ISSN 2007-9737

Advances and Applications of Artificial Intelligence in Wastewater Treatment: A Bibliometric Analysis and Systematic Review

Enzo Mayta-Laura[1], Alberto Yauri-Olivares[1], Javier Gamboa-Cruzado[2], Patricia Huarancca-Contreras[3], Jaime Mayorga-Rojas[2], Angel Nuñez Meza[4], Jorge Nolasco Valenzuela[5], Cesar Jesús Núñez-Prado[6, *]

[1] Universidad Nacional Federico Villarreal,
Peru

[2] Universidad Nacional Mayor de San Marcos,
Peru

[3] Universidad Nacional “San Luis Gonzaga”,
Peru

[4] Universidad Nacional Daniel Alcides Carrión,
Peru

[5] Universidad San Ignacio de Loyola,
Peru

[6] Instituto Politécnico Nacional, ESIMEZ,
Mexico

{2019231022, 2021018292} @unfv.edu.pe, {jgamboac, jaime.mayorga} @unmsm.edu.pe, patricia.huarancca@unica.edu.pe, anunezm@undac.edu.pe, jorge.nolasco@usil.pe, cesar.jnprado@gmail.com

Abstract. Artificial intelligence (AI) has emerged as a key technology for optimizing wastewater treatment by improving monitoring, contaminant prediction, and operational efficiency. However, its application remains fragmented due to the variety of models and metrics employed, which limits a comprehensive understanding of its impact and effectiveness. The objective of this paper was to analyze the approaches, models, and metrics used in recent scientific literature. The methodology followed PRISMA guidelines, analyzing 5,487 studies from five databases (Scopus, Web of Science, IEEE Xplore, EBSCOhost, and ProQuest) between 2017 and 2025, and selecting 63 papers based on quality and relevance criteria. The findings revealed that the most common metrics used to assess AI effectiveness are RMSE, MAE, and R^2; that the highest-impact journals are concentrated in quartiles Q1 and Q2, reflecting a high level of scientific rigor; and that China, the United States, and India lead the collaboration and co-occurrence networks. Moreover, emerging topics revolve around deep learning, the Internet of Things (IoT), and environmental sustainability. Overall, the review confirms that AI represents a strategic tool to enhance operational efficiency and decision-making in wastewater management, opening new perspectives toward intelligent, sustainable, and adaptive systems.

Keywords. Artificial intelligence, neural networks, expert systems, wastewater treatment, sewage treatment, wastewater management, systematic review.

1 Introduction

Artificial intelligence (AI) has become a key tool for automating, optimizing, and accelerating complex tasks that require processing large volumes of data. Thanks to these capabilities, AI promotes the development of sustainable solutions aimed at environmental protection. In wastewater treatment, its application is essential, as it enables the simulation and optimization of processes without

ISSN 2007-9737

compromising ecosystems, fostering more efficient and sustainable decision-making. The use of machine learning (ML) models and artificial neural networks (ANN) has demonstrated high effectiveness in predicting and improving results at different stages of wastewater treatment. In recent years, this line of research has gained remarkable relevance, driven by access to AI technologies and the growing need for smarter and more environmentally friendly water management. In the field of water quality parameter prediction, authors [2, 3, 5] have demonstrated the effectiveness of machine learning models such as MLP, GFFR, and neural networks for estimating conductivity, biological oxygen demand, and pollutant removal efficiency, achieving correlations higher than 0.99 and highlighting the influence of variables such as total suspended solids. Meanwhile, other investigations [7, 21] applied feedforward neural networks and random forests to predict BOD_5 concentration and effluent quality, identifying temperature and conductivity as critical factors influencing treatment performance.

Regarding the implementation of virtual sensors, recent contributions [8, 11, 13] developed hybrid architectures based on GRU, CNN, and LSTM to predict parameters such as ammonium and viruses in wastewater, improving the adaptive capacity of the system through integration with knowledge dictionaries and generative adversarial networks. Likewise, works [9, 12, 23] optimized process control in treatment plants using reinforcement learning and policy gradient algorithms, reducing time and energy consumption by up to 24.25% and increasing operational efficiency. Concerning energy prediction and optimization, studies [16, 96] employed LSTM and BiGRU models together with KPI-based methodologies to predict energy consumption and improve eco-efficiency, achieving very low MAPE error values compared to traditional approaches.

In the treatment of specific contaminants, research [18, 20, 104] applied fuzzy modeling, marine predator algorithms, and nature-inspired techniques to degrade pharmaceutical compounds such as amoxicillin and reduce N_2O emissions, improving the efficiency of the activated sludge process. In the field of real-time monitoring, authors [4, 24, 71] implemented deep learning systems such as MANet and GLFMN to classify flocculating bacteria and segment treatment plants from microscopic and satellite images, while drones and robots facilitated marine pollution monitoring with an increasing global focus. Additionally, several investigations [92, 101] integrated IoT sensors with artificial intelligence to predict optimal chemical dosing and monitor physicochemical parameters in aquaculture, reducing human errors and improving growth rates. In evaluating challenges and trends, research by [17, 78, 99] analyzed how AI addresses the complexity of wastewater treatment, identifying obstacles such as data scarcity, limited industrial adoption, and the need for adaptive solutions in the face of climate change. On the other hand, other studies [100, 102, 105] compared traditional and emerging methods, highlighting the predominance of neural networks, random forest, and SVM in environmental remediation, although noting challenges related to interpretability and data quality.

In terms of geographical and technological contexts, studies [14, 94] documented the transformation of treatment plants in Poland following European integration and identified research hotspots in India on ecology, biogas production, and modeling, while works [10, 25] applied neural networks and DenseNet to predict odor properties of sludge and model anaerobic bioreactors, achieving accuracies of 97.44% in performance evaluation. Despite the remarkable progress in applying artificial intelligence (AI) to wastewater treatment, the scientific literature reveals methodological and conceptual fragmentation. The reviewed studies focus mainly on specific predictions of physicochemical parameters or the optimization of individual processes but lack an integrative vision that articulates the different AI approaches, such as deep learning, reinforcement learning, and hybrid models, based on their comparative effectiveness, adaptability, and operational sustainability. Likewise, there are limitations in the standardization of datasets, performance metrics, and validation protocols, which hinders reproducibility and comparative analysis among studies. This gap highlights the need for systematic syntheses and rigorous bibliometric analyses to identify trends, consolidate methodological taxonomies, and assess the impact of AI on integrated water management. Similarly, a geographical and technological gap persists: most studies are concentrated in regions with high computational capacity, while developing countries show incipient adoption.

Therefore, this study is justified by integrating the systematic review and bibliometric analysis as complementary tools to map the state of the art, quantify scientific evolution, and propose a critical view of the challenges and opportunities of AI in wastewater treatment, thus contributing to more efficient, resilient, and sustainable water resource management.

The present study aims to comprehensively analyze the advances, applications, and trends of artificial intelligence in wastewater treatment through a systematic literature review (SLR) complemented by a bibliometric analysis of recent scientific production.

Accordingly, this SLR paper is organized as follows: Section 2 presents the theoretical background, describing the conceptual foundations and main previous contributions regarding the application of artificial intelligence in wastewater treatment. Section 3 presents the methodology, detailing the stages of the review process, inclusion and exclusion criteria, and bibliometric tools used. Section 4 develops the results and discussion, integrating quantitative and qualitative analysis of the reviewed studies. Finally, Section 5 presents the conclusions and recommendations for future research, aimed at strengthening the use of artificial intelligence in the sustainable management of water resources.

2 Theoretical Background

Given the constant evolution and diversification of artificial intelligence (AI) applications in wastewater treatment, it is essential to understand the conceptual and technological foundations that support this field before delving into recent research. This section aims to contextualize the most relevant AI principles, techniques, and approaches applied in wastewater treatment, providing a theoretical basis that enables a more accurate interpretation of the findings and trends analyzed in this systematic review.

2.1 Artificial Intelligence

Artificial intelligence (AI) is a branch of computer science that seeks to develop systems capable of imitating human cognitive functions such as learning, decision-making, and problem-solving [1]. Its usefulness lies in optimizing complex tasks through the analysis of large volumes of data [91], simulating intelligent behaviors by means of computational algorithms and models [7]. A core component is machine learning, which allows machines to identify patterns and improve their performance without human intervention [12], particularly valuable in dynamic environments that demand real-time decision-making [16]. AI has proven effective in industrial and environmental sectors by facilitating the modeling of nonlinear phenomena, predicting critical variables, and enabling predictive maintenance through artificial neural networks [15, 17]. Its adaptive capability has reduced reliance on physical sensors and improved the estimation of key parameters [31]. Due to its efficiency and adaptability, AI is positioned as a strategic tool across multiple sectors [26]. Within this context, machine learning (ML) emerges with the goal of creating agents that interact with their environment and adjust their behavior through continuous feedback [103].

2.2. Wastewater Treatment

Wastewater treatment is a physical, chemical, and biological process designed to remove contaminants and microorganisms from wastewater, allowing for its safe discharge into the environment or its reuse [10]. This process is fundamental for protecting public health and the environment [20] and typically involves several technical purification stages [30]. Moreover, it is considered a key tool to address the increasing pressure on water resources [36]. Wastewater treatment plants (WWTPs) are essential infrastructures that process water from residential and industrial sources before its return to the natural environment [3]. These facilities improve water quality [16], ensure compliance with environmental standards [23], and contribute to the sustainable management of the water cycle [42]. They are also vital for preserving water resources in the face of urban growth [59].

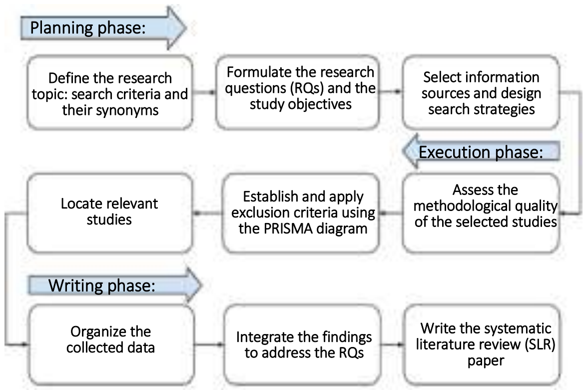


Fig. 1. Development phases of the SLR

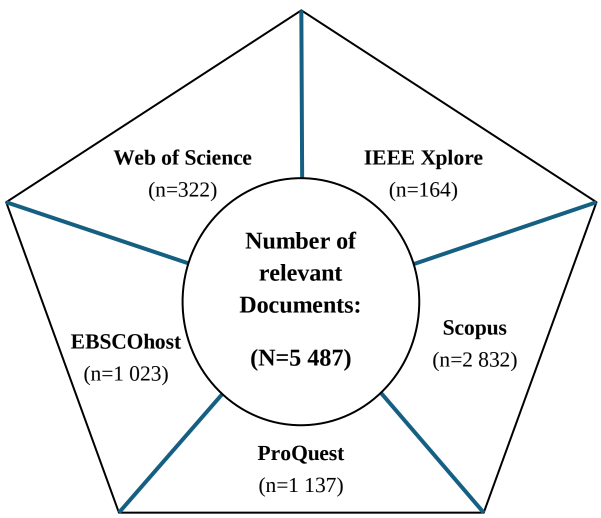


Fig. 2. Studies identified by source

3 Review Method

The Systematic Literature Review (SLR) was applied with the purpose of examining, in a comprehensive, structured, and verifiable manner, the research related to artificial intelligence in wastewater treatment, allowing the study to rigorously address the previously defined research questions (Kitchenham, B. [64]; Philipp, S. [65]; Petersen, K. and colleagues [67]) in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines [108]. In turn, the Systematic Mapping Study (SMS) was used as a complementary approach to identify predominant trends and emerging topics within the research area, through the construction of a keyword map that facilitated thematic classification and the recognition of research patterns (Petersen, K. [66]; Linnenluecke, M. K. [68]; Okoli, C. and his team [69]) (see Figure 1).

3.1 Research Problems

Considering the broad body of previous research on the application of Artificial Intelligence (AI) in wastewater treatment, this study seeks to address a set of Research Questions (RQs) that guide the review process, allowing for the structuring of available scientific evidence and the identification of the main lines of contribution:

RQ1: What criteria are used to evaluate the effectiveness of Artificial Intelligence?
RQ2: What models are currently applied in the development of Artificial Intelligence?
RQ3: What quartile levels do the journals that have published research on the impact of Artificial Intelligence in Wastewater Treatment present?
RQ4: What definitions and theoretical frameworks have been established in studies related to Artificial Intelligence and Wastewater Treatment?
RQ5: Which countries most frequently show co-occurrence relationships in studies on Artificial Intelligence and its influence on Wastewater Treatment?
RQ6: What thematic categories are identified in research addressing Artificial Intelligence and its influence on Wastewater Treatment?

3.2 Information Sources and Search Equations

To ensure the comprehensiveness and scientific quality of the study, internationally recognized academic databases were selected, including Web of Science, IEEE Xplore, EBSCOhost, ProQuest, and Scopus, all of which provide extensive coverage of indexed publications in the fields of engineering, environmental sciences, and information technologies. The comprehensive database search was completed on August 7, 2025.

ISSN 2007-9737

Table 1. Sources and search equations

Source	Search Equation
Web of Science	TI=((“artificial intelligence” OR “neural networks” OR “deep learning” OR “expert systems” OR “cognitive computing”) AND (“wastewater treatment” OR “sewage treatment” OR “purification of wastewater” OR “wastewater management” OR “wastewater purification”)) OR AK=((“artificial intelligence” OR “neural networks” OR “deep learning” OR “expert systems” OR “cognitive computing”) AND (“wastewater treatment” OR “sewage treatment” OR “purification of wastewater”OR “wastewater management” OR “wastewater purification”))
IEEE Xplore	("Document Title":“artificial intelligence” OR "Document Title":“neural networks” OR "Document Title":“deep learning”OR "Document Title":“expert systems”OR "Document Title":“cognitive computing”) AND ("Document Title":“wastewater treatment” OR "Document Title":“sewage treatment” OR "Document Title":“purification of wastewater” OR "Document Title":“wastewater management” OR "Document Title":“wastewater purification”) OR ("Author Keywords":“artificial intelligence” OR "Author Keywords":“neural networks” OR "Author Keywords":“deep learning”OR "Author Keywords":“expert systems”OR "Author Keywords":“cognitive computing”) AND ("Author Keywords":“wastewater treatment” OR "Author Keywords":“sewage treatment” OR "Author Keywords":“purification of wastewater” OR "Author Keywords":“wastewater management” OR "Author Keywords":“wastewater purification”)
EBSCOhost	TI ((“artificial intelligence” OR “neural networks” OR “deep learning" OR “expert systems” OR “cognitive computing”) AND (“wastewater treatment” OR “sewage treatment” OR “purification of wastewater” OR “wastewater management” OR “wastewater purification”)) OR SU ((“artificial intelligence” OR “neural networks” OR “deep learning” OR “expert systems” OR “cognitive computing”) AND (“wastewater treatment” OR “sewage treatment” OR “purification of wastewater” OR “wastewater management” OR “wastewater purification”))
ProQuest	(TI("artificial intelligence" OR "deep learning" OR "neural networks" OR "expert systems" OR "cognitive computing") AND TI ("wastewater treatment" OR "sewage treatment" OR "purification of wastewater" OR "wastewater management" OR "wastewater purification")) OR (SU("artificial intelligence" OR "deep learning" OR "neural networks" OR "expert systems" OR "cognitive computing") AND SU("wastewater treatment" OR "sewage treatment" OR "purification of wastewater" OR "wastewater management" OR "wastewater purification"))
Scopus	((TITLE("artificial intelligence") OR TITLE("deep learning") OR TITLE("neural networks") OR TITLE("expert systems")OR TITLE("cognitive computing")) AND (TITLE("wastewater treatment") OR TITLE("sewage treatment") OR TITLE("purification of wastewater") OR TITLE("wastewater management") OR TITLE("wastewater purification"))) OR ((KEY("artificial intelligence") OR KEY("deep learning") OR KEY("neural networks") OR KEY("expert systems" OR KEY("cognitive computings") AND (KEY("wastewater treatment") OR KEY("sewage treatment") OR KEY("purification of wastewater") OR KEY("wastewater management") OR KEY("wastewater purification")))

The search strategy was based on the use of standardized descriptors and their synonyms, directly linked to the independent variable (Artificial Intelligence) and the dependent variable (Wastewater Treatment).

These descriptors were combined using the logical operator “OR” (represented by “/”) to broaden the scope of relevant studies retrieved and avoid bias in the identification of literature.

The main terms used were:

- artificial intelligence / neural networks / deep learning / expert systems / cognitive computing
- wastewater treatment / sewage treatment / purification of wastewater / wastewater management / wastewater purification

Table 1 presents the selected information sources along with the search equations used in each of them, formulated based on the previously defined descriptors. These equations enabled the construction of precise and reproducible queries, optimizing the retrieval of relevant studies.

The descriptors were combined through Boolean operators (AND, OR) to ensure semantic coherence between the variables and maximize the relevance of the obtained results.

3.3 Identified Studies

In the initial stage of the search, various studies related to the application of Artificial Intelligence in wastewater treatment were identified, obtained from the selected databases.

The number of studies varied depending on the source, reflecting its coverage and thematic scope. Figure 2 shows the distribution of studies found in each database.

ISSN 2007-9737

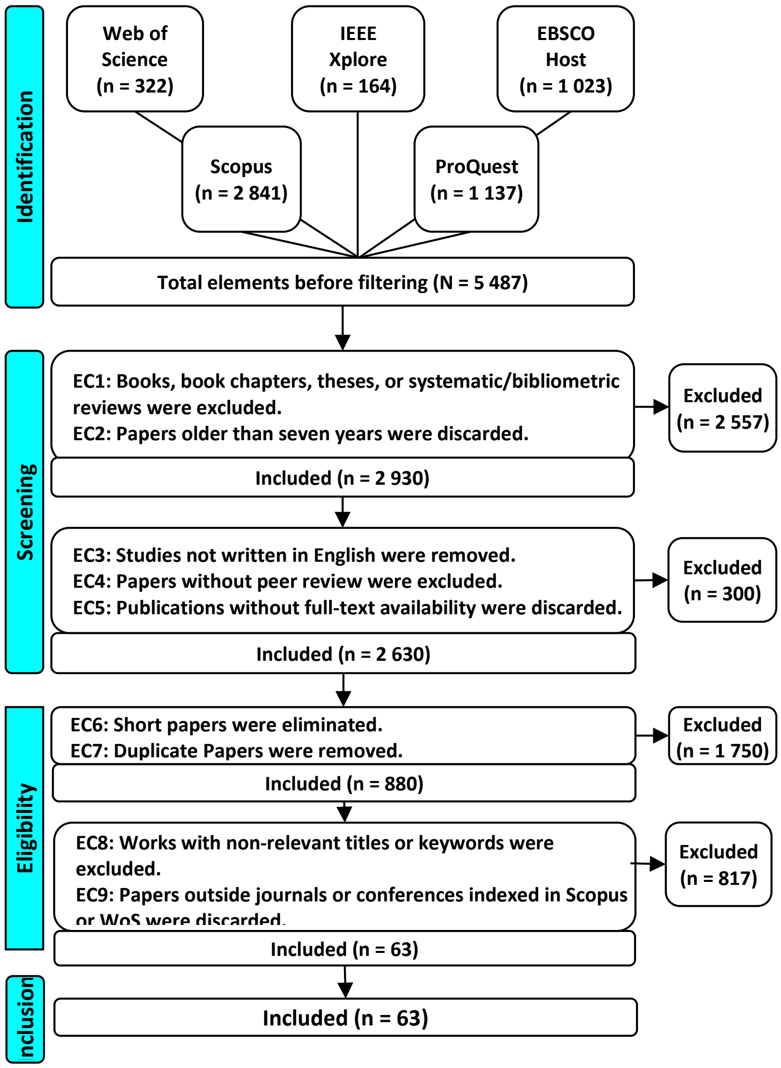


Fig. 3. PRISMA flow diagram

3.4 Study Selection

The study selection was carried out by applying exclusion criteria (EC) designed to ensure the scientific quality, thematic relevance, and methodological validity of the publications included in the review.

These criteria were applied sequentially, allowing the filtering of non-relevant or low-rigor documents. The main criteria considered included: (i) exclusion of books, theses, and previous reviews; (ii) publication date older than seven years; (iii) language other than English; (iv) absence of peer review; (v) lack of access to the full text; (vi) short papers; (vii) duplicate records; (viii) low adequacy of the title or keywords; and (ix) publications not indexed in Scopus or Web of Science.

The selection process was represented through the PRISMA diagram (Preferred Reporting Items for Systematic Reviews and Meta-Analyses), which synthesizes the phases of identification, screening, eligibility, and inclusion. In the final stage, 63 papers that met all established criteria and were deemed relevant for the development of this research were selected (see Figure 3).

ISSN 2007-9737

Table 2. Quality assessment results

Reference	Type	QA1	QA2	QA3	QA4	QA5	QA6	QA7	Score
[1]	Journal	2	2	3	2	2	2	2	15
[2]	Journal	3	2	2	3	2	3	1	16
[3]	Journal	2	3	2	3	1	3	2	16
[4]	Journal	2	3	2	3	2	2	1	15
[5]	Journal	1	2	1	2	3	2	3	14
[6]	Journal	2	3	2	2	1	2	1	13
[7]	Journal	2	2	1	2	3	3	2	15
[8]	Journal	2	2	3	3	1	2	1	14
[9]	Journal	1	2	2	1	2	3	2	13
[10]	Journal	3	3	1	2	2	2	2	15
[11]	Journal	1	3	1	2	3	2	1	13
[12]	Journal	2	2	3	2	1	2	2	14
[13]	Journal	2	2	3	3	2	3	2	17
[14]	Journal	2	3	3	3	2	3	3	19
[15]	Journal	1	2	3	3	2	3	2	16
[16]	Journal	2	2	3	3	2	1	2	15
[17]	Journal	3	3	2	3	3	3	3	20
[18]	Journal	2	2	3	2	2	3	2	16
[19]	Journal	2	2	2	2	3	1	1	13
[20]	Journal	3	2	2	3	2	3	2	17
[21]	Journal	3	2	1	3	3	3	3	18
[22]	Journal	1	3	2	2	3	3	2	16
[23]	Journal	2	2	2	2	1	3	3	15
[24]	Journal	3	2	3	3	2	1	2	16
[25]	Journal	2	3	1	3	2	3	1	15
[26]	Journal	2	3	1	2	3	3	1	15
[27]	Journal	3	1	3	3	1	2	2	15
[28]	Journal	3	2	2	3	1	3	1	15
[29]	Journal	3	2	2	2	2	3	3	17
[30]	Journal	3	3	3	2	3	3	2	19
[31]	Journal	2	2	3	2	3	3	2	17
[32]	Journal	2	3	2	1	2	3	3	16
[33]	Journal	3	1	3	2	3	2	1	15
[34]	Journal	3	2	3	2	3	3	2	18
[35]	Journal	3	2	3	3	1	3	1	16
[36]	Journal	3	3	2	3	1	2	2	16
[37]	Journal	2	3	3	2	3	2	2	17
[38]	Journal	1	3	3	2	3	3	3	18
[39]	Journal	3	2	2	3	1	3	3	17
[40]	Journal	3	2	2	2	2	1	2	14
[41]	Journal	2	3	1	2	2	2	2	14
[42]	Journal	2	3	2	3	3	3	1	17
[43]	Journal	3	2	3	2	1	3	2	16
[44]	Journal	3	3	3	2	3	3	2	19
[45]	Journal	3	3	2	2	3	3	2	18
[46]	Journal	2	3	2	2	3	3	3	18
[47]	Journal	3	2	3	2	3	1	2	16
[48]	Journal	2	3	3	1	2	2	3	16
[49]	Journal	2	2	1	2	2	2	2	13
[50]	Journal	2	3	1	3	3	3	2	17
[51]	Journal	3	3	1	3	2	3	1	16
[52]	Journal	2	2	2	1	2	2	1	12
[53]	Journal	1	2	3	3	2	2	2	15
[54]	Journal	2	1	3	2	2	2	2	14
[55]	Journal	1	2	1	3	3	2	3	15
[56]	Journal	1	2	1	2	3	3	2	14
[57]	Journal	2	3	1	2	3	3	1	15
[58]	Journal	3	2	3	1	2	2	3	16
[59]	Journal	2	2	1	3	1	2	2	13
[60]	Journal	3	2	3	1	2	2	1	14
[61]	Journal	3	2	2	2	3	3	2	17
[62]	Journal	2	3	3	2	3	3	1	17
[63]	Journal	3	2	3	3	3	3	2	19

ISSN 2007-9737

3.6 Quality Assessment

During this phase, the selected papers were evaluated using seven quality assessment criteria (QA), with the purpose of ensuring the clarity, consistency, and methodological validity of the included studies. Each paper was analyzed in detail to determine the level of compliance with the following aspects:

QA1: Are the study objectives clearly stated?
QA2: Does the methodological design coherently address the stated objectives?
QA3: Are the applied techniques accurately explained and justified?
QA4: Are the data collection procedures adequately described?
QA5: Is the purpose of the data analysis clearly stated?
QA6: Are potential limitations to the validity of the results acknowledged and discussed?
QA7: Are the relationships between data, analysis, and interpretation clearly evidenced?

A three-level evaluation scale was applied to each paper (1 = Poor, 2 = Fair, 3 = Good), establishing a minimum inclusion threshold of 11 points. Only studies that exceeded this score were considered for the final analysis. The results of this evaluation are presented in Table 2.

The values show that the 63 evaluated papers exceeded the minimum quality threshold, meeting the methodological and scientific rigor criteria established. This thorough evaluation process ensured the reliability of the selected studies and accurately determined the publications ultimately included in the systematic analysis.

3.6 Data Extraction Strategies

In this phase, once the final collection of 63 papers that passed the quality assessment was consolidated, the relevant information was systematically extracted to address the defined research questions. Data extraction was performed in a structured manner, ensuring traceability and consistency across sources. The elements considered included: paper title, URL, indexing source, year of publication, country of origin, ISSN, publication type, journal name, authors, institutional affiliation, publication quartile, number of citations, methodology used, abstract, and keywords. This strategy enabled the construction of a complete and homogeneous information base, essential for subsequent bibliometric and systematic review analyses.

Outcomes: (1) structured bibliographic metadata per paper (title, source, year, country, ISSN, publication type, journal, authors, and institutional affiliation); (2) scientometric indicators (indexing database, quartile classification, citation count); (3) methodological descriptors (approach, abstract, keywords); and (4) a unified dataset ensuring transparency, traceability, and consistency across sources.

3.7 Synthesis of Findings

During this stage, an exhaustive and precise search was conducted across the papers that could address each of the research questions (RQ1–RQ6). Based on the findings obtained, statistical comparisons were carried out among the results associated with each question, identifying patterns, trends, and significant differences. All selected papers provided a comprehensive and well-founded view of the field of study, forming the analytical foundation for the development of this work.

4 Results and Discussion

This section presents the main results and analyses derived from the systematic review, considering both the theoretical context and the RQs established in the research. The review and detailed examination of the selected papers were carried out manually and rigorously, using unstructured data processing techniques and following the methodological phases defined in the study. This comprehensive process made it possible to identify patterns, trends, and gaps in the literature, which are summarized in Figure 4.

4.1 General Description of the Studies

The review and analysis of the unstructured texts were carried out manually, following a systematic process that ensured traceability at each stage of evaluation and analysis of the selected documents.

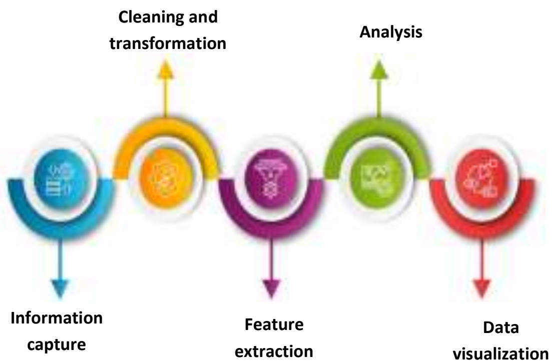


Fig. 4. Paper processing

As a result, a final corpus of 63 scientific papers was compiled and examined in detail to address the established RQs. This set of studies constitutes the empirical basis on which the results and discussion sections presented below were developed.

Figure 5 represents the main associations among the most frequent keywords in the analyzed papers, showing the thematic relationship between artificial intelligence (AI) approaches and wastewater treatment processes. The strongest connections indicate the frequency and strength of co-occurrence between terms related to deep learning models and water quality parameters.

The results reveal a strong association among artificial neural networks, deep learning, and wastewater treatment, reflecting the predominance of these methods in process modeling and the prediction of critical variables. Likewise, the connections with LSTM and soft sensors suggest a growing interest in solutions based on sequential learning and intelligent monitoring. The links with parameters such as biological oxygen demand, total nitrogen, and total phosphorus demonstrate that AI is primarily applied to contaminant control and treatment efficiency optimization. Moreover, the presence of photocatalysis and nitrous oxide indicates progress toward more sustainable approaches focused on emission reduction. Overall, the co-occurrences reveal an interdisciplinary research ecosystem that integrates AI, environmental engineering, and operational sustainability.

The comparative studies provide a convergent perspective on the relevance of key terms in the literature on Artificial Intelligence applied to wastewater treatment. Adeoba and colleagues [71] emphasize that machine learning constitutes the most influential conceptual axis, serving as the semantic core of research in the field. In contrast, Yu and collaborators [76] identified four thematic clusters: machine learning, effluent, models, and prediction, highlighting that the first one is the most recurrent and therefore has the greatest articulating capacity within the co-occurrence network. It is worth noting that Lqbal and their team [97] expand this perspective by demonstrating the strong interrelationship among the terms remote sensing, image processing, and computer vision, which form principal nodes with high thematic connectivity.

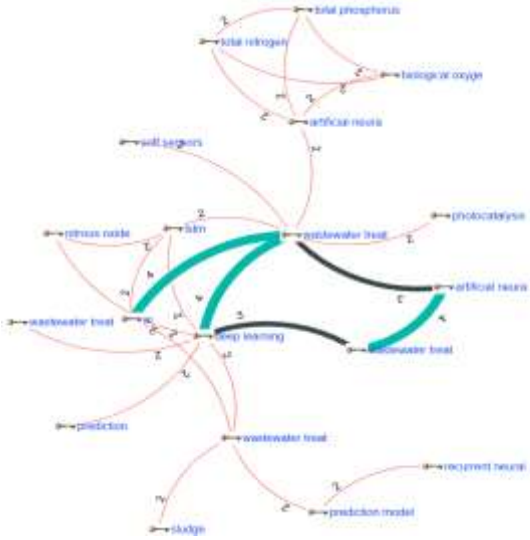
Although the study by Hu and other authors [98] shifts the focus toward climate change, their findings maintain the link with water quality and expert systems, revealing an expansion of the conceptual framework toward global environmental issues.

Finally, Baarimah and collaborators [80] organized three groups of keywords: wastewater treatment, artificial intelligence, and artificial neural network, with Artificial Intelligence being the one with the highest centrality; consequently, it is consolidated as the backbone of the scientific discourse on wastewater management and optimization. Altogether, the analyzed papers display a cohesive thematic structure in which machine learning and artificial intelligence act as conceptual catalysts of the domain.

These findings suggest opportunities to apply AI in other sectors such as agriculture, mining, and energy management, where environmental monitoring is essential. They also open the possibility of implementing predictive models in different geographic areas with variable hydrological conditions. Finally, the identified trend could evolve toward global systems for automated control and real-time prediction of contaminants.

Figure 6 and Table 3 present the global distribution of publications on Artificial Intelligence applied to wastewater treatment, combining heat maps, bar charts, and collaboration networks by continent. This representation allows the identification of countries with the highest scientific output and their impact in terms of citations, H-Index, and quality of contributions.

ISSN 2007-9737



Keyword1	Keyword2	Weight
ai	wastewater treatment	4
artificial neural networks	wastewater treatment plants	4
deep learning	wastewater treatment	4
artificial neural networks	wastewater treatment	3
deep learning	wastewater treatment plants	3
ai	lstm	2
ai	wastewater treatment plant	2
artificial neural network	biological oxygen demand	2
artificial neural network	total nitrogen	2

Fig. 5. Keyword co-occurrence

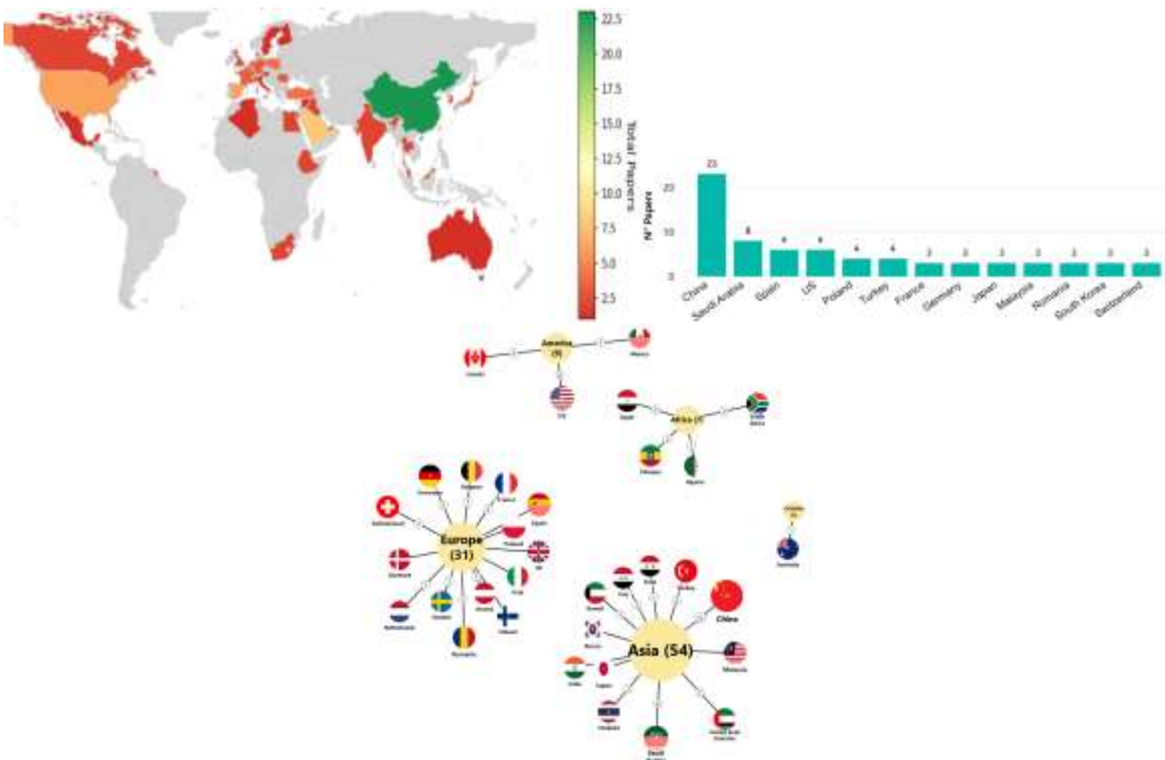


Fig. 6. Distribution of papers by continent and country

China leads production with 23 papers (22.5%), 387 citations (21.2%), and a remarkable H-Index of 37, consolidating itself as the main global reference in AI applied to wastewater treatment.

ISSN 2007-9737

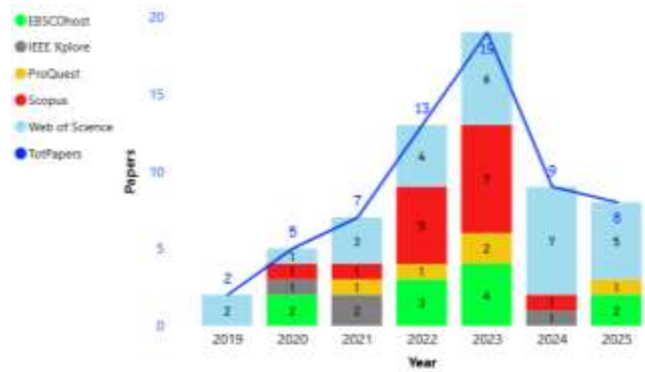


Fig. 7. Distribution of papers by year

Table 3. Impact of countries in research studies

Country	N° Papers	% of Papers	No. of Citations	% of Citations	H-Index	Citations per Paper
China	23	22,5	387	21,2	3717	16,8
Saudi Arabia	8	7,8	179	9,8	1043	22,4
Spain	6	5,9	159	8,7	984	26,5
US	6	5,9	103	5,6	862	17,2
Poland	4	3,9	129	7,1	995	32,3
Turkey	4	3,9	12	0,7	273	3,0
France	3	2,9	109	6,0	544	36,3
Germany	3	2,9	32	1,8	691	10,7
Japan	3	2,9	74	4,1	979	24,7
Korea	3	2,9	58	3,2	616	19,3
Malaysia	3	2,9	20	1,1	267	6,7
Romania	3	2,9	2	0,1	383	0,7
Switzerland	3	2,9	66	3,6	589	22,0
Canada	2	2,0	11	0,6	273	5,5
Egypt	2	2,0	3	0,4	161	1,5
Ethiopia	2	2,0	8	0,4	125	4,0
Total	102	100,0	1827	100,0	16086	17,9

It is followed by Saudi Arabia with 8 publications (7.8%) and 179 citations (9.8%), standing out for its high efficiency with 22.4 citations per paper. Spain and the United States each contribute 6 papers (5.9%), although Spain surpasses the U.S. in impact with 159 citations (8.7%) and 26.5 citations per paper, compared to 103 citations (5.6%) and 17.2 citations per paper in the U.S.

In Europe, Poland (4 papers, 129 citations, 32.3 citations per paper) and France (3 papers, 109 citations, 36.3 citations per paper) show strong research efficiency, while Japan (3 papers, 74 citations, 24.7 citations per paper) and Switzerland (3 papers, 66 citations, 22.0 citations per paper) maintain high impact averages.

In total, the 102 studies gather 1,827 citations, with a general average of 17.9 citations per paper, showing leadership concentrated in Asia (54 papers) and Europe (31), compared to lower participation from the Americas, Africa, and Oceania.

Overall, scientific production is concentrated in countries with strong investment in environmental innovation and technological development.

ISSN 2007-9737

Table 4. Impact of institutions in publications

Affiliation	No. of Papers	No. of Citations	H-Index	Citations per Paper
Universitat Autònoma de Barcelona	4	113	717	28
Chinese Academy of Sciences	3	129	568	43
King Abdullah University of Science and Tech.	3	137	393	46
Tongji University	3	54	369	18
United Arab Emirates University	3	31	241	10
Gdańsk University of Technology	2	32	497	16
Karadeniz Technical University	2	1	48	1
Mahidol University	2	45	330	23
Prince Sattam Bin Abdulaziz University	2	29	362	15
Shanghai Urban Construction Design Research	2	10	246	5
The University of Manchester	2	59	456	30
Umm Al-Qura University	2	3	119	2
University of Malaya	2	15	152	8
University of Wisconsin-Madison	2	45	330	23
Dunarea de Jos' University of Galati	1	1	273	1
Total	156	2522	24693	16

Comparative studies agree that the scientific production on Artificial Intelligence applied to wastewater treatment exhibits broad geographic distribution, although dominated by certain academic poles. Adeoba and colleagues [71] show that all continents have contributed with at least one publication, highlighting the United States and China as the main centers of scientific output. In contrast, Yu and collaborators [76] identify India as the country with the highest number of papers, followed by China, also emphasizing that the latter leads in international collaborations, reinforcing its strategic role in the global research network. Complementarily, Kun Zhou, Boran Wu, and Xin Zhang [77] confirm that China, the United States, and India account for most of the publications on the adoption of AI techniques for water management and treatment. While Lqbal and their team [97] also highlight the United States and China as the countries exceeding 100 publications, they include England, Germany, and India as relevant secondary actors. Meanwhile, Mukonza and Chiang [87] provide a different view by positioning the European Union as the bloc with the highest number of studies, followed by the United States and China. Altogether, the results confirm a shared scientific hegemony among Asia, North America, and Europe, forming a tri-continental axis of leadership in this emerging field.

These findings reflect the concentration of scientific leadership in Asia and Europe, opening opportunities for cooperation with underrepresented regions such as Africa and Latin America. Furthermore, the evidence suggests that the successful models of China, Spain, and France could be replicated in the water, mining, and clean energy industries. Finally, it is projected that the expansion of environmental AI in the coming years will foster a more equitable and sustainable global network for water resource management.

Figure 7 shows the temporal evolution of the publications collected and analyzed from the EBSCOhost, IEEE Xplore, ProQuest, Scopus, and Web of Science databases. This analysis allows identifying trends of growth and consolidation of research during the 2019–2025 period.

A steady growth is observed from 2019 (2 papers) to a peak in 2023 (19 papers), demonstrating the rise of scientific interest in the topic. The upward trend begins in 2020 (5) and

Table 5. Relevant results of the publications

Method Category	Methods used	Research Areas	Performance	Key Contributions	Limitations	Refs.	Qty. (%)
Deep Learning for Prediction & Control	CNN; LSTM; GRU; BiLSTM; Denoising Autoencoder; Seq2Seq; Attention	Wastewater treatment modeling; process prediction & control	Acc=97.44%; RMSE=3.75 mg/ L (BOD); NSE=0.99; MAPE≈1.36%–1.44%	Robust predictors for effluent/ quality variables; real-time/ near-real-time feasibility; reduced manual monitoring	Data scarcity and drift; sensitivity to noise and preprocessing; generalization beyond site-specific settings	[4-6] [13] [25] [44]	6 (9.5)
Reinforcement Learning & Advanced Control	DRL (PPO, DQN); RL (TD3); ANN-IMC; Transfer Learning; Adaptive filtering	Intelligent control; aeration/ energy optimization; sensor/ actuator control	Energy cost −20%; Compliance 95%; IAE↑68.5%; ISE↑80.5% (transfer)	Multi-objective control under variable weather/ loads; sample-efficient transfer; improved robustness to noise	Convergence time and data intensity; risk of negative transfer; simulation-to-reality gap	[9] [23] [32] [36] [40] [41] [46] [47]	8 (12.7)
Soft Sensors & Virtual Sensors	LSTM soft-sensor; GRU+CNN hybrid; Kalman-Elman; ANN	Online estimation (NH4, TN, BOD, COD); process monitoring	Violation detection 86–94%; RMSE better than Elman variants; 37% RMSE improvement vs. baseline	Cost-effective virtual sensing; improved reliability for hard-to-measure analytes	Calibration may raise false positives; seasonal effects; dependence on historical data	[7] [8] [13] [26] [31] [42]	5 (7.9)
Hybrid / Ensemble / Metaheuristics	ANN-PSO; ANN-GA; AHMPSO-LSTM-AM; Coevolutionary SS-VAE; Hybrid mechanistic+DL (CNN/ LSTM); Transformer+meta-classifier; GA-assisted IFFNN; PCA-CNN-LSTM+GA	Performance prediction; planning; noise-tolerant learning; resource optimization	Error ↓ (RMSE/ MAE/ MAPE) 9–28%; Energy/ Materials −10–15%	Hybridization improves accuracy/ robustness vs. single models; uncertainty-aware training	Hyperparameter burden; limited transparency; calibration/uncertainty quantification gaps	[28] [30] [33] [39] [53] [54] [55] [59]	8 (12.7)
Time Series & Forecasting	Distributional ML (GAMLSS); LSTM; Koopman+DNN; RNN; Lifelong LMPNet	Inflow forecasting; multi-step dynamics; quality trajectory prediction	RMSE=15.2 L/ s; MAE=12.5 L/ s; MSE low; R^2=0.95–0.98	Quantified uncertainty; stable intraday forecasts; physics-guided operators	Needs accurate exogenous inputs (rain); complexity hinders deployment	[51] [52] [56-58]	5 (7.9)

2021 (7), reaching a notable increase in 2022 (13) and a peak in 2023, mainly driven by publications in Scopus (6) and Web of Science (4).

In 2024, production decreases to 9 papers, and in 2025, 8 papers are recorded, suggesting a phase of stabilization rather than decline.

The most active databases were Scopus and IEEE Xplore, standing out for their consistency between 2021 and 2024. This pattern indicates maturity in the research line, with methodological consolidation and diversification of approaches.

The analyzed studies reveal a sustained upward trend in scientific production on Artificial Intelligence applied to wastewater treatment, with temporal variations across regions and databases.

Adeoba and colleagues [71] report that the number of publications reached its peak in 2023, with around 2,200 papers, reflecting the consolidation of academic interest in this field. Similarly, Yu and collaborators [76] indicate that growth began in 2018 and reached 254 publications in 2022, confirming a recent acceleration of the field.

In turn, Kun Zhou, Boran Wu, and Xin Zhang [77] specify that, globally, the increase started in 2016, while in China it began a year later, also

Table 5. Continuation

Method Category	Methods used	Research Areas	Performance	Key Contributions	Limitations	Refs.	Qty. (%)
Remote Sensing & Facility Mapping	MANet; GLFMN; Joint Deep Learning (residual attention; fine-tuned detectors)	WWTP facility identification; LULC+WWTP detection	Acc=80.1%; Recall=90.4%; F1=90.4%	Automated mapping/ segmentation of WWTPs for planning & oversight	Scale/shape variability; reliance on high-res imagery and labels	[24] [27]	2 (3.2)
Photocatalysis & Advanced Oxidation Modeling	Fuzzy modeling; Marine Predators Algorithm; RSM+ANN	AOP process optimization; textile/ antibiotic removal	RMSE=0.3148; COD ↓90%; Decolorization ↓88%	Data-driven optimization of AOP parameters; better degradation yields	Small samples; scalability and CAPEX constraints	[18] [22]	2 (3.2)
Sludge, Odor & Biogas Process Modeling	ANN; Decision Trees; Adsorption studies + AI	Odor intensity/ hedonics; eco-adsorbents; biogas yield	MAE=0.57; COD ↓70.5%; TSS ↓65.4%; NR	Soft sensing for odor; green adsorbent effectiveness; biogas optimization	Limited compound coverage; transferability across effluents	[10] [19] [44]	3 (4.8)
Emissions & GHG Modeling	PLO-CNN-BiLSTM-Attention; LSTM	N_2O emissions prediction; GHG-aware control	R^2=0.99; MAE=0.028; RMSE=0.054	High-fidelity N_2O forecasting; identification of key drivers (e.g., temperature)	Limited to specific sites; longer-horizon accuracy needs process inputs	[20] [48] [62]	3 (4.8)
IoT & Spectroscopy Sensing	AOA-SSDAE; Deep-UV LIRFS + CNN	IoT-enabled prediction; micropollutant monitoring	MAE(T-N)=3.11; nitrate assignment >95%	Low-signal detection via spectroscopy+AI; IoT prediction pipelines	Small samples; model interpretability in complex mixtures	[38] [43]	2 (3.2)
Supervised ML (non-deep) for WWTP Performance	RF; SVR; XGBoost; KNN; GEP; Polynomial; ANN (shallow)	Effluent/ quality prediction; sludge output; coagulation	R^2 up to 0.97; RMSE=2.12; MAPE(COD)=16.8%; Turbidity ↓72%	Competitive baselines; interpretable features; rapid deployment	Underfitting complex dynamics; site specificity	[1] [2] [3] [21] [35] [49] [50]	7 (11.1)
Fault & Anomaly Detection	LSTM (fault); N-BEATS autoencoder; SVM	Sensor fault detection; anomaly prediction	Recall=92% (faults); Anomaly acc=98%	First automatic LSTM fault detector; unified prediction+anomaly framework	Requires extensive preprocessing; generalization to extreme noise	[34] [63]	2 (3.2)
Mathematical Modeling & Strategic Perspectives	Mathematical modeling; data mining; ANN optimization	Compliance/ control frameworks; strategic AI deployment	NR	Integration of control + real-time monitoring for regulatory goals	Data access/cost barriers (small WWTPs)	[14] [17]	2 (3.1)
Microalgae & Valorization	ANN; CNN; LSTM; kNN; RF	Biomass/ biorefinery from wastewater; biofuels	MSE=0.0028	Predictive bioresource modeling (18 variables)	Incomplete variables; monitoring challenges	[15]	1 (1.6)
NLP / Chatbots (out-of-domain in corpus)	LSTM; GRU; Seq2Seq; Attention; CoCoSo-AHP-SVNS	Chatbots; response optimization; decision-making	BLEU-4=0.8537; validated via sensitivity analysis	Benchmarking sequential/ NLP models; MCDM under uncertainty	Domain mismatch vs. WWTP; usability not addressed	11] [12] 29] [37] 45] [60] [61]	7 (11.1)

placing the productivity peak in 2023, reaffirming the country's leadership in scientific research on the subject.

Complementarily, Hu and other authors [98] identify a progressive increase starting in 2015, with significant continuity after 2020. Consistently, Mukonza and Chiang [87] highlight a constant evolution of publication volume between 2015 and 2023. Overall, the findings coincide in evidencing exponential growth over the last decade, associated with technological progress, the expansion of AI applications, and the urgent need for sustainable solutions in water management.

The growth between 2020 and 2023 reflects the rapid adoption of AI in environmental projects, which could be replicated in sectors such as energy, agriculture, or waste management. The recent stabilization opens opportunities for longitudinal and comparative studies. In other regions and time periods, these trends may guide investment strategies in research focused on sustainability and intelligent environmental management.

Table 4 presents the performance and impact of the main institutions conducting research on Artificial Intelligence applied to wastewater treatment, considering the number of publications, citations, H-Index, and average citations per paper. This analysis makes it possible to identify the institutions with the greatest academic and scientific leadership at the global level.

The Universitat Autònoma de Barcelona leads in productivity with 4 publications, 113 citations, and an H-Index of 717, reflecting a high degree of influence and scientific consolidation. It is followed by the Chinese Academy of Sciences, with 3 papers, 129 citations, and an outstanding average of 43 citations per paper, evidencing research excellence.

The King Abdullah University of Science and Technology shows the highest average performance, with 46 citations per paper supported by a total of 137 citations. Meanwhile, Tongji University and the United Arab Emirates University achieve 18 and 10 citations per paper, respectively, consolidating Asia's role in this line of research. In Europe, institutions such as The University of Manchester (30 citations per paper) and Gdańsk University of Technology (16 citations per paper) demonstrate significant contributions. In total, the 156 publications gather 2,522 citations, with a global average of 16 citations per paper, reflecting a field of growing maturity and international collaboration.

The institutional leadership concentrated in Europe and Asia highlights the potential for cooperation with Latin American and African universities to reduce the scientific gap. This collaborative model could extend to sectors such as energy, urban sanitation, and industrial waste management. Likewise, the accumulated experience in environmental AI could be adapted to new geographies and productive contexts, fostering sustainable innovation on a global scale.

Table 5 synthesizes the methodological approaches identified in the corpus, grouping the research according to application categories, techniques used, achieved performance, theoretical and practical contributions, and main limitations. This structure makes it possible to demonstrate the evolution and maturity of Artificial Intelligence in the management and optimization of wastewater treatment processes.

Method category. The distribution reveals the predominance of *Reinforcement Learning & Advanced Control* (12.7%), *Hybrid/Ensemble/Metaheuristics* (12.7%), *Supervised ML* (11.1%), and *Deep Learning for Prediction & Control* (9.5%), which together account for 46% of all studies. These categories reflect the consolidation of AI as a tool for optimization and prediction in complex processes. In contrast, emerging approaches such as *Remote Sensing* (3.2%), *Photocatalysis* (3.2%), and *Mathematical Modeling* (3.1%) show lower representation, suggesting research lines still under development but with high potential for interdisciplinary expansion.

Methods used. The most recurrent methods, such as CNN, LSTM, GRU, BiLSTM, and Attention, applied in 9.5% of the studies, stand out for their predictive capability regarding critical variables (BOD, TN, COD). Hybrid and metaheuristic strategies, representing 12.7%, confirm the pursuit of robustness through combinations such as *ANN-PSO*, *ANN-GA*, and *Transformer + meta-classifier*. About 11.1% of the papers employ classical algorithms like *RF*, *SVR*, *KNN*, or *XGBoost*, validating their usefulness in control environments and rapid diagnostics. The coexistence of deep, hybrid, and supervised techniques demonstrates a mature methodological convergence oriented toward computational efficiency and interpretability.

Research areas. The most represented research areas correspond to process modeling, prediction, and control (≈40%), driven by categories such as *Deep Learning* (9.5%) and *Reinforcement Learning* (12.7%). They are followed by studies on energy optimization and virtual sensing (≈20%), including *Soft Sensors* (7.9%) and *IoT & Spectroscopy* (3.2%). Research on emissions and environmental sustainability (4.8%), together with microalgae valorization (1.6%), expands the frontier toward a circular economy. Altogether, this thematic diversity reflects a balanced ecosystem between intelligent control, operational efficiency, and environmental sustainability.

Performance. The results show high average accuracy, with R^2 values between 0.95 and 0.99, RMSE up to 3.75 mg/L (BOD), and energy reductions of 20%. The categories with the best metrics, *Deep Learning* (9.5%), *Hybrid/Metaheuristics* (12.7%), and *Reinforcement Learning* (12.7%), achieve accuracies above 97% and error reductions (RMSE/MAE/MAPE) of up to 28%. The *Soft Sensors* (7.9%) and *Time Series Forecasting* (7.9%) lines complement this robustness by offering stable and reliable predictions. Overall, the quantitative performance confirms the maturity and effectiveness of AI models applied to wastewater treatment.

Key contributions. The main contributions come from the most represented approaches: *Reinforcement Learning* (12.7%), *Hybrid/Ensemble* (12.7%), and *Supervised ML* (11.1%), which together provide more than 35% of technical innovations. These include energy control automation (−20%), real-time prediction, and anomaly detection with 92% recall. Likewise, *Soft Sensor* studies (7.9%) introduce low-cost virtual sensors for variables such as NH_4, TN, BOD, and COD, while *Remote Sensing* (3.2%) and *IoT* (3.2%) strengthen distributed supervision. Overall, these contributions reflect an evolution toward autonomous and sustainable cyber-physical systems.

Limitations. The most frequently reported limitations are associated with data scarcity (12.7%), generalization issues (11.1%), and simulation–reality gaps (9.5%). Deep learning–based models exhibit high sensitivity to noise and preprocessing, while hybrid models face high computational loads and calibration difficulties. In lower-weight categories such as *Photocatalysis* (3.2%) or *Mathematical Modeling* (3.1%), scalability and sample limitation challenges predominate, restricting experimental validation. This combination of weaknesses underscores the need to standardize metrics and strengthen interoperability among models.

The analysis reveals significant technical consolidation, with nearly 50% of studies focused on deep and hybrid learning, driving a new generation of intelligent plants. Their application can extend to sectors such as mining, agriculture, energy, and public health, optimizing resources and mitigating pollution.

Geographically, the developed models can be transferred to countries with limited water infrastructure, strengthening sustainable management.

At the enterprise level, AI offers opportunities to design predictive monitoring and adaptive control systems. Finally, the global standardization of data and metrics emerges as a key requirement for the future expansion of this technology.

4.2 Research Question Responses

This section presents the answers to the research questions formulated within the framework of the systematic literature review, integrating the main findings, discussions, and projections toward future lines of research.

A rigorous discussion of the results obtained for each RQ is also conducted, contrasting the evidence and trends reported in the analyzed studies.

The identification and selection of relevant information were carried out using publications indexed in specialized journals and conferences, ensuring thematic relevance, methodological quality, and scientific robustness of the corpus considered.

RQ1: What criteria are used to evaluate the effectiveness of Artificial Intelligence?

Table 6 and Figure 8 present the main criteria used to evaluate the effectiveness of artificial intelligence models in wastewater treatment.

The colors in the chart reflect usage frequency: green tones represent the most frequently used metrics, while red and orange tones indicate lower application.

This visualization helps illustrate the methodological preference for error indicators and predictive performance metrics.

ISSN 2007-9737

Table 6. Criteria used in the studies

Criterion	Reference	Qty. (%)
RMSE (Root Mean Squared Error)	[1-8] [11] [13] [15] [16] [18] [20] [25] [26] [28] [30] [31] [34] [35] [41] [45] [48] [49] [51] [53] [55] [57] [58] [62] [63]	31 (35)
MAE (Mean Absolute Error)	[6] [8] [10] [11] [13] [16] [20] [25] [30] [47] [49] [51] [53] [54] [56-58] [62] [63]	19 (22)
MAPE (Mean Absolute Percentage Error)	[5] [6] [10] [11] [13] [16] [20] [28] [30] [42] [47] [49] [54] [56]	14 (16)
SMAPE (Symmetric Mean Absolute Percentage Error)	[11]	1 (1)
R-SQUARED (R^2)	[16] [18] [21] [57] [62]	5 (6)
MSE (Mean Squared Error)	[6] [8] [13] [15] [16] [21] [32] [36] [37] [41] [46] [48] [49] [51] [54] [57] [63]	17 (19)
MSLE (Mean Squared Logarithmic Error)	[16]	1 (1)

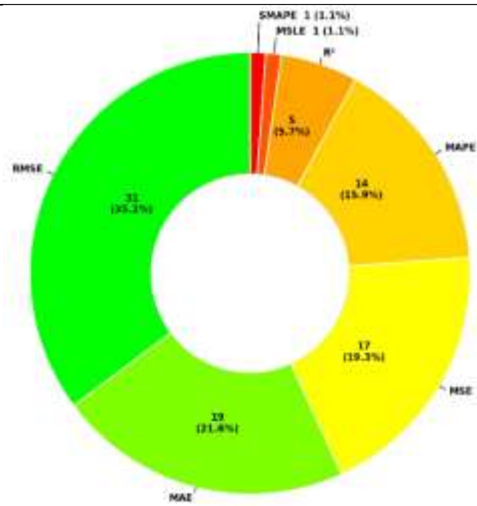


Fig. 8. Distribution of criteria in the studies

The results show a clear predominance of RMSE (35.2%), followed by MAE (21.6%) and MSE (19.3%), confirming a trend toward the use of metrics based on the magnitude of absolute and squared errors. MAPE (15.9%) maintains a relevant presence, especially in studies that prioritize the percentage interpretability of error. In contrast, metrics such as R^2 (5.7%), SMAPE (1.1%), and MSLE (1.1%) exhibit limited adoption, restricted to specific experimental contexts.

The preference for direct error indicators reflects the need for precise and quantifiable models in industrial environments, where small deviations can lead to high operational costs. Altogether, these metrics establish a robust framework for validating the effectiveness of AI systems in real contexts of prediction and control.

Comparative studies agree that the most used metrics for evaluating the effectiveness of Artificial Intelligence focus on statistical indicators of error and correlation, although they differ in frequency and methodological emphasis. Sharma and collaborators [75] highlight that R-squared, RAE, and RMSE constitute the predominant criteria, covering 48% of the analyzed papers, in contrast with MAE, correlation coefficient, and relative root square error (RRSE), which account for only 18%. Conversely, Lima and colleagues [74] confirm the

ISSN 2007-9737

Table 7. Learning models in the studies

Model	Reference	Qty. (%)
LSTM (Long Short-Term Memory)	[2] [8] [11] [13] [15] [16] [20] [26] [30] [32-36] [39-42] [48-49] [51] [53-57] [60] [62-63]	29 (24)
GRU (Gated Recurrent Unit)	[8] [11] [13] [16] [30] [54]	7 (6)
BILSTM (Bidirectional LSTM)	[13]	1 (1)
ANN (Artificial Neural Network)	[1-8] [10] [11] [13] [15-17] [19] [21] [22] [28] [29] [32] [33] [35-37] [40] [41] [42] [44] [45] [46] [48] [49] [50] [51] [52] [54] [56] [57] [58] [59] [61] [62]	41 (34)
CNN (Convolutional Neural Network)	[1] [4] [8] [15] [16] [17] [20] [21] [24] [25] [27] [39] [41] [43] [53] [55] [60] [63]	18 (15)
RNN (Recurrent Neural Network)	[1] [8] [13] [16] [21] [26] [30] [31] [33] [34] [36] [40] [41] [42] [48] [53] [54] [55] [57] [58] [59] [60] [62] [63]	24 (20)

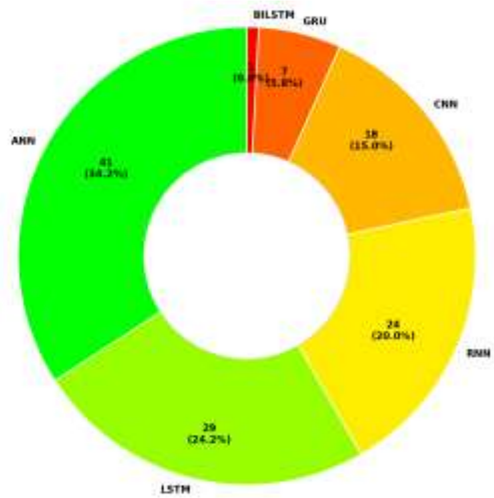


Fig. 9. Distribution of learning models

predominance of RMSE, followed by MSE and MAE among nine identified metrics, reinforcing their usefulness for validating predictive models. Complementarily, Tosan and their team [82] emphasize that RMSE, MAE, and R^2 are the most recurrent measures in the reviewed studies, demonstrating a preference for metrics based on mean squared error and linear correlation. Although Baruah and collaborators [85] report a different orientation by prioritizing ACC, F1-Score, and Precision as performance indicators, they relegate MSE and R^2 to a marginal presence, with only one publication each. Meanwhile, Olu-Ajayi and other authors [83] reaffirm the prevalence of RMSE, used eight times, followed by MSE and MAPE. Overall, the findings suggest that RMSE is the most widely adopted metric for estimating model accuracy, reinforcing its value as a comparative standard in evaluating the performance of artificial intelligence algorithms.

The predominance of metrics such as RMSE and MAE can be extrapolated to other sectors, such as energy, manufacturing, or environmental management, where predictive accuracy is crucial for automated decision-making. Future studies could compare the stability of these metrics across different geographic areas and operational scales. Likewise, incorporating computational efficiency and sustainability indicators is recommended to strengthen the comprehensive evaluation of AI in new temporal and industrial contexts.

RQ2: What models are currently applied in the development of Artificial Intelligence?

Table 7 and Figure 9 present the main learning models used in the development of artificial

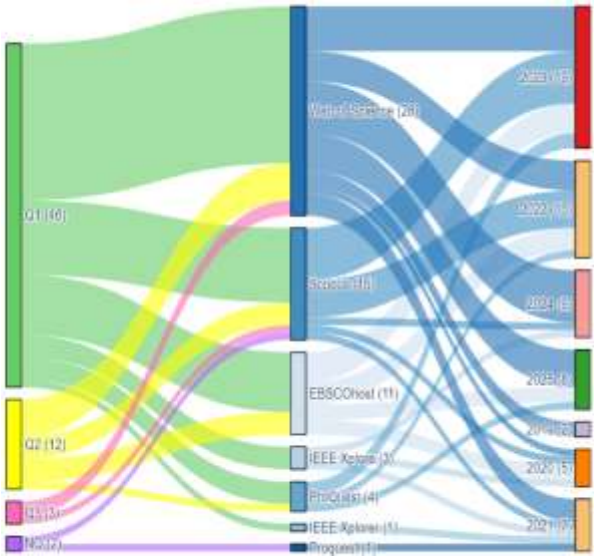


Fig. 10. Papers by quartile and publication source

Table 8. Impact of quartiles in the studies

Quartile	No. of Papers	No. of Citations	Citations per Paper	H-Index
Q1	46	1033	22	9197
Q2	12	196	16	1115
Q3	3	4	1	210
NQ	2	7	4	144
Total	63	1240	20	10666

intelligence applied to wastewater treatment. The colors in the pie chart reflect the relative frequency of use: green tones indicate the most frequently applied models, while reddish tones represent lower representation within the analyzed studies.

The results show a clear predominance of artificial neural networks (ANN) with 34.2%, followed by LSTM models with 24.2%, which stand out for their ability to process sequential data and capture temporal dependencies.

In third place, RNN models (20%) remain relevant as baseline structures for dynamic predictions, while CNN models (15%) are used in classification and spatial analysis tasks. In contrast, GRU (5.8%) and BiLSTM (0.8%) models show low adoption, suggesting that their advantages in efficiency and bidirectionality are not yet widely exploited in this domain.

Overall, the distribution indicates a predominant focus on deep recurrent and hierarchical architectures that facilitate the prediction of complex variables in wastewater treatment environments.

The analyzed studies reveal a clear convergence in the use of machine learning models for the development of Artificial Intelligence, although with variations in frequency and focus. Abdahlla and collaborators [89] highlight the predominance of ANN, GA, and LR models, emphasizing the robustness of the first due to its error tolerance and capacity to model nonlinear relationships in multivariate systems. Conversely, Ajayi and their team [84] identify SVM as the most frequently used model, with a 23% incidence, followed by ANN (19%) and RF (13%), while DNN is the least common, with only 3%. Complementarily, Amaro and colleagues [86] confirm the predominance of SVM, together with CNN and ANN, reinforcing the consistency of these algorithms across diverse contexts.

Likewise, Mukonza and Chiang [87] group the most frequently used models, SVM, neural networks, and their variants RNN, CNN, and LSTM, reflecting the evolution toward deeper and more specialized architectures. Meanwhile, Nazirun and other authors [88] position Random Forest as the most applied model, followed by SVM, consolidating the preference for hybrid and ensemble methods.

Additionally, León-Pérez and colleagues [106] highlight the predominance of DNN and Naive Bayes, valued for their high classification accuracy. It is worth noting that this trend aligns with findings in other domains: Gamboa-Cruzado and collaborators [81] and Cárdenas-Quispe and colleagues [70] report similar usage frequencies of ANN, RF, and decision trees, demonstrating the versatility and stability of these models in different areas of artificial intelligence.

Overall, the results confirm that ANN, SVM, and RF constitute a consolidated and adaptable methodological core capable of offering an optimal balance between accuracy, interpretability, and computational efficiency.

The widespread application of ANN and LSTM can be extended to other sectors such as energy management, smart agriculture, or environmental health, where temporal patterns are crucial. Future studies could evaluate the adaptation of GRU and BiLSTM models in geographic contexts with climatic variability or in decentralized urban infrastructures. Moreover, the combined use of hybrid neural networks could optimize computational efficiency and enhance predictive capacity across different time horizons.

RQ3: What quartile levels are presented by the journals that have published research on the impact of Artificial Intelligence in wastewater treatment?

Figure 10 and Table 8 show the distribution of studies according to the quartile level of the journals, indexing databases, and years of publication. The Sankey diagram illustrates the flow among these three dimensions, allowing the visualization of the concentration of high-impact studies in recognized sources and their temporal evolution. This representation helps identify both the quality and the recency of the analyzed scientific production.

The results indicate a clear predominance of publications in Q1 journals (46 papers; 73.0%), which account for 1,033 citations and an H-Index of 91, reflecting the high quality and visibility of the studies. In second place, Q2 journals (12 papers; 19.0%) maintain a remarkable performance with 196 citations and an H-Index of 11, reaffirming their significant contribution. In contrast, Q3 journals (3 papers; 4.8%) and non-indexed journals (NQ, 2 papers; 3.2%) show limited representation, evidencing the preference for high-impact publications. Temporally, 2023 (19 papers) and 2022 (13 papers) mark productivity peaks, while Web of Science (28) and Scopus (15) are the most used databases. This concentration in Q1–Q2 journals suggests sustained scientific consolidation in the field.

The reviewed results show a clear trend toward publishing research on the impact of Artificial Intelligence in wastewater treatment in high-impact journals, mainly in the Q1 quartile, with slight variations depending on the databases used. Kun Zhou, Boran Wu, and Xin Zhang [77] found that seven of the ten journals with the highest number of publications belong to Q1, reflecting the high quality and academic recognition of studies in this area. Similarly, De La Cruz and collaborators [90] report that among articles indexed in WoS and Scopus, nine correspond to Q1, one to Q2, and only one lacks a quartile classification, confirming the preference for journals of greater prestige. Conversely, Rosário and Dias [95] identify that among fourteen analyzed articles, nine are in Q1, three in S1, and one in Science Citation (SC), reinforcing the predominance of the highest indexing levels. Meanwhile, Nikhar and colleagues [93] indicate that 27% of the papers were published in Q1 journals according to WoS and 24% according to Scopus, evidencing coherence between both sources. The results suggest that most relevant works in this line are disseminated in Q1 journals, denoting the consolidation of the topic within elite scientific research and its growing recognition in international academic communities.

The strong trend toward Q1 and Q2 journals reinforces the maturity of the field and its international recognition, laying the groundwork for expansion into related areas such as clean energy, sustainability, and environmental management. Future studies could explore the evolution of these patterns in emerging regions or indexing databases. Furthermore, the temporal projection highlights opportunities to expand institutional collaboration and strengthen intercontinental networks of applied research.

RQ4: What definitions and theoretical frameworks have been established in studies related to Artificial Intelligence and wastewater treatment?

Table 9 summarizes the most frequently used definitions in the analyzed studies on Artificial Intelligence (AI) and its impact on wastewater treatment. These definitions are grouped into three main categories: functional, technical, and applied, which reflect different levels of abstraction and usage approaches. Their analysis provides insight into how the conceptualization of AI has evolved within the environmental and water engineering fields.

The functional category is the most representative, with 36 studies (54%), focusing on AI as a discipline that emulates human cognitive processes such as learning, problem-solving, and decision-making. It is followed by the technical definition, present in 20 studies (34%), which conceives AI from its algorithmic dimension, emphasizing the use of neural networks and

ISSN 2007-9737

Table 9. Definitions used in the studies

Category	Definition used	Reference	Qty. (%)
Functional	Development of systems and algorithms that simulate human cognitive processes such as learning, problem-solving, and decision-making.	[1-4] [9-13] [18] [23] [26] [30] [34–37] [42] [43] [44] [46-51] [53] [56] [57] [58] [59] [61]	36 (54)
Technical	Use of computational models, neural networks, and machine learning techniques to simulate intelligent behavior and process complex data.	[5] [7] [8] [14] [16] [17] [21] [24] [25] [28] [29] [31] [32] [39] [40] [41] [52] [54] [55] [63]	20 (34)
Applied	Application of Artificial Intelligence to optimize environmental, industrial, and wastewater treatment processes through predictive modeling and control.	[15] [20] [33] [39] [52] [54] [62]	7 (12)

Table 10. Theoretical foundations in the studies

Category	Synthesized theoretical basis	Reference	Qty.(%)
Paradigms	Artificial Intelligence, Simulation of Human Cognitive Processes, Computational Intelligence, Theory of Computation, Mathematical Logic	[1] [2] [5] [7] [9] [11] [13] [14] [15] [16] [17] [20] [21] [25] [27] [28] [30] [34] [37] [38] [43] [44] [49] [50] [53] [54] [57] [58] [59] [61] [62] [63]	32 (58)
Technical	Artificial Neural Networks (ANNs), Recurrent Neural Networks, LSTM, Deep Learning, Model Predictive Control (MPC), Metaheuristic Algorithms	[8] [17] [27] [29] [31] [32] [35] [36] [39] [40] [41] [42] [52] [59] [63]	15 (27)
Applied areas	Modeling and Optimization of Environmental and Industrial Processes, Wastewater Treatment, Emission Prediction, System Monitoring, Anomaly Detection, Biological Process Control	[3] [4] [10] [12] [18] [23] [26] [47] [48] [56]	20 (36)
Indeterminate	Not determined in the provided context	[3] [4] [10] [12] [18] [23] [26] [47] [48] [56]	10 (18)

machine learning models to process large volumes of data. To a lesser extent, the applied definition appears in 7 studies (12%), highlighting the use of AI to optimize environmental and industrial processes through predictive control and efficient monitoring.

This distribution reveals a progressive transition from the conceptual to the operational, evidencing the interdisciplinary maturity reached by the field.

This paper positions itself as a pioneering study in reviewing the definitions related to Artificial Intelligence and its impact on wastewater treatment.

As one of the first studies to address this type of systematic analysis, no sufficient prior works were found for direct comparison. This reflects the need for further exploration of the topic to consolidate a common and robust theoretical framework for future research.

The predominance of the functional approach suggests a strong theoretical foundation that can be transferred to other sectors such as energy, agriculture, or environmental health, where automated decision-making is key. Internationally, the growing adoption of the technical approach will enable performance comparisons across contexts with different levels of digital infrastructure. Finally, the expansion of applied definitions could enhance technology transfer and the formulation of sustainable policies in future scenarios.

Table 10 and Figure 11 present the theoretical foundations identified in studies on Artificial Intelligence applied to wastewater treatment, grouped into four main categories: paradigms,

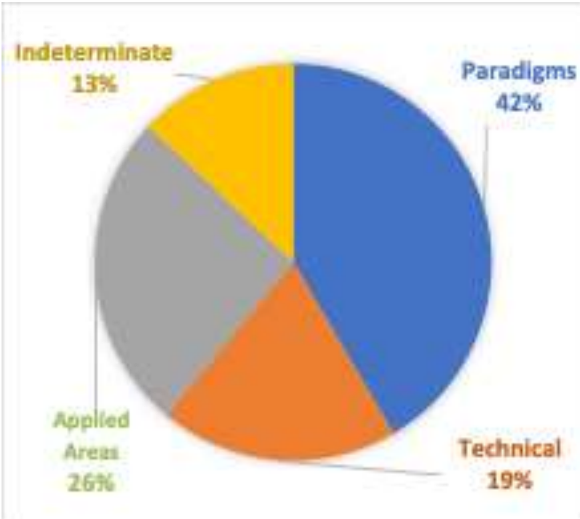


Fig. 11. Percentage of definition categories

technical, applied areas, and indeterminate. This analysis highlights the conceptual and methodological bases that support research and guide the development of predictive models, control systems, and optimization methodologies.

The paradigm category dominates with 32 studies (58%), highlighting theories such as computational intelligence, the simulation of human cognitive processes, and mathematical logic, which serve as conceptual pillars of AI. It is followed by the technical category, with 15 studies (27%), focused on neural architectures such as ANN, RNN, and LSTM, as well as predictive control algorithms and metaheuristics applied to complex data processing. In third place, the applied areas encompass 20 studies (36%), centered on the modeling, monitoring, and control of environmental and industrial processes, reinforcing the connection between theory and practical implementation. Finally, the indeterminate group (18%) suggests the absence of an explicit theoretical framework in some studies, reflecting the empirical orientation of certain recent works. Overall, the evidence demonstrates theoretical and technical maturity in the field, though with remaining opportunities for conceptual standardization.

This paper stands as one of the first efforts to systematize the foundational elements that explain the application of Artificial Intelligence in wastewater treatment. The scarcity of consolidated literature in this area has limited theoretical comparisons with prior research, emphasizing the need to develop stronger and more coherent conceptual frameworks in future studies.

The predominance of theoretical paradigms reveals a solid basis for expanding AI into sectors such as energy, mining, or industrial waste management. Strengthening theoretical frameworks in technical studies could enhance result comparability across different geographic areas and operational contexts. Likewise, the development of integrated theoretical foundations will optimize the use of AI in future scenarios, promoting sustainability and evidence-based decision-making.

RQ5: Which countries most frequently show co-occurrence relationships in research on Artificial Intelligence and its influence on wastewater treatment?

Figure 12 presents the bibliometric network and flow of co-occurrence among countries collaborating in research on Artificial Intelligence applied to wastewater treatment. The nodes and connecting lines indicate the frequency and direction of scientific cooperation, reflecting the international structure of knowledge networks within this emerging field.

To construct the bibliometric co-occurrence network of countries, the association strength between nations was computed to identify collaborative patterns in research output related to artificial intelligence applied to wastewater treatment. In this case, the cosine similarity measure was used to quantify the degree of co-occurrence between countries, based on the frequency with which they appeared jointly in the author affiliations of the analyzed papers. This metric evaluates the closeness between country vectors in a multidimensional space, where higher similarity values indicate stronger research collaboration. The cosine measure between two countries c_i and c_j is obtained using the following equation:

$$\cos(c_i, c_j) = \frac{\sum_{k=1}^{m}(c_{ik} * c_{jk})}{\sqrt{(\sum_{k=1}^{m} c_{ik}^2) * (\sum_{k=1}^{m} c_{jk}^2)}}. \quad (1)$$

ISSN 2007-9737

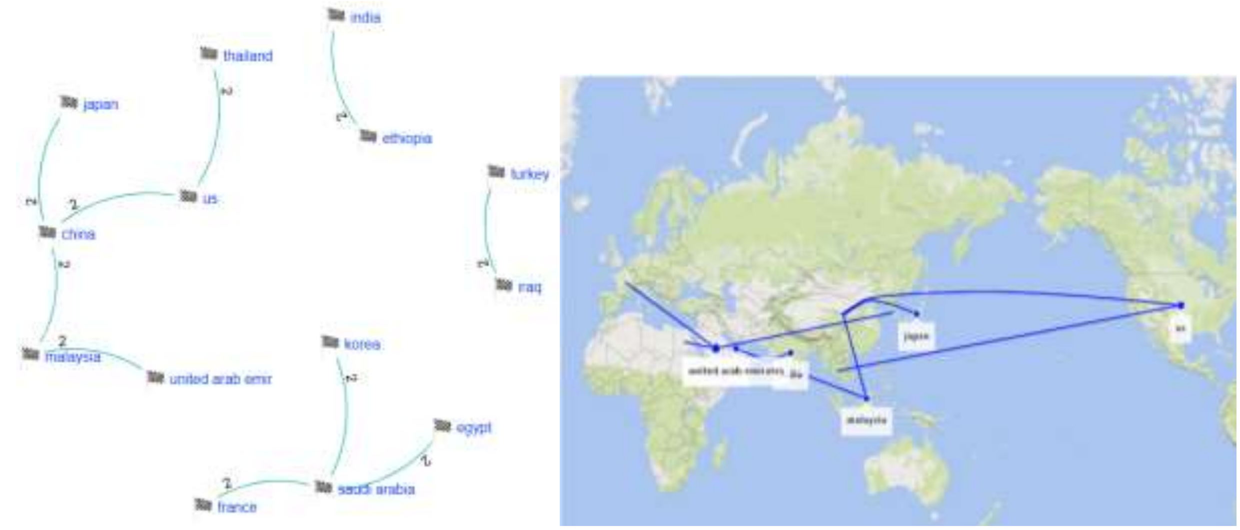


Fig. 12. Collaboration map by countries

Where c_{ik} represents the weight of the co-authorship or publication count feature *k* associated with country c_i.

A notable concentration of collaborations is observed among China, the United States, and Malaysia, which act as central hubs of scientific exchange. The connections between Japan–U.S., China–Japan, and Malaysia–United Arab Emirates stand out for their frequency (weight 2), evidencing consolidated technological alliances across Asia and North America.

Likewise, the links India–Ethiopia and Saudi Arabia–Egypt reflect growing interest in applying AI to water management and arid-region contexts. Europe participates through connections such as France–Korea and Turkey–Iraq, strengthening intercontinental cooperation.

The global map reveals that the flow of knowledge is concentrated along the Asia–America axis, with extensions toward Africa and the Middle East, suggesting a distributed scientific network with highly productive centers.

The analyzed studies reveal a marked concentration of scientific production and international collaboration networks around a small group of countries, with China and the United States as the main poles of research in Artificial Intelligence applied to wastewater treatment. Kun Zhou, Boran Wu, and Xin Zhang [77] highlight China as the most prominent country, further emphasizing Asia's leadership in this field. Conversely, Yu and collaborators [76] position the United States as the country with the highest contribution, followed by Spain, China, Iran, and Italy, revealing a more balanced distribution across regions. In line with this, De la Hoz and colleagues [78] confirm the joint hegemony of China and the United States as the primary drivers of research in this area.

Similarly, Altowayti and collaborators [79] emphasize the relevance of the U.S., Spain, and China as the countries maintaining the strongest co-occurrence links, reflecting active cooperation networks. Finally, Bayhan and colleagues [107] underline that the United States represents the center with the most international collaborations, especially with India, England, Australia, and China, the latter ranking second in scientific interaction. Altogether, the findings confirm the consolidation of the United States and China as strategic axes of global research, around which European and Asian countries gravitate, evidencing a dynamic and expanding international ecosystem in the application of Artificial Intelligence to wastewater treatment.

The co-occurrence network reveals the potential to expand collaborations toward Latin American and African regions, where water-related challenges are similar. The integration of these countries could foster technological transfer and

ISSN 2007-9737

the development of local AI-based solutions. Moreover, the strengthening of intercontinental cooperation could extend to other sectors such as renewable energy and environmental management, consolidating a global network oriented toward sustainability.

RQ6: What thematic categories are identified in research addressing Artificial Intelligence and its influence on wastewater treatment?

Figure 13 and Table 11 present the thematic map of studies on Artificial Intelligence applied to wastewater treatment, constructed from the analyzed keywords. Each theme is characterized by two parameters: centrality, which reflects its relevance within the field (horizontal axis), and density, which measures its level of development (vertical axis). This analysis enables the identification of motor, basic, specialized, and marginal themes within the investigated domain.

To construct the thematic map, the Callon centrality–density method was applied to identify the structural and developmental relevance of research themes related to Artificial Intelligence applied to wastewater treatment. This approach measures the degree of internal cohesion (density) and external connectivity (centrality) of each cluster, thus determining their position within the thematic network.

The density (D) of a theme quantifies its internal development based on the strength of links among the keywords that compose it, while centrality (C) reflects its interaction with other clusters in the network. These two metrics are computed using the following equations:

$$D_i = 100 \times \frac{\sum e_{jk}}{n_i(n_i - 1)/2}, \quad C_i = 10 \times \sum e_{ij}. \tag{2}$$

Where e_{jk} represents the link strength between keywords *j* and *k* within cluster *i*, n_i is the number of keywords in cluster *i*, and e_{ij} indicates the link strength between cluster *i* and any other cluster *j*.

The motor themes, characterized by high density and centrality, include Nutrient Load (0.98; 0.98), Neural Prediction (0.89; 0.98), and Neural Wastewater (0.58; 0.80), which stand out for their theoretical consolidation and high impact (306 citations across 15 documents). These represent mature research lines linked to neural prediction and nutrient load modeling, key pillars in wastewater treatment optimization.

In contrast, marginal themes such as AI Wastewater (0.30; 0.35) and Intelligent Wastewater (0.21; 0.23) exhibit low density and centrality, indicating emerging or still exploratory areas. Concepts like Deep Learning Wastewater and Neural-Based Wastewater (0.05–0.27) suggest incipient trends toward the integration of deep AI and intelligent monitoring.

Overall, the map evidences an evolution from consolidated neural approaches toward new applications in intelligent control and monitoring.

The comparative findings show a diversity of thematic categories reflecting the maturity and expansion of the field of Artificial Intelligence applied to wastewater treatment, with a clear orientation toward machine learning and intelligent water management.

Bayhan and collaborators [107] highlight research lines such as groundwater level, geostatistics, data-driven models, and SVM, emphasizing the growing adoption of machine learning approaches for predictive analysis and environmental system modeling.

Conversely, Flores-Iwasaki and colleagues [72] identify water quality, IoT, and aquaculture as the main axes of research, along with emerging topics such as sensors, water parameters, and freshwater, which, though less prominent at present, represent areas with high development potential.

Likewise, García and collaborators [73] reveal thematic linkages among water treatment, pollution, and groundwater resources, highlighting the use of algorithms such as decision trees and artificial neural networks, while identifying sensors as a line still under consolidation.

These findings confirm that research has evolved from traditional monitoring and quality control themes toward more integrated and technological approaches, where the synergy among IoT, machine learning, and deep learning defines a new paradigm for sustainability and optimization of water treatment systems.

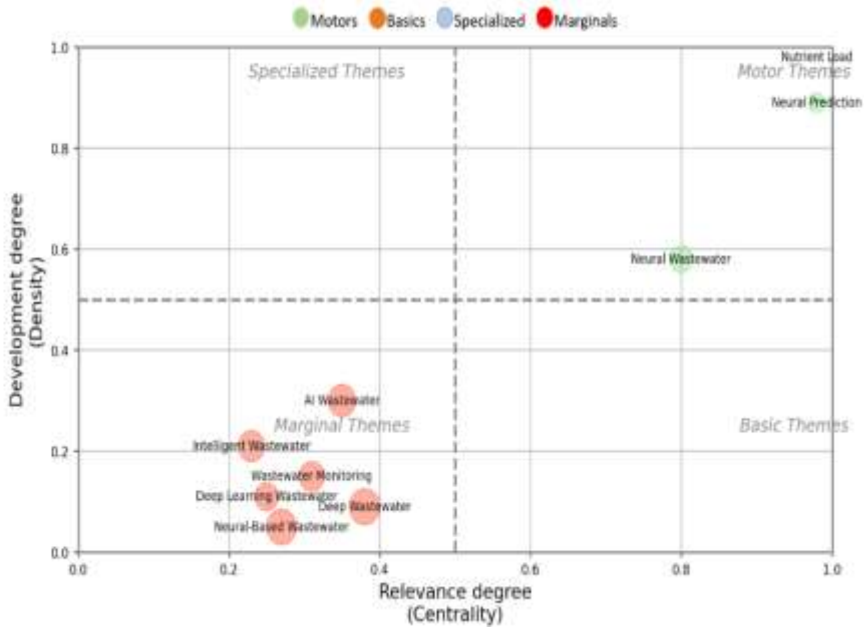


Figure 13. Topics identified in the papers.

Table 11. Centrality and density of theme categories.

Topic	Density	Centrality	No. Of Citations	No. Of Documents	Category
Nutrient Load	0,98	0,98	3	6	Motor
Neural Prediction	0,89	0,98	99	10	Motor
Neural Wastewater	0,58	0,80	306	15	Motor
AI wastewater	0,30	0,35	434	19	Marginal
Intelligent Wastewater	0,21	0,23	395	23	Marginal
Wastewater Monitoring	0,15	0,31	371	22	Marginal
Deep Learning Wastewater	0,11	0,25	317	18	Marginal
Deep Wastewater	0,09	0,38	537	31	Marginal
Neural-Based Wastewater	0,05	0,27	517	23	Marginal

The predominance of motor themes demonstrates progress toward precise and sustainable predictive systems, also applicable in other sectors such as agriculture, energy, and environmental management.

The marginal lines offer potential for future cross-sectoral research and interdisciplinary collaborations. Moreover, their development across different regions could enhance the capacity to address global water challenges and contribute to smarter, more resilient water management.

5 Conclusions and Future Research

The results of this systematic review reveal a scenario of scientific maturity and methodological diversification surrounding the use of Artificial Intelligence applied to wastewater treatment.

First, regarding RQ1, the most commonly employed criteria to evaluate the effectiveness of AI models are Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-Squared (R^2), indicators that reflect a methodological consensus oriented toward measuring the precision and predictive capacity of algorithms. This trend confirms an evolution toward the standardization of quantitative and reproducible metrics, reinforcing the rigor and comparability of studies.

In relation to RQ2, the most applied models correspond to robust architectures such as Artificial Neural Networks (ANN), Support Vector Machine (SVM), and Random Forest (RF), whose high frequency of use responds to their ability to handle complex data, reduce errors, and optimize the operational efficiency of treatment systems.

Meanwhile, RQ4 evidences a solid theoretical foundation supported by hybrid and fuzzy approaches that integrate deep learning techniques with traditional statistical methodologies, consolidating an interdisciplinary vision of the field. It is worth noting that these conceptual frameworks facilitate the interpretation of nonlinear phenomena and favor the design of adaptive intelligent systems.

Finally, RQ6 reveals the existence of consolidated thematic categories around machine learning, deep learning, water quality, and IoT, with emerging lines linked to real-time monitoring and energy efficiency. Altogether, the findings reflect an expanding field characterized by analytical rigor, empirical validation, and the pursuit of sustainable solutions to the environmental and operational challenges of treatment plants.

In summary, Artificial Intelligence is consolidated as a strategic tool to optimize processes, reduce uncertainties, and advance toward smarter, more resilient, and more efficient water management.

Future research should focus on integrating advanced Artificial Intelligence algorithms with IoT platforms and energy-sustainable predictive models capable of operating in environments with incomplete or noisy data.

Likewise, it is advisable to strengthen interdisciplinary collaboration and the creation of open repositories that promote data traceability and metric standardization. Altogether, these directions will make it possible to consolidate smarter, more transparent, and environmentally sustainable water management.

References

1. **Abdalrahman, G., Hin, S., Kumar, P., Ahmed, A. N. (2022).** Modeling the infiltration rate of wastewater infiltration basins considering water. Engineering Applications of Computational Fluid Mechanics, 16(1), 397–421.
2. **Afan, H. A., Melini Wan Mohtar, W. H., Khaleel, F., Kamel, A. H., Mansoor, S. S., Alsultani, R., Ahmed, A. N., Sherif, M., El-Shafie, A. (2024).** Data-driven water quality prediction for wastewater treatment plants. Heliyon, 10(18), 1-14.
3. **Aghdam, E., Mohandes, S. R., Manu, P., Cheung, C., Yunusa-Kaltungo, A., Zayed, T. (2023).** Predicting quality parameters of wastewater treatment plants using artificial intelligence techniques. Journal of Cleaner Production, 405, 1-12.
4. **Al-Ani, S., Guo, H., Fyfe, S., Long, Z., Donnaz, S., Kim, Y. (2024).** Deep learning-based image analysis for filamentous and floc-forming bacteria in wastewater treatment. Journal of Water Process Engineering, 65, 105772.
5. **Alnajjar, H. Y. H., Üçüncü, O. (2023).** Performance prediction and control for wastewater treatment plants using artificial neural network modeling of mechanical and biological treatment. Archives of Environmental Protection, 49(2), 16–29.
6. **Alnajjar, H. Y. H., Üçüncü, O. (2023).** Removal efficiency prediction model based on the artificial neural network for pollution prevention in wastewater treatment plants. Arabian Gulf Journal of Scientific Research, 41(4), 610–626.
7. **Alsulaili, A., Refaie, A. (2021).** Artificial neural network modeling approach for the prediction of five-day biological oxygen demand and

wastewater treatment plant performance. Water Supply, 21(5), 1861–1877.

8. **Alvi, M., French, T., Cardell-Oliver, R., Keymer, P., Ward, A. (2022).** Cost-effective soft sensing for wastewater treatment facilities. IEEE Access, 10, 55694–55708.
9. **Aponte-Rengifo, O., Francisco, M., Vilanova, R., Vega, P., Revollar, S. (2023).** Intelligent control of wastewater treatment plants based on model-free deep reinforcement learning. Processes, 11(8), 2269.
10. **Byliński, H., Sobecki, A., Gębicki, J. (2019).** The use of artificial neural networks and decision trees to predict the degree of odor nuisance of post-digestion sludge in the sewage treatment plant process. Sustainability, 11(16), 4407.
11. **Chen, J., N'Doye, I., Myshkevych, Y., Aljehani, F., Monjed, M. K., Laleg-Kirati, T. M., Hong, P. Y. (2025).** Viral particle prediction in wastewater treatment plants using nonlinear lifelong learning models. npj Clean Water, 8(1), 1–13.
12. **Cheng, Q., Chunhong, Z., Qianglin, L. (2023).** Development and application of random forest regression soft sensor model for treating domestic wastewater in a sequencing batch reactor. Scientific Reports, 13(1), 9149.
13. **Cheng, T., Harrou, F., Kadri, F., Sun, Y., Leiknes, T. (2020).** Forecasting of wastewater treatment plant key features using deep learning-based models: A case study. IEEE Access, 8, 184475–184485.
14. **Drewnowski, J., Szeląg, B., Sabba, F., Piłat-Rożek, M., Piotrowicz, A., Łagód, G. (2023).** Innovations in wastewater treatment: Harnessing mathematical modeling and computer simulations with cutting-edge technologies and advanced control systems. Journal of Ecological Engineering, 24(12), 208–222.
15. **Harrou, F., Dairi, A., Dorbane, A., Sun, Y. (2023).** Energy consumption prediction in water treatment plants using deep learning with data augmentation. Results in Engineering, 20, 101428.
16. **Hasan, M. A. (2021).** An emergent addition for the optimal systemization of wastewater utilization plants using artificial intelligence. Water Science and Technology, 84(10–11), 2805–2817.
17. **Hassan, M. K., Cotfas, D. T., Rezk, H., Youssef, H., Shehata, A. S., El-Bary, A. A. (2024).** Ideal parameter estimation of photocatalysis process to boost amoxicillin degradation efficiency using marine predators optimization algorithm. International Journal of Photoenergy, 2024(1).
18. **Hassan, R. (2024).** Efficient pulp and paper wastewater treatment using green tea-synthesized magnetite: Artificial intelligence and kinetic analysis. Desalination and Water Treatment, 320, 100804.
19. **Huang, Z., Bai, Y., Liu, H. (2025).** Symmetry-inspired prediction of nitrous oxide emissions in wastewater treatment using deep learning and explainable analysis. Symmetry, 17(2).
20. **Jafar, R., Awad, A., Jafar, K., Shahrour, I. (2022).** Predicting effluent quality in full-scale wastewater treatment plants using shallow and deep artificial neural networks. Sustainability, 14(23), 15598.
21. **Jayakumar, M., Sundramurthy, V. P., Gebeyehu, K. B., Selvakumar, K. V., Emana, A. N., Manivannan, S., Mohanasundaram, S., Sagadevan, S., Baskar, G. (2024).** Artificial neural network-guided optimization of limiting factors for enhancing photocatalytic treatment of textile wastewater using UV/TiO_2 and kinetic studies. Desalination and Water Treatment, 320, 100828.
22. **Khan, S. N., Loukil, F., Ghedira-Guegan, C., Benkhelifa, E., Bani-Hani, A. (2021).** Blockchain smart contracts: Applications, challenges, and future trends. Peer-to-Peer Networking and Applications, 14(5), 2901–2925.
23. **Klawikowska, Z., Grochowski, M. (2024).** Optimizing control of wastewater treatment plant with reinforcement learning: Technical evaluation of twin-delayed deep deterministic policy gradient agent. IEEE Access, 12, 130318–130333.

24. **Li, C., Chen, Z., Huang, Z., Shuai, Y., Wang, S., Qi, X., Zheng, J. (2025).** An assembled feature attentive algorithm for automatic detection of wastewater treatment plants based on multiple neural networks. Remote Sensing, 17(9), 1645.
25. **Li, G., Ji, J., Ni, J., Wang, S., Guo, Y., Hu, Y., Liu, S., Huang, S. F., Li, Y. Y. (2022).** Application of deep learning for predicting the treatment performance of real municipal wastewater based on one-year operation of two anaerobic membrane bioreactors. Science of the Total Environment, 813, 151920.
26. **Li, H., Liu, C., Guo, X., Sun, H., Li, X., Jiang, H., Miao, S. (2025).** Applying machine learning approach to design operational control strategies for a wastewater treatment plant in typical scenarios. Water, 17(3), 310.
27. **Li, H., Zech, J., Hong, D., Ghamisi, P., Schultz, M., Zipf, A. (2022).** Leveraging OpenStreetMap and multimodal remote sensing data with joint deep learning for wastewater treatment plant detection. International Journal of Applied Earth Observation and Geoinformation, 110, 102804.
28. **Lin, Q., Luo, A., Zhang, Y., Wang, Y., Liang, Z., Yuan, P. (2021).** Employing artificial neural networks to predict the performance of domestic sewage treatment terminals in the rural region. Mathematical Problems in Engineering, 1–18.
29. **Liu, X., Nie, Y., Wu, X. L. (2023).** Predicting microbial community compositions in wastewater treatment plants using artificial neural networks. Microbiome, 11(1), 1–15.
30. **Liu, X., Shi, Q., Liu, Z., Yuan, J. (2021).** Using LSTM neural network based on improved PSO and attention mechanism for predicting the effluent COD in a wastewater treatment plant. IEEE Access, 9, 146082–146096.
31. **Liu, Y., Tian, W., Xie, J., Huang, W., Xin, K. (2023).** LSTM-based model-predictive control with rationality verification for bioreactors in wastewater treatment. Water, 15(9), 1779.
32. **Liu, Y., Yuan, L., Li, D., Li, Y., Huang, D. (2021).** Process monitoring of quality-related variables in wastewater treatment using Kalman–Elman neural network-based soft-sensor modeling. Water, 13(24), 1–18.
33. **Ma, S., Zhu, Z., Liu, Y., Zheng, Y., Lu, J., Xu, J. (2024).** Artificial intelligence-enabled predictive planning for sewage treatment based on improved DNN and LSTM. Computers & Industrial Engineering, 198, 110636.
34. **Mamandipoor, B., Majd, M., Sheikhalishahi, S., Modena, C., Osmani, V. (2020).** Monitoring and detecting faults in wastewater treatment plants using deep learning. Environmental Monitoring and Assessment, 192(2), 148.
35. **McCormick, M. (2022).** An artificial neural network for simulation of an upflow anaerobic filter wastewater treatment process. Sustainability, 14(13), 1–18.
36. **Mohammadi, E., Stokholm-Bjerregaard, M., Hansen, A. A., Nielsen, P. H., Ortiz-Arroyo, D., Durdevic, P. (2024).** Deep learning-based simulators for the phosphorus removal process control in wastewater treatment via deep reinforcement learning algorithms. Engineering Applications of Artificial Intelligence, 133, 107992.
37. **Muloiwa, M., Dinka, M., Nyende-Byakika, S. (2023).** Application of artificial neural network for predicting biomass growth during domestic wastewater treatment through a biological process. Engineering in Life Sciences, 23(5), 1–18.
38. **Neelakandan, S., Reddy, N. V. R., Ghfar, A. A., Pandey, S., Kiran, S., Thillai Arasu, P. (2023).** Internet of Things with nanomaterials-based predictive model for wastewater treatment using stacked sparse denoising auto-encoder. Water Reuse, 13(2), 233–249.
39. **Pattengale, N. D., Hudson, C. M. (2020).** Decentralized genomics audit logging via permissioned blockchain ledgering. BMC Medical Genomics, 13, 1–12.
40. **Peng, Y., Li, E. (2025).** A generative model-based coevolutionary training framework for noise-tolerant soft sensors in wastewater treatment processes. Complex & Intelligent Systems, 11(5), 1–25.

41. **Pisa, I., Morell, A., Vicario, J. L., Vilanova, R. (2020).** Denoising autoencoders and LSTM-based artificial neural networks data processing for its application to internal model control in industrial environments: The wastewater treatment plant control case. Sensors, 20(13), 3743.
42. **Pisa, I., Morell, A., Vicario, J. L., Vilanova, R. (2023).** Transfer learning in wastewater treatment plants control: Measuring the transfer suitability. Journal of Process Control, 124, 36–53.
43. **Pisa, I., Santín, I., Vicario, J. L., Morell, A., Vilanova, R. (2019).** ANN-based soft sensor to predict effluent violations in wastewater treatment plants. Sensors, 19(6), 1280.
44. **Post, C., Heyden, N., Reinartz, A., Foerderer, A., Bruelisauer, S., Linnemann, V., Hug, W., Amann, F. (2022).** Possibilities of real-time monitoring of micropollutants in wastewater using laser-induced Raman and fluorescence spectroscopy and artificial intelligence. Sensors, 22(13), 4668.
45. **Rodríguez-Rángel, H., Arias, D. M., Morales-Rosales, L. A., Gonzalez-Huitron, V., Valenzuela Partida, M., García, J. (2022).** Machine learning methods modeling carbohydrate-enriched cyanobacteria biomass production in wastewater treatment systems. Energies, 15(7), 2500.
46. **Sakiewicz, P., Piotrowski, K., Ober, J., Karwot, J. (2020).** Innovative artificial neural network approach for integrated biogas–wastewater treatment system modelling: Effect of plant operating parameters on process intensification. Renewable and Sustainable Energy Reviews, 124, 109784.
47. **Saleh, B. A., Kayi, H. (2021).** Prediction of chemical oxygen demand from the chemical composition of wastewater by artificial neural networks. Journal of Physics: Conference Series, 1818(1), 012035.
48. **Santín, I., Meneses, M., Pedret, C., Barbu, M., Vilanova, R. (2025).** Improving hierarchical ammonium control operation in wastewater treatment with non-ideal sensors and actuators using filters, adaptive control and measurement error prediction. Systems Science & Control Engineering, 13(1), 1–15.
49. **Seo, G., Yoon, S., Kim, M., Mun, C., Hwang, E. (2021).** Deep reinforcement learning-based smart joint control scheme for on/off pumping systems in wastewater treatment plants. IEEE Access, 9, 95360–95371.
50. **Seshan, S., Poinapen, J., Zandvoort, M. H., van Lier, J. B., Kapelan, Z. (2025).** Forecasting nitrous oxide emissions from a full-scale wastewater treatment plant using LSTM-based deep learning models. Water Research, 268, 122754.
51. **Shao, S., Fu, D., Yang, T., Mu, H., Gao, Q., Zhang, Y. (2023).** Analysis of machine learning models for wastewater treatment plant sludge output prediction. Sustainability, 15(18), 1–17.
52. **Sibiya, N. P., Amo-Duodu, G., Tetteh, E. K., Rathilal, S. (2022).** Model prediction of coagulation by magnetised rice starch for wastewater treatment using response surface methodology with artificial neural network. Scientific African, 17, e01282.
53. **Sonnenschein, B., Ziel, F. (2023).** Probabilistic intraday wastewater treatment plant inflow forecast utilizing rain forecast data and sewer network sensor data. Water Resources Research, 59(11), 1–21.
54. **Tian, W., Liu, Y., Xie, J., Huang, W., Chen, W., Tao, T., Xin, K. (2023).** Simulation and dynamic properties analysis of the anaerobic–anoxic–oxic process in a wastewater treatment plant based on Koopman operator and deep learning. Water, 15(10), 1960.
55. **Verhaeghe, L., Verwaeren, J., Kirim, G., Daneshgar, S., Vanrolleghem, P. A., Torfs, E. (2024).** Towards good modelling practice for parallel hybrid models for wastewater treatment processes. Water Science and Technology, 89(11), 2971–2990.
56. **Voipan, D., Voipan, A. E., Barbu, M. (2025).** Evaluating machine learning-based soft sensors for effluent quality prediction in wastewater treatment under variable weather conditions. Sensors, 25(6), 1692.
57. **Wang, J., Wan, K., Gao, X., Cheng, X., Shen, Y., Wen, Z., Tariq, U., Piran, M. J. (2020).** Energy and materials-saving management via

deep learning for wastewater treatment plants. IEEE Access, 8, 191694–191705.

58. **Wei, Z., Wu, N., Zou, Q., Zou, H., Zhu, L., Wei, J., Huang, H. (2023).** Data modeling of sewage treatment plant based on long short-term memory with multilayer perceptron network. Water, 15(8), 1472.
59. **Wongburi, P., Park, J. K. (2022).** Prediction of sludge volume index in a wastewater treatment plant using recurrent neural network. Sustainability, 14(10), 6276.
60. **Wongburi, P., Park, J. K. (2023).** Prediction of wastewater treatment plant effluent water quality using recurrent neural network models. Water, 15(19), 3325.
61. **Xie, Y., Chen, Y., Lian, Q., Yin, H., Peng, J., Sheng, M., Wang, Y. (2022).** Enhancing real-time prediction of effluent water quality of wastewater treatment plant based on improved feedforward neural network coupled with optimization algorithm. Water, 14(7), 1053.
62. **Xu, A., Zou, X., Wang, C. (2023).** Prediction of the wastewater's pH based on deep learning incorporating sliding windows. Computer Systems Science and Engineering, 47(1), 1043–1059.
63. **Zhang, Y., Suzuki, G., Shioya, H. (2022).** Prediction and detection of sewage treatment process using N-BEATS autoencoder network. IEEE Access, 10, 112594–112608.
64. **Kitchenham, B., Charters, S. (2007).** Guidelines for performing systematic literature reviews in software engineering. Evidence-Based Software Engineering Technical Report, EBSE-2007-01, Keele University and University of Durham.
65. **Sauer, P. C., Seuring, S. (2023).** How to conduct systematic literature reviews in management research: A guide in six steps and fourteen decisions. Review of Managerial Science, 17, 1899–1933.
66. **Petersen, K., Feldt, R., Mujtaba, S., Mattsson, M. (2015).** Systematic mapping studies in software engineering. Information and Software Technology, 64, 1–18.
67. **Petersen, K., Vakkalanka, S., Kuzniarz, L. (2015).** Guidelines for conducting systematic mapping studies in software engineering: An update. Information and Software Technology, 64, 1–8.
68. **Okoli, C., Schabram, K. (2010).** Towards a framework for conducting systematic mapping studies. Journal of Software Engineering, 56(4), 271–280.
69. **Linnenluecke, M. K., Marrone, M., Singh, A. K. (2020).** Conducting systematic literature reviews and bibliometric analyses. Australian Journal of Management, 45(2), 175–194.
70. **Cárdenas-Quispe, A., Vergaray-Mezarina, R., Gamboa-Cruzado, J. (2021).** Machine learning para la detección de malware en Android: Revisión sistemática de la literatura. Revista Ibérica de Sistemas e Tecnologias de Informação (RISTI), E45, 318–331.
71. **Adeoba, M. I., Pandelani, T., Ngwangwa, H., Masebe, T. (2025).** The role of artificial intelligence in sustainable ocean waste tracking and management: A bibliometric analysis. Sustainability, 17(9), 1–20.
72. **Flores-Iwasaki, M., Guadalupe, G. A., Pachas-Caycho, M., Chapa-Gonza, S., Mori-Zabarburú, R. C., Guerrero-Abad, J. C. (2025).** Internet of Things (IoT) sensors for water quality monitoring in aquaculture systems: A systematic review and bibliometric analysis. AgriEngineering, 7(3), 78.
73. **García, J., Leiva-Araos, A., Diaz-Saavedra, E., Moraga, P., Pinto, H., Yepes, V. (2023).** Relevance of machine learning techniques in water infrastructure integrity and quality: A review powered by natural language processing. Applied Sciences, 13(22), 1–40.
74. **Lima, M., Neto, M., Filho, T. S., Roberta, R. A. (2022).** Learning under concept drift for regression: A systematic literature review. IEEE Access, 10, 45410–45429.
75. **Sharma, R. K., Kaur, J., Feng, G., Huang, Y., Kumar, C., Wang, Y., Sharma, S., Jenkins, J., Dhillon, J. (2025).** Maize and soybean yield prediction using machine learning methods: A systematic literature review. Discover Agriculture, 3(1), 1–25.
76. **Yu, X., Chen, S., Zhang, X., Wu, H., Guo, Y., Guan, J. (2023).** Research progress of the artificial intelligence application in wastewater

treatment during 2012–2022: A bibliometric analysis. Water Science and Technology, 88(7), 1750–1766.

77. **Zhou, K., Wu, B., Zhang, X. (2025).** Worldwide research progress and trends in the application of machine learning to wastewater treatment: A bibliometric analysis. Water, 17(9), 1–17.
78. **De la Hoz-M, J., Ariza-Echeverri, E. A., Vergara, D. (2024).** Exploring the role of artificial intelligence in wastewater treatment: A dynamic analysis of emerging research trends. Resources, 13, 171.
79. **Coy-Aceves, L. E., Corona-Vasquez, B. (2022).** Bibliometric analysis of artificial intelligence algorithms used for microbial fuel cell research. Water Practice and Technology, 17(10), 2071–2082.
80. **Baarimah, A., Bazel, A., Alaloul, S., Alazaiza, M., Al-Zghoul, T., Almuhaya, B., Khan, A., Mushtaha, A. (2024).** Artificial intelligence in wastewater treatment: Research trends and future perspectives through bibliometric analysis. Processes, 14(2), 136–159.
81. **Gamboa-Cruzado, J., Crisostomio-Castro, R., Vila-Buleje, J., López-Goycochea, J., Nolasco-Valenzuela, J. (2024).** Heart attack prediction using machine learning: A comprehensive systematic review and bibliometric analysis. Journal of Theoretical and Applied Information Technology (JATIT), 102(5), 1930–1944.
82. **Tosan, M., Gharib, M. R., Attar, N. F., Maroosi, A. (2025).** Enhancing evapotranspiration estimation: A bibliometric and systematic review of hybrid neural networks in water resource management. Computer Modeling in Engineering & Sciences, 142(2), 1–35.
83. **Olu-Ajayi, R., Alaka, H., Sunmola, F., Ajayi, S., Mporas, I. (2024).** Statistical and artificial intelligence-based tools for building energy prediction: A systematic literature review. IEEE Transactions on Engineering Management, 71, 14733–14753.
84. **Daza Vergaray, A., Apaza Pérez, O. G., Zagaceta Daza, J. A., Samanez Torres, K. C., Benites Noriega, J., Llanos Gonzales, O. T., Cochachin Cochachin, A. D. (2024).** Machine learning and deep learning techniques to predict software defects: A bibliometric analysis, systematic review, challenges and future works. Computers & Electrical Engineering, 119, 109452.
85. **Baruah, A., Changkakati, B., Sharma, U. (2025).** Examination of fuzzy logic applications for short-term stock market forecasting: A systematic review. Interdisciplinary Nexus: Law, Humanities and Management, January issue.
86. **Amaro, L. C., Reyes, S. R., Meléndez, M. A., Ovalle, C. (2024).** Systematic review of artificial intelligence with near-infrared in blueberries. IAES International Journal of Artificial Intelligence, 13(4), 3761–3771.
87. **Mukonza, S. S., Chiang, J. L. (2023).** Meta-analysis of satellite observations for United Nations sustainable development goals: Exploring the potential of machine learning for water quality monitoring. Environments, 10(10), 1–25.
88. **Nazirun, N. N., Abdul Wahab, A., Selamat, A., Fujita, H., Krejcar, O., Kuca, K., Hong Seng, G. (2024).** Prediction models for type 2 diabetes progression: A systematic review. IEEE Access, 12, 161595.
89. **Abdallah, M., Abu Talib, M., Feroz, S., Nasir, Q., Abdalla, H., Mahfood, B. (2020).** Artificial intelligence applications in solid waste management: A systematic research review. Waste Management, 109, 231–246.
90. **De La Cruz, J. E. C., Mamani, W. A., Pineda, F., Yana-Mamani, V., Cruz, R. S., Maldonado-Ramírez, R., Pérez-Astonitas, R., Morales-Rojas, E. (2025).** A review on artificial intelligence for water quality prediction in Amazonian countries. Natural Environment and Pollution Technology, 24(2), 1–15.
91. **Andrade-Mogollon, T., Gamboa-Cruzado, J., Amayo-Gamboa, F. (2025).** Systematic literature review of generative AI and IoT as key technologies for precision agriculture. Computación y Sistemas, 29(2), 857–882.
92. **Flores-Iwasaki, M., Guadalupe, G. A., Pachas-Caycho, M., Chapa-Gonza, S., Mori-**

Zabarburú, R. C., Guerrero-Abad, J. C. (2025). Internet of Things (IoT) sensors for water quality monitoring in aquaculture systems: A systematic review and bibliometric analysis. AgriEngineering, 7, 78.

93. **Nikhar, C. K., Vyas, G. S., Dalvi, R. S., Bhoye, D. Y. (2025).** A critical review on applications of machine learning in wastewater treatment: Insights and implications for distillery wastewater. Water Quality Research Journal, 60(1), 1–18.
94. **Basu, A., Hazra, A. K., Chaudhury, S., Ross, A. B., Balachandran, S. (2021).** State of the art research on sustainable use of water hyacinth: A bibliometric and text mining analysis. Informatics, 8(2), 1–24.
95. **Rosário, A. T., Dias, J. C. (2022).** Sustainability and the digital transition: A literature review. Sustainability, 14(7), 1–18.
96. **De Matos, B., Salles, R., Mendes, J., Gouveia, J. R., Baptista, A. J., Moura, P. (2023).** A review of energy and sustainability KPI-based monitoring and control methodologies on WWTPs. Mathematics, 11, 173.
97. **Iqbal, U., Riaz, M. Z. B., Barthelemy, J., Perez, P., Idrees, M. B. (2023).** The last two decades of computer vision technologies in water resource management: A bibliometric analysis. Water and Environment Journal, 37(3), 373–389.
98. **Hu, J., Miao, C., Wu, Y., Su, J. (2025).** Advances in hydrological research in China over the past two decades: Insights from advanced large language model and topic modeling. Fundamental Research, in press.
99. **Abdallah, M., Abu Talib, M., Feroz, S., Nasir, Q., Abdalla, H., Mahfood, B. (2020).** Artificial intelligence applications in solid waste management: A systematic research review. Waste Management, 109, 231–246.
100. **Altowayti, W. A. H., Shahir, S., Othman, N., Eisa, T. A. E., Yafooz, W. M. S., Al-Dhaqm, A., Soon, C. Y., Yahya, I. B., Che Rahim, N. A. N. B., Abaker, M., Ali, A. (2022).** The role of conventional methods and artificial intelligence in wastewater treatment: A comprehensive review. Processes, 10, 1832.
101. **Cinthya, M., Vinarti, R. A., Aini Rakhmawati, N. (2023).** A systematic literature review on determining optimal coagulant for water treatment: Artificial intelligence techniques and variable selection. In 2023 IEEE 8th International Conference for Convergence in Technology (I2CT) (pp. 1–6). IEEE.
102. **Da Costa, M. F. P., Araújo, R. S., Silva, A. R., Pereira, L., Silva, G. M. M. (2025).** Predictive artificial neural networks as applied tools in the remediation of dyes by adsorption: A review. Applied Sciences, 15, 2310.
103. **Gamboa-Cruzado, J., Mosqueira, T., Torre, A., Mendoza, R., Raymundo, N., García, J., López, C. (2025).** Systematic literature review of generative AI and IoT as key technologies for precision agriculture. Computación y Sistemas, 29(2), 925–953.
104. **Deepak, M., Rustum, R. (2023).** Review of latest advances in nature-inspired algorithms for optimization of activated sludge processes. Processes, 11, 77.
105. **El Alaoui El Fels, A., Mandi, L., Kammoun, A., Ouazzani, N., Monga, O., Hbid, M. L. (2023).** Artificial intelligence and wastewater treatment: A global scientific perspective through text mining. Water, 15, 3487.
106. **León-Pérez, J., Bocangel-Rivera, E., Niño-Montero, J., Gamboa-Cruzado, J., Soto-Soto, L., Oseda-Gago, D. (2021).** Revisión sistemática de la literatura sobre redes neuronales artificiales: Detección de ataques cardíacos. Revista Ibérica de Sistemas e Tecnologias de Informação (RISTI), E45, 303–317.
107. **Bayhan, K., Başakın, E. E., Ekmekcioğlu, Ö., Pham, Q. B., Chu, H. J. (2025).** Overviewing the machine learning utilization on groundwater research using bibliometric analysis. Water, 17(7), 1–25.
108. **PRISMA. (2025).** Preferred Reporting Items for Systematic Reviews and Meta-Analyses. Disponible en: https://www.prismastatement.org (accedido el 15 de julio de 2025).

Article received on 084/07/2025; accepted on 25/10/2025.
**Corresponding author is Cesar Jesús Núñez-Prado.*