ReGeo ISSN: 2177-3246

ARTIFICIAL INTELLIGENCE AND GEOSPATIAL TECHNOLOGIES FOR WATER QUALITY MONITORING: A BIBLIOMETRIC ANALYSIS OF TRENDS, METHODS, AND RESEARCH GAPS

INTELIGÊNCIA ARTIFICIAL E TECNOLOGIAS GEOESPACIAIS PARA O MONITORAMENTO DA QUALIDADE DA ÁGUA: UMA ANÁLISE BIBLIOMÉTRICA DE TENDÊNCIAS, MÉTODOS E LACUNAS DE PESQUISA

INTELIGENCIA ARTIFICIAL Y TECNOLOGÍAS GEOESPACIALES PARA EL MONITOREO DE LA CALIDAD DEL AGUA: UN ANÁLISIS BIBLIOMÉTRICO DE TENDENCIAS, MÉTODOS Y BRECHAS DE INVESTIGACIÓN

https://doi.org/10.56238/revgeov17n10-030

Data de submissão: 08/09/2026 **Data de publicação:** 08/10/2026

Gabriela Luiza Almeida Silva[1], Paulo Tibúrcio Pereira[2], Ana Clara Cassin Cordeiro[3], Helen Marques Martins[4], Felipe Tadeu Fiorini Gomide[5]

ABSTRACT

Water quality monitoring is essential for ensuring the sustainable use of aquatic resources and preserving ecosystem health. The increasing application of artificial intelligence (AI) in this field has provided innovative approaches to modeling and predicting key water quality parameters with greater efficiency. This study aimed to map the landscape of scientific production, understand the correlation network among identified themes, and analyze the temporal evolution of publication output. We conducted a bibliometric analysis of scientific publications related to the use of AI in water quality assessment, employing tools such as Bibliometrix and VOSviewer. A total of 130 peer-reviewed articles were retrieved from the Web of Science and Scopus databases, covering the period from 2005 to 2025, and analyzed to identify research trends, key themes, and collaboration networks. Thematic clustering of keywords revealed four main research areas: (i) remote sensing, (ii) machine learning, (iii) water quality indicators, and (iv) hydrodynamic and watershed modeling. The results show that the term “machine learning” is central to the analyzed literature, connecting clusters focused on hydrological modeling, remote sensing, supervised learning algorithms, and the prediction of physicochemical variables. The findings highlight the interdisciplinary nature of the field and underscore the growing role of AI and geotechnologies as cost-effective tools for supporting water quality management, particularly in hard-to-access areas. These insights

[1] Master's student. Universidade Federal de São João del-Rei (UFSJ). Programa de Pós-Graduação em Engenharia Química (PPGEQ). E-mail: gabi-lsilva@hotmail.com Orcid: https://orcid.org/0009-0002-6834-6887 Lattes: http://lattes.cnpq.br/6594736090145487

[2] Doctor. Universidade Federal de São João del-Rei (UFSJ). Programa de Pós-Graduação em Engenharia Química (PPGEQ). E-mail: paulotiburcio@ufsj.edu.br Orcid: https://orcid.org/0000-0003-2344-5777 Lattes: http://lattes.cnpq.br/9589927393728395

[3] Master Universidade Federal de São João del-Rei (UFSJ). Programa de Pós-Graduação em Engenharia Química (PPGEQ). E-mail: anacassin26@gmail.com Orcid: https://orcid.org/0000-0001-5977-8544 Lattes: https://lattes.cnpq.br/8383353778513705

[4] Master's student. Universidade Federal de São João del-Rei (UFSJ). Programa de Pós-Graduação em Engenharia Química (PPGEQ). E-mail: helenmmartins@gmail.com Orcid: https://orcid.org/0009-0006-7812-2835 Lattes: http://lattes.cnpq.br/7365637195489056

[5] Doctor. Universidade Federal de São João del-Rei (UFSJ). Centro de Inovação e Empreendedorismo Universitário (CIEU). E-mail: felipetfgomide@gmail.com Orcid: https://orcid.org/0000-0003-1276-4326 Lattes: http://lattes.cnpq.br/1510482922012122

offer guidance for future research and help environmental agencies develop more efficient strategies for monitoring and preserving water resources.

Keywords: Artificial Intelligence. Machine Learning. Water Quality Monitoring. Remote Sensing. Hydrological Modeling. Bibliometric Analysis. Environmental Monitoring. VOSviewer.

RESUMO

O monitoramento da qualidade da água é essencial para garantir o uso sustentável dos recursos hídricos e preservar a saúde dos ecossistemas. A crescente aplicação da inteligência artificial (IA) nesta área tem proporcionado abordagens inovadoras para modelar e prever parâmetros-chave da qualidade da água com maior eficiência. Este estudo teve como objetivo mapear o panorama da produção científica, compreender a rede de correlação entre os temas identificados e analisar a evolução temporal das publicações. Realizou-se uma análise bibliométrica de publicações científicas relacionadas ao uso de IA na avaliação da qualidade da água, utilizando ferramentas como Bibliometrix e VOSviewer. Um total de 130 artigos revisados por pares foi recuperado das bases de dados Web of Science e Scopus, abrangendo o período de 2005 a 2025, e analisado para identificar tendências de pesquisa, temas principais e redes de colaboração. O agrupamento temático de palavras-chave revelou quatro áreas de pesquisa principais: (i) sensoriamento remoto, (ii) aprendizado de máquina, (iii) indicadores de qualidade da água e (iv) modelagem hidrodinâmica e de bacias hidrográficas. Os resultados mostram que o termo "aprendizado de máquina" é central na literatura analisada, conectando agrupamentos focados em modelagem hidrológica, sensoriamento remoto, algoritmos de aprendizado supervisionado e previsão de variáveis físico-químicas. As conclusões destacam a natureza interdisciplinar da área e ressaltam o papel crescente da IA e das geotecnologias como ferramentas de custo-benefício favorável para apoiar a gestão da qualidade da água, particularmente em áreas de difícil acesso. Essas informações oferecem orientações para pesquisas futuras e auxiliam órgãos ambientais no desenvolvimento de estratégias mais eficientes para o monitoramento e a preservação dos recursos hídricos.

Palavras-chave: Inteligência Artificial. Aprendizado de Máquina. Monitoramento da Qualidade da Água. Sensoriamento Remoto. Modelagem Hidrológica. Análise Bibliométrica. Monitoramento Ambiental. VOSviewer.

RESUMEN

El monitoreo de la calidad del agua es esencial para garantizar el uso sostenible de los recursos acuáticos y preservar la salud de los ecosistemas. La creciente aplicación de la inteligencia artificial (IA) en este campo ha proporcionado enfoques innovadores para modelar y predecir parámetros clave de la calidad del agua con mayor eficiencia. Este estudio tuvo como objetivo mapear el panorama de la producción científica, comprender la red de correlaciones entre los temas identificados y analizar la evolución temporal de las publicaciones. Se realizó un análisis bibliométrico de publicaciones científicas relacionadas con el uso de la IA en la evaluación de la calidad del agua, empleando herramientas como Bibliometrix y VOSviewer. Se recopilaron 130 artículos revisados por pares de las bases de datos Web of Science y Scopus, abarcando el periodo de 2005 a 2025, los cuales fueron analizados para identificar tendencias de investigación, temas clave y redes de colaboración. La agrupación temática de palabras clave reveló cuatro áreas de investigación principales: (i) teledetección, (ii) aprendizaje automático (*machine learning*), (iii) indicadores de calidad

del agua y (iv) modelado hidrodinámico y de cuencas. Los resultados muestran que el término "aprendizaje automático" es central en la literatura analizada, conectando grupos centrados en el modelado hidrológico, la teledetección, los algoritmos de aprendizaje supervisado y la predicción de variables fisicoquímicas. Los hallazgos destacan la naturaleza interdisciplinaria del campo y subrayan el papel creciente de la IA y las geotecnologías como herramientas rentables para apoyar la gestión de la calidad del agua, particularmente en zonas de difícil acceso. Estas perspectivas ofrecen orientación para futuras investigaciones y ayudan a las agencias ambientales a desarrollar estrategias más eficientes para el monitoreo y la preservación de los recursos hídricos.

Palabras clave: Inteligencia Artificial. Aprendizaje Automático. Monitoreo de la Calidad del Agua. Teledetección. Modelado Hidrológico. Análisis Bibliométrico. Monitoreo Ambiental. VOSviewer.

1 INTRODUCTION

Pollution levels in water bodies have increased due to the discharge of untreated sewage, agricultural, and industrial effluents without proper treatment [1], [2]. Efforts are required to improve water quality, reduce pollution, and promote safe water recycling, as well as to conserve and sustainably use freshwater ecosystems, according to the United Nations Sustainable Development Goals [3].

An alternative to ensure good water quality is to increase the monitoring of water bodies to map the points where the water body is undergoing significant changes [4]. However, conventional monitoring methods, which require *in situ* sampling followed by laboratory analysis of the collected samples, have restrictions to keeping a continuous process, such as difficulties in accessing specific areas due to distance, rugged terrain, or the presence of dense vegetation [5]. In this context, the development of more efficient methods for monitoring water quality becomes necessary, as increased monitoring allows environmental agencies to identify pollution sources more accurately and implement timely and targeted mitigation measures.

The use of artificial intelligence (AI) and geotechnologies, such as satellite remote sensing and UAV-based hyperspectral imaging, facilitates water quality monitoring without the need for conventional, labor-intensive techniques. This approach allows the detection of pollution points in water bodies more efficiently, even in hard-to-reach locations [6], [7], [8]. Moreover, it enables environmental agencies to monitor the recovery or further degradation of aquatic ecosystems over time [9].

Water quality models play a crucial role in this scenario by mathematically representing the behavior of aquatic systems and predicting the concentration and distribution of pollutants over time. These models are developed by integrating hydrodynamic principles, water chemistry, and biological processes, calibrated with observed data to reflect real-world conditions [10], [11].

One of the earliest and most well-known models is the Streeter-Phelps model (1925), which describes the oxygen sag curve in rivers due to organic pollution and has served as the foundation for water quality modeling [12]. In contrast, more recent models, such as the hybrid PCWA-ResCNN model (2025), incorporate AI and deep learning techniques combined with satellite data to enhance prediction accuracy and spatial resolution in complex watersheds [13]. Compared to early models, which were highly simplified and limited to specific parameters, modern models can handle large datasets, simulate multiple variables

simultaneously, and adapt to dynamic environmental conditions. Applications of these models include predicting dissolved oxygen levels, assessing nutrient loading, evaluating the impact of land-use changes, and supporting decision-making for sustainable water resource management [14], [15].

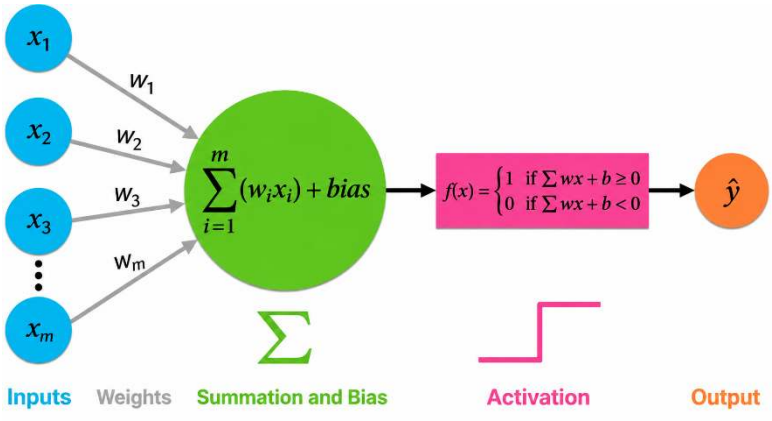
It is crucial to understand the development of Artificial Intelligence (AI) to grasp why it's being so widely employed in modeling studies today. The first Artificial Intelligence (AI) systems, developed between the 1950s and 1970s, were based on programs designed to solve specific problems in restricted domains. This approach, known as symbolic artificial intelligence or Good Old-Fashioned Artificial Intelligence (GOFAI), was based on the idea that intelligence could be represented by symbols operated through explicit logical rules. In this way, the AI was designed to perform tasks such as playing chess or checkers, proving mathematical theorems, and conducting medical diagnoses in specific contexts [16].

Over time, significant progress has been made in both computing power and big data, today High-performance computers, combined with large databases, can be used to train neural networks and solve problems without the need to know the underlying equations or to create a specific program for the task [17].

The working principles of an AI indicate how important that variable is for obtaining the result Figure 1. Next, the system computes a weighted sum of the inputs, and a function defines the activation threshold, producing the output [17], [18], [19].

Figure 1

Perceptron operating model [20].



Artificial intelligence (AI) is a field of computer science that aims to develop systems capable of performing tasks that would normally require human intelligence, such as pattern

recognition, decision-making, and scenario prediction [19]. Among the main types of AI used in environmental applications are supervised learning, unsupervised learning, and reinforcement learning [17]. Models such as artificial neural networks (ANNs), support vector machines (SVMs), and random forests have been widely employed in modeling environmental variables, including physicochemical parameters of water [18].

In this context, considering the need for the development of accurate methodologies to monitor water bodies over time, with the potential to predict their behavior across the years, AI appears as a viable alternative. The relevance of AI is particularly higher considering the availability of a substantial historical dataset, which could be combined with artificial intelligence (AI) tools to model water quality, thereby enabling the identification of patterns and the forecasting of future scenarios [13], [14].

This study proposes a bibliometric review to quantify scientific production, map the correlation network among identified themes, and analyze the temporal evolution of publications on water quality modeling in aquatic ecosystems using artificial intelligence tools.

2 CONCEPTUAL FRAMEWORK

The conceptual framework adopted in this study integrates three core components: artificial intelligence (AI), remote sensing technologies, and hydrological modeling, forming a unified data-driven approach for water quality monitoring.

As illustrated in Figure 2, remote sensing technologies provide large-scale environmental observations through satellite imagery and UAV-based sensors, enabling the acquisition of high-resolution spatial and temporal data related to water bodies. These data include key environmental variables such as turbidity, chlorophyll concentration, suspended sediments, and surface temperature.

Hydrological modeling contributes to the representation of physical and chemical processes governing aquatic systems, including flow dynamics, pollutant transport, and nutrient cycling. These models provide a process-based understanding of water quality variations and serve as a complementary component to data-driven approaches.

Artificial intelligence techniques, particularly machine learning algorithms, act as the central analytical component of the framework. These methods process and integrate heterogeneous datasets obtained from remote sensing and hydrological models, enabling the identification of patterns, prediction of water quality parameters, and generation of decision-support information.

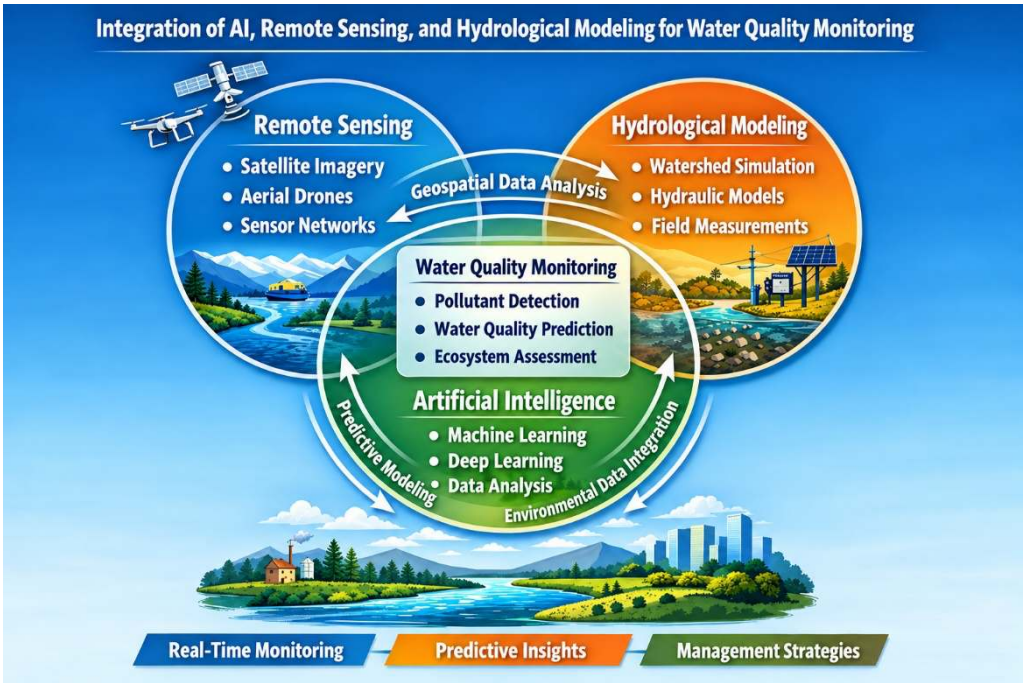
The interaction among these three components enables the development of advanced environmental monitoring systems capable of:

- Detecting pollution events.
- Predicting water quality dynamics.
- Supporting environmental decision-making.
- Enabling large-scale monitoring in data-scarce or inaccessible regions.

Overall, the proposed framework highlights the convergence of data-driven and process-based approaches, demonstrating how artificial intelligence and geospatial technologies can enhance the efficiency, scalability, and accuracy of water quality monitoring systems.

Figure 2

Conceptual framework integrating artificial intelligence, remote sensing technologies, and hydrological modeling for water quality monitoring.



The framework illustrates the interaction between data acquisition (remote sensing), process-based simulation (hydrological modeling), and data-driven analysis (artificial intelligence) to support environmental monitoring, prediction of water quality parameters, and decision-making processes.

3 MATERIALS AND METHODS

The analysis was carried out using R, RStudio, Bibliometrix [21], and VOSviewer [22] software packages to examine literature sourced from the Scopus and Web of Science databases, since these databases offer the largest number of articles specific to this topic. A temporal window from 2005 to 2025 was considered for this investigation, starting from the year the first articles on the topic were published.

The first article presents a new operational algorithm for the simultaneous retrieval of water quality parameters, such as chlorophyll content (chl), suspended minerals (sm), and dissolved organics (doc), from satellite data, applicable in both clear and turbid waters [23].

Initially, articles that covered the themes of machine learning and water quality were searched by utilizing a specific query. Keywords and their synonyms related to the theme of this study were selected. To do so, artificial intelligence tools like Gemini 2.5 Flash and ChatGPT, based on OpenAI's GPT-4 model, were used to generate synonyms associated with the initially conceived query syntax. Several attempts were made to refine the initially proposed syntax, utilizing the word cloud feature available within Bibliometrix to help filter out irrelevant keywords, exclude those without a true correlation to the topic, and keep highly relevant terms. Boolean operators were used to group synonyms and incorporate complementary terms, ensuring convergence with the study's theme. The final syntax consisted of:

("Neural Network*" OR "Multicriteria Analysis*" OR "Multiple-criteria analysis" OR "MCA" OR "Learning algorithms" OR "Machine Learning" OR "Deep learning") AND ("water quality" OR "water condition" OR "Aquatic Health" OR "Water Assessment" OR "Environmental Water Status") AND (river* OR basin) AND ("geo-referencing" OR "GIS operation*" OR "satellite remote sensing" OR "remote sensing" OR "satellite data" OR "satellite imagery") AND ((optimization) OR (flow) OR ("convolutional network*") OR (simulation) OR (database)).

Following the syntax definition and search, databases were downloaded from Scopus and Web of Science. Then, the downloaded datasets were merged with the aid of Zotero, and duplicate records were removed to return the original articles published on the subject. Articles were retained only if they were published in English, Portuguese, or Spanish due to the language limitations of the authors who conducted the initial article screening and data extraction. Consequently, seven articles written in Mandarin and Korean were excluded from

this study. Additionally, works associated with conferences were removed, with only publications and theses being retained for analysis.

In the next step, the VOSviewer and Bibliometrix software were employed to conduct bibliometric analyses. From this analysis, specific visualizations were generated to map the landscape of publications in the area, identify thematic trends, analyze the evolution of the number of publications over time, and map collaboration networks.

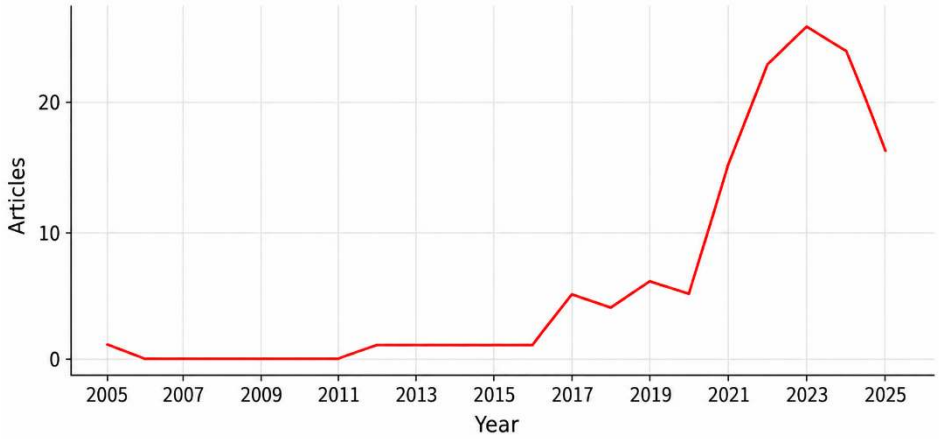
4 RESULTS AND DISCUSSION

We obtained 85 results from Scopus and 104 from the Web of Science databases in a temporal window from 2005 to 2025. After the refinement process, the results were reduced to 130 after filtering steps.

The results indicated that studies in this area are recent, as publications, as exemplified by the graph generated with the aid of Bibliometrix, which showed modest growth until 2017, followed by a significant increase from 2020 onwards, peaking around 2023. The lower number of articles in 2025 is due to the fact that the year has not yet concluded, and data collection for the period is still ongoing Figure 3.

Figure 3

Annual scientific production.

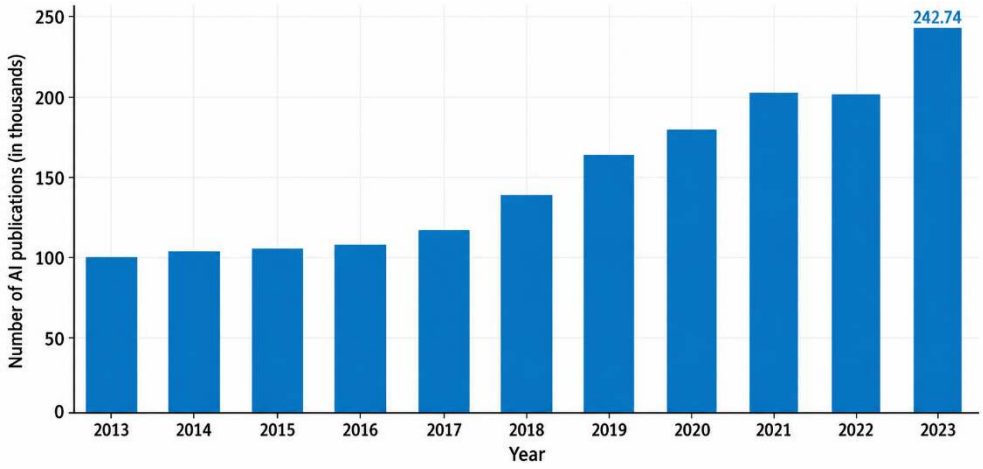


Starting in 2013, the global Artificial Intelligence (AI) research landscape witnessed exponential growth, as evidenced by the total number of publications nearly tripling between 2013 and 2023, according to Figure 4, increasing from approximately 100,000 to over 242,000 in computer science and related disciplines [24]. This surge in global AI research likely paved the way for a subsequent intensification of its applications in specific domains. A clear

example of this trend is observed in hydrological modeling, where the application of AI became significantly more evident and frequent from 2018 to 2020 onwards.

Figure 4

Number of AI publications in CS worldwide, 2013–23 [24].



Regarding the most relevant sources (Table 1), the journal Remote Sensing stands out, contributing 17 articles, followed by Water (9), Journal of Hydrology (8), and Sustainability (6). These journals reflect the interdisciplinary nature of the field, combining remote sensing, hydrology, and environmental sustainability.

Table 1

Most Relevant Sources.

Sources	Articles
REMOTE SENSING	17
WATER	9
JOURNAL OF HYDROLOGY	8
SUSTAINABILITY	6
SCIENCE OF THE TOTAL ENVIRONMENT	5
ECOLOGICAL INDICATORS	4
IEEE TRANSACTIONS ON GEOSCIENCE AND REMOTE SENSING	4
JOURNAL OF ENVIRONMENTAL MANAGEMENT	4
REMOTE SENSING APPLICATIONS: SOCIETY AND ENVIRONMENT	3

The bibliometric analysis performed using the VOSviewer software enabled the identification of thematic patterns in the literature related to water quality monitoring using artificial intelligence techniques. The generated visualizations revealed the structure of the

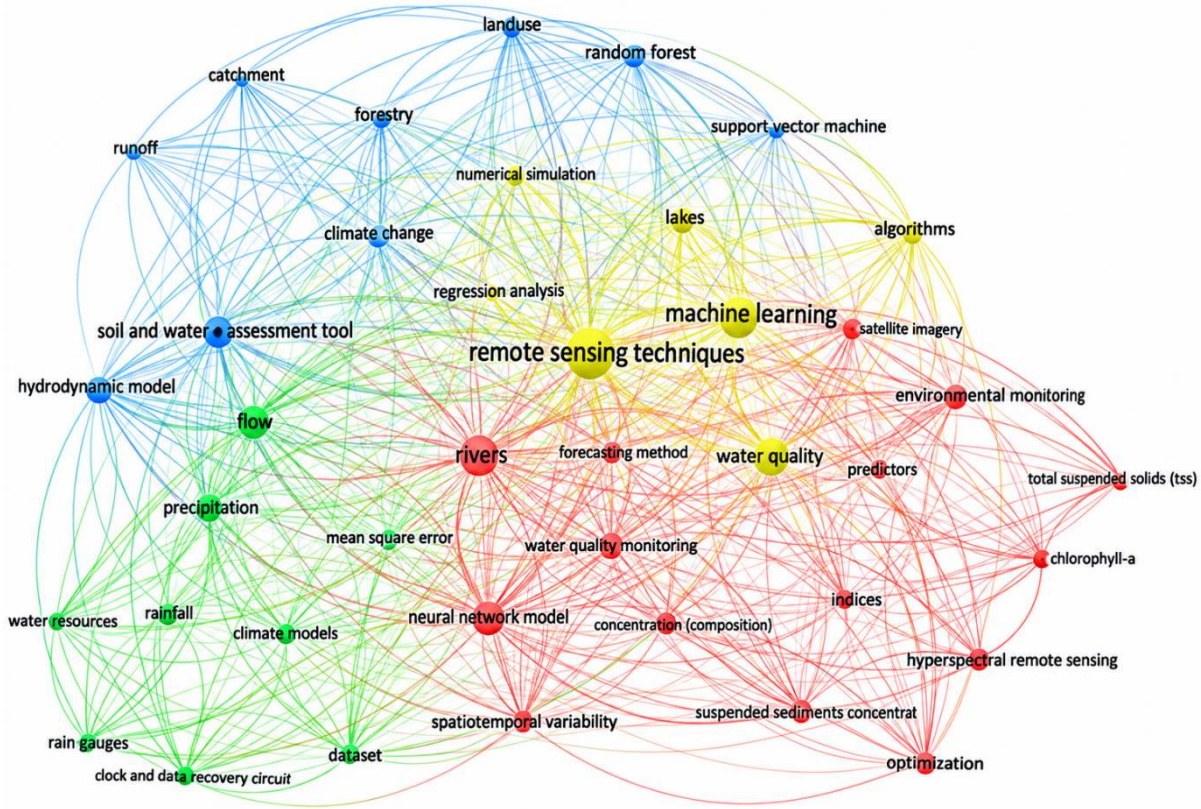
field based on keyword co-occurrence from the articles, allowing the identification of core themes, semantic clusters, and relevant conceptual connections.

To generate the keyword co-occurrence map, the analysis type selected was co-occurrence, with the unit of analysis set to keywords. The counting method chosen was fractional counting, which mitigates the disproportionate influence of documents with large numbers of keywords by distributing the weight [25]. A minimum threshold of 6 keyword occurrences was applied so that the map generated would not be too polluted. The layout of items was computed using the VOS layout algorithm with association strength normalization, standard attraction (3), and repulsion (0) parameters. For clustering, a resolution parameter of 1.75 was applied, with a minimum cluster size of 7.

The division of terms into four thematic clusters, represented by distinct colors, offers a structured view of articles (Figure 5). The green cluster is related to hydrological modeling, flow simulation, the Soil and Water Assessment Tool (SWAT), and environmental monitoring, with a thematic focus on modeling tools for environmental monitoring.

Figure 5

VOSviewer Co-occurrence Analysis.



The red cluster brings together concepts associated with artificial intelligence applied to the prediction of physicochemical variables, such as neural network models, chlorophyll a, Total Suspended Solids (TSS), and suspended sediment concentration. The blue cluster focuses on water quality and climate factors, including water quality, rivers and precipitation, indicative of studies evaluating water quality under climatic and hydrological influences. At last, the yellow cluster highlights remote sensing techniques and encompasses studies applying satellite and aerial imagery to aquatic environments, as evidenced by its high number of links to other clusters and its prominence as the term with the highest occurrence.

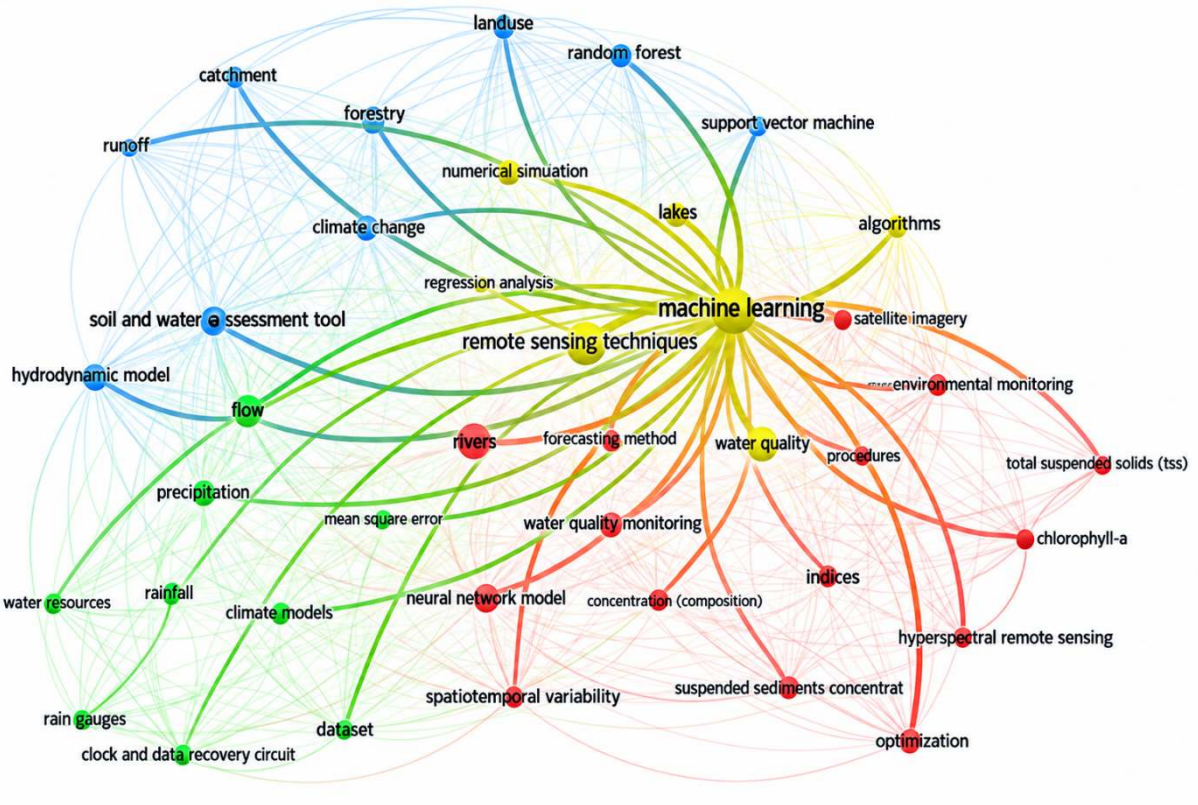
These groupings across all visualizations demonstrate the robustness of thematic associations and point to a research field that is consolidated and expanding. The clusters indicate that studies are not limited to a single methodological approach but instead combine multiple tools, from sensors to predictive algorithms, to address the challenges of water quality monitoring. This trend indicates the advancement of environmental science toward data-driven solutions with potential for public policy, early warning systems, and water quality management strategies.

In all visualizations, the term “machine learning”, in yellow, stands out as the central node in the network, indicating its frequency and strong interconnection with other keywords. This prominence is consistent with the increasing use of machine learning methods in environmental monitoring studies and with the query syntax. Strong associations were observed between machine learning and terms such as remote sensing techniques, water quality, neural network models, environmental monitoring, and rivers, demonstrating that AI applications are linked to both spatial data collection and predictive analysis of environmental parameters.

In visualizations where line thickness is proportional to link strength between terms, pairs such as machine learning–remote sensing techniques display intense connections (Figure 6). When filtering the database in Zotero by the keywords machine learning and remote sensing techniques, a total of 52 articles are found to be linked by both terms. Considering that the full database consists of 130 articles, this means that nearly half of the analyzed publications are connected through the combination of these two concepts, highlighting their central role in the field. This suggests that these concept combinations co-occur in the analyzed articles, indicating a strong relationship between remote sensing technologies and AI algorithms in practical environmental monitoring applications [25].

Figure 6

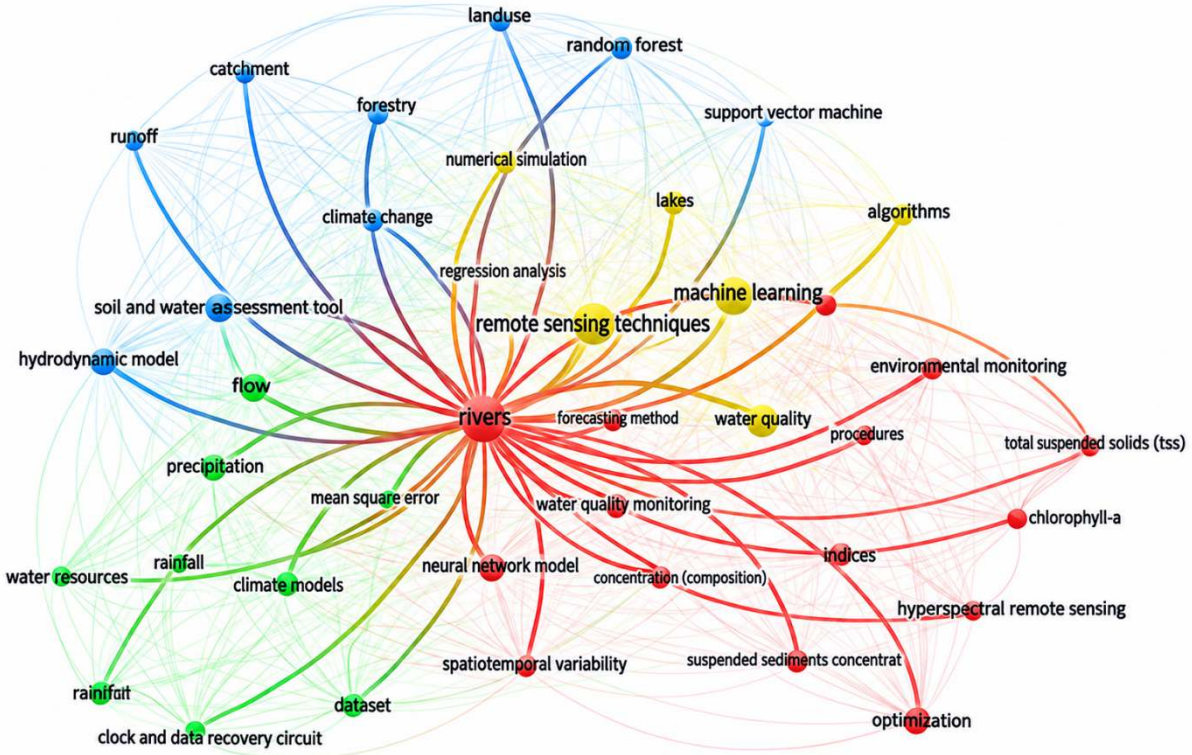
VOSviewer Co-occurrence Analysis, focus on Machine Learning.



The term “rivers”, in red, for example (Figure 7), demonstrates how it functions as an anchor between themes, underscoring that rivers are widely studied in this context. Particularly, 40 out of the 130 articles in the analyzed database include rivers as a keyword, covering a variety of applications such as water quality monitoring in rivers using UAV-based hyperspectral imaging [8], prediction of river pollution and sediment load through machine learning models [7], hydrological modeling of river basins under land-use change [26], and detection of suspended matter and algal blooms in rivers via satellite remote sensing [15]. This prominence reinforces the central role of rivers as key environmental and hydrological systems studied across different methodological approaches.

Figure 7

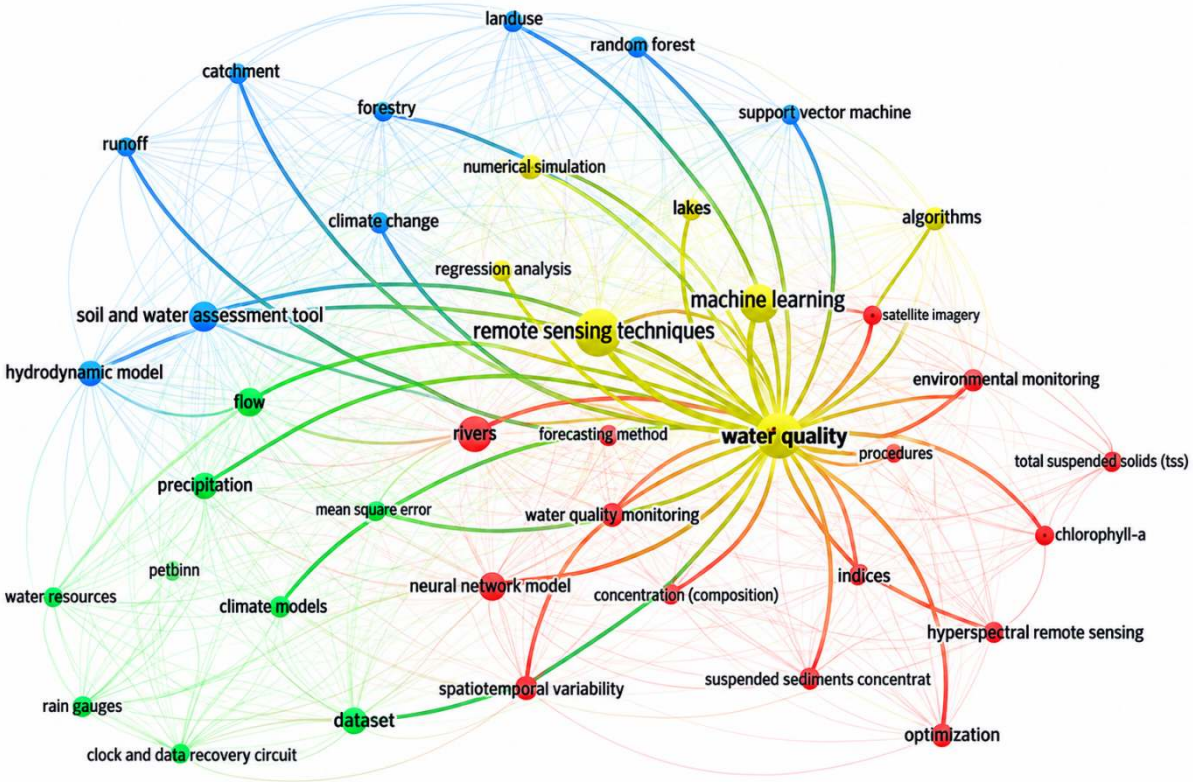
VOSviewer Co-occurrence Analysis, focus on rivers.



The term “water quality”, in yellow, stands out in some representations (Figure 8), signaling it as a main theme, even amid the technological emphasis on AI. In the database analyzed, water quality appears as a keyword in 38 out of 130 articles, evidencing its presence in research on rivers and aquatic systems. These studies employ a variety of approaches, from evaluation of optical water quality indicators [14], water quality prediction with neural networks [9], [27], prediction of nitrates in rivers [28], to monitoring water quality in rivers with UAVs [7], [8]. This centrality underscores that, while computational tools and advanced techniques enhance the analytical capacity of environmental studies, the ultimate focus remains on understanding, preserving, and improving water quality in natural and anthropogenic contexts.

Figure 8

VOSviewer Co-occurrence Analysis, focus on water quality.



5 RESEARCH GAPS AND FUTURE DIRECTIONS

Despite the rapid growth of research on artificial intelligence applications in water quality monitoring, several critical gaps remain that limit the full potential of data-driven environmental monitoring systems.

First, many studies rely on single-source datasets, typically derived either from remote sensing or in situ measurements. Although these approaches provide valuable insights, they often fail to capture the full complexity of aquatic systems. Future research should focus on the integration of heterogeneous data sources, including satellite imagery, UAV-based observations, hydrological models, and field measurements. Such multi-source integration is essential for improving model robustness, spatial coverage, and predictive accuracy.

Second, the issue of model interpretability remains a major challenge. While advanced machine learning and deep learning models offer high predictive performance, they are often considered “black-box” systems. This lack of transparency can limit their adoption in environmental management and policy-making, where explainability is crucial. Future studies

should explore explainable artificial intelligence (XAI) approaches to improve model transparency and support decision-making processes.

Third, the geographic distribution of case studies is uneven. A significant portion of the literature is concentrated in regions with well-established research infrastructure, such as North America, Europe, and parts of Asia. In contrast, regions such as South America and Africa remain underrepresented, despite facing significant water quality challenges. Expanding research efforts to these regions is essential for developing more inclusive and globally relevant environmental monitoring frameworks.

Fourth, although many studies focus on predictive modeling, relatively few address real-time monitoring and operational deployment. The integration of artificial intelligence with Internet of Things (IoT) technologies, sensor networks, and cloud-based platforms represents a promising direction for future research. These systems could enable continuous monitoring, early warning of pollution events, and adaptive environmental management.

Fifth, there is still limited integration between data-driven models and process-based hydrological models. While machine learning approaches excel at pattern recognition, hydrological models provide valuable physical insights into water system dynamics. Future research should explore hybrid modeling approaches that combine the strengths of both methodologies, enhancing both interpretability and predictive capability.

Finally, issues related to data quality, standardization, and reproducibility remain insufficiently addressed. Differences in data acquisition methods, preprocessing techniques, and model evaluation metrics make it difficult to compare results across studies. The development of standardized datasets, benchmarking frameworks, and reproducible workflows is essential for advancing the field and ensuring scientific rigor.

In summary, addressing these research gaps will be critical for advancing the integration of artificial intelligence and geospatial technologies in water quality monitoring. Future research should prioritize interdisciplinary approaches, scalable monitoring systems, and the development of transparent and reliable models capable of supporting sustainable water resource management.

6 IMPLICATIONS FOR ENVIRONMENTAL MONITORING AND POLICY

The findings of this study have important implications for environmental monitoring practices, technological development, and water resource management policies.

First, the increasing integration of artificial intelligence with remote sensing technologies demonstrates the potential for transforming traditional water quality monitoring systems into scalable, data-driven frameworks. These approaches enable continuous observation of aquatic environments over large spatial extents, reducing reliance on labor-intensive field sampling and improving the temporal resolution of environmental data.

Second, the strong centrality of machine learning identified in the bibliometric analysis highlights its role as a key analytical tool for processing complex environmental datasets. The ability of machine learning algorithms to model nonlinear relationships and integrate heterogeneous data sources supports more accurate predictions of water quality parameters. This capability is particularly valuable for early detection of pollution events and for anticipating environmental risks.

Third, the convergence of artificial intelligence, geospatial technologies, and hydrological modeling offers new opportunities for the development of decision-support systems. These systems can assist environmental agencies and policymakers in evaluating water quality scenarios, optimizing resource allocation, and implementing evidence-based management strategies.

From a policy perspective, the adoption of data-driven monitoring systems can significantly enhance the effectiveness of environmental governance. Real-time monitoring platforms, supported by sensor networks and artificial intelligence models, can provide timely information for regulatory enforcement and rapid response to environmental incidents. This is particularly relevant in regions facing increasing water stress and pollution pressures.

Furthermore, the application of artificial intelligence in water quality monitoring can contribute to achieving international sustainability targets, including those established by the Sustainable Development Goals. Improved monitoring capabilities support more accurate assessment of water resources and facilitate the implementation of sustainable management practices.

However, the deployment of these technologies also raises important challenges. Issues related to data quality, model transparency, and technological accessibility must be addressed to ensure equitable and reliable implementation. Policymakers should consider establishing standards for data sharing, model validation, and interoperability across monitoring systems.

Finally, the results of this study suggest that future environmental monitoring strategies should adopt interdisciplinary approaches that integrate data-driven methods with process-

based knowledge. Such integration is essential for developing robust, transparent, and scalable monitoring systems capable of supporting long-term sustainability and adaptive water resource management.

7 CONCLUSION

This study presented a comprehensive bibliometric analysis of scientific publications addressing the application of artificial intelligence in water quality monitoring between 2005 and 2025. The results reveal a clear and sustained increase in scientific production, particularly after 2018, reflecting the growing importance of data-driven approaches in environmental sciences.

The analysis identified four main thematic clusters that structure the research field: machine learning techniques, remote sensing applications, hydrological processes, and water quality assessment. Among these, machine learning emerged as the central analytical component, strongly interconnected with geospatial technologies and environmental monitoring frameworks. The increasing integration of machine learning with remote sensing data demonstrates a significant shift toward scalable and efficient monitoring solutions capable of addressing the complexity of aquatic ecosystems.

From a methodological perspective, the findings highlight the transition from traditional monitoring approaches toward hybrid systems that combine data-driven models with process-based hydrological understanding. This convergence enables improved predictive performance, enhanced spatial coverage, and more effective detection of environmental changes.

Despite these advances, the study also identified important limitations and research gaps. Challenges related to data integration, model interpretability, and real-time monitoring capabilities remain significant barriers to the widespread adoption of artificial intelligence in environmental monitoring. Additionally, the uneven geographic distribution of research efforts underscores the need for broader global participation, particularly in regions facing critical water quality challenges.

The implications of this study extend beyond academic research, contributing to the development of more effective environmental monitoring systems and supporting evidence-based decision-making in water resource management. The integration of artificial intelligence, geospatial technologies, and hydrological modeling has the potential to

significantly enhance the efficiency, scalability, and responsiveness of water quality monitoring frameworks.

Future research should focus on the development of integrated multi-source data systems, the application of explainable artificial intelligence techniques, and the implementation of real-time monitoring platforms supported by sensor networks and cloud-based infrastructures. These advancements will be essential for addressing emerging environmental challenges and for promoting sustainable water resource management in a rapidly changing world.

In conclusion, this study provides a structured and comprehensive overview of the scientific landscape of artificial intelligence applications in water quality monitoring, offering valuable insights into current trends, challenges, and future directions. The results demonstrate that the convergence of artificial intelligence and geospatial technologies represents a transformative paradigm in environmental monitoring, with significant potential to support sustainable development and global water security.

ACKNOWLEDGEMENTS

The authors would like to thank the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) for the financial support provided for this work, under the research project "Fish restocking in the Doce River Basin by aerial robotics" (Grant No. APQ-05441-23).

REFERENCES

[1] R. A. Rather, S. Ara, S. A. Padder, S. Sharma, S. P. Pathak, and T. R. Baba, “Seasonal fluctuation of water quality and ecogenomic phylogeny of novel potential microbial pollution indicators of Veshaw River Kashmir-Western Himalaya”, *Environ. Pollut.*, vol. 320, p. 121104, 2023, doi: https://doi.org/10.1016/j.envpol.2023.121104.

[2] Y. Yang, Y. Yang, S. Yang, and J. Zhang, “Pollution control mechanism of national water network with water quality differences based on differential game and case study”, *Sci. Rep.*, vol. 15, nº 1, p. 5588, feb. 2025, doi: 10.1038/s41598-025-89282-9.

[3] “Goal 6 | Department of Economic and Social Affairs”. Acesso em: 1º de junho de 2025. [Online]. Disponível em: https://sdgs.un.org/goals/goal6

[4] L. Mu, C. Zhang, X. Zeng, R. Ma, Y. Li, and H. Liu, “The impact of the river chief system on transboundary water pollution”, *Sci. Rep.*, vol. 15, nº 1, p. 8192, Mar. 2025, doi: 10.1038/s41598-025-92503-w.

[5] Z. Yigit Avdan, G. Kaplan, S. Goncu, e U. Avdan, “Monitoring the Water Quality of Small Water Bodies Using High-Resolution Remote Sensing Data”, *ISPRS Int. J. Geo-Inf.*, vol. 8, nº 12, Art. nº 12, dez. 2019, doi: 10.3390/ijgi8120553.

[6] J. Moon, S. Jung, S. Suh, and J. Pyo, “Development of deep learning quantization framework for remote sensing edge device to estimate inland water quality in South Korea”, *Water Res.*, vol. 283, p. 123760, 2025, doi: https://doi.org/10.1016/j.watres.2025.123760.

[7] M. Wang *et al.*, “Inversion of Water Quality Parameters from UAV Hyperspectral Data Based on Intelligent Algorithm Optimized Backpropagation Neural Networks of a Small Rural River”, *Remote Sensing*, vol. 17, nº 1. Multidisciplinary Digital Publishing Institute (MDPI), p. 119, 2 de janeiro de 2025. doi: 10.3390/rs17010119.

[8] X. Bai *et al.*, “Research progress of inland river water quality monitoring technology based on unmanned aerial vehicle hyperspectral imaging technology”, *Environmental Research*, vol. 257. Academic Press Inc., p. 119254, setembro de 2024. doi: 10.1016/j.envres.2024.119254.

[9] A. Krishnaraj and R. Honnasiddaiah, “Remote sensing and machine learning based framework for the assessment of spatio-temporal water quality in the Middle Ganga Basin”, *Environmental Science and Pollution Research*, vol. 29, nº 43. Springer Science and Business Media Deutschland GmbH, p. 64939–64958, setembro de 2022. doi: 10.1007/s11356-022-20386-9.

[10] X. Sun *et al.*, “Monitoring water quality using proximal remote sensing technology”, *Science of The Total Environment*, vol. 803. Elsevier B.V., p. 149805, janeiro de 2022. doi: 10.1016/j.scitotenv.2021.149805.

[11] S. H. Rahat *et al.*, “Remote sensing-enabled machine learning for river water quality modeling under multidimensional uncertainty”, *Science of The Total Environment*, vol. 898. Elsevier B.V., p. 165504, novembro de 2023. doi: 10.1016/j.scitotenv.2023.165504.

[12] H. W. Streeter and E. B. Phelps, “A study of the pollution and natural purification of the Ohio River”, *Public Health Bull.*, nº 146, 1925.

[13] Q. Li, Z. Guo, X. Li, e B. Ban, “Application of residual convolutional neural network based on mixed attention mechanism to water quality inversion: a case study of the yellow river basin in Ningxia”, *Proc. SPIE - Int. Soc. Opt. Eng.*, 2025, doi: doi: 10.1117/12.3054418.

[14] S. Gaur, R. Bhattacharjee, S. Chander, A. Ohri, e P. K. Srivastava, “Globally validated non-unique inversion framework to estimate optically active water quality indicators using in situ and space-borne hyperspectral data sets”, *Frontiers of Environmental Science & Engineering*, vol. 19, nº 1. Higher Education Press Limited Company, janeiro de 2025. doi: 10.1007/s11783-025-1930-x.

[15] R. Bhattacharjee, A. Gupta, N. Das, A. K. Agnihotri, A. Ohri, and S. Gaur, “Analysis of algal bloom intensification in mid-Ganga river, India, using satellite data and neural network techniques”, *Environmental Monitoring and Assessment*, vol. 194, nº 8. SPRINGER, VAN GODEWIJCKSTRAAT 30, 3311 GZ DORDRECHT, NETHERLANDS, agosto de 2022. doi: 10.1007/s10661-022-10213-6.

[16] A. Newell and H. A. Simon, “Computer Science as Empirical inquiry: symbols and search”, *Commun. ACM*, vol. 19, nº 3, p. 113–126, 1976, doi: https://doi.org/10.1145/360018.360022.

[17] T. J. Sejnowski, *A revolução do aprendizado profundo*, 1ª. Rio de Janeiro: Alta Books, 2019.

[18] F. M. Silva, M. L. Lenz, P. H. C. Freitas, e S. C. B. Santos, *Inteligência Artificial*. Porto Alegre: SAGAH, 2023.

[19] S. J. Russell e P. Norvig, *Inteligência Artificial: Uma Abordagem Moderna*, 3º ed. Rio de Janeiro: Elsevier, 2013.

[20] G. Kaur, “Neural Networks: A journey into the weights and biases”. Acesso em: 23 de junho de 2025. [Online]. Disponível em: https://medium.com/@gurmeharkaur01/neural-networks-a-journey-into-the-weights-and-biases-39a00ac9265b

[21] M. Aria e C. Cuccurullo, “bibliometrix: An R-tool for comprehensive science mapping analysis”, *J. Informetr.*, vol. 11, nº 4, p. 959–975, 2017, doi: https://doi.org/10.1016/j.joi.2017.08.007.

[22] N. J. van Eck and L. Waltman, “Software survey: VOSviewer, a computer program for bibliometric mapping”, *Scientometrics*, vol. 84, nº 2, p. 523–538, ago. 2010, doi: 10.1007/s11192-009-0146-3.

[23] D. Pozdnyakov, R. Shuchman, A. Korosov, and C. Hatt, “Operational algorithm for the retrieval of water quality in the Great Lakes”, *Remote Sensing of Environment*, vol. 97, nº 3. p. 352–370, 15 de agosto de 2005. doi: 10.1016/j.rse.2005.04.018.

[24] N. Maslej, “Artificial intelligence index report 2025”, Stanford University, California, 2025.

[25] N. J. van Eck and L. Waltman, “VOSviewer manual”.

[26] P. Rajput, M. K. Sinha, and R. Taram, “Assessing Future hydrological response of an urban watershed using machine learning based LULC forecasting models”, *Disaster Advances*, vol. 17, nº 11. World Researchers Associations, p. 35–48, 30 de setembro de 2024. doi: 10.25303/1711da035048.

[27] Y. Kang, J. Song, K. Li, X. Zhai, e Y. Li, “Research on water quality prediction model based on echo state network”, *Journal of Computational Methods in Sciences and Engineering*, vol. 22, nº 3. IOS PRESS, NIEUWE HEMWEG 6B, 1013 BG AMSTERDAM, NETHERLANDS, p. 901–910, 13 de maio de 2022. doi: 10.3233/jcm-225954.

[28] A. Sharifi *et al.*, “Effect of Water Quality Sampling Approaches on Nitrate Load Predictions of a Prominent Regression-Based Model”, *Water*, vol. 9, nº 11. MDPI AG, ST ALBAN-ANLAGE 66, CH-4052 BASEL, SWITZERLAND, p. 895, 16 de novembro de 2017. doi: 10.3390/w9110895.