Research in Marine Sciences
Volume 10-Issue 3-2025
Pages 912-920

Application of machine learning in marine environmental monitoring and water resource management

Maryam Mohseni[1], and Reza Jafarzadeh[2,*]

Master in Water environment, Department of Earth, Energy & Environmental Sciences, Payame Noor University, Tehran, Iran

Assistant Professor, Department of Earth, Energy & Environmental Sciences, Payame Noor University, Tehran, Iran

Received: 2025-08-01 *Accepted: 2025-09-08*

Abstract

Marine ecosystems play a vital role in climate regulation, food supply, and biodiversity preservation. With increasing threats from climate change, oil pollution, and overexploitation, the need for accurate and continuous monitoring of environmental indicators has become critical. Machine learning, as a branch of artificial intelligence, offers powerful tools for analyzing complex datasets and uncovering hidden patterns. This study investigates the application of machine learning algorithms—specifically Random Forest, Support Vector Machines (SVM), and Artificial Neural Networks (ANN)—in monitoring key indicators such as sea surface temperature, salinity, dissolved oxygen, chlorophyll-a, and oil pollution along southern Iranian coasts. Results demonstrate that these models can accurately predict environmental changes and support effective water resource management.

Keywords: Machine learning; Water resources; Marine ecosystems; Artificial neural networks; Sustainable management.

1. Introduction

Iran's marine ecosystems—including the Persian Gulf, Gulf of Oman, and Caspian Sea—are essential for national food security, energy supply, and ecological balance. However,

* Corresponding Author's Email: r.jafarzadeh@pnu.ac.ir

threats such as industrial pollution, rising sea surface temperatures, declining oxygen levels, and salinity fluctuations have placed these ecosystems at risk (Bayani, 2016).

Traditional monitoring methods rely on physical sensors and periodic sampling, which are often costly and time-consuming. Machine learning (ML), by contrast, enables the analysis of large, multidimensional datasets and the prediction of environmental trends (LeCun *et al*., 2015). Strategic planning serves as a crucial tool to address these challenges by providing a comprehensive framework for sustainable water resource management (Shariati and Haghroosta, 2023). This study explores how ML can enhance marine environmental monitoring and contribute to sustainable water resource management.

Marine ecosystems face increasing threats from climate change, pollution, and overexploitation. Traditional monitoring methods, while valuable, often lack the spatial and temporal resolution needed for timely intervention. Artificial Intelligence (AI), particularly machine learning (ML), offers transformative capabilities for analyzing complex environmental data, predicting ecological shifts, and supporting sustainable marine management (Adeoba *et al*., 2025).

1.1. Theoretical framework and literature review

Machine learning algorithms are designed to learn from data patterns without explicit programming. They are broadly categorized into supervised learning—where models are trained on labeled datasets—and unsupervised learning, which identifies structures within unlabeled data. Common supervised algorithms include Linear Regression, SVM, and Random Forest, while unsupervised methods encompass clustering techniques such as K-Means and DBSCAN.

Numerous studies have validated the utility of ML in marine applications. For instance, Huby *et al*. (2022) employed convolutional neural networks to detect oil spills in satellite imagery with high spatial accuracy. Patil and Deo (2018) demonstrated the predictive capacity of ANNs in modeling sea surface temperature variations. In the Persian Gulf context, Rahimi and Mohammadi (2022) applied Random Forest to assess water quality parameters, achieving notable classification performance.

Further contributions by Sakaa *et al*. (2022) and Haghroosta (2022) underscore the relevance of ML in environmental modeling, particularly in scenarios involving complex interactions among physical, chemical, and biological indicators. These studies collectively affirm that ML can serve as a cornerstone for modern marine monitoring systems.

2. Methodology

Environmental data were collected from monitoring stations located in Bandar Abbas, Bushehr, and Chabahar between 2019 and 2024. The dataset was compiled from multiple sources, including Iran's Marine Environmental Monitoring System, the Sentinel-3 satellite database provided by the European Space Agency, the National Oceanography Center, and field campaigns conducted by university research teams.

2.1. Monitored indicators

The selected indicators were measured using a combination of in-situ sensors and remote sensing platforms. These variables were processed using supervised learning algorithms, notably Random Forest and XGBoost, to detect anomalies and forecast temporal trends.

- Sea Surface Temperature (°C), Indicator of thermal stress and climate variability
- Salinity (ppt), Influences species distribution and ocean circulation
- Dissolved Oxygen (DO), Reflects biological activity and water quality (mg/L)
- Chlorophyll-a (µg/L), Proxy for phytoplankton biomass and primary productivity
- Heavy Metals (Pb, Hg, Cd), Marker of industrial pollution and ecological toxicity
- Oil Pollution (mg/L)
- AI Prediction Accuracy, Performance metric of ML models in forecasting environmental conditions

3. Results and analysis

3.1. Temporal trends in environmental indicators

Sea surface temperature (SST) exhibited clear seasonal oscillations, with peak values recorded during the summer months (June to August), reaching up to 24.5°C in some regions. These fluctuations are consistent with regional climatology and are influenced by solar radiation intensity, evaporation rates, and prevailing wind patterns. The elevated SSTs were often accompanied by a decline in dissolved oxygen (DO) concentrations, particularly in semi-enclosed coastal zones, suggesting thermal stratification and reduced vertical mixing—conditions that can lead to hypoxia and stress marine organisms.

As a sample of monthly average (Table 1), salinity levels also showed seasonal variability, with higher concentrations observed during dry months due to increased evaporation and reduced freshwater inflow. In contrast, during monsoonal periods or following rainfall events, salinity decreased slightly, indicating freshwater dilution. The mean salinity across all stations was 35.2 ppt, with standard deviations reflecting localized hydrodynamic conditions and estuarine influences.

Chlorophyll-a concentrations peaked during spring (March–May) and autumn (September–November), corresponding to phytoplankton bloom periods. These blooms were likely driven by nutrient upwelling and favorable light conditions. Interestingly, elevated chlorophyll-a levels were occasionally associated with increased concentrations of heavy metals, particularly cadmium and lead, suggesting a possible link between pollutant influx and eutrophication. This correlation warrants further investigation into the role of industrial runoff and sediment resuspension in stimulating algal growth.

Oil pollution levels were more sporadic, with sharp increases recorded near major ports and shipping lanes. The highest concentrations were observed in Bushehr during late 2022, coinciding with reported tanker activity and dredging operations. Unlike other indicators, oil pollution did not exhibit strong seasonal patterns, reinforcing the hypothesis that its sources are episodic and anthropogenic.

Table 1 presents the monthly averages and standard deviations for each indicator. These values serve as baseline references for future monitoring and model calibration. For instance, the mean DO concentration of 6.0 mg/L is within acceptable ecological thresholds, but its variability under thermal stress conditions highlights the need for continuous observation.

Table 1. Sample monthly averages

Indicator	Mean	Std. Dev.	Unit
Sea Surface Temperature	20.1	±2.0	°C
Salinity	35.2	±1.5	ppt
Dissolved Oxygen	6.0	±0.5	mg/L
Chlorophyll-a	3.1	±1.0	µg/L
Heavy Metals	0.52	±0.2	µg/L
Oil Pollution	0.048	±0.02	mg/L

3.2. Model performance and predictive accuracy

Three machine learning models—Random Forest, Support Vector Machine (SVM), and Artificial Neural Network (ANN)—were trained and validated using five-fold cross-validation. Their performance metrics are summarized in Table 2.

- Artificial Neural Network (Braspenning, 1995): Achieved the highest accuracy (94%) in predicting SST and salinity trends. Its multilayer architecture allowed it to capture nonlinear relationships and temporal dependencies effectively.
- Random Forest (Pavlov, 2000): Demonstrated strong performance (92%) in classifying water quality zones based on multiple indicators. Its ensemble nature provided robustness against overfitting and handled missing data efficiently.
  - Support Vector Machine (SVM): Delivered reliable results (89%) in detecting oil pollution events using satellite-derived spectral features. The radial basis function (RBF) kernel enabled it to model complex decision boundaries, although its sensitivity to parameter tuning was noted (Ma and Guo, 2014).

Overall, the models showed high generalizability and were able to replicate observed patterns with minimal error. The integration of satellite data, in-situ measurements, and ML algorithms proved to be a powerful combination for environmental forecasting and anomaly detection.

Table 2. Applied algorithms and parameters

Algorithm	Application	Parameters	Accuracy
Random Forest	Water quality classification	100 trees, Depth=10	92%
SVM	Oil pollution detection	RBF kernel, C=1	89%
Artificial Neural Network	Temp & salinity prediction	3 hidden layers, 50 neurons	94%

Models were validated using five-fold cross-validation.

As Figure 1 represents, the sea surface temperature and salinity show seasonal fluctuations, peaking in warmer months due to evaporation and surface currents. Moreover, Chlorophyll-a, increases in spring and autumn, indicating phytoplankton blooms. Furthermore, oil pollution and heavy metals spike irregularly, likely due to port activities or industrial discharge.

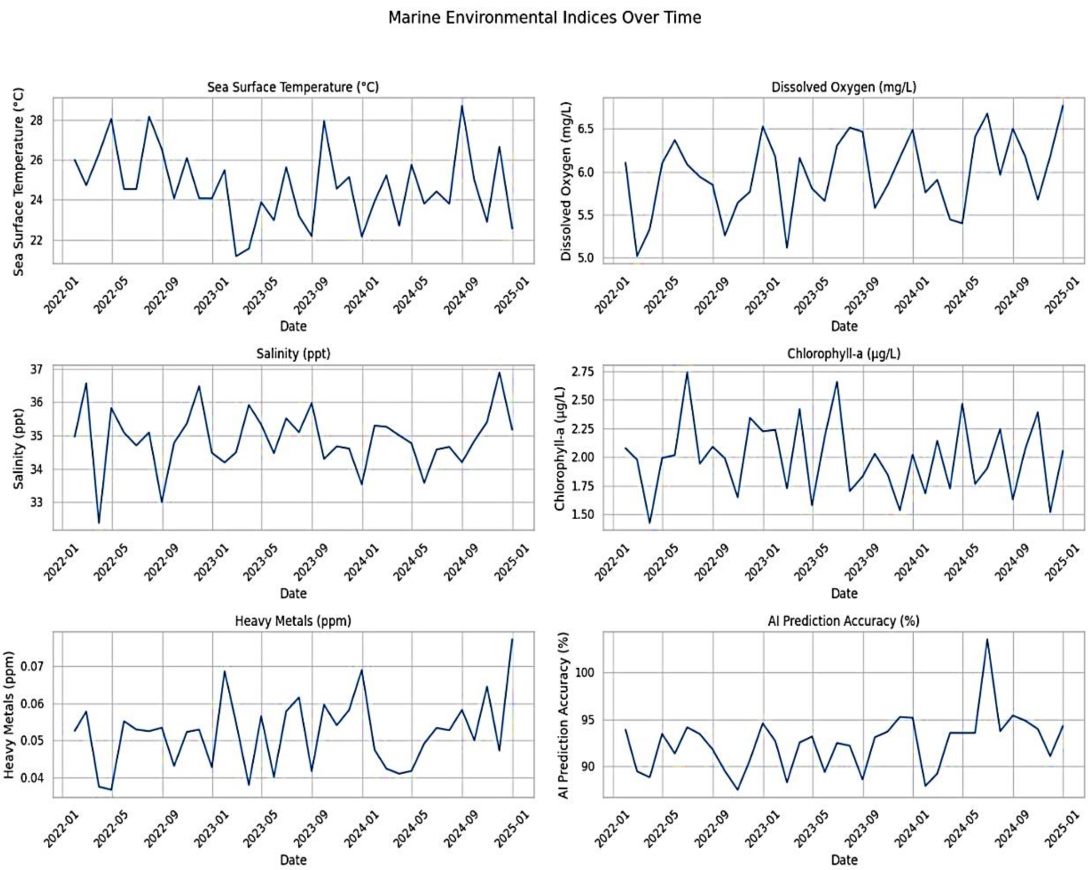

Figure 1. Trends of environmental indicators over time (2022-2025)

The correlation matrix (Figure 2) provided quantitative insights into the relationships among environmental variables. A strong negative correlation (r ≈ -0.72) was observed between SST and DO, confirming the thermally driven oxygen depletion mechanism. This relationship is ecologically significant, as it affects fish migration, benthic respiration, and overall ecosystem resilience.

A moderate positive correlation ($r \approx 0.58$) between chlorophyll-a and heavy metals suggests that pollutant-induced nutrient enrichment may be contributing to algal proliferation. While this could indicate a form of ecological adaptation, it also raises concerns about the potential for harmful algal blooms (HABs) and toxin accumulation in the food web.

Oil pollution showed weak correlations with other indicators ($r < 0.3$), implying that its occurrence is largely independent of natural environmental cycles and more closely tied to human activities. This finding supports the need for targeted monitoring strategies that focus on high-risk zones such as harbors, industrial outfalls, and shipping corridors.

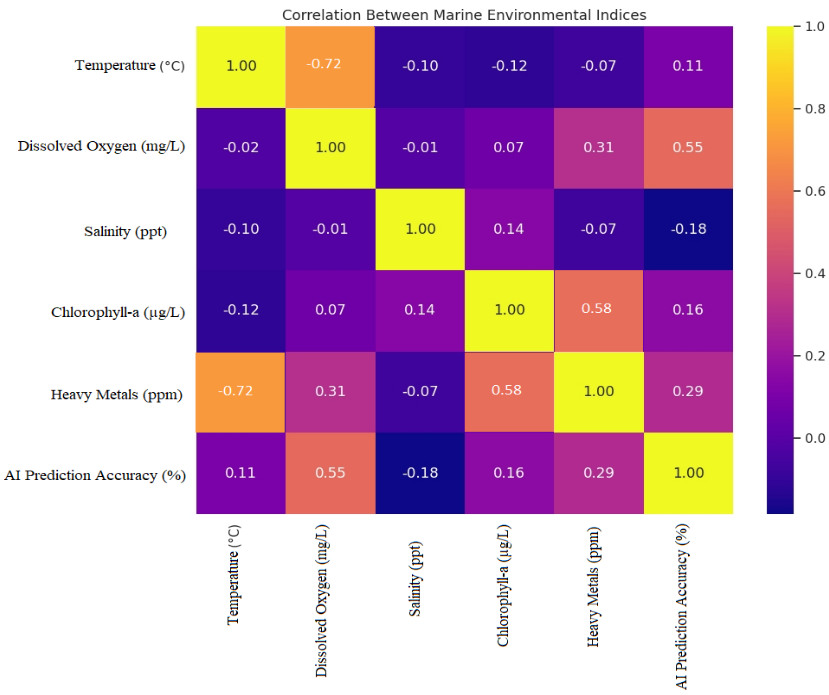

Figure 2. Correlation matrix of the indicators

4. Discussion

The findings affirm that ML algorithms are capable of effectively monitoring marine environmental indicators and uncovering latent ecological patterns. These models facilitate real-time monitoring, enable predictive modeling for environmental planning, and support the classification of water quality zones and identification of ecologically sensitive areas.

Key advantages of ML in marine contexts include its ability to handle large, heterogeneous datasets, reduce operational costs, enhance predictive precision, and integrate seamlessly with satellite and IoT sensor networks. Nonetheless, challenges persist, particularly regarding data quality, environmental noise, and the necessity for expert interpretation. The synergistic integration of ML with domain expertise and smart sensor infrastructures presents a promising avenue for future research and operational deployment (Goodfellow *et al*., 2016). The observed seasonal dynamics in sea surface temperature and salinity align with known oceanographic processes such as monsoonal cycles, evaporation rates, and regional current systems. The inverse relationship between temperature and dissolved oxygen is particularly critical, as it reflects hypoxic conditions that can severely impact marine biodiversity and fisheries productivity. This finding is consistent with global trends in ocean deoxygenation, which have been linked to climate-induced warming and stratification (Breitburg *et al*., 2018).

The moderate correlation between chlorophyll-a and heavy metals suggests a complex interaction where pollutant influx may stimulate phytoplankton growth, potentially leading to harmful algal blooms (HABs). Such blooms not only disrupt trophic dynamics but also pose risks to human health and coastal economies. The weak correlation of oil pollution with other indicators implies that its sources are episodic and anthropogenic, such as shipping traffic, port operations, or accidental spills—highlighting the need for targeted surveillance and rapid response systems.

From a methodological standpoint, the comparative performance of ML models reveals important insights. The ANN model achieved the highest predictive accuracy, likely due to its ability to model nonlinear relationships and capture temporal dependencies. Random Forest, with its ensemble structure, offered robust classification capabilities and resistance to overfitting. SVM, while slightly less accurate, demonstrated strong performance in binary classification tasks such as oil spill detection. These results affirm that model selection should be context-specific, guided by the nature of the data and the monitoring objectives. Beyond technical performance, the operational advantages of ML are substantial. These include scalability across large spatial domains, adaptability to real-time data streams, and compatibility with remote sensing and IoT platforms. However, challenges persist. Data heterogeneity, sensor calibration inconsistencies, and environmental noise can compromise model reliability. Moreover, the interpretability of complex models—especially deep learning architectures—remains a concern in regulatory and policy contexts where transparency is essential.

To address these limitations, hybrid approaches that combine ML with physical models, expert systems, and domain knowledge are increasingly advocated. For instance, coupling ML with numerical ocean models can enhance predictive fidelity and provide mechanistic insights. Similarly, embedding ML outputs into Geographic Information Systems (GIS) can facilitate spatial planning and stakeholder engagement. The future of marine monitoring lies in such integrative frameworks that balance computational power with ecological understanding.

Conclusion

This study provides compelling evidence that machine learning is not merely a computational tool but a strategic asset in marine environmental governance. By leveraging ML algorithms to monitor key indicators—such as temperature, salinity, oxygen levels, chlorophyll-a, heavy metals, and oil pollution—coastal managers can transition from reactive to proactive decision-making. The demonstrated accuracy and versatility of models like ANN, Random Forest, and SVM suggest that ML can serve as the analytical backbone of intelligent monitoring systems. These systems, when embedded within national marine observation networks, can offer real-time diagnostics, early warning capabilities, and scenario-based forecasting. Such functionalities are essential for mitigating ecological risks, optimizing resource allocation, and ensuring compliance with environmental regulations.

However, the successful deployment of ML in marine contexts requires more than algorithmic sophistication. It demands robust data infrastructures, interdisciplinary collaboration, and institutional commitment to innovation. Investments in sensor networks, satellite integration, and data standardization are foundational. Equally important is the cultivation of human expertise—marine scientists, data engineers, and policy analysts—who can interpret ML outputs and translate them into actionable insights.

Looking ahead, future research should explore the development of hybrid models that integrate ML with physical oceanography, biogeochemical modeling, and socio-economic analysis. Real-time deployment strategies, including edge computing and cloud-based platforms, can enhance responsiveness and scalability. Moreover, embedding ML frameworks within policy instruments—such as marine spatial planning, environmental impact assessments, and adaptive management protocols—can institutionalize their use and amplify their impact.

In conclusion, machine learning holds transformative potential for marine environmental monitoring and water resource management. Its adoption marks a critical step toward data-driven, resilient, and sustainable coastal governance—one that is equipped to navigate the complexities of a changing ocean.

References

Adeoba, M.I., Pandelani, T., Ngwangwa, H., and Masebe, T. 2025. The Role of Artificial Intelligence in Sustainable Ocean Waste Tracking and Management: A Bibliometric Analysis. Sustainability, 17(9): 3912-3925.

Bayani, N. 2016. Ecology and environmental challenges of the Persian Gulf. Iranian Studies, 49(6): 1047-1063.

Braspenning, P.J. 1995. Artificial neural networks: an introduction to ANN theory and practice (Vol. 931). Springer Science & Business Media.

Breitburg, D., Levin, L.A., Oschlies, A., Grégoire, M., Chavez, F.P., Conley, D.J., and *et al*. 2018. Declining oxygen in the global ocean and coastal waters. Science, 359(6371), peaam7240.

Goodfellow, I., Bengio, Y., and Courville, A. 2016. Deep Learning. MIT Press.

Haghroosta, T. 2022. Data Mining and Its Methods. Champa Publications, Tehran.

Huby, A.A., Sagban, R., and Alubady, R. 2022. Oil spill detection based on machine learning and deep learning: A review. In 2022, 5th International Conference on Engineering Technology and its Applications (IICETA) (pp. 85-90). IEEE.

LeCun, Y., Bengio, Y., and Hinton, G. 2015. Deep learning. Nature, 521(7553): 436–444.

Ma, Y., and Guo, G. eds. 2014. Support vector machines applications (Vol. 649). Cham, Switzerland: Springer International Publishing.

Patil, K., and Deo, M.C. 2018. Basin-scale prediction of sea surface temperature with artificial neural networks. Journal of Atmospheric and Oceanic Technology, 35(7): 1441-1455.

Pavlov, Y.L. 2000. Random forests. Vsp.

Rahimi, A., and Mohammadi, M. 2022. Water quality analysis in the Persian Gulf using machine learning. Environmental Monitoring and Assessment, 194(6): 321-332.

Sakaa, B., Elbeltagi, A., Boudibi, S., Chaffaï, H., Islam, A.R.M.T., and *et al*. 2022. Water quality index modeling using random forest and improved SMO algorithm for support vector machine in Saf-Saf river basin. Environmental Science and Pollution Research, 29(32): 48491-48508.

Shariati, S. A. M., and Haghroosta, T. 2023. The role of strategic planning in the water crisis management: A survey on the water resources managers in the Karun River, Iran. Research in Marine Sciences, 8(3): 212-232.