia de las características, como las explicaciones contrafactuales, que son directamente operativas para los gestores de recursos hídricos. Para que el campo madure, se requiere el desarrollo de *benchmark datasets* y protocolos de validación estandarizados, así como la evaluación de los modelos en entornos operativos del mundo real.

En definitiva, la aplicación de la IA explicable en la predicción de la calidad del agua ha demostrado un potencial transformador. Sin embargo, para pasar de la promesa académica a la implementación robusta y fiable, es indispensable una mayor profundización metodológica en XAI, la estandarización de las prácticas y una colaboración interdisciplinaria más estrecha. Solo así se podrá garantizar que estas tecnologías avanzadas contribuyan de manera efectiva y transparente a la protección de nuestros recursos hídricos.

Financiamiento

Este trabajo de investigación no recibió financiamiento de ninguna organización pública, comercial o sin fines de lucro.

Disponibilidad de datos

Este estudio constituye una revisión teórica que no generó ni utilizó conjuntos de datos originales. Todos los materiales consultados están debidamente referenciados en la sección bibliográfica.

Contribuciones de autoría CRediT

- **Mathias Augusto Yanqui-Cheneaux:** Conceptualización, Metodología, Análisis formal, Supervisión, Redacción - borrador original.
- **Mathias Joaquín Medina-Grados:** Curación de información, Validación, Análisis formal, Visualización - revisión y edición.

Conflictos de interés

Los autores certifican que no existen intereses económicos, personales o profesionales que puedan comprometer la objetividad del presente trabajo académico. No existen relaciones financieras ni compromisos institucionales que puedan interpretarse como influencia en los resultados o conclusiones de esta investigación.

Reconocimiento

Los autores expresan su agradecimiento a la Dra. Ana Daneida Villanueva Llapa por su valiosa orientación académica durante el desarrollo del curso *Taller de Investigación* en la Escuela Profesional de Ingeniería Industrial de la Universidad Católica San Pablo (Arequipa, Perú).

Referencias

- Alfwzan, Wafa F. et al. (jun. de 2024). «Water quality assessment using Bi-LSTM and computational fluid dynamics (CFD) techniques». En: *Alexandria Engineering Journal* 97, págs. 346-359. ISSN: 1110-0168. DOI: [10.1016/J.AEJ.2024.04.030](https://doi.org/10.1016/J.AEJ.2024.04.030).
- Anand, Vicky, Bakimchandra Oinam y Silke Wieprecht (dic. de 2024). «Machine learning approach for water quality predictions based on multispectral satellite images». En: *Ecological Informatics* 84, pág. 102868. ISSN: 1574-9541. DOI: [10.1016/J.ECOINF.2024.102868](https://doi.org/10.1016/J.ECOINF.2024.102868).
- Bolick, Madeleine M et al. (2023). «Comparison of machine learning algorithms to predict dissolved oxygen in an urban stream». En: *Environmental Science and Pollution Research* 30.32, págs. 78075-78096.
- Castillo, Alberto Fernández del et al. (sep. de 2024). «Improving river water quality prediction with hybrid machine learning and temporal analysis». En: *Ecological Informatics* 82, pág. 102655. ISSN: 1574-9541. DOI: [10.1016/J.ECOINF.2024.102655](https://doi.org/10.1016/J.ECOINF.2024.102655).
- Chen, Shengyue et al. (ene. de 2024). «A coupled model to improve river water quality prediction towards addressing non-stationarity and data limitation». En: *Water Research* 248, pág. 120895. ISSN: 0043-1354. DOI: [10.1016/J.WATRES.2023.120895](https://doi.org/10.1016/J.WATRES.2023.120895).
- Commission, European (2024). *AI Act Explorer – Interactive Guide to the Artificial Intelligence Act*. <https://artificialintelligenceact.eu/ai-act-explorer/>. Accessed: 2025-06-04.
- Das, Abhijeet (jun. de 2025). «Spatiotemporal evaluation and impact of superficial factors on surface water quality for drinking using innovative techniques in Mahanadi River Basin, Odisha, India». En: *Journal of Hydrology: Regional Studies* 59, pág. 102366. ISSN: 2214-5818. DOI: [10.1016/J.EJRH.2025.102366](https://doi.org/10.1016/J.EJRH.2025.102366). URL: <https://linkinghub.elsevier.com/retrieve/pii/S2214581825001910>.
- Ding, Fei et al. (mar. de 2025). «Using multiple machine learning algorithms to optimize the water quality index model and their applicability». En: *Ecological Indicators* 172, pág. 113299. ISSN: 1470-160X. DOI: [10.1016/J.ECOLIND.2025.113299](https://doi.org/10.1016/J.ECOLIND.2025.113299).
- Ehteram, Mohammad et al. (mar. de 2024). «An advanced deep learning model for predicting water quality index». En: *Ecological Indicators* 160, pág. 111806. ISSN: 1470-160X. DOI: [10.1016/J.ECOLIND.2024.111806](https://doi.org/10.1016/J.ECOLIND.2024.111806).
- Eid, Mohamed Hamdy et al. (dic. de 2024). «An advanced approach for drinking water quality indexing and health risk assessment supported by machine learning modelling in Siwa Oasis, Egypt». En: *Journal of Hydrology: Regional Studies* 56, pág. 101967. ISSN: 2214-5818. DOI: [10.1016/J.EJRH.2024.101967](https://doi.org/10.1016/J.EJRH.2024.101967).
- Geng, Mingming et al. (2024). «Assessing the impact of water-sediment factors on water quality to guide river-connected lake water environment improvement». En: *Science of The Total Environment* 912, pág. 168866.
- Gupta, Neetu, Surendra Yadav y Neha Chaudhary (ene. de 2024). «Time Series Analysis and Forecasting of Water Quality Parameters along Yamuna River in Delhi». En: *Procedia Computer Science* 235, págs. 3191-3206. ISSN: 1877-0509. DOI: [10.1016/J.PROCS.2024.04.302](https://doi.org/10.1016/J.PROCS.2024.04.302).
- Han, Zhoulin et al. (mayo de 2025). «Predicting and investigating water quality index by robust machine learning methods». En: *Journal of Environmental Management* 381, pág. 125156. ISSN: 0301-4797. DOI: [10.1016/J.JENVMAN.2025.125156](https://doi.org/10.1016/J.JENVMAN.2025.125156).
- Irwan, Dani et al. (jun. de 2025). «River water quality monitoring using machine learning with multiple possible in-situ scenarios». En: *Environmental and Sustainable Indicators* 26, pág. 100620. ISSN: 2665-9727. DOI: [10.1016/J.INDIC.2025.100620](https://doi.org/10.1016/J.INDIC.2025.100620).
- Islam, Md Sahidul, Hailong Yin y Mustafizur Rahman (ene. de 2024). «Long-term trend prediction of surface water quality of two main river basins of China using Machine Learning Method». En: *Procedia Computer Science* 236, págs. 257-264. ISSN: 1877-0509. DOI: [10.1016/J.PROCS.2024.05.029](https://doi.org/10.1016/J.PROCS.2024.05.029).
- Islam, Md Touhidul et al. (dic. de 2024). «Regional irrigation water quality index for the Old Brahmaputra River, Bangladesh: A multivariate and GIS-based spatiotemporal assessment». En: *Results in Engineering* 24, pág. 103667. ISSN: 2590-1230. DOI: [10.1016/J.RINENG.2024.103667](https://doi.org/10.1016/J.RINENG.2024.103667).
- Kandasamy, Lavanya et al. (mar. de 2025). «Enhanced remote sensing and deep learning aided water quality detection in the Ganges River, India supporting monitoring of aquatic environments». En: *Results in Engineering* 25, pág. 103604. ISSN: 2590-1230. DOI: [10.1016/J.RINENG.2024.103604](https://doi.org/10.1016/J.RINENG.2024.103604).
- Karpagam, V., S. Christy y M.O. Edeh (2024). «Deep learning-based water quality index classification using stacked ensemble variational mode decomposition». En: *Environmental Research Communications* 6 (6). DOI: [10.1088/2515-7620/ad549e](https://doi.org/10.1088/2515-7620/ad549e).
- Kundu, Sourav et al. (feb. de 2025). «Unveiling the hidden connections: Using explainable artificial intelligence to assess water quality criteria in nine giant rivers». En: *Journal of Cleaner Production* 492, pág. 144861. ISSN: 0959-6526. DOI: [10.1016/J.JCLEPRO.2025.144861](https://doi.org/10.1016/J.JCLEPRO.2025.144861).

- Kushwaha, Nand Lal et al. (mayo de 2024). «Stacked hybridization to enhance the performance of artificial neural networks (ANN) for prediction of water quality index in the Bagh river basin, India». En: *Heliyon* 10 (10), e31085. ISSN: 2405-8440. DOI: [10.1016/J.HELİYON.2024.E31085](https://doi.org/10.1016/J.HELİYON.2024.E31085).
- Lausch, Angela et al. (ene. de 2025). «Monitoring and modelling landscape structure, land use intensity and landscape change as drivers of water quality using remote sensing». En: *Science of The Total Environment* 960 (December 2024), pág. 178347. ISSN: 0048-9697. DOI: [10.1016/J.SCITOTENV.2024.178347](https://doi.org/10.1016/J.SCITOTENV.2024.178347). URL: <https://doi.org/10.1016/j.scitotenv.2024.178347>.
- Liu, Dong y Xiaolong Cheng (jun. de 2025). «FNN-BiLSTM-Attention-DA: A hybrid fuzzy neural network and BiLSTM with multi-sensor information fusion for water quality monitoring and warning». En: *Alexandria Engineering Journal* 124, págs. 624-639. ISSN: 1110-0168. DOI: [10.1016/J.AEJ.2025.04.011](https://doi.org/10.1016/J.AEJ.2025.04.011).
- Lloyd, Simon D. et al. (mayo de 2024). «Predicting recreational water quality and public health safety in urban estuaries using Bayesian Networks». En: *Water Research* 254, pág. 121319. ISSN: 0043-1354. DOI: [10.1016/J.WATRES.2024.121319](https://doi.org/10.1016/J.WATRES.2024.121319).
- Lu, Yanbo, Tong Li y Ying Deng and (2024). «Predication of Water Pollution Peak Concentrations by Hybrid BP Artificial Neural Network Coupled with Genetic Algorithm». En: *Applied Artificial Intelligence* 38.1, pág. 2341356. DOI: [10.1080/08839514.2024.2341356](https://doi.org/10.1080/08839514.2024.2341356). eprint: <https://doi.org/10.1080/08839514.2024.2341356>. URL: <https://doi.org/10.1080/08839514.2024.2341356>.
- Makumbura, R.K. et al. (2024). «Advancing water quality assessment and prediction using machine learning models, coupled with explainable artificial intelligence (XAI) techniques like shapley additive explanations (SHAP) for interpreting the black-box nature». En: *Results in Engineering* 23. DOI: [10.1016/j.rineng.2024.102831](https://doi.org/10.1016/j.rineng.2024.102831).
- Maußner, Claudia et al. (ene. de 2025). «Explainable artificial intelligence for reliable water demand forecasting to increase trust in predictions». En: *Water Research* 268, pág. 122779. ISSN: 0043-1354. DOI: [10.1016/J.WATRES.2024.122779](https://doi.org/10.1016/J.WATRES.2024.122779).
- Nallakuruppan, M. K. et al. (2024). «Reliable water quality prediction and parametric analysis using explainable AI models». En: *Scientific Reports* 14.1, pág. 7520. ISSN: 2045-2322. DOI: [10.1038/s41598-024-56775-y](https://doi.org/10.1038/s41598-024-56775-y). URL: <https://doi.org/10.1038/s41598-024-56775-y>.
- Nasir, Nida et al. (2022). «Water quality classification using machine learning algorithms». En: *Journal of Water Process Engineering* 48, pág. 102920. ISSN: 2214-7144. DOI: <https://doi.org/10.1016/j.jwpe.2022.102920>. URL: <https://www.sciencedirect.com/science/article/pii/S2214714422003646>.
- Nong, Xizhi et al. (nov. de 2024). «A novel coupling interpretable machine learning framework for water quality prediction and environmental effect understanding in different flow discharge regulations of hydro-projects». En: *Science of The Total Environment* 950, pág. 175281. ISSN: 0048-9697. DOI: [10.1016/J.SCITOTENV.2024.175281](https://doi.org/10.1016/J.SCITOTENV.2024.175281).
- Page, Matthew J et al. (2021). «The PRISMA 2020 statement: an updated guideline for reporting systematic reviews». En: *BMJ* 372. DOI: [10.1136/bmj.n71](https://doi.org/10.1136/bmj.n71). eprint: <https://www.bmj.com/content/372/bmj.n71.full.pdf>. URL: <https://www.bmj.com/content/372/bmj.n71>.
- Park, Jungsu et al. (2022). «Interpretation of ensemble learning to predict water quality using explainable artificial intelligence». En: *Science of the Total Environment* 832, pág. 155070.
- Pazo, M. et al. (jun. de 2024). «Enhancing water quality prediction for fluctuating missing data scenarios: A dynamic Bayesian network-based processing system to monitor cyanobacteria proliferation». En: *Science of The Total Environment* 927, pág. 172340. ISSN: 0048-9697. DOI: [10.1016/J.SCITOTENV.2024.172340](https://doi.org/10.1016/J.SCITOTENV.2024.172340).
- Ramya, S., S. Srinath y P. Tuppad (2024). «Comprehensive analysis of multiple classifiers for enhanced river water quality monitoring with explainable AI». En: *Case Studies in Chemical and Environmental Engineering* 10. DOI: [10.1016/j.csee.2024.100822](https://doi.org/10.1016/j.csee.2024.100822).
- Rigueira, Xurxo et al. (mayo de 2025). «Multivariate functional data analysis and machine learning methods for anomaly detection in water quality sensor data». En: *Environmental Modelling & Software* 190, pág. 106443. ISSN: 1364-8152. DOI: [10.1016/J.ENVSOFT.2025.106443](https://doi.org/10.1016/J.ENVSOFT.2025.106443).
- Roy, Bimol Nath et al. (ago. de 2024). «Principal component analysis incorporated water quality index modeling for Dhaka-based rivers». En: *City and Environment Interactions* 23, pág. 100150. ISSN: 2590-2520. DOI: [10.1016/J.CACINT.2024.100150](https://doi.org/10.1016/J.CACINT.2024.100150).
- Rudin, Cynthia (mayo de 2019). «Stop Explaining Black Box Machine Learning Models for High Stakes Decisions and Use Interpretable Models Instead». En: *Nature Machine Intelligence* 1, págs. 206-215. DOI: [10.1038/s42256-019-0048-x](https://doi.org/10.1038/s42256-019-0048-x).
- Sajib, Abdul Majed et al. (mayo de 2024). «Assessing water quality of an ecologically critical urban canal incorporating machine learning approaches». En: *Ecological Informatics* 80, pág. 102514. ISSN: 1574-9541. DOI: [10.1016/J.ECOINF.2024.102514](https://doi.org/10.1016/J.ECOINF.2024.102514).
- Sajib, Abdul Majed et al. (mar. de 2025). «Novel groundwater quality index (GWQI) model: A reliable approach for the assessment of groundwater». En: *Results in Engineering* 25 (February), pág. 104265. ISSN: 2590-1230. DOI: [10.1016/J.RINENG.2025.104265](https://doi.org/10.1016/J.RINENG.2025.104265). URL: <https://doi.org/10.1016/j.rineng.2025.104265>.
- Savoy, Philip y Judson W. Harvey (2023). «Predicting Daily River Chlorophyll Concentrations at a Continental Scale». En: *Water Resources Research* 59.11. e2022WR034215. DOI: <https://doi.org/10.1029/2022WR034215>. eprint: <https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/2022WR034215>. URL: <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2022WR034215>.
- Song, Yang et al. (oct. de 2024). «Quantitative prediction of water quality in Dongjiang Lake watershed based on LUCC». En: *Ecotoxicology and Environmental Safety* 284, pág. 117005. ISSN: 0147-6513. DOI: [10.1016/J.ECOENV.2024.117005](https://doi.org/10.1016/J.ECOENV.2024.117005).
- Tian, Peilong et al. (oct. de 2024). «Exploring the effects of climate change and urban policies on lake water quality using remote sensing and explainable artificial intelligence». En: *Journal of Cleaner Production* 475, pág. 143649. ISSN: 0959-6526. DOI: [10.1016/J.JCLEPRO.2024.143649](https://doi.org/10.1016/J.JCLEPRO.2024.143649). URL: <https://linkinghub.elsevier.com/retrieve/pii/S0959652624030981>.
- Tsegay, Biniam Abbrha y Nicolás M. Peleato (ene. de 2025). «Enhancing long-term water quality modeling by addressing base demand, demand patterns, and temperature uncertainty using unsupervised machine learning techniques». En: *Water Research* 268, pág. 122701. ISSN: 0043-1354. DOI: [10.1016/J.WATRES.2024.122701](https://doi.org/10.1016/J.WATRES.2024.122701).
- Uddin, Md Galal et al. (mayo de 2024). «Data-driven evolution of water quality models: An in-depth investigation of innovative outlier detection approaches-A case study of Irish Water Quality Index (IEWQI) model». En: *Water Research* 255, pág. 121499. ISSN: 0043-1354. DOI: [10.1016/J.WATRES.2024.121499](https://doi.org/10.1016/J.WATRES.2024.121499).
- Uddin, Md Galal et al. (mayo de 2025). «The role of optimizers in developing data-driven model for predicting lake water quality incorporating advanced water quality models». En: *Alexandria Engineering Journal* 122, págs. 411-435. ISSN: 1110-0168. DOI: [10.1016/J.AEJ.2025.03.022](https://doi.org/10.1016/J.AEJ.2025.03.022).
- Usang, Richard O., Bamidele I. Olu-Owolabi y Kayode O. Adebawale (feb. de 2025). «Integrating principal component analysis, fuzzy inference systems, and advanced neural networks for enhanced estuarine water quality assessment». En: *Journal of Hydrology: Regional Studies* 57, pág. 102182. ISSN: 2214-5818. DOI: [10.1016/J.EJRH.2025.102182](https://doi.org/10.1016/J.EJRH.2025.102182).
- Wan, Hang et al. (ago. de 2025). «Temporal and spatial feature extraction using graph neural networks for multi-point water quality prediction in river network areas». En: *Water Research* 281, pág. 123561. ISSN: 0043-1354. DOI: [10.1016/J.WATRES.2025.123561](https://doi.org/10.1016/J.WATRES.2025.123561).
- Xiao, Weifeng et al. (feb. de 2025). «Spatiotemporal dynamics and optimization of water quality assessment in the Nantong section of the Yangtze River Basin: A WQImIn approach». En: *Journal of Hydrology: Regional Studies* 57, pág. 102106. ISSN: 2214-5818. DOI: [10.1016/J.EJRH.2024.102106](https://doi.org/10.1016/J.EJRH.2024.102106).
- Xu, Jing et al. (jul. de 2024). «Assessing and predicting water quality index with key water parameters by machine learning models in coastal cities, China». En: *Heliyon* 10 (13), e33695. ISSN: 2405-8440. DOI: [10.1016/J.HELİYON.2024.E33695](https://doi.org/10.1016/J.HELİYON.2024.E33695).
- Xue, Jingyuan et al. (feb. de 2024). «Predictive modeling of nitrogen and phosphorus concentrations in rivers using a machine learning framework: A case study in an urban-rural transitional area in Wenzhou China». En: *Science of The Total Environment* 910, pág. 168521. ISSN: 0048-9697. DOI: [10.1016/J.SCITOTENV.2023.168521](https://doi.org/10.1016/J.SCITOTENV.2023.168521).
- Yousefi, Mahmood et al. (dic. de 2024). «An innovative combination of extra trees within adaboost for accurate prediction of agricultural water quality indices». En: *Results in Engineering* 24, pág. 103534. ISSN: 2590-1230. DOI: [10.1016/J.RINENG.2024.103534](https://doi.org/10.1016/J.RINENG.2024.103534).