Journal of Watershed Ecology

SCIEPublish

Perspective

Bridging the Interpretability Gap in Machine Learning for Water Quality

Xinghong Wu, Sangar Khan, Collins Oduro and Naicheng Wu *

School of Geography and Remote Sensing, Zhejiang-Germany Joint Laboratory on Remote Sensing of Coastal Ecosystem, Ningbo University, Ningbo 315211, China; 2511110244@nbu.edu.cn (X.W.); sangar@nbu.edu.cn (S.K.); collins@nbu.edu.cn (C.O.)

* Corresponding author. E-mail: wunaicheng@nbu.edu.cn or naichengwu88@gmail.com (N.W.)

Received: 31 July 2026; Revised: 14 September 2026; Accepted: 13 October 2026; Available online: 30 October 2026

ABSTRACT: Machine learning (ML) has been widely used to predict nonlinear relationships among water quality variables and to address complex ecological problems. However, the ML models' results are sometimes interpreted as real ecological findings. ML tools like explainable artificial intelligence explain model behavior but not the natural processes occurring in the water system. We define this limitation as an "interpretability gap" between the output of ML models and ecological causality. For example, in a random forest model of benthic chlorophyll-a, colloidal phosphorus was the most influential predictor, but algae cannot take up colloidal phosphorus directly. A practical rule is to check whether a variable can affect the response directly before calling it a driver. This can lead to incorrect conclusions, especially when input variables are highly collinear, as in nutrient cycling and across physical and biological processes. This interpretability gap is not only a technical issue in ML models but also a major scientific concern that may lead a new researcher down the wrong track. To address this limitation, we proposed a process-based hydraulic and biogeochemical model alongside ML models, because high accuracy must be supported by ecological understanding. Closing this gap is essential for making ML results more reliable for science and decision-making. We frame this challenge from a watershed perspective, since the watershed is widely recognized as the fundamental spatial unit for integrating the hydrological, biogeochemical, and land-use processes that jointly determine water quality outcomes.

Keywords: Feature importance; Random forest; XGBoost; Water quality; Artificial intelligence

1. Introduction

Clean, adequate water is essential for the survival of living organisms. However, anthropogenic and natural activities have significantly reduced water quality and supply globally [1]. Ignoring water quality assessment threatens the survival of living organisms worldwide [2]. Advanced tools can address this threat by providing low-cost solutions and strategies.

Today, most researchers in environmental science and technology use machine learning (ML) techniques for ecological risk assessment and water health [3]. Several reviews have summarized the use of machine learning in hydrology and water resources [4,5]. ML is an advanced technique that relies less on prior knowledge and can solve complex ecological problems [6,7]. Compared with traditional modeling,

© 2026 The authors. This is an open access article under the Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/).

ML is more accurate for complex problems and provides clearer explanations of the input-output relationships [8]. Recently, ML has been used to predict water quality (Table 1) by predicting chlorophyll-a (a proxy for algal biomass) in freshwater bodies, urban rivers, and lakes. This Perspective therefore focuses on inland waters at the watershed scale. Different ML models have been used. For instance, Wang et al. (2025) [9] used the random forest (RF), artificial neural network (ANN), support vector regression (SVR), and extreme gradient boosting (XGBoost) for water quality prediction in freshwater streams. In addition, Khan et al. (2025a) [8] used the RF and XGBoost for water quality assessment in the Thousand Island catchment area. ML has also been used to assess heavy metals in water bodies. Ullah et al. (2025) [10], Lu et al. (2025) [11], and El Bilali et al. (2021) [12] also investigated emerging pollutants in water bodies and the impact of climate change on groundwater quality using ML models. In addition, Hussain et al. (2025) [13] used ML to control the negative effect of agricultural non-point source pollution on water resources and promote the coordinated development of water resource protection and sustainable agriculture. While ML is a promising technique in water quality assessment, several fundamental challenges need more attention. Many ML models still struggle with explainability and interpretability despite their predictive accuracy.

Table 1. Use of machine learning models in water quality assessment.

Study Area	Machine Learning Models	Prediction Performance (R^2)	Target Variable	References
Imha Reservoir, South Korea	Multiple linear regression (MLR), support vector machine (SVM), Artificial neural network (ANN)	0.80 (SVM, validation)	Chlorophyll-a and Secchi depth	[14]
Babol-Rood River, Iran	M5 Prime (M5P), Random forest (RF), and their hybrids with bagging, random subspace (RS), random committee (RC) and additive regression (AR)	0.995 (AR-M5P, testing)	Electrical conductivity (salinity)	[15]
Hwanggujicheon Stream (urban river), South Korea	Random Forest, Gradient Boosting, adaptive boosting (AdaBoost)	0.912 (AdaBoost, test set)	Dissolved oxygen	[16]
Santiago-Guadalajara River	MLR, Ridge, least absolute shrinkage and selection operator (Lasso), generalized additive model (GAM), Logistic Regression, linear discriminant analysis (LDA)	0.9992 (GAM)	Santiago–Guadalajara River water quality index (SGR-WQI)	[17]
Lake Erie	Bayesian additive regression trees	0.88	Chlorophyll-a	[18]
Water resources of Lianyungang City	Quad-Modality Deep Neural Network	0.73	Chlorophyll-a	[19]
Reservoirs	XGBoost and RF	0.90 (average for both)	Dissolved organic matter (DOM)-associated bacteria and fungi	[20]
Lakes and reservoirs	Deep learning neural network	0.67 for lakes and 0.79 for reservoirs	Chlorophyll a	[21]
Lake Müggelsee	XGBoost; long short-term memory neural network (NN)	0.8 for XGBoost and 0.71 for long short-term memory NN	Cyanobacterial harmful algal bloom (CyanoHAB)	[22]
Water bodies	RF, SVR, XGBoost, Catboost	0.88 (RF), 0.84 (SVR), 0.98 (XGBoost), 0.90 (Catboost)	Emerging inorganic pollutants	[10]

The studies in Table 1 are from 2020 to 2025. Some early studies compared simple models such as MLR, SVM, and ANN, while later studies more often used tree-based models such as random forest and XGBoost. Study areas shifted from single rivers or reservoirs to urban rivers. Target variables have also diversified to include chlorophyll-a, dissolved oxygen, electrical conductivity, and a water quality index.

In water quality research, the shift from traditional modeling to ML and deep learning models has increased significantly, providing highly accurate results [23]. Tree-based and neural network review articles have summarized the use of machine learning in hydrology and water resource modeling, including nonlinear input-output relationships and high-dimensional relationships that traditional models cannot capture [24]. ML and deep learning models are increasingly used for surface, ground, and coastal water quality prediction; however, because of their complexity, their interpretability remains in question [25]. For this reason, interpretability has attracted much attention in geoscience and hydrology [26,27]. Meanwhile, researchers need to pay more attention to this problem. Many studies indicate that ML and deep learning models provide accurate and precise predictions (Table 1) but cannot explain the physical and chemical rationale behind them [28].

These feature-importance results may explain the high prediction performance but offer limited transparency into the model's rationale [29]. Without this distinction, predictive attribution may be interpreted as an ecological explanation. The lack of transparency also has broader implications. Many mechanical processes occur in the water quality system, such as nutrient cycling, hydrological flushing, sediment-water exchange, and biological uptake. When ML models identify predictors (such as phosphorus, temperature, or flow velocity) for water quality (often using chlorophyll-a as a proxy), it is important to determine whether the outputs reflect mechanisms or only statistical correlations. Because of this difference, ML results may not provide ecological explanations. This lack of transparency carries profound consequences for future studies and planning [30]. ML models have been widely used in water purification technology; however, stakeholders need scientifically sound, interpretable answers [31]. Many concerns remain about ML decisions, which may lack transparency and logical coherence [29]. These limitations do not contradict the value of ML in water quality modeling. Rather, they highlight the need for structured interpretability frameworks. Many researchers remain cautious when using ML in contexts where interpreting results matters [32]. ML is important in water quality assessment; however, its limited transparency should be addressed [30].

2. The Interpretability Gap: Attribution Versus Causality

Machine learning models are often called "black boxes" because of their complex internal structures. But for water quality research, what really matters isn't the complexity of the structure itself; it's the lack of transparency in how the results are explained. These black box models often confuse statistical and ecological causality [29]. Most of the literature relies on these black-box models and focuses on performance metrics like R^2, root mean square error (RMSE), and mean absolute error (MAE), which provide limited information about the relationship between model results and hydro-geochemical processes (Figure 1) [33]. Many studies used feature importance ranking, but these analyses rarely explained predictor collinearity, seasonal variability, or stability across cross-validation stages [33]. When machine learning models identify temperature, phosphorus [34], or flow as influential predictors of chlorophyll-a, it is essential to clarify whether the model is capturing mechanistic dominance or statistical association.

Explainable artificial intelligence (XAI) tools identify each predictor's contribution to the model's output. However, XAI did not explain any ecological mechanism [29]. This difference is especially important. Which variables did the ML models use to make their predictions? The feature importance analysis can answer this question. Another question is how the physical, chemical, and biological variables explain ecological processes. In water science, hydro-biogeochemical processes occur in which each variable or predictor depends on the others during a complex process [34]. These questions are not

interchangeable. If this difference between model attribution and causality cannot be explained, it may lead to overconfident policy conclusions.

For example, if Shapley additive explanations (SHAP) values suggest that temperature affects chlorophyll-a [8], this may indicate collinearity between temperature and nutrient stratification rather than direct control by temperature variation. For example, in a random forest model of benthic chlorophyll-a in the Thousand Island Lake watershed, colloidal phosphorus was the most influential predictor (R^2 = 0.88), but algae cannot take up colloidal phosphorus directly, and it must first be released as dissolved inorganic phosphorus [35]. This shows that the model relied on colloidal phosphorus, but it does not mean colloidal phosphorus directly affects algal biomass. A similar problem occurs with temperature. Temperature often ranks as the most important predictor of chlorophyll-a, but it does not supply nutrients on its own; its importance mainly reflects covariation with season, light, and stratification [8]. A practical rule is to check whether a variable can affect the response directly before calling it a driver. Statistical associations from ML models cannot establish causality without mechanistic understanding.

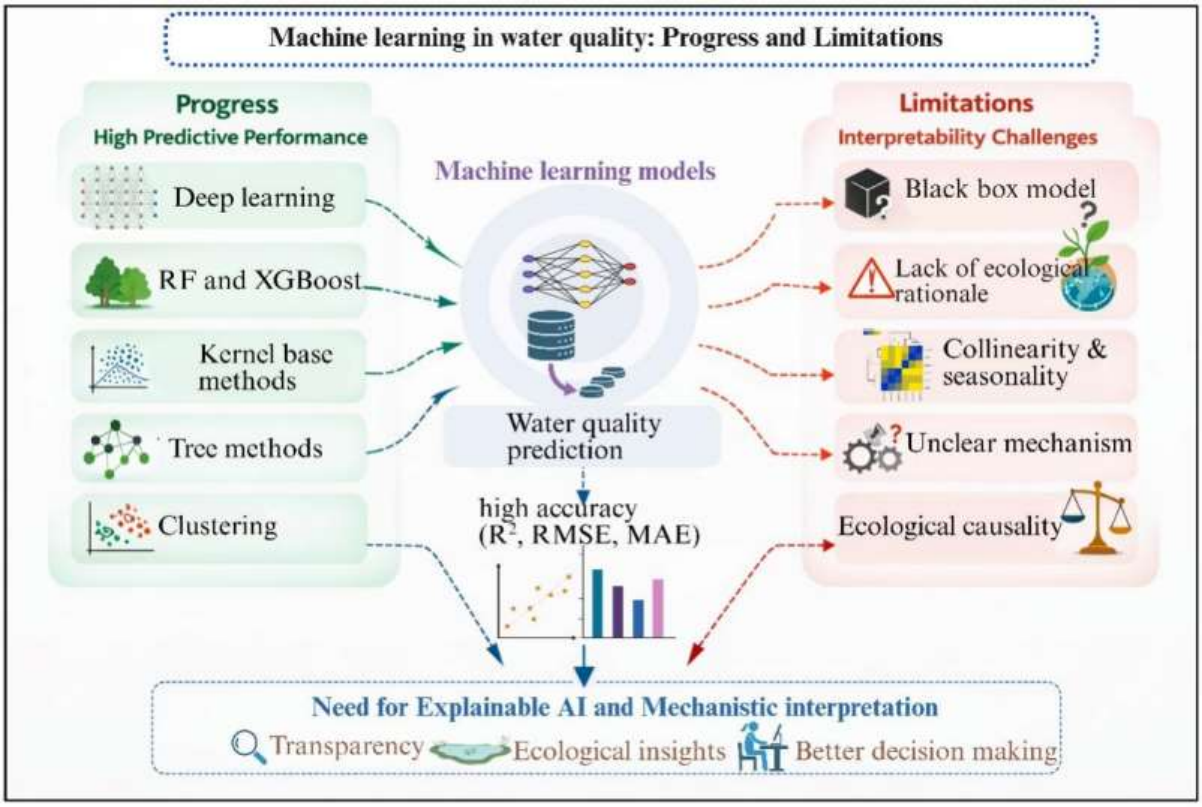


Figure 1. Progress and limitations of machine learning in water quality.

3. Explainable Artificial Intelligence for Water Quality Modeling

The integration of explainable artificial intelligence (XAI) into water quality modeling has emerged as a response to growing concerns about transparency in complex machine learning systems. While ML models can achieve high predictive performance, decision-makers and researchers require insight into how predictions are generated to evaluate their reliability and applicability [36]. Interpretability is therefore not primarily about improving accuracy, but about strengthening transparency, confidence, and scientific defensibility [29]. Some studies have also proposed easy-to-understand models, such as physically interpretable neural networks [37]. Adding XAI to water quality monitoring is important not only for model accuracy but also for improving the interpretability of predictions [36]. Water resource managers and researchers must understand how ML makes predictions to take appropriate action and build confidence in

the results. Among XAI tools, Shapley additive explanations (SHAP) can assign each input feature an exact, additive contribution, making it a powerful tool for transparency. SHAP uses game theory to ensure the input feature's impact on ML predictions [38]. SHAP can explain predictions for the overall dataset or individual inputs by assigning the prediction to each variable. In water studies, SHAP has been used extensively to predict chlorophyll-a and identify threshold levels of nutrients (Nitrogen and Phosphorus) in lakes and other water bodies [9]. For example, SHAP values have recently been used to investigate how P fractions influence algal biomass prediction. SHAP analysis can increase transparency by showing how model inputs contribute to the model output under different conditions [35]. The SHAP values depend on the methodological conditions. More specifically, the contribution estimated by SHAP depends on how the ML models handle missing values or excluded features [39]. In hydro-biogeochemical systems, the variables are highly collinear, whereby importance may be shared among them. As a result, SHAP indicates how the model made the prediction using the input, rather than showing the system's true drivers [39]. So, SHAP interpretations must recognize the distinction between model reliance and system causality. SHAP values explain how ML models behave by showing positive and negative contributions, but they do not validate any mechanistic processes within the hydro-biogeochemical system. Partial dependence plots (PDPs) also play a vital role in water quality assessments, even though they focus on the overall behavior of ML models. PDPs show the relationship between two input features while holding the other variable constant and provide the predicted outcome [38]. Two-way PDP plots can explain interactions between two or three variables and clarify the threshold region of the predictors [24]. Although PDPs are a useful ML tool, typical PDPs rely on marginalization over input and output variables and capture limited interactions between features [10]. As we already mentioned, in the environmental system, the variables are strongly correlated to each other (like temperature and nutrients or the flow velocity, depth, and nutrients), the PDP may produce imaginary combinations and lead to imaginary results. This limitation can be mitigated with accumulated local effect plots, but interpreting the results remains crucial. Like SHAP, PDP explains the model's statistical assumptions rather than the real mechanism in the ecological system [40]. Local Interpretable Model-agnostic Explanations (LIME) can explain a complex ML model using a single observation, helping identify which variable affects a specific model output. In water quality studies, LIME explains bloom alerts, unusual events, or predictions at specific sites [39]. However, LIME focuses on a single prediction and cannot explain how the neighborhood is defined; therefore, it does not provide a complete understanding of the ML model and should not be used to interpret results that explain the ecological system [29].

4. Hybrid Modeling and Governance Implications for Accountable Machine Learning

Here, we do not propose a hybrid model as a novel approach. However, we discuss hybrid models as a potential pathway to explain and reduce the interpretation gap between ML models and real ecological processes. The future of water quality should be defined within a hydrobiological model framework that uses ML models to explain hydrodynamic and biogeochemical processes. This process-based model will provide a mechanistic understanding of ecological processes, explaining the complex, nonlinear interactions between inputs and outputs [41]. The hybrid model framework can reduce the false behavior of ML models, while ML models can improve predictive accuracy under complex or poor ecological conditions [40]. In this kind of modeling, interpretability plays a crucial role, rather than relying only on XAI tool visualizations [30]. The feature importance plot from these models can be used to check whether the relationships built by the hybrid framework are consistent with known hydro-biogeochemical processes [40]. For example, if a variable is especially important in the hybrid model but has no mechanistic interpretation, this may indicate model bias, data imbalance, or model misspecification. In such cases, model interpretability can be assessed by checking whether the outputs are consistent with known, published ecological processes [30]. Combining the hybrid model with interpretability can improve stakeholders' trust in the model assumptions and reduce model uncertainty (Figure 2). Ecological systems have measurement or human error,

and combining ML models can reduce error and improve prediction performance [42]. When interpretability tools are used alongside overfitting or sensitivity analyses, they can improve prediction performance and make model outputs more scientifically meaningful. Water is a fundamental natural resource underpinning life, ecosystems, and human society [43,44], so water quality management requires both technical and institutional alignment in ML. The hybrid ML framework can increase the reliability and interpretability of results, and together they support a shift away from black-box prediction.

Such hybrid frameworks are not only theoretical; they have already been applied in real water bodies. Chen et al. (2024) [45] combined a physics-based model with machine learning (long short-term memory (LSTM), random forest, and SVM) to predict chlorophyll-a in a freshwater lake. They also used Bayesian model averaging to improve the forecast. Huang et al. (2024) [46] coupled a hydrodynamic-based water quality model with an attention-based gated recurrent unit (GRU) network to estimate daily chemical oxygen demand (COD) and total phosphorus loads in the middle-lower Yangtze River. The coupled model performed better than the machine learning model alone (Kling–Gupta efficiency (KGE) = 0.77–0.91 for COD) and reduced parameter uncertainty. These two examples show that combining process-based models with machine learning can improve accuracy in real water bodies. However, the authors used the process model mainly to improve predictions and did not check whether the machine-learning results align with the process model. This Perspective addresses this gap.

However, hybrid frameworks also face practical challenges. Errors from the process-based model can enter the machine-learning model, and process-based models require many parameters that are hard to measure. Running both models together also requires more time, costs more, and requires more data. Other studies discuss these problems [47,48]. So, hybrid modeling may not suit catchments with limited data, but it can be useful when data and resources are sufficient.

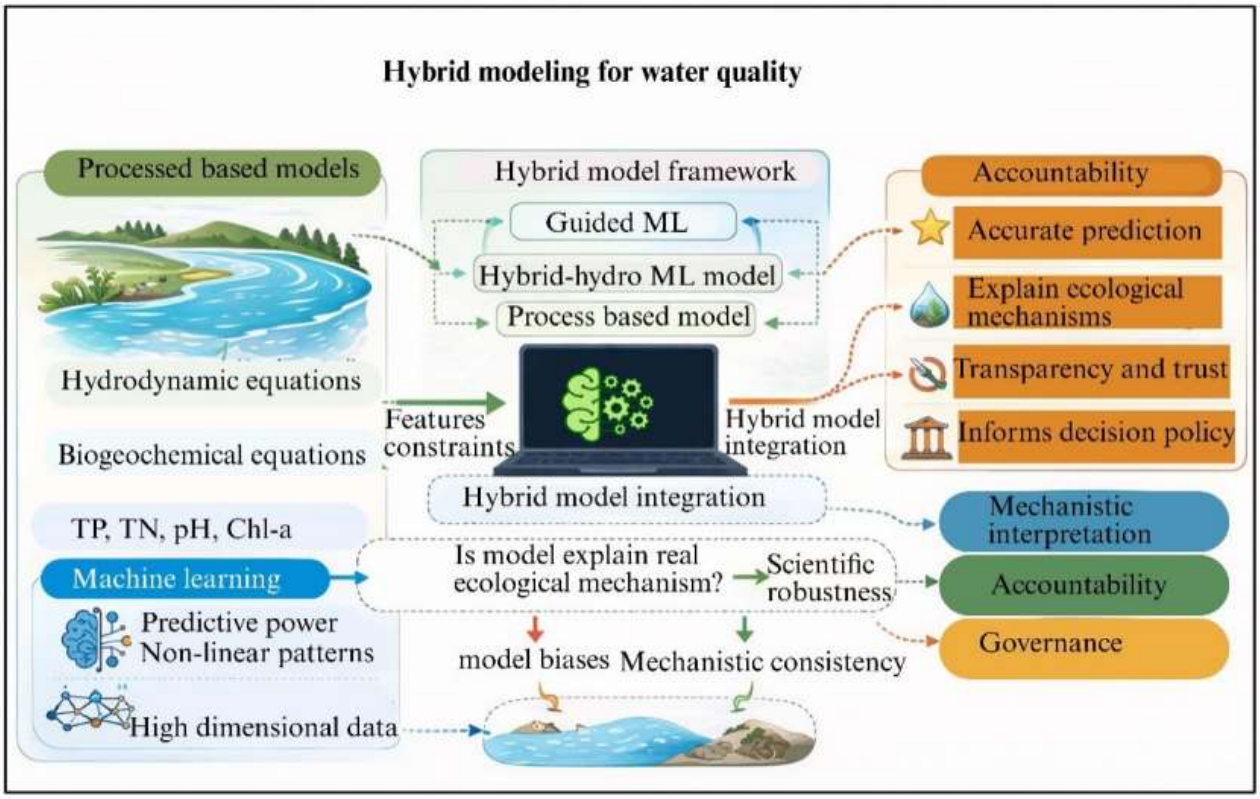


Figure 2. Conceptual framework for hybrid modeling of water quality by integrating process-based knowledge with machine learning (ML). Process-based models provide hydrodynamic and biogeochemical knowledge and impose process-informed feature constraints, whereas ML provides predictive power for nonlinear patterns and high-dimensional data. The hybrid framework integrates these complementary components through guided ML, hybrid hydro-ML modeling, and process-based

modeling to improve the interpretation of model outputs. The framework explicitly evaluates whether model predictions align with ecological mechanisms, including potential model bias and mechanistic consistency, linking predictive performance to mechanistic interpretation, transparency, accountability, and governance. Green arrows indicate process-based knowledge and constraints; blue arrows indicate ML inputs or model integration; orange arrows indicate links from the hybrid model to accountability and decision-relevant outcomes; dashed arrows indicate conceptual or indirect relationships rather than direct causal pathways. TP, total phosphorus; TN, total nitrogen; pH, potential of hydrogen (acidity/alkalinity); Chl-a, chlorophyll-a; ML, machine learning.

5. Conclusions

This Perspective highlights a conceptual challenge that is often overlooked in ML studies: the tendency to interpret model results as ecological causality. ML is a popular tool for predicting water quality and can solve complex environmental problems with high accuracy. However, ML models also have disadvantages in ecological studies. Many ML models cannot interpret their outputs, raising concerns about their transparency and credibility. SHAP, PDP, and LIME have been used to address this limitation, but these XAI tools explain model behavior rather than providing interpretation. We argue that we need hybrid models that combine process-based models and machine learning. Such a hybrid model can increase ML accuracy and improve result interpretability. Using mechanistic data prediction with interpretability practices can improve model reliability and can gain the attention and trust of policymakers. We also suggest that future studies report not only model accuracy, but also whether the important variables can truly affect the response. These findings are particularly relevant for watershed-scale management, where reliable and mechanistically grounded predictions are needed to support sustainable, evidence-based decision-making across the catchment.

Statement of the Use of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this manuscript, the author(s) used Claud to assist with the translation of the article. After using this tool/service, the author(s) reviewed.

Acknowledgments

We thank Zeeshan Zafar (Wuhan University) for his guidance.

Author Contributions

Writing—Original Draft Preparation, X.W.; Writing—Review & Editing, S.K.; Writing—Review & Editing, C.O.; Supervision, N.W.

Ethics Statement

Not Applicable.

Informed Consent Statement

Not Applicable.

Data Availability Statement

Not Applicable.

Funding

The National Natural Science Foundation of China (Grant Nos. 52279068 and 42350410434) provided funding to support the study.

Declaration of Competing Interest

The authors declare no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Mishra S, Chauhan MS, Sundaramurthy S. Assessing groundwater quality dynamics in Madhya Pradesh: Chemical contaminants and their temporal patterns. *Environ. Res.* **2024**, *252*, 118887. DOI:10.1016/j.envres.2024.118887
2. Karthikeyan P, Vennila G, Venkatachalapathy R, Subramani T, Prakash R, Aswini MK. Assessment of heavy metals in the surface sediments of the Emerald Lake using of spatial distribution and multivariate techniques. *Environ. Monit. Assess.* **2018**, *190*, 668. DOI:10.1007/s10661-018-7037-0
3. Ai H, Zhang K, Sun J, Zhang H. Short-term Lake Erie algal bloom prediction by classification and regression models. *Water Res.* **2023**, *232*, 119710. DOI:10.1016/j.watres.2023.119710
4. Shen C. A Transdisciplinary Review of Deep Learning Research and Its Relevance for Water Resources Scientists. *Water Resour. Res.* **2018**, *54*, 8558–8593. DOI:10.1029/2018WR022643
5. Xu T, Liang F. Machine learning for hydrologic sciences: An introductory overview. *WIREs Water* **2021**, *8*, e1533. DOI:10.1002/wat2.1533
6. Liang Z, Soranno PA, Wagner T. The role of phosphorus and nitrogen on chlorophyll a: Evidence from hundreds of lakes. *Water Res.* **2020**, *185*, 116236. DOI:10.1016/j.watres.2020.116236
7. Xia Z, Chen J, Jeppesen E, Guo K, Su G, Liu G, et al. Functional similarity shapes global patterns of fish naturalization and invasion richness. *Proc. Natl. Acad. Sci. USA* **2026**, *123*, e2610486123. DOI:10.1073/pnas.2610486123
8. Khan S, Juvigny-Khenafou NPD, Dalu T, Milham PJ, Hamid Y, Eltohamy KM, et al. The influence of pH and temperature on benthic chlorophyll-a: Insights from SHAP-XGBoost and random forest models. *Ecol. Inform.* **2025**, *91*, 103355. DOI:10.1016/j.ecoinf.2025.103355
9. Wang Y, Khan S, Milham PJ, Lin Z, Qi X, Gao H, et al. Sediment-associated ammonium-N increased benthic chlorophyll-a in streams: Evidence from SHAP integrated machine learning analysis. *Ecol. Indic.* **2025**, *178*, 113933. DOI:10.1016/j.ecolind.2025.113933
10. Ullah H, Khan S, Zhu X, Chen B, Rao Z, Wu N, et al. Machine learning-aided biochar design for the adsorptive removal of emerging inorganic pollutants in water. *Sep. Purif. Technol.* **2025**, *362*, 131421. DOI:10.1016/j.seppur.2025.131421
11. Lu X, Wang Z, Zhao M, Peng S, Geng S, Ghorbani H. Data-Driven Insights into Climate Change Effects on Groundwater Levels Using Machine Learning. *Water Resour. Manag.* **2025**, *39*, 3521–3536. DOI:10.1007/s11269-025-04120-x
12. El Bilali A, Taleb A, Brouziyne Y. Groundwater quality forecasting using machine learning algorithms for irrigation purposes. *Agric. Water Manag.* **2021**, *245*, 106625. DOI:10.1016/j.agwat.2020.106625
13. Hussain F, Ahmed S, Muhammad Zaigham Abbas Naqvi S, Awais M, Zhang Y, Zhang H, et al. Agricultural Non-Point Source Pollution: Comprehensive Analysis of Sources and Assessment Methods. *Agriculture* **2025**, *15*, 531. DOI:10.3390/agriculture15050531
14. Mamun M, Kim JJ, Alam MA, An KG. Prediction of Algal Chlorophyll-a and Water Clarity in Monsoon-Region Reservoir Using Machine Learning Approaches. *Water* **2019**, *12*, 30. DOI:10.3390/w12010030
15. Melesse AM, Khosravi K, Tiefenbacher JP, Heddam S, Kim S, Mosavi A, et al. River Water Salinity Prediction Using Hybrid Machine Learning Models. *Water* **2020**, *12*, 2951. DOI:10.3390/w12102951
16. Moon J, Lee J, Lee S, Yun H. Urban River Dissolved Oxygen Prediction Model Using Machine Learning. *Water* **2022**, *14*, 1899. DOI:10.3390/w14121899
17. Fernández del Castillo A, Yebra-Montes C, Verduzco Garibay M, de Anda J, Garcia-Gonzalez A, Gradilla-Hernández MS. Simple Prediction of an Ecosystem-Specific Water Quality Index and the Water Quality Classification of a Highly Polluted River through Supervised Machine Learning. *Water* **2022**, *14*, 1235. DOI:10.3390/w14081235
18. Isabwe A, Maguire TJ, Stow CA, Godwin CM. Lake Erie summer chlorophyll phenology: A Bayesian additive regression trees comparison of growth and decay phases. *Water Res.* **2025**, *282*, 123770. DOI:10.1016/j.watres.2025.123770
19. He H, Li X, Wang D, Qiao W, Sun Y, Han Y, et al. A novel quad-modality deep neural network for estimating chlorophyll-a concentrations in Lianyungang's lakes and reservoirs using Sentinel-2 MSI data. *Water Res.* **2025**, *286*, 124246. DOI:10.1016/j.watres.2025.124246
20. Shi K, Zhang J, Wu C, Zhao Y, Li W, Huo R, et al. Vertical dynamics of DOM-specialized bacteria and fungi drive stability in stratified reservoirs: Mechanisms revealed by machine learning. *Water Res.* **2025**, *287*, 124334. DOI:10.1016/j.watres.2025.124334

21. Ma Y, Wang J, Huo S, Wang D, Wang Y, Li J, et al. Explainable machine learning reveals climate warming increases risk of algal blooms in lakes and reservoirs. *Water Res.* **2025**, *287*, 124460. DOI:10.1016/j.watres.2025.124460
22. Kim HG, Cho KH, Recknagel F. Bibliometric network analysis of scientific research on early warning signals for cyanobacterial blooms in lakes and rivers. *Ecol. Inform.* **2024**, *80*, 102503. DOI:10.1016/j.ecoinf.2024.102503
23. Zhu M, Wang J, Yang X, Zhang Y, Zhang L, Ren H, et al. A review of the application of machine learning in water quality evaluation. *Eco-Environ. Health* **2022**, *1*, 107–116. DOI:10.1016/j.eehl.2022.06.001
24. Khan S, Gao H, Milham P, Eltohamy KM, Ullah H, Mu H, et al. Predicting the governing factors for the release of colloidal phosphorus using machine learning. *Chemosphere* **2024**, *362*, 142699. DOI:10.1016/j.chemosphere.2024.142699
25. Sun AY, Scanlon BR. How can Big Data and machine learning benefit environment and water management: A survey of methods, applications, and future directions. *Environ. Res. Lett.* **2019**, *14*, 073001. DOI:10.1088/1748-9326/ab1b7d
26. McGovern A, Lagerquist R, John Gagne D, Jergensen GE, Elmore KL, Homeyer CR, et al. Making the Black Box More Transparent: Understanding the Physical Implications of Machine Learning. *Bull. Am. Meteorol. Soc.* **2019**, *100*, 2175–2199. DOI:10.1175/BAMS-D-18-0195.1
27. Zounemat-Kermani M, Kheimi M. Explainable Artificial Intelligence in Hydrology: A Review. *Water Resour. Manag.* **2026**, *40*, 106. DOI:10.1007/s11269-025-04435-9
28. Pyo J, Pachepsky Y, Kim S, Abbas A, Kim M, Kwon YS, et al. Long short-term memory models of water quality in inland water environments. *Water Res. X* **2023**, *21*, 100207. DOI:10.1016/j.wroa.2023.100207
29. Rudin C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **2019**, *1*, 206–215. DOI:10.1038/s42256-019-0048-x
30. Aderemi IA, Kehinde TO, Daniel Okwor U, Ahmad KH, Adjei KY, Cyriacus Ekechi C. Explainable AI for Water Quality Monitoring: A Systematic Review of Transparency, Interpretability, and Trust. *IEEE Sens. Rev.* **2025**, *2*, 419–443. DOI:10.1109/SR.2025.3595500
31. Preece A, Harborne D, Braines D, Tomsett R, Chakraborty A. Stakeholders in Explainable AI. *arXiv* **2018**, arXiv:1810.00184. DOI:10.48550/arXiv.1810.00184
32. Imrie F, Davis R, van der Schaar M. Multiple stakeholders drive diverse interpretability requirements for machine learning in healthcare. *Nat. Mach. Intell.* **2023**, *5*, 824–829. DOI:10.1038/s42256-023-00698-2
33. Zheng H, Liu Y, Wan W, Zhao J, Xie G. Large-scale prediction of stream water quality using an interpretable deep learning approach. *J. Environ. Manag.* **2023**, *331*, 117309. DOI:10.1016/j.jenvman.2023.117309
34. Khan S, Wang Y, Milham PJ, Lin Z, Qi X, Gao H, et al. Effect of stormwater and land use on biogeochemical transformations of dissolved and particulate phosphorus in freshwater systems. *Water Res.* **2025**, *284*, 124007. DOI:10.1016/j.watres.2025.124007
35. Wang Y, Khan S, Lin Z, Qi X, Eltohamy KM, Oduro C, et al. Effect of phosphorus fractions on benthic chlorophyll-a: Insight from the machine learning models. *Ecol. Inform.* **2025**, *85*, 102990. DOI:10.1016/j.ecoinf.2024.102990
36. Baroni G, Francke T. An effective strategy for combining variance- and distribution-based global sensitivity analysis. *Environ. Model. Softw.* **2020**, *134*, 104851. DOI:10.1016/j.envsoft.2020.104851
37. Toms BA, Barnes EA, Ebert-Uphoff I. Physically Interpretable Neural Networks for the Geosciences: Applications to Earth System Variability. *J. Adv. Model. Earth Syst.* **2020**, *12*, e2019MS002002. DOI:10.1029/2019MS002002
38. Infant SS, Vickram S, Saravanan A, Mathan Muthu CM, Yuarajan D. Explainable artificial intelligence for sustainable urban water systems engineering. *Results Eng.* **2025**, *25*, 104349. DOI:10.1016/j.rineng.2025.104349
39. Chen H, Covert IC, Lundberg SM, Lee SI. Algorithms to estimate Shapley value feature attributions. *Nat. Mach. Intell.* **2023**, *5*, 590–601. DOI:10.1038/s42256-023-00657-x
40. Başağaoğlu H, Chakraborty D, Lago CD, Gutierrez L, Şahinli MA, Giacomoni M, et al. A Review on Interpretable and Explainable Artificial Intelligence in Hydroclimatic Applications. *Water* **2022**, *14*, 1230. DOI:10.3390/w14081230
41. Zanoni MG, Majone B, Bellin A. A catchment-scale model of river water quality by Machine Learning. *Sci. Total Environ.* **2022**, *838*, 156377. DOI:10.1016/j.scitotenv.2022.156377
42. Nearing GS, Kratzert F, Sampson AK, Pelissier CS, Klotz D, Frame JM, et al. What Role Does Hydrological Science Play in the Age of Machine Learning? *Water Resour. Res.* **2021**, *57*, e2020WR028091. DOI:10.1029/2020WR028091
43. Cai QH. Why Watershed Ecology? *J. Watershed Ecol.* **2026**, *1*, 10001. DOI:10.70322/jwe.2026.10001
44. Xu YY. What Should Ecology Study for Watershed? *J. Watershed Ecol.* **2026**, *1*, 10011. DOI:10.70322/jwe.2026.10011
45. Chen C, Chen Q, Yao S, He M, Zhang J, Li G, et al. Combining physical-based model and machine learning to forecast chlorophyll-a concentration in freshwater lakes. *Sci. Total Environ.* **2024**, *907*, 168097. DOI:10.1016/j.scitotenv.2023.168097
46. Huang S, Xia J, Wang Y, Wang G, She D, Lei J. Pollution loads in the middle-lower Yangtze river by coupling water quality models with machine learning. *Water Res.* **2024**, *263*, 122191. DOI:10.1016/j.watres.2024.122191

47. Karpatne A, Ebert-Uphoff I, Ravela S, Babaie HA, Kumar V. Machine Learning for the Geosciences: Challenges and Opportunities. *IEEE Trans. Knowl. Data Eng.* **2019**, *31*, 1544–1554. DOI:10.1109/TKDE.2018.2861006
48. Willard J, Jia X, Xu S, Steinbach M, Kumar V. Integrating Scientific Knowledge with Machine Learning for Engineering and Environmental Systems. *ACM Comput. Surv.* **2022**, *55*, 1–37. DOI:10.1145/3514228