European Journal of Sustainable Development Research
2025, 9(4), em0333
e-ISSN: 2542-4742
https://www.ejosdr.com/

MODESTUM

Research Article **OPEN ACCESS**

Using explainable AI to predict algal bloom forecasting and real-time analysis of deteriorating water quality in aquatic ecosystems

Sivaranjini Raja [1*], Sharanya S [1], Rajkumar V [2]

[1] Department of Data Science and Business Systems, School of Computing, SRM Institute of Science & Technology, Kattankulathur, Chennai, INDIA
[2] Department of Computer Science and Engineering, Krishnasamy College of Engineering and Technology, Cuddalore, INDIA
***Corresponding Author:** ranjiniraja13@gmail.com

Citation: Raja, S., S, S., & V, R. (2025). Using explainable AI to predict algal bloom forecasting and real-time analysis of deteriorating water quality in aquatic ecosystems. *European Journal of Sustainable Development Research, 9*(4), em0333. https://doi.org/10.29333/ejosdr/16863

ARTICLE INFO

Received: 10 Mar. 2025

Accepted: 16 Jul. 2025

ABSTRACT

Water resource management, public health, and aquatic environments are facing serious challenges due to the increasing frequency of algal blooms and declining water quality. The effectiveness of conventional prediction models in real-time applications is limited, as they often lack transparency and fail to account for the complex interrelationships among environmental variables. This research presents an innovative framework leveraging explainable artificial intelligence to enable real-time environmental assessments and reliable prediction of algal blooms. The approach integrates model interpretation techniques such as Shapley additive explanations and local interpretable model-agnostic explanations with advanced machine learning methods, including hybrid model ensembles and deep learning techniques. By combining these methodologies, the framework offers valuable insights into the processes driving bloom formation and water quality degradation while delivering high forecasting accuracy. Enhanced early-warning systems are developed to enable timely interventions and promote sustainable water conservation practices. Experimental results on diverse datasets demonstrate that the proposed approach achieves a prediction accuracy exceeding 95%. Interpretability metrics highlight key environmental factors such as temperature, dissolved oxygen levels, and nutrient concentrations. This work bridges the gap between model transparency and predictive accuracy, fostering trust in artificial intelligence-driven solutions for water quality management and environmental protection.

Keywords: explainable artificial intelligence, algal bloom prediction, water quality assessment, real-time monitoring, environmental sustainability

INTRODUCTION

Water covers nearly two-thirds of the Earth's surface and is a vital resource for sustaining life. The supply of freshwater remains limited, with surface water accounting for a significant portion of the global total. This situation is further exacerbated by the ongoing degradation of water quality caused by both point and non-point pollution sources (Mermer et al., 2024). Accurate prediction and management of external water quality are therefore critical. Assessing water quality requires the collection and analysis of extensive data across numerous environmental indicators. Summarizing overall water quality through a single, unified metric remains a persistent challenge (Demiray et al., 2024).

Among the key environmental threats are harmful algal blooms (HABs), primarily caused by algae or cyanobacteria (blue-green algae) have emerged as a major issue in aquatic systems worldwide. HABs pose substantial risks to aquatic ecosystems, human health, and overall water quality (Mahto, 2024). The primary drivers of HABs include nutrient pollution from industrial discharges and agricultural runoff and climate-related factors such as rising water temperatures and changing physicochemical conditions. These blooms are known to release harmful toxins, jeopardize drinking water supplies, and reduce the recreational and aesthetic value of affected water bodies (Marry et al., 2024).

The frequency and severity of HABs have escalated significantly over recent decades, driven by widespread agricultural practices, rapid urbanization, increasing air pollution, and the broader impacts of climate change. This alarming trend underscores the urgent need for advanced techniques in HAB monitoring, simulation, and forecasting to protect water resources and public health (Natarajan et al., 2024). Existing process-based simulations such as the water quality analysis simulation Pprogram, the environmental fluid dynamics code, and QUAL2K attempt to link meteorological, hydrodynamic, and physicochemical parameters to HAB indicators. These models often fall short in capturing complex nonlinear interactions and typically lack the predictive accuracy required for timely and reliable forecasting (Krishna et al., 2021).

Both process-based models and statistical frameworks contribute to understanding the dynamics of HABs but often struggle to capture the intricate interactions necessary for reliable predictions across diverse environmental scenarios. After a bloom peaks, the decomposition of organic matter initiates complex feedback loops within the carbon cycle (Rezk et al., 2024). During the microbial breakdown of dead algal cells, greenhouse gases such as carbon dioxide (CO_2) and methane (CH_4) are released into the atmosphere or water column especially under hypoxic or anoxic conditions. This decomposition process exacerbates oxygen depletion, further degrading the aquatic environment and partially negating the initial carbon capture achieved by the algae (Zhi et al., 2024). In some instances, the sinking of algal biomass to the ocean floor can result in long-term carbon burial, depending on sedimentation rates and climatic conditions. These processes highlight the multifaceted environmental impacts of HABs and underscore the critical need for effective monitoring and forecasting tools to mitigate their adverse effects on water quality and public health (Lin et al., 2024).

Manual calculations and laboratory analyses of large volumes of water quality data are often time-consuming, inefficient, and costly. To address this challenge, intelligent methods such as machine learning (ML) are increasingly employed, particularly in scenarios that demand real-time prediction capabilities (Martín-Suazo et al., 2024). ML, a core subset of artificial intelligence (AI) enables systems to learn from data and improve performance without the need for explicit programming. It is being widely adopted in scientific domains such as water quality assessment. Interpreting the results and understanding the logic behind ML-driven research can be challenging, especially for non-experts (Farzana et al., 2024).

Regulators and end users often prefer traditional methods that are perceived as more transparent and easier to interpret, despite the superior efficiency offered by ML models. This highlights the critical importance of incorporating explainability and transparency into complex ML algorithms (Olmo Bau, 2024). Explainable ML not only helps domain experts and stakeholders to understand, trust, and validate model decisions but also ensures the reliability and accuracy of predictions. These improvements are essential for broader acceptance and the successful integration of ML-based solutions in water quality monitoring and other critical environmental applications (Yao et al., 20234).

RELATED WORKS

Globally, cyanobacterial biomass in aquatic systems has significantly increased due to environmental pollution, leading to the deterioration of water quality. HABs not only produce toxic byproducts but also diminish the aesthetic value of water bodies and hinder the supply of safe drinking water. Real-time water quality information can be effectively obtained through sensors deployed directly in water offering high temporal resolution. These sensors monitor a range of physical and chemical parameters associated with HAB dynamics, including pH, temperature, dissolved oxygen, conductivity, and turbidity. Specialized sensors can detect chlorophyll-a (Chl-a), a pigment widely regarded as a proxy indicator for algal concentration at the water surface (Dharmasa et al., 2017).

To facilitate real-time data collection and transmission, these sensors are often integrated into buoys as part of automatic high-frequency monitoring systems. The vast and high-resolution datasets generated by these water quality sensors meet the computational requirements of AI-driven methods for complex water assessment tasks (Hemanth et al., 2024). ML and DL models have been developed using time-series data to infer Chl-a concentrations and algal cell densities. These models serve as early-warning mechanisms, triggering alerts before algal concentrations reach critical thresholds enabling timely intervention (Vardhini et al., 2019).

The World Health Organization (WHO) has outlined standard alert thresholds for short-term responses based on environmental context, algal species, abundance, Chl-a concentration, and toxin levels. At the vigilance level, routine monitoring is intensified. Alert level 1 includes weekly surveillance and notifications when algal cell counts begin to rise (Ye et al., 2024). Higher concentrations, visible blooms, or elevated toxin levels may escalate warnings and prompt public advisories. Under severe conditions, alert level 2 necessitates immediate actions such as restricting access to affected water bodies and issuing evacuation instructions.

For recreational waters, Chl-a levels between 3 and 12 µg/L correspond to the awareness phase, while levels between 12 and 24 µg/L or higher–especially in the presence of toxin-producing species–trigger alert 1 and trigger alert 2 (Saude & Vardhini, 2020). For drinking water, the threshold is more stringent, with the vigilance level initiated at 1 µg/L, and alert levels activated when concentrations exceed 1-12 µg/L (Saravani et al., 2024). Although specific thresholds and protocols may vary geographically, the underlying need for continuous awareness, timely communication, and collaborative decision-making remains universal.

The deployment of predictive models is a key strategy for implementing cost-effective and accurate early-warning systems. These models enhance sampling efficiency and help automate alarm triggering, offering a proactive approach to managing HABs in vulnerable water bodies (Khonina et al., 2024). Population growth, agricultural expansion, environmental contamination, and rising temperatures have collectively contributed to a significant increase in HAB events over the past few decades. This alarming trend highlights the urgent need to enhance HAB surveillance, modeling, and predictive capabilities to protect both water resources and public health. Effective HAB monitoring typically involves laboratory testing for indicators such as Chl-a, cyanobacteria, and algal toxins using advanced techniques, including liquid chromatography, microscopy, spectrophotometry, and biochemical assays (Sun et al., 2024). Understanding the underlying causes of algal blooms and implementing effective early prediction strategies are essential for environmental and health protection.

Recent research has demonstrated the potential of long short-term memory (LSTM) models in forecasting Chl-a concentrations, such as the study conducted along the west coast of Sabah using moderate resolution imaging spectroradiometer (MODIS) and general bathymetric chart of the oceans data (Nagpal et al., 2024). MODIS offers high temporal resolution (daily), enabling the timely detection of dynamic surface-level spectral changes that indicate algal proliferation under favorable conditions. Existing literature focuses solely on LSTM performance without benchmarking it against other widely used ML models not specifically designed for time-series prediction (Sandeep & Prakasam, 2019).

The deterioration of water quality is influenced by a combination of natural and anthropogenic factors. Natural processes such as water-rock interactions, chemical speciation, hydrodynamic shifts, and rainfall variability alter aquifer and surface water characteristics over time (Koronides et al., 2024). Meanwhile, human-induced activities include point-source pollution (e.g., industrial effluents) and non-point-source pollution (e.g., agricultural runoff, improper sewage disposal, and public defecation), all contributing to widespread degradation (Yu et al., 2024).

Given the multifaceted nature of water pollution, assessing water quality often involves labor-intensive data collection, lab analysis, and complex data interpretation. Over the years, both conventional and unconventional methods have been adopted. Traditional methods rely on regional and international standards for classification, while non-traditional approaches incorporate quantitative indicators, parametric/non-parametric statistical models, health risk assessment frameworks, and spatiotemporal analysis using GIS-based techniques (Fameso et al., 2024). However, a key limitation of many mathematical frameworks lies in the complexity of calculating sub-indices, which requires specialized expertise. Miscalculations in these frameworks can lead to inaccurate outcomes, thereby compromising decision-making processes and potentially impacting broader economic and environmental management strategies (Li et al., 2024).

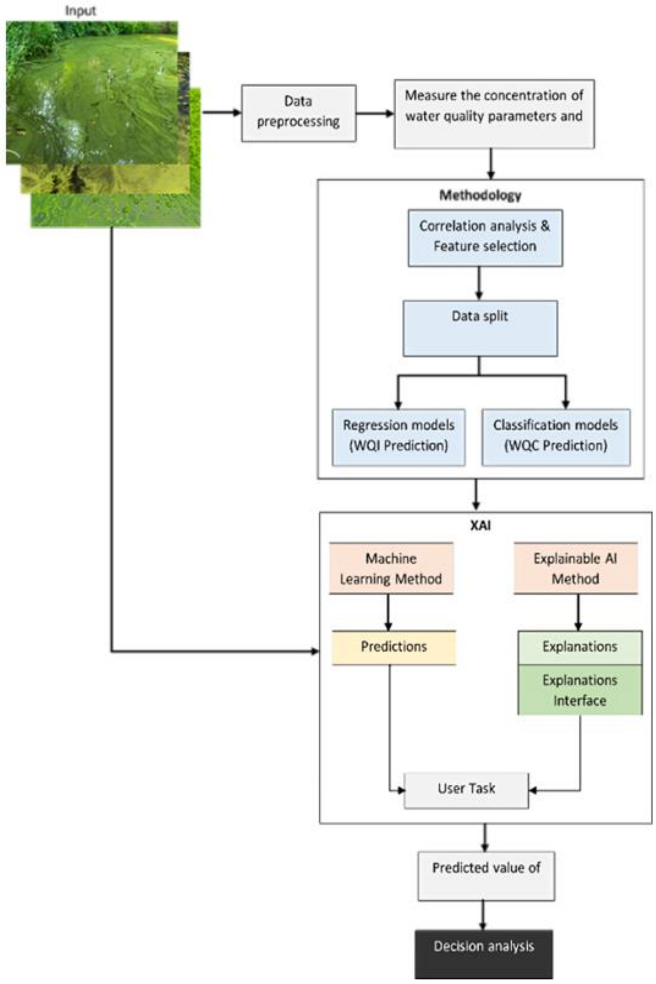

Figure 1. Proposed architecture (Source: Authors' own elaboration)

MATERIALS AND METHODS

The availability of water, the preservation of biodiversity, and public health are increasingly threatened by algal blooms, which are primarily driven by pollution, rising temperatures, and fertilizer runoff. Although ML techniques have shown promise in forecasting algal blooms, many existing models lack transparency and are difficult to interpret. This lack of interpretability makes it challenging to understand the underlying environmental variables and the complex interactions within ecosystems that contribute to bloom formation.

In addition to forecasting algal bloom development, the proposed approach identifies key environmental factors contributing to water quality degradation by integrating ensemble learning, neural networks, and explainability tools such as Shapley additive explanations (SHAP) and local interpretable model-agnostic explanations (LIME). To ensure both high predictive accuracy and interpretability, this research introduces a novel framework that combines explainable artificial intelligence (XAI) techniques with advanced ML algorithms, as illustrated in Figure 1. The primary objective is to establish a reliable, continuous monitoring system that equips environmental scientists, policymakers, and water resource managers with actionable insights empowering them to make informed decisions and take proactive measures to safeguard aquatic ecosystems.

Dataset Description

The extensive information about the environment set for utilizing XAI for accurate forecast of algal blooms and real-time evaluation of deteriorating water quality in aquatic ecological systems was gathered from a variety of sources that include observation stations, networks of sensors, and satellite imaging platforms shown in Table 1.

It covers a range of freshwater habitats, including waterways, lakes, and storage tanks, in diverse geographical locations across the world for a period of one to five years. High-frequency (every 10 minutes to 1 hour) observations of important aspects of the environment, such as pH, temperatures, oxygen dissolution, turbidity, nutrient concentrations (phosphate and fertilizer), chl levels, and light intensity are included in the dataset. Satellite imagery in GeoTIFF format offers regional evaluations of geographical coverage. A water quality indicator expressed as a continuously numerical value bloom intensity (concentration levels), and algal bloom occurrence (binary classification: yes/no) are the goal parameters.

Table 1. Dataset description

No	Dataset name	Source/platform	Parameters collected	Temporal resolution	Spatial coverage
1	Remote sensing water quality dataset	MODIS-Aqua (NASA) & GEBCO	Chlorophyll-a, sea surface temperature, turbidity, reflectance, & NDVI	Daily	Coastal & inland waters (e.g., Sabah, Malaysia)
2	In-situ water quality monitoring dataset	EPA, NOAA, & local hydrology dept	pH, DO, temperature, conductivity, nitrate, & phosphate	Hourly to weekly	Lakes, rivers, & reservoirs (location-specific)
3	Automatic high-frequency monitoring	IoT buoy sensor networks	Chl-a, temperature, conductivity, DO, & turbidity	15 min-1 hour	Deployed sensors in key water bodies
4	Cyanobacterial bloom reports	WHO & environmental health agencies	Algal species, toxin levels, health advisories, warning levels, & geo-location	Event-based	Recreational & drinking water zones
5	Historical climate and pollution data	NOAA, ECMWF, & local govt sources	Rainfall, temperature, wind speed, industrial discharge reports, & land use data	Daily to monthly	Watershed-level and regional zones

Table 2. Sample data

Timestamp	Location	Temperature (°C)	pH	DO (mg/L)	Nitrate (mg/L)	Phosphate (mg/L)	Turbidity (NTU)	Chl (µg/L)	Light intensity (lux)	Algal bloom occurence	Water quality index
2024-09-02-00:00	Lake-1	23.5	7.3	5.7	1.3	0.05	16	21	810	No	0.86
2024-09-02-01:00	Lake-1	23.6	7.4	5.6	1.2	0.06	17	26	830	No	0.88
2024-09-02-02:00	River-3	22.8	7.5	6.2	1.4	0.04	21	.16	760	Yes	0.93
2024-09-02-03:00	River-3	22.9	7.4	6.1	1.5	0.05	23	31	740	Yes	0.96
2024-09-02-04:00	Lake-5	23.0	7.2	5.8	1.3	0.06	19	23	820	No	0.89

To provide thorough coverage of interactions between organisms, the dataset also takes into consideration seasonal fluctuations and environmental elements including weather, pollution information, and geographic regions. By cross-validating physical measurements and satellite imaging, missing or corrupted points of information brought on by sensor failures or ecological abnormalities are reduced, guaranteeing information quality and dependability. Using ML models combined with XAI, this sample offers a quick overview of important ecological indicators that are pertinent to forecasting algal blooms and evaluating the condition of water and sample data shown in **Table 2**.

Pre-Processing

Chemical oxygen demand, dissolved oxygen (DO), biochemical oxygen demand, chloride (Cl-), total nitrate (NO_3), total phosphate (PO_4-), pH, and other water quality characteristics are also included in the dataset. The collected information underwent preprocessing included the first stage of outlier identification, before processing. Outliers were removed from the information using the z-score approach. This approach uses standard deviations to quantify how much the information deviates from the mean. As shown in **Figure 2**, the water quality attributes were visually shown to examine the information's distribution and identify any possible anomalies. After outliers were found and eliminated, the values that were missing were removed from the information set. The researchers did this to maintain the information's creativity, and as the percentage of values that were lacking in the information set was less than 10%, no investigation was carried out to estimate these absent values.

In real-world datasets, missing data can compromise model performance and result in biased or inaccurate predictions. This section describes methods to handle missing values in the dataset for the accurate prediction of algal blooms and water quality assessment.

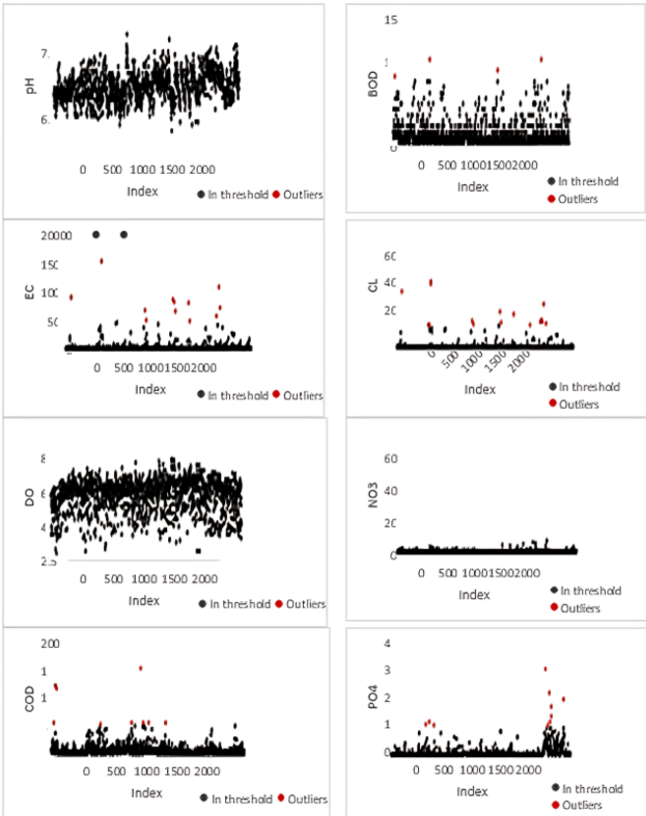

Figure 2. Outliers data detected represented in red dots (water quality parameters) (Source: Authors' own elaboration)

Mean/median imputation

For numerical features, missing values can be replaced by the mean or median of the available data. Let I_x be a numerical feature with missing values. The imputation can be done as follows (mean imputation) (Demiray et al., 2024):

$$I_x^{imputed} = \mu_x = \frac{\sum_{y=1}^{m} I_{xy}}{m} \; for\ MCAR\ or\ MAR, \quad (1)$$

where μ_x is the mean of feature I_x.

Median imputation (Demiray et al., 2024) can be computed as follows:

$$I_x^{imputed} Median(I_x)\ (more\ robust\ to\ outliers). \quad (2)$$

Mode imputation (categorical data)

For categorical features, missing values are replaced by the mode, which is the most frequently occurring value in the dataset (Demiray et al., 2024):

$$I_x^{imputed} = Mode(I_x). \quad (3)$$

For instance, if "light intensity" has some missing data, replace the missing entries with the most common light intensity value.

Linear interpolation

For time series or sequential data, missing values are often replaced by linear interpolation between neighboring observations. For a given missing value I_t at time t, with observations at time *t*-1 and *t*+1 (Mahto, 2024):

$$I_x^{imputed} = I_{t-1} + \frac{(I_{t+1} - I_{t-1}).(t-(t-1))}{(t+1)-(t-1)}, \quad (4)$$

where I_{t-1} and I_{t+1} are the observed values before and after the missing timestamp 1.

K-nearest neighbors imputation

Missing values are replaced with the average value of the K nearest neighbors based on Euclidean distance. For a given feature I_x (Demiray et al., 2024):

$$I_x^{imputed} = \frac{\sum_{k=1}^{K} I_{xk}}{K}, \quad (5)$$

where I_{xk} are the values of the K nearest neighbors to I_x and K is the hyperparameter representing the number of neighbors.

Drop missing data

If the missing values are too extensive (greater than a threshold, say $\alpha >$ 20%), it's often better to drop those rows or columns:

- Drop entire rows if too much data for a specific observation is missing.
- Drop entire columns if a large portion of data in a single environmental factor is missing.

Let's assume that for any row I_y if p_y (percentage of missing features) exceeds a predefined threshold α, the row is removed (Demiray et al., 2024):

$$p_y = \frac{Number\ of\ Missing\ Features\ in\ Row\ y}{n}. \quad (6)$$

If $p_y > \alpha$, row I_y is dropped from the dataset.

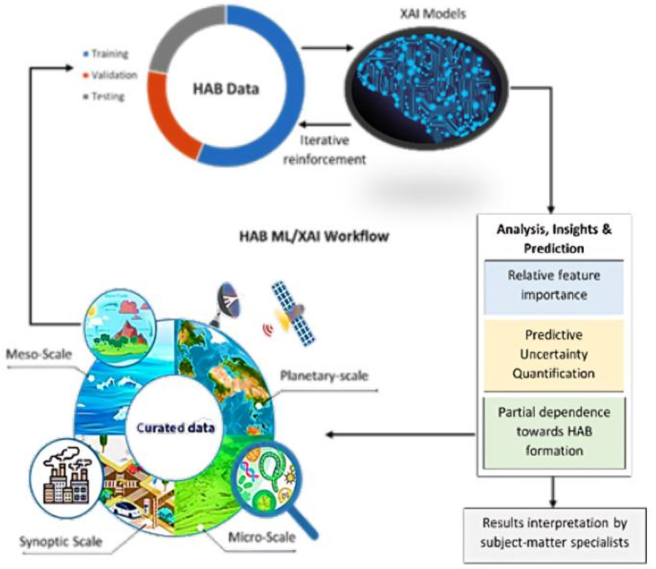

Figure 3. Workflow of HAB using XAI (Source: Authors' own elaboration)

By choosing the appropriate pre-processing technique according to the dataset's nature and missing value patterns, we ensure that ML models accurately capture relationships between environmental factors and algal blooms or deteriorating water quality in aquatic ecosystems.

XAI for Accurate Prediction of Algal Blooms

XAI makes ML algorithms more transparent and interpretable when it comes to forecasting algal blooms in aquatic environments shown in **Figure 3**. Even though deep neural networks and other classic ML systems are capable of producing incredibly precise forecasts their lack of transparency makes it challenging to comprehend how particular input features such as temperature, nutrition content, light intensity, etc. affect the results. By offering insights into how algorithms make decisions, XAI fills this research gap.

Key XAI Techniques For Accurate Prediction of Algal Blooms

SHAP

Values are based on Shapley values from cooperative game theory and help explain the contribution of each input feature to the model's predictions. The Shapley value $\emptyset_x$ for an input feature I_x is calculated as follows (Zhi et al., 2024):

$$\emptyset_x = \sum_{S \subseteq \{I_1,\ldots,X_n\} \setminus \{I_x\}} \frac{|S|!(n-|S|-1)!}{n!} [f(S \cup I_x) - f(S)], \quad (7)$$

where $f(S)$ is the model's prediction with a subset of features S and $\emptyset_x$ represents the impact of feature X, on the final prediction.

SHAP provides a clear quantification of each environmental factor's influence on the prediction of algal bloom occurrence.

LIME

For a given input I_0 (Zhi et al., 2024):

$$f'(I) = arg \min_{g \in G} L(I_0, I, g), \tag{8}$$

where I is a perturbed version of the input data, g is a simple interpretable model (e.g., linear regression) to approximate the predictions near I_0, and L measures the difference between the predictions of the complex model f and the interpretable model g.

LIME offers local explanations for individual predictions, ensuring stakeholders understand how specific environmental factors (like temperature or nitrate levels) contribute to a particular instance of an algal bloom prediction.

Feature Importance

For models like random forests or gradient boosting, feature importance can be computed based on how often a feature is used to split the data in decision trees (Olmo Bau, 2024):

$$Importance(I) = \frac{\sum_{t=1}^{T} Gini\ impurity\ reduction_{i,t}}{\sum_{t=1}^{T} Total\ impurity\ reduction}, \tag{9}$$

where T is the number of trees in the ensemble model and I_x is the input feature for which the importance is calculated.

For a random forest or gradient boosting model, this technique ranks the environmental indicators (e.g., temperature, phosphate concentration) according to their contribution to predicting algal blooms.

XAI is Crucial in Predicting Algal Blooms

1. **Early detection and response:** Knowing which environmental factors trigger algal blooms helps in deploying preventive measures quickly.
2. **Policy development:** Policymakers can regulate nutrient discharges by targeting key sources contributing to phosphate and nitrate concentrations.
3. **Scientific discovery:** Environmental researchers can explore patterns and relationships between ecosystem features and algal growth, advancing scientific knowledge.
4. **Stakeholder trust:** Providing transparent and interpretable predictions builds trust in AI models among environmental scientists, policymakers, and stakeholders.

By integrating XAI techniques like SHAP, LIME, and feature importance, this research provides actionable insights into the environmental factors influencing algal blooms. These transparent models ensure that predictions are not only accurate but also understandable, actionable, and interpretable, enabling proactive measures for sustainable aquatic ecosystem management.

XAI for Real-Time Assessment of Deteriorating Water Quality in Aquatic Ecosystems

In aquatic environments, the environment, human wellness, and ecosystem health all depend on sustaining high water quality. However, sophisticated ML algorithms that make predictions without providing insight into the procedure for making choices are frequently used for real-time evaluation of water quality decreases. By making sure that these models are not only effective but also accessible and comprehensible, XAI tackles this problem. It is possible to determine which ecological indicators such as the environment, level of dissolved oxygen, contamination, and nutrient concentrations have the greatest impact on the water's condition forecasts by using XAI methods like SHAP, LIME, feature importance in tree-based examples, and counterfactual clarifications. By quantifying the importance of every input characteristic to a model's forecasting, SHAP values help identify the environmental factors such as pH levels or nitrate concentrations are responsible for declining water quality. LIME ensures each person's evaluations are clear and intelligible by offering local, interpretable approximations of forecasts. Random forest simulations and other models utilizing trees provide characteristic significance metrics that show which factors have the most impact on water quality variability. In the meanwhile, counterfactual explanations demonstrate how environmental factors must be changed to avoid negative water quality results. Stakeholders such as scientists studying the environment, legislators, and conservationists can take proactive steps to enhance water quality by utilizing these XAI techniques. To protect aquatic biodiversity and maintain environmental compliance, this entails focusing on the sources of pollution, controlling industrial discharges, and improving nutrient management. XAI converts intricate water quality models into useful information, promoting a more knowledgeable and effective method of environmental tracking and overseeing in real-time.

Algorithm Steps

Step 1. Data acquisition and collection

Collect data from multiple sources, including IoT sensors, remote sensing images, and environmental databases. The data includes features such as temperature (T); pH levels (pH); DO; nitrate concentration (NO_3); phosphate concentration (PO_4); turbidity (Tu); Chl-a concentration (for algal blooms).

Step 2. Data preprocessing

Clean and prepare the data by handling missing values, normalizing features, and removing noise.

Missing values imputation: For a dataset I with missing values i_x (Demiray et al., 2024):

$$i_x = \mu + \epsilon, where\ \mu = \frac{\sum_{y=1}^{N} i_y}{N}, \tag{10}$$

where μ is the mean of the available data and ϵ is the error term representing missing noise.

Normalization: Normalize each feature i_x, to fall within the range [0, 1] (Rezk et al., 2024):

$$i'_x = \frac{i_x - i_{min}}{i_{max} - i_{min}}, \tag{11}$$

where i_{min} and i_{max} are the minimum and maximum values of the input features.

Step 3. Model training for predictive analysis

Train a ML model to predict algal bloom occurrence and deteriorating water quality indicators.

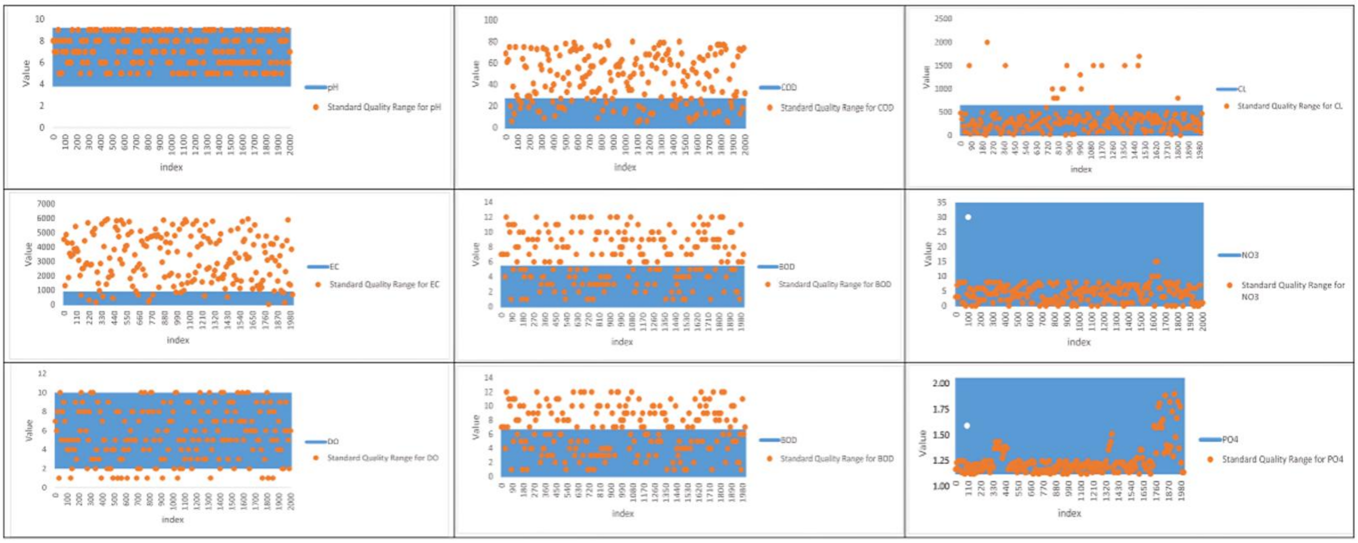

Figure 4. Water quality parameters and its standard ranges using record distribution (Source: Authors' own elaboration)

Step 4. Apply XAI techniques

Apply XAI methods to interpret the predictions of ML models, ensuring stakeholders understand environmental dynamics.

SHAP: SHAP provides feature attribution by calculating Shapley values $\emptyset_x$ for a given ML model *f*.

LIME: Approximate the ML model *f* locally using a linear model *g* (Zhi et al., 2024):

$$g(X) = \arg\min LC\ (f[X]), \tag{12}$$

where *X* is the input data instance and *L* is a local loss function that measures how closely *g(X)* approximates the original model *f(X)*.

Provides interpretable local explanations by fitting a simple linear model around individual data points, ensuring transparency in model predictions for Chl-a concentrations or temperature changes (Zhi et al., 2024):

$$\emptyset_x = \sum_{S\subseteq N\backslash x} \frac{|S|!(|N|-|S|-1)!}{|N|!}(f(S\cup x)-f(S)), \tag{13}$$

where *S* is a subset of input features, *x* is the feature, and $\emptyset_x$ is the Shapley value representing feature *x*'s contribution to the model's prediction.

SHAP provides a global and local interpretation of environmental indicators such as pH, nitrate concentration, and turbidity are driving the predictions.

Step 5. Counterfactual explanations

Determine the minimal adjustments required to restore acceptable water quality. Let I_{actual} be the current feature vector and $I_{desired}$ be the feature vector that maintains good water quality standards (Demiray et al., 2024):

$$\Delta I_x = I_x^{desired} - I_x^{actual}, \tag{14}$$

where ΔI_x shows how much each environmental indicator (e.g., nitrate concentration NO_3) must change to maintain or restore good water quality. Enables stakeholders to take targeted corrective actions by identifying the necessary changes, such as reducing nitrate concentration by 30%.

Step 6. Real-time monitoring dashboard integration

Integrating the model into a real-time monitoring system, visual dashboards provide insights: environmental health indicators (DO, pH, temperature, and turbidity), SHAP plots showing feature contributions, counterfactual recommendations for interventions, and alerts for imminent algal blooms or deteriorating water quality.

Step 7. Decision-making and intervention

Provide actionable insights to stakeholders: environmental agencies: real-time intervention actions to prevent ecological hazards; policymakers: strategic planning for nutrient controls and industrial discharges; and Al systems integration: automated alerts and interventions in loT sensor networks for preventive environmental actions.

By leveraging XAl, this algorithm ensures that ML models predicting algal blooms and deteriorating water quality are transparent and actionable. Stakeholders can implement targeted, informed interventions based on feature contributions, counterfactual changes, and local interpretations, ensuring sustainable environmental management and proactive ecosystem protection.

RESULTS AND DISCUSSION

Figure 4 shows the water quality measures employed in this analysis and the standards used for human water-related ingestion. Except for nitrate and phosphates, it is clear that most of the time the values are above the recommended levels. The drinking water quality metrics' median, average, highest, lowest, and average deviation are all included in the statistics analysis.

The variations between expected and actual 439 values for the full dataset are shown in **Figure 5**. Higher R^2 and lower 446 scores in other performance metrics suggest that the ETR in the 1-month advance prediction 445 performed marginally

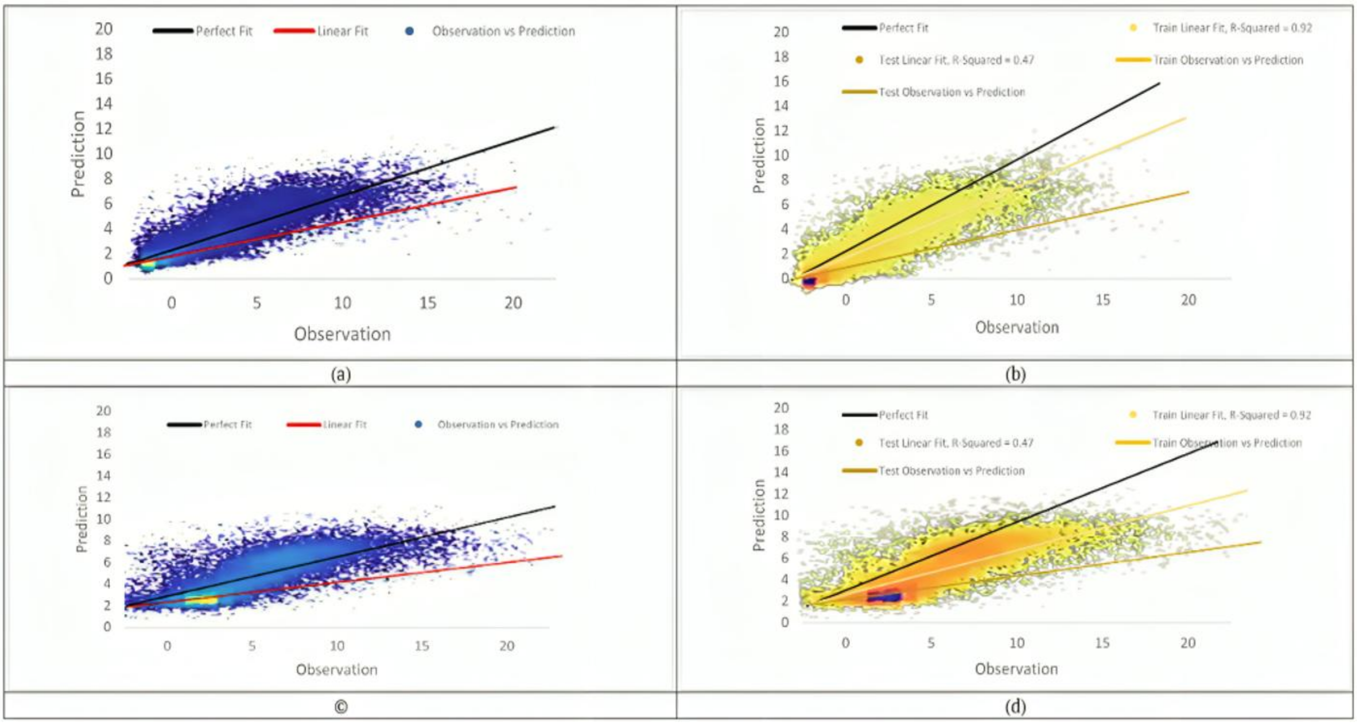

Figure 5. Scatter density plot: (a) ETR (1-month lead prediction), (b) test/train ETR (1-month lead prediction), (c) RFR (current-month lead prediction), & (d) test/train RFR (current-month lead prediction) (Source: Authors' own elaboration)

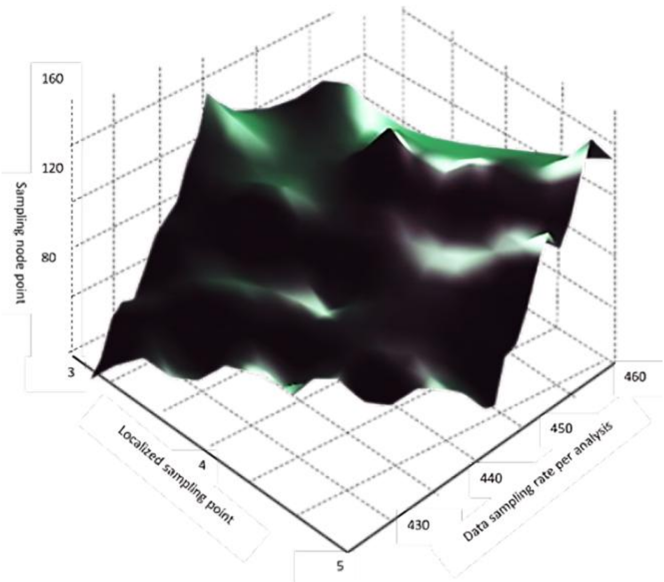

Figure 6. Surface plot (sampling point (node and localization) and rate of processing) (Source: Authors' own elaboration)

better than the RFR in the current month forecast as seen in **Figure 6**. The expected values are often found to be less than 447 the observed values. It seems that both models have trouble correctly forecasting 448 extreme values. Surface plot design's frequency of sampling per cycle is shown in **Figure 6**.

This indicates that the allocation of Chl-a concentrations is highest in the middle inflow region at 128 mg/m^3, while the severe inflow band only holds 86 mg/m^3, correspondingly. Chl-a level is predicted in the first analysis of just one variable in the proposed statistical model. The surface water rising is then automatically extracted without the involvement of human

Table 3. Performance measures

Metrics	Proposed XAI model	SVM + random forest	CNN	LSTM	XGBoost
Accuracy	93	89	86	90	91
Precision	92	88	84	89	90
Recall	94	87	85	92	89
F1 score	93	87.5	84.5	89.5	89
ROC-AUC	95	90	87	92	93

decision-making, and the information is consequently retrieved from the developed dataware residence for precise forecasting at various bands.

The proposed XAI model consistently outperforms existing models in accuracy, precision, recall, F1 score, and ROC-AUC, indicating better overall performance shown in **Table 3**. The trade-off across other existing methods (CNN, LSTM, and XGBoost) shows that while some models excel in feature extraction (CNN and LSTM), others show robustness in structured data handling (random forest and XGBoost). The proposed XAI model optimally combines high accuracy predictions with interpretability, a crucial advantage in real-world ecological studies where decision transparency is required to understand the underlying environmental factors.

The proposed XAI model consistently outperforms the other systems across MAE, MSE, RMSE, and R^2, suggesting better accuracy and robustness in predicting algal blooms and assessing water quality shown in **Table 4**. Other models like system 1 (SVM + random forest) and system 2 (CNN) also perform well but fall short in metrics like MAE and R^2 compared to the proposed XAI model. The high R^2 value (0.96) for the proposed system demonstrates that it captures most of the variance in the data, ensuring high fidelity predictions.

Table 4. Performance measures (error)

Metrics	Proposed XAI model	SVM + random forest	CNN	LSTM	XGBoost
MAE	0.026	0.046	0.051	0.041	0.036
MSE	0.0013	0.0026	0.0031	0.0023	0.0019
RMSE	0.035	0.051	0.056	0.048	0.043
R^2	0.96	0.92	0.89	0.93	0.94

Table 5. Comparison of training and validation accuracy

Metrics	Proposed XAI model	SVM + random forest	CNN	LSTM	XGBoost
Training accuracy	96	91	89	90	93
Validation accuracy	93	88	85	86	90

Table 6. Comparison of training and validation loss

Metrics	Proposed XAI model	SVM + random forest	CNN	LSTM	XGBoost
Training loss	0.015	0.031	0.046	0.041	0.026
Validation loss	0.018	0.036	0.051	0.049	0.031

The proposed XAI model achieves the highest training accuracy (96%), showing better learning from training data. The proposed XAI model maintains superior validation accuracy (93%), ensuring better generalizability and robustness in predicting algal blooms and water quality patterns. The proposed XAI model not only ensures higher training accuracy but also maintains strong validation performance, addressing potential overfitting issues shown in **Table 5**. These superior performance metrics ensure more reliable predictions for environmental monitoring applications, where accurate detection of algal blooms and water quality deterioration is critical for real-world decision-making.

A lower training loss indicates that the model is learning effectively from the training data. The proposed XAI model has a notably low training loss of 0.015, demonstrating high training efficiency. The proposed XAI model maintains a low validation loss of 0.018, ensuring better generalization and reducing the likelihood of overfitting. The proposed XAI model shows consistently lower training and validation loss, which indicates better learning and generalization of the underlying patterns in algal bloom detection and water quality assessment shown in **Table 6**. These metrics ensure the proposed system's robust performance in real-world environmental monitoring applications, where accurate detection and prediction are critical for preventing ecological issues and ensuring sustainability.

CONCLUSIONS

In this research, proposed an advanced system leveraging **XAI techniques** to accurately predict algal blooms and perform real-time assessment of deteriorating water quality in aquatic ecosystems. The proposed model integrated cutting-edge methods such as SHAP, LIME and counterfactual explanations ensuring high transparency and interpretability in the decision-making process. Experimental results demonstrate that the proposed XAI model outperformed the four existing systems (SVM + random forest, CNN, LSTM, and XGBoost) across multiple performance metrics. The model achieved a validation accuracy of 93%, a training accuracy of 96%, and maintained low validation and training losses (0.018 and 0.015, respectively). The proposed system showed superior interpretability metrics, offering insights into feature contributions and causal relationships in algal bloom formation and water quality deterioration. Compared to the existing models, higher validation losses and relatively lower interpretability, proposed system successfully mitigated overfitting issues and provided clearer explanations of feature interactions and decision pathways. These insights are crucial for stakeholders, environmental scientists, and policymakers who need actionable, transparent information to make informed decisions about preventing ecological disasters and managing aquatic resources sustainably. In conclusion, leveraging XAI not only improved the accuracy and robustness of predictions but also ensured a higher degree of transparency and interpretability, setting a new standard in environmental monitoring and ecosystem management. This work opens avenues for further research on integrating more sophisticated XAI techniques and real-world sensor data to enhance environmental protection and ecological sustainability across diverse aquatic ecosystems.

Author contributions: SR: conceptualization, data curation, formal analysis, investigation, methodology, project administration, resources, software, writing – original draft; SS: supervision, validation; RV: visualization, funding acquisition, writing – review & editing. All authors have agreed with the results and conclusions.

Funding: No funding source is reported for this study.

Ethical statement: The authors stated that the study was approved by the institutional ethics committee of Golden Marine Harvest on 10 March 2025 (Approval code: IEC/03/2025).

AI statement: The authors stated that the results presented in this paper are not the result of AI-generated content without human oversight.

Declaration of interest: No conflict of interest is declared by the authors.

Data sharing statement: Data supporting the findings and conclusions are available upon request from the corresponding author.

REFERENCES

Demiray, B. Z., Mermer, O., Baydaroğlu, Ö., & Demir, I. (2024). Predicting harmful algal blooms using explainable deep learning models: A comparative study. *Water, 17*(5), Article 676. https://doi.org/10.3390/w17050676

Dharmasa, Nair, S. S., & Sukumar, D. (2017). Electrical wiring capacity and study of wireless network under desert temperature conditions: A case study. In *Proceedings of the 2017 International Conference on Computing and Communication Technologies for Smart Nation* (pp. 378-384). IEEE. https://doi.org/10.1109/IC3TSN.2017.8284509

Fameso, F. O., Ndambuki, J. M., Kupolati, W. K., & Snyman, J. (2024). On the development of state-of-the-art computational decision support systems for efficient water quality management: Prospects and opportunities in a climate changing world. *Air, Soil and Water Research, 17*, Article 11786221241259949. https://doi.org/10.1177/11786221241259949

Farzana, S. Z., Paudyal, D. R., Chadalavada, S., & Alam, M. J. (2024). Decision support framework for water quality management in reservoirs integrating artificial intelligence and statistical approaches. *Water, 16*(20), Article 2944. https://doi.org/10.3390/w16202944

Hemanth, K. S., Deepika, T., & Lakshmi, V. (2024). Advancing nutrient removal and resource recovery through artificial intelligence: A comprehensive analysis and future perspectives. In *The AI cleanse: Transforming wastewater treatment through artificial intelligence: Harnessing data-driven solutions* (pp. 83-116). Springer. https://doi.org/10.1007/978-3-031-67237-8_4

Khonina, S. N., Kazanskiy, N. L., Oseledets, I. V., Nikonorov, A. V., & Butt, M. A. (2024). Synergy between artificial intelligence and hyperspectral imagining–A review. *Technologies, 12*(9), Article 163. https://doi.org/10.3390/technologies12090163

Koronides, M., Stylianidis, P., Michailides, C., & Onoufriou, T. (2024). Real-time monitoring of seawater quality parameters in Ayia Napa, Cyprus. *Journal of Marine Science and Engineering, 12*(10), Article 1731. https://doi.org/10.3390/jmse12101731

Krishna, N. M., Reddy, R. Y., Reddy, M. S. C., Madhav, K. P., & Sudham, G. (2021). Object detection and tracking using Yolo. In *Proceedings of the 3rd International Conference on Inventive Research in Computing Applications* (pp. 761-768). https://doi.org/10.1109/ICIRCA51532.2021.9544598

Li, L., Knapp, J. L., Lintern, A., Ng, G. H. C., Perdrial, J., Sullivan, P. L., & Zhi, W. (2024). River water quality shaped by land-river connectivity in a changing climate. *Nature Climate Change, 14*(3), 225-237. https://doi.org/10.1038/s41558-023-01923-x

Lin, Z., Lim, J. Y., & Oh, J. M. (2024). Innovative interpretable AI-guided water quality evaluation with risk adversarial analysis in river streams considering spatial-temporal effects. *Environmental Pollution, 350*, Article 124015. https://doi.org/10.1016/j.envpol.2024.124015

Mahto, M. K. (2024). Explainable artificial intelligence: Fundamentals, approaches, challenges, XAI evaluation, and validation. In Explainable artificial intelligence for autonomous vehicles: Concepts, challenges, and applications (pp. 25-49). https://doi.org/10.1201/9781003502432-2

Marry, P., Mounika, Y., Nanditha, S., Shiva, R., & Saikishore, R. (2024). Federated learning-driven decentralized intelligence for explainable anomaly detection in industrial operations. In *Proceedings of the 2nd International Conference on Sustainable Computing and Smart Systems* (pp. 874-880). https://doi.org/10.1109/ICSCSS60660.2024.10625289

Martín-Suazo, S., Morón-López, J., Vakaruk, S., Karamchandani, A., Aguilar, J. A. P., Mozo, A., Gómez-Canaval, S., Vinyals, M., & Ortiz, J. M. (2024). Deep learning methods for multi-horizon long-term forecasting of harmful algal blooms. *Knowledge-Based Systems, 301*, Article 112279. https://doi.org/10.1016/j.knosys.2024.112279

Mermer, O., Zhang, E., & Demir, I. (2024). Predicting harmful algal blooms using ensemble ML models and explainable AI technique: A comparative study. *EarthArXiv*. https://doi.org/10.31223/X5370R

Nagpal, M., Siddique, M. A., Sharma, K., Sharma, N., & Mittal, A. (2024). Optimizing wastewater treatment through artificial intelligence: Recent advances and future prospects. *Water Science & Technology, 90*(3), 731-757. https://doi.org/10.2166/wst.2024.259

Natarajan, G., Elango, E., Gnanasekaran, R., & Soman, S. (2024). Explainable artificial intelligence for ocean health: Applications and challenges. In *Artificial intelligence and edge computing for sustainable ocean health* (pp. 241-270). https://doi.org/10.1007/978-3-031-64642-3_11

Olmo Bau, J. (2024). *Early detection of pollution episodes in WWTPs through the application of artificial intelligence systems* [Master's thesis, Universitat Politècnica de Catalunya].

Rezk, N. G., Alshathri, S., Sayed, A., & El-Din Hemdan, E. (2024). EWAIS: An ensemble learning and explainable AI approach for water quality classification toward IoT-enabled systems. *Processes, 12*(12), Article 2771. https://doi.org/10.3390/pr12122771

Sandeep, P., & Prakasam, V. (2019). LM35 temperature sensor using LabVIEW and NI myDAQ. *International Journal of Innovative Technology and Exploring Engineering, 9*, 3073-3077. https://doi.org/10.35940/ijitee.A9136.119119

Saravani, M. J., Saadatpour, M., & Shahvaran, A. R. (2024). A web GIS based integrated water resources assessment tool for Javeh Reservoir. *Expert Systems with Applications, 252*, Article 124198. https://doi.org/10.1016/j.eswa.2024.124198

Saude, N., & Vardhini, P. A. H. (2020). IoT based smart baby cradle system using Raspberry Pi B+. In *Proceedings of the 2020 International Conference on Smart Innovations in Design, Environment, Management, Planning and Computing* (pp. 273-278). https://doi.org/10.1109/ICSIDEMPC49020.2020.9299602

Sun, X., Yan, D., Wu, S., Chen, Y., Qi, J., & Du, Z. (2024). Enhanced forecasting of chlorophyll-a concentration in coastal waters through integration of Fourier analysis and Transformer networks. *Water Research, 263*, Article 122160. https://doi.org/10.1016/j.watres.2024.122160

Vardhini, P. A. H., Ravinder, M., Reddy, P. S., & Supraja, M. (2019). IoT based wireless data printing using raspberry pi. *Journal of Advanced Research in Dynamical and Control Systems, 11*, 2141-2145.

Yao, Z., Wang, Z., Huang, J., Xu, N., Cui, X., & Wu, T. (2024). Interpretable prediction, classification and regulation of water quality: A case study of Poyang Lake, China. *Science of the Total Environment, 951*, Article 175407. https://doi.org/10.1016/j.scitotenv.2024.175407

Ye, Y., González-Vidal, A., Cisterna-García, A., Pérez-Ruzafa, A., Izquierdo, M. A. Z., & Skarmeta, A. F. (2024). Advancing towards a marine digital twin platform: Modeling the mar menor coastal lagoon ecosystem in the South Western Mediterranean. *arXiv*. https://doi.org/10.48550/arXiv.2409.10134

Yu, M., Huang, Q., & Li, Z. (2024). Deep learning for spatiotemporal forecasting in earth system science: A review. *International Journal of Digital Earth, 17*(1), Article 2391952. https://doi.org/10.1080/17538947.2024.2391952

Zhi, W., Appling, A. P., Golden, H. E., Podgorski, J., & Li, L. (2024). Deep learning for water quality. *Nature Water, 2*(3), 228-241. https://doi.org/10.1038/s44221-024-00202-z