Ethiopian Journal of Environmental Studies & Management 19(2): 194 – 212, 2026
ISSN:1998-0507 doi: https://ejesm.org/doi/v19i2.6
Submitted: February 5, 2026 Accepted: April 1, 2026

MACHINE LEARNING APPLICATIONS IN ENVIRONMENTAL MONITORING: A NARRATIVE REVIEW

***UGEGE, P.E., AGBENU, D.O., AKOUN, M. AND BADEJO, D.K.**
Forestry Research Institute of Nigeria, Ibadan, Nigeria.
*Corresponding author: pugege@frin.gov.ng

Abstract

Environmental monitoring is essential for understanding ecosystem dynamics, detecting environmental change, and supporting evidence-based decision-making for sustainable development. Increasing pressures from climate change, urbanization, deforestation, pollution, and biodiversity loss have intensified the demand for monitoring approaches that are accurate, scalable, and capable of handling large and complex datasets. In this context, machine learning (ML) has emerged as a powerful analytical tool due to its ability to model non-linear relationships, process high-dimensional data, and generate reliable predictions. This paper presents a narrative review of ML applications in environmental monitoring, outlining major paradigms: supervised, unsupervised, semi-supervised, and reinforcement learning, and examining their relevance across key environmental domains. Core application areas reviewed include land use and land cover mapping, climate and weather analysis, air and water quality assessment, and biodiversity monitoring. The review also evaluates common environmental data sources such as remote sensing imagery, in situ sensor networks, environmental databases, and citizen science datasets, as well as emerging data processing platforms. In addition, it discusses challenges related to data quality, model interpretability, generalization across regions, and computational constraints, and highlights future directions emphasizing data integration, explainable artificial intelligence, and operational deployment for sustainable environmental management.

Keywords: *Machine learning, Environmental monitoring, Remote sensing, Environmental data, Artificial intelligence, Sustainability*

Introduction

Environmental monitoring plays a critical role in understanding ecosystem dynamics, detecting environmental change, and supporting evidence-based decision-making for sustainable development (Clark and Dickson, 2003; Wang *et al*., 2025). Increasing pressures from climate change, urbanization, deforestation, pollution, and biodiversity loss have intensified the need for timely, accurate, and scalable monitoring approaches (Nakhle *et al*., 2024). Conventional environmental monitoring methods, which often rely on field-based surveys and manual data interpretation, are resource-intensive, time-consuming, and spatially limited (Olawade *et al*., 2024). As a result, they are increasingly complemented by computational techniques capable of handling large, complex, and heterogeneous datasets (Wang *et al*., 2025).

In recent years, machine learning (ML) has emerged as a powerful tool for environmental monitoring due to its ability to identify patterns, model non-linear relationships, and make predictions from large volumes of data (Pichler AND Hartig, 2023; Zhang *et al*., 2025). Machine learning techniques are particularly well suited to environmental applications where data are noisy, high-dimensional, and often incomplete (Caiafa *et al*., 2021; Zhong *et al*., 2021). By leveraging advances in computational power, remote sensing technologies, and data availability, ML methods have been widely adopted across various environmental domains, including land use and land cover mapping, climate analysis, air and water quality assessment, disaster risk monitoring, and biodiversity conservation (Alotaibi and Nassif, 2024).

The integration of machine learning into environmental monitoring has been facilitated by the rapid growth of Earth observation data from satellite platforms, unmanned aerial vehicles, sensor networks, and citizen science initiatives (Antonelli *et al*., 2023; Demissie *et al*., 2026; Espíndola *et al*., 2025). These data sources generate continuous streams of information that exceed the capacity of traditional analytical methods. Machine learning algorithms, in contrast, can process and analyze such data efficiently, enabling near-real-time monitoring and improved predictive performance (Brown *et al*., 2025).

Despite the growing body of literature on machine learning applications in environmental sciences, many studies remain fragmented across disciplines and application areas, limiting synthesis and cross-domain learning (Valencia-Arias *et al*., 2025). There is therefore a need for integrative reviews that synthesize existing knowledge, highlight common methodological trends, and identify opportunities for future research. Narrative reviews are particularly useful in this context, as they provide a broad overview of concepts, tools, and applications without the restrictive requirements of systematic review protocols (Grant and Booth, 2009; Grønstad, 2026).

This review provides a narrative synthesis of machine learning applications in environmental monitoring. It focuses on commonly used algorithms, key application areas, data sources, and emerging challenges. By summarizing current practices and research directions, the paper aims to serve as a reference for researchers, practitioners, and policymakers interested in leveraging machine learning for environmental monitoring and management.

Metadata Analysis of the Reviewed Literature

To provide an analytical overview of the reviewed studies and enhance the technical depth of this narrative review, a metadata assessment was conducted on all cited references (n = 69). The literature was manually classified according to publication year, application domain, machine learning (ML) paradigm, and geographic focus. The objective was not to perform a full bibliometric analysis, but rather to identify general trends that contextualize the evolution and scope of ML applications in environmental monitoring.

The temporal distribution of publications (Figure 1) reveals a marked increase in research output after 2020, with a strong concentration of studies published between 2022 and 2025. This trend reflects the rapid integration of advanced ML techniques with expanding environmental datasets and computational platforms.

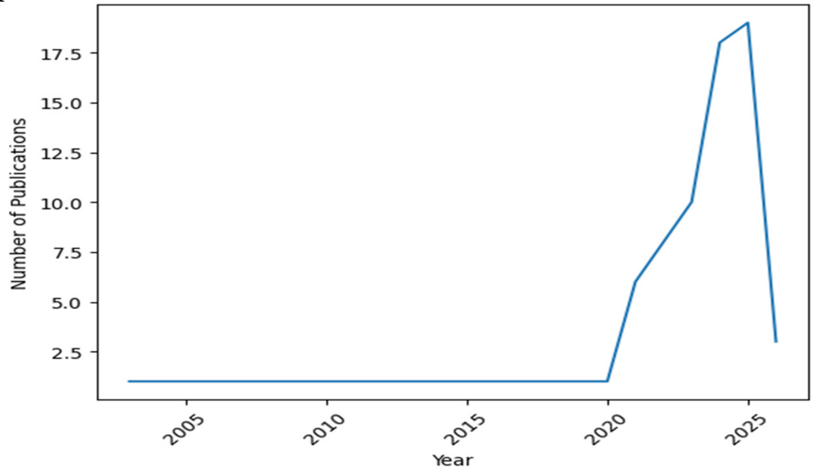

Fig. 1: Annual publication trend of reviewed studies (n = 69) on machine learning applications in environmental monitoring

Analysis by application domain (Figure 2) indicates that land use and land cover (LULC) mapping and remote sensing–based monitoring constitutes the largest proportion of studies, followed by water quality and hydrological applications, and climate and weather monitoring. Biodiversity and ecosystem studies also represent a significant share, while urban sustainability and cross-domain environmental AI reviews form smaller but growing categories.

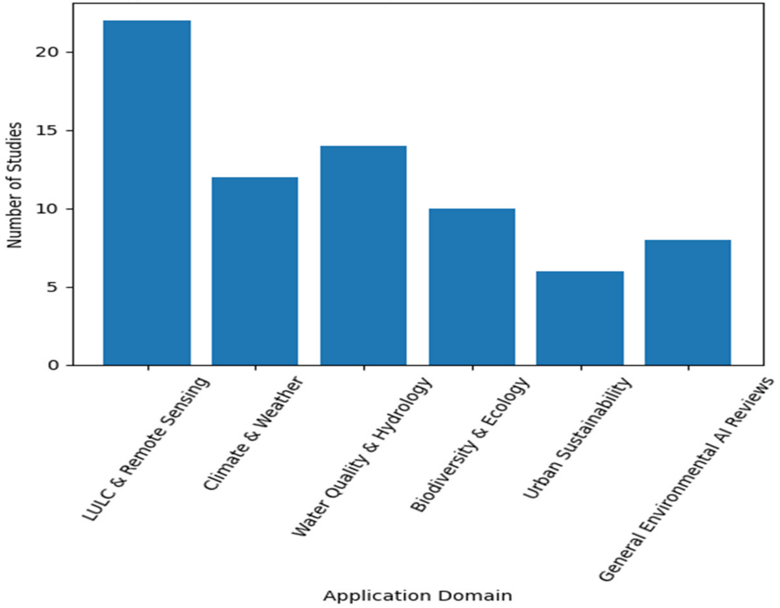

Fig. 2: Distribution of reviewed studies by environmental application domain

With respect to methodological trends (Figure 3), supervised learning approaches dominate the literature, particularly ensemble methods and neural networks, while deep learning methods show increasing adoption. Semi-supervised and reinforcement learning applications remain comparatively limited. Geographic distribution (Figure 4) demonstrates broad global engagement, with notable concentrations in Asia and Europe.

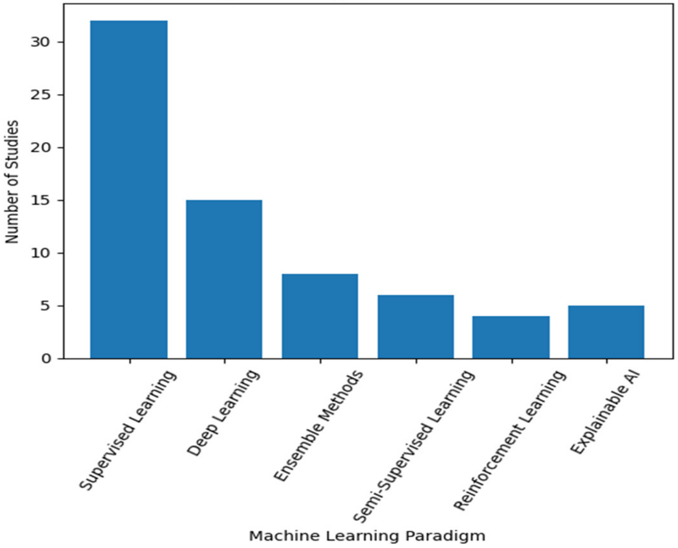

Fig. 3: Distribution of reviewed studies by dominant machine learning paradigm

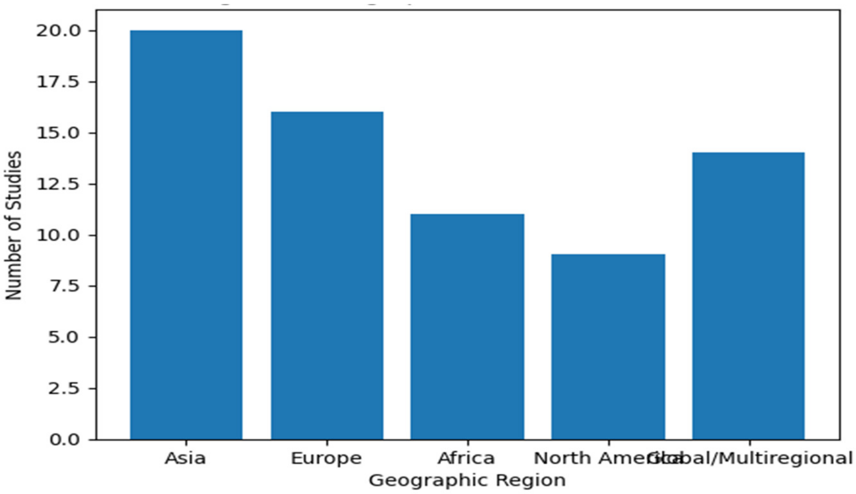

Fig. 4: Geographic distribution of reviewed studies by primary study region.

Overview of Machine Learning Paradigms

Machine learning refers to a class of computational techniques that enable systems to learn patterns and relationships from data without being explicitly programmed (Janiesch *et al.*, 2021). In environmental monitoring, ML algorithms are employed to classify, regress, cluster, and predict environmental variables based on observed data (Jemeļjanova *et al.*, 2024; Koldasbayeva *et al.*, 2024; Olawade *et al*., 2024; Roy *et al*., 2025). Depending on the learning strategy and availability of labeled data, machine learning approaches are broadly categorized into supervised, unsupervised, semi-supervised, and reinforcement learning (Obaido *et al*., 2024).

Supervised Learning

Supervised learning is the most widely used machine learning paradigm in environmental monitoring. It involves training algorithms using labeled datasets, where both input features and corresponding target outputs are known. Common supervised learning tasks include classification and regression.

In environmental applications, supervised learning has been extensively applied to land use and land cover classification, species distribution modeling, air quality prediction, and surface temperature estimation (Li *et al*., 2024; Olawade *et al*., 2024. Algorithms such as linear regression, decision trees, random forests, support vector machines, and artificial neural networks are frequently employed (Ikuemonisan *et al*., 2025; Loukika *et al*., 2021; Maurya and Pandey, 2026). Among these, ensemble methods such as random forests are particularly popular due to their robustness to noise, ability to handle non-linear relationships, and relatively low sensitivity to overfitting (Ojwang *et al*., 2024; Tahir *et al*., 2025).

Supervised learning models rely heavily on the quality and representativeness of training data. In many environmental contexts, collecting labeled data through field surveys or expert annotation can be challenging and expensive (Song *et al*., 2025). Nevertheless, supervised approaches remain dominant due to their interpretability and strong predictive performance when adequate training data are available (Cacciarelli and Kulahci, 2024).

Unsupervised Learning

Unsupervised learning techniques are used when labeled data are unavailable or limited. These methods aim to identify inherent structures, patterns, or groupings within the data. Common unsupervised learning algorithms include k-means clustering, hierarchical clustering, self-organizing maps, and principal component analysis.

In environmental monitoring, unsupervised learning is often applied to exploratory data analysis, anomaly detection, and dimensionality reduction (Russo *et al*., 2020). For example, clustering techniques have been used to identify vegetation types, climate zones, and pollution hotspots without prior class labels (Drogkoula *et al*., 2023; Khanam *et al*., 2025; Wang and Biljecki, 2022). Dimensionality reduction methods are also employed to simplify complex environmental datasets by reducing redundancy among correlated variables, particularly in multispectral and hyperspectral remote sensing data (Jarocińska *et al*., 2024). While unsupervised learning provides valuable insights, its outcomes can be difficult to interpret and validate due to the absence of ground truth data. As a result, unsupervised methods are frequently used in combination with supervised approaches to improve interpretability and performance (Montavon *et al*., 2022).

Semi-Supervised Learning

Semi-supervised learning represents a hybrid approach that combines a small amount of labeled data with a large volume of unlabeled data. This paradigm is particularly relevant in environmental monitoring, where labeled datasets are often scarce, but unlabeled data from remote sensing platforms and sensor networks are abundant (Cipriano *et al*., 2025). Semi-supervised learning techniques leverage the structure of unlabeled data to improve model performance and generalization. These methods have been applied in land cover classification, habitat mapping, and environmental change detection (Cipriano *et al*., 2025). By reducing reliance on extensive labeled datasets, semi-supervised learning offers a cost-effective alternative for large-scale environmental monitoring (Sun *et al*., 2022).

Reinforcement Learning

Reinforcement learning involves training an agent to make decisions through interaction with an environment by maximizing a cumulative reward. Although less commonly applied than supervised and unsupervised learning, reinforcement learning has shown potential in environmental management and monitoring tasks that involve sequential decision-making (Alotaibi and Nassif, 2024). Examples include adaptive environmental sampling, ecosystem management, and optimization of monitoring strategies (Zuccotto *et al*., 2024). Reinforcement learning models can learn optimal policies over time, making them suitable for dynamic environmental systems. However, their application in environmental monitoring remains limited due to high computational demands and the complexity of defining appropriate reward functions (Zuccotto *et al*., 2024).

Applications of Machine Learning in Environmental Monitoring

Machine learning techniques have been applied across a wide range of environmental monitoring domains. Their versatility allows them to address diverse challenges related to spatial analysis, temporal prediction, and pattern recognition in complex environmental systems (Alotaibi and Nassif, 2024; Olawade *et al*., 2024).

Land Use and Land Cover Mapping

Land use and land cover (LULC) mapping is one of the most established applications of machine learning in environmental monitoring (Alegbeleye *et al*., 2024; Bayable *et al*., 2025). Accurate LULC information is essential for urban planning, ecosystem assessment, climate modeling, and natural resource management (Gaur and Singh, 2023). Machine learning algorithms are widely used to classify satellite imagery into different land cover categories based on spectral, spatial, and temporal features derived from remote sensing data (Zafar *et al*., 2024).

Supervised classification methods such as random forests, support vector machines, and neural networks have demonstrated high accuracy in LULC mapping across diverse geographic regions. These models are capable of handling high-dimensional remote sensing data and capturing complex relationships between spectral signatures and land cover classes (Azedou *et al*., 2023). Time-series machine learning approaches have also been employed to detect land cover changes and monitor urban expansion and deforestation over time (Azedou *et al*., 2023).

Climate and Weather Monitoring

Machine learning has increasingly been used in climate and weather monitoring to model complex atmospheric processes and improve prediction accuracy (Zhang *et al*., 2025). ML algorithms have been applied to temperature forecasting, rainfall estimation, drought monitoring, and extreme weather event detection (Ojo and Ogunjo, 2022; Yang *et al*., 2024).

By integrating historical climate data with remote sensing observations, machine learning models can identify non-linear patterns that are difficult to capture using traditional statistical methods (Ojo and Ogunjo, 2022). These approaches have been particularly useful in downscaling climate models and filling gaps in observational datasets, thereby enhancing spatial and temporal resolution of climate information (Zhu *et al*., 2025).

Air and Water Quality Assessment

Monitoring air and water quality is critical for protecting public health and ecosystems (Zou *et al*., 2025). Machine learning models have been widely used to predict pollutant concentrations, identify pollution sources, and assess water quality parameters using both in situ and remotely sensed data (Li *et al*., 2025). Algorithms such as regression models, neural networks, and ensemble methods are commonly employed to model relationships between environmental variables and pollution indicators (Zou *et al*., 2025).

In urban environments, machine learning has been applied to integrate data from ground-based sensors, satellite observations, and meteorological variables to improve air quality monitoring and forecasting (Li *et al.,* 2025). Similarly, water quality assessments have benefited from ML-based analysis of physicochemical parameters and remote sensing data, enabling more efficient and scalable monitoring of aquatic system (Anand *et al*., 2024).

Biodiversity and Ecosystem Monitoring

Machine learning plays an important role in biodiversity monitoring and ecosystem assessment. Applications include species distribution modeling, habitat suitability analysis, and automated species identification (Moradi *et al*., 2025; Zare *et al*., 2024). By combining species occurrence records with environmental variables, ML models can predict species distributions and assess the impacts of environmental change on biodiversity patterns (Moradi *et al*., 2025).

Advances in image and acoustic recognition have also enabled automated monitoring of wildlife using camera traps and audio recordings (Buxton *et al*., 2018). These approaches reduce the need for manual data processing and enhance the scalability and efficiency of biodiversity monitoring efforts, particularly in large or remote landscapes (Buxton *et al*., 2018).

Environmental Data Sources and Platforms

The effectiveness of machine learning applications in environmental monitoring is strongly influenced by the availability, quality, and diversity of data sources (Olawade *et al*., 2024). Advances in Earth observation technologies, sensor networks, and data-sharing platforms have significantly expanded access to environmental data (Kumar *et al*., 2024). These datasets vary in spatial resolution, temporal frequency, and thematic content, making them suitable for different monitoring objectives. Machine learning techniques are increasingly used to integrate and analyze these heterogeneous data sources to generate actionable environmental insights (Kumar *et al*., 2024).

Remote Sensing Data

Remote sensing is one of the most important data sources for machine learning–based environmental monitoring (Nguyen *et al*., 2021). Satellite and airborne sensors provide continuous, large-scale observations of the Earth's surface and atmosphere. These data are particularly valuable for monitoring land cover, vegetation dynamics, surface temperature, water bodies, and atmospheric conditions (Andriambololonaharisoamalala *et al*., 2025).

Optical satellite imagery from multispectral and hyperspectral sensors is widely used for land use and land cover classification, vegetation analysis, and environmental change detection. Machine learning models are trained on spectral bands, vegetation indices, and texture features derived from these images to classify environmental features and monitor temporal changes. Hyperspectral data, with their high spectral resolution, enable detailed characterization of surface materials but also introduce high dimensionality, which is effectively handled using machine learning and dimensionality reduction techniques (Li *et al*., 2022).

Radar remote sensing, including synthetic aperture radar (SAR), provides complementary information to optical imagery. SAR data are particularly useful in regions with frequent cloud cover and for applications such as forest structure assessment, flood mapping, and soil moisture estimation (Liu, 2025). Machine learning algorithms have been applied to extract meaningful features from radar backscatter and polarization data, improving classification accuracy and robustness under varying environmental conditions (Shen *et al*., 2023).

Thermal remote sensing data are commonly used in surface temperature analysis and urban heat island studies (Liu, 2025). Machine learning models integrate thermal bands with land cover and meteorological variables to estimate land surface temperature and assess heat-related environmental risks (Shen *et al*., 2023). The increasing availability of high-temporal-resolution satellite data has further enhanced the applicability of machine learning for near-real-time environmental monitoring (Nguyen *et al*., 2021).

In-situ and Sensor-Based Data

In situ observations and sensor networks provide ground-based measurements that are essential for calibrating, validating, and complementing remote sensing data (Pause *et al*., 2016. These datasets include meteorological observations, air and water quality measurements, soil properties, and hydrological variables. Although spatially limited compared to satellite data, in situ measurements offer high accuracy and temporal resolution (Ortenzi *et al*., 2024).

Machine learning techniques are widely used to integrate sensor data with remote sensing observations to improve prediction accuracy (Li *et al*., 2022). For example, air quality monitoring systems combine data from ground-based sensors, satellite observations, and weather stations to model pollutant concentrations across urban and regional scales (Wu *et al*., 2024). Similarly, hydrological monitoring benefits from machine learning–based analysis of streamflow, precipitation, and water quality data collected from sensor network (Zhu *et al*., 2022).

The growth of Internet of Things (IoT) technologies has further expanded the availability of real-time environmental sensor data (Zhu *et al*., 2022). Machine learning models can process continuous data streams from sensor networks to detect anomalies, forecast environmental conditions, and support early warning systems. These applications are particularly relevant for disaster monitoring, pollution detection, and climate-related risk assessment (Olawade *et al*., 2024).

Climate and Environmental Databases

Global and regional climate databases provide long-term records of environmental variables that are essential for trend analysis and predictive modeling (Olawade *et al*., 2024). These datasets include temperature, precipitation, humidity, wind speed, and other climatic parameters derived from observational records and reanalysis products. Machine learning models use these data to analyze historical trends, predict future conditions, and assess climate variability and extremes (Zhu *et al*., 2022).

Environmental databases also include land cover maps, soil datasets, biodiversity records, and hydrological datasets compiled by research institutions and international organizations (Alegbeleye *et al*., 2024). The availability of standardized and openly accessible databases has facilitated the application of machine learning across different environmental domains (Olawade *et al*., 2024; Valencia-Arias *et al*., 2025). By combining multiple datasets, ML models can capture complex interactions among environmental variables and improve predictive performance.

However, environmental databases often contain missing values, inconsistencies, and uncertainties. Machine learning techniques such as data imputation, feature selection, and ensemble modeling are commonly employed to address these challenges (Hasan *et al*., 2021). The ability of ML algorithms to handle imperfect data makes them particularly suitable for environmental applications where data quality may vary across space and time (Zhu *et al*., 2022).

Citizen Science and Crowdsourced Data

Citizen science initiatives and crowdsourced data platforms have emerged as valuable sources of environmental information (Alfonso *et al*., 2022). These data are generated by volunteers and community members through observations, mobile applications, and online platforms. Examples include species occurrence records, pollution reports, and environmental hazard observations (Hognogi *et al*., 2023).

Machine learning plays a critical role in processing and validating citizen science data, which are often characterized by varying levels of accuracy and completeness (Alfonso *et al*., 2022. ML algorithms are used to filter noise, identify outliers, and assess data reliability. In biodiversity monitoring, for instance, machine learning–based image recognition tools assist in species identification from photographs submitted by citizen scientists (Hognogi *et al*., 2023).

The integration of citizen science data with traditional monitoring datasets enhances spatial coverage and supports participatory environmental monitoring (Alfonso *et al*., 2022). While challenges related to data quality and standardization remain, machine learning provides effective tools for extracting meaningful information from crowdsourced (Zhu *et al*., 2022).

Data Processing and Analysis Platforms

The increasing volume and complexity of environmental data have led to the development of cloud-based platforms and computational tools that support machine learning–based analysis (Alotaibi and Nassif, 2024; Olawade *et al*., 2024). These platforms enable users to access, process, and analyze large datasets without the need for extensive local computing resources.

Cloud-based geospatial platforms facilitate the integration of satellite imagery, climate data, and machine learning algorithms for large-scale environmental monitoring (Olawade *et al*., 2024). Machine learning workflows implemented on these platforms support tasks such as image classification, time-series analysis, and environmental modeling. The availability of application programming interfaces (APIs) and open-source libraries has further lowered the barrier to entry for researchers and practitioners (Davison *et al*., 2025). In addition to cloud platforms, open-source programming environments and machine learning libraries provide flexible tools for environmental data analysis. These tools support a wide range of algorithms and visualization techniques, enabling researchers to customize models for specific monitoring objectives. The combination of accessible data platforms and machine learning tools has significantly accelerated research and operational applications in environmental monitoring (Zhu *et al*., 2022).

Challenges, Limitations, and Future Directions

Despite the growing adoption of machine learning in environmental monitoring, several challenges and limitations continue to constrain its effectiveness and broader application. These challenges relate to data quality, model interpretability, computational demands, and the practical integration of machine learning outputs into environmental decision-making processes (Alotaibi and Nassif, 2024; Olawade *et al*., 2024).

Data Quality and Availability

One of the primary challenges in applying machine learning to environmental monitoring is the quality and availability of data. Environmental datasets are often characterized by missing values, measurement errors, spatial and temporal inconsistencies, and varying resolutions (Zhu *et al*., 2022). In many regions, particularly in developing countries, long-term and high-resolution environmental data remain limited (Ayanlade *et al*., 2022).

Machine learning models are highly dependent on the quality of input data, and poor data quality can lead to biased or unreliable predictions (Mohammed *et al*., 2025). While techniques such as data preprocessing, imputation, and normalization can mitigate some of these issues, they do not fully eliminate underlying data limitations. Improving data collection infrastructure, promoting open data policies, and enhancing data standardization remain critical priorities for advancing ML-based environmental monitoring (Zhu *et al*., 2022).

Model Interpretability and Transparency

Another significant limitation of machine learning models, especially complex algorithms such as deep neural networks, is their lack of interpretability. Many ML models operate as “black boxes,” making it difficult to understand how input variables influence predictions. This lack of transparency poses challenges for environmental decision-making, where interpretability and accountability are essential (Cipriano *et al*., 2025).

In environmental monitoring applications, stakeholders often require clear explanations of model outputs to support policy formulation and management actions. Although interpretable machine learning techniques and model explanation tools have been developed, their adoption in environmental studies remains limited (Zhu *et al*., 2022). Future research should prioritize the development and application of explainable machine learning approaches that balance predictive accuracy with interpretability (Mohammed *et al*., 2025.

Generalization and Transferability

Machine learning models trained in one geographic region or environmental context may not perform well when applied elsewhere. Differences in climate, land cover, sensor characteristics, and socio-environmental conditions can limit model transferability. This challenge is particularly relevant for global and regional environmental monitoring initiatives.

Addressing generalization issues requires the development of models that are robust to spatial and temporal variability. Techniques such as transfer learning, domain adaptation, and ensemble modeling offer promising solutions, but their practical implementation in environmental monitoring is still evolving (Zhu *et al*., 2022). Ensuring that machine learning models can generalize across diverse environmental conditions remains an important research direction (Mohammed *et al*., 2025).

Computational and Technical Constraints

The application of advanced machine learning models often requires substantial computational resources and technical expertise. High-resolution remote sensing data, deep learning architectures, and large-scale environmental datasets can be computationally intensive to process and analyze. These requirements may limit the adoption of ML-based approaches in resource-constrained settings.

While cloud computing platforms and open-source tools have improved accessibility, challenges related to internet connectivity, data storage, and technical capacity persist. Capacity building, training, and the development of user-friendly tools are essential for enabling wider adoption of machine learning in environmental monitoring, particularly in low- and middle-income regions (Zhu *et al*., 2022).

Future Directions

Future research in machine learning–based environmental monitoring is expected to focus on improved data integration, model transparency, and operational deployment. The integration of multi-source and multi-scale data, including remote sensing, sensor networks, and socio-economic datasets, will enhance the ability of ML models to capture complex environmental processes (Brown *et al*., 2025; Olawade *et al*., 2024).

Advances in explainable artificial intelligence are likely to play an important role in improving trust and usability of machine learning models in environmental. In addition, the incorporation of domain knowledge and hybrid modeling approaches that combine physical models with machine learning techniques may improve model robustness and interpretability.

The development of real-time and near-real-time monitoring systems represents another promising direction. By leveraging streaming data and adaptive machine learning models, environmental monitoring systems can support early warning, rapid response, and adaptive management. As computational tools and data availability continue to improve, machine learning is expected to become an increasingly integral component of environmental monitoring frameworks (Zhu *et al*., 2022).

Conclusion

Machine learning has emerged as a transformative tool in environmental monitoring, offering powerful capabilities for analyzing complex, large-scale, and heterogeneous datasets. Its applications span a wide range of environmental domains, including land use and land cover mapping, climate and weather monitoring, air and water quality assessment, and biodiversity conservation. By enabling more accurate predictions and efficient data analysis, machine learning supports improved understanding and management of environmental systems.

This narrative review has provided an overview of key machine learning paradigms, major application areas, and commonly used environmental data sources. It has also highlighted challenges related to data quality, model interpretability, generalization, and computational constraints. While these limitations remain significant, ongoing advances in data collection, algorithm development, and computational infrastructure continue to expand the potential of machine learning in environmental monitoring.

As environmental challenges become increasingly complex and interconnected, the role of machine learning in supporting evidence-based decision-making is likely to grow. Continued collaboration between environmental scientists, data scientists, and policymakers will be essential for translating machine learning research into practical monitoring and management solutions. By addressing current limitations and embracing emerging opportunities, machine learning can contribute meaningfully to more effective and sustainable environmental monitoring systems.

References

Alegbeleye, O. M., Rotimi, Y. O., Shomide, P., Oyediran, A., Ogundipe, O., & Akintunde-Alo, A. (2024). Land use land cover (LULC) analysis in Nigeria: A systematic review of data, methods, and platforms with future prospects. *Bulletin of the National Research Centre*, *48*(1), 127. https://doi.org/10.1186/s42269-024-01286-z

Alfonso, L., Gharesifard, M. and Wehn, U. (2022). Analysing the value of environmental citizen-generated data: Complementarity and cost per observation. *Journal of Environmental Management*, 303: 114157. https://doi.org/10.1016/j.jenvman.2021.114157

Alotaibi, E. and Nassif, N. (2024). Artificial intelligence in environmental monitoring: In-depth analysis. *Discover Artificial Intelligence*, 4(1): 84. https://doi.org/10.1007/s44163-024-00198-1

Anand, V., Oinam, B. and Wieprecht, S. (2024). Machine learning approach for water quality predictions based on multispectral satellite imageries. *Ecological Informatics*, 84: 102868. https://doi.org/10.1016/j.ecoinf.2024.102868

Andriambololonaharisoamalala, R., Helmholz, P., Ivánová, I., Jones, E., Soon, S., Bulatov, D. and Song, Y. (2025). Enhancing the spatial and temporal resolution of satellite-derived land surface temperature in urban environments: A systematic literature review. *Urban Climate*, 60: 102345. https://doi.org/10.1016/j.uclim.2025.102345

Antonelli, A., Dhanjal-Adams, K.L. and Silvestro, D. (2023). Integrating machine learning, remote sensing and citizen science to create an early warning system for biodiversity. *Plants, People, Planet*, 5(3): 307–316. https://doi.org/10.1002/ppp3.10337

Ayanlade, A., Oluwaranti, A., Ayanlade, O.S., Borderon, M., Sterly, H., Sakdapolrak, P., Jegede, M.O., Weldemariam, L.F. and Ayinde, A. F.O. (2022). Extreme climate events in sub-Saharan Africa: A call for improving agricultural technology transfer to enhance adaptive capacity. *Climate Services*, 27: 100311. https://doi.org/10.1016/j.cliser.2022.100311

Azedou, A., Amine, A., Kisekka, I., Lahssini, S., Bouziani, Y. and Moukrim, S. (2023). Enhancing Land Cover/Land Use (LCLU) classification through a comparative analysis of hyperparameters optimization approaches for deep neural network (DNN). *Ecological Informatics*, 78: 102333. https://doi.org/10.1016/j.ecoinf.2023.102333

Bayable, G., Gebrie, G., Melese, T. and Melaku, A. (2025). Land use/cover classification using machine learning algorithms and their impacts on land surface temperature and soil moisture in the Alawuha Watershed, Ethiopia. *Environmental and Sustainability Indicators*, 27: 100797. https://doi.org/10.1016/j.indic.2025.100797

Brown, L.E., Maavara, T., Zhang, J., Chen, X., Klaar, M., Moshe, F. O., Ben-Zur, E., Stein, S., Grayson, R., Carter, L., Levintal, E., Gal, G., Ziv, P., Tarkowski, F., Pathak, D., Khamis, K., Barquín, J., Philamore, H., Gradilla-Hernández, M.S. and Arnon, S. (2025). Integrating sensor data and machine learning to advance the science and management of river carbon emissions. *Critical Reviews in Environmental Science and Technology*, 55(9): 600–623. https://doi.org/10.1080/10643389.2024.2429912

Buxton, R.T., Lendrum, P.E., Crooks, K.R. and Wittemyer, G. (2018). Pairing camera traps and acoustic recorders to monitor the ecological impact of human disturbance. *Global Ecology and Conservation*, 16: e00493. https://doi.org/10.1016/j.gecco.2018.e00493

Cacciarelli, D. and Kulahci, M. (2024). Active learning for data streams: A survey. *Machine Learning*, 113(1): 185–239. https://doi.org/10.1007/s10994-023-06454-2

Caiafa, C.F., Sun, Z., Tanaka, T., Marti-Puig, P. and Solé-Casals, J. (2021). Machine Learning Methods with Noisy, Incomplete or Small Datasets. *Applied Sciences*, 11(9): 4132. https://doi.org/10.3390/app11094132

Cipriano, C., Noce, S., Mereu, S. and Santini, M. (2025). Algorithms going wild – A review of machine learning techniques for terrestrial ecology. *Ecological Modelling*, 506: 111164. https://doi.org/10.1016/j.ecolmodel.2025.111164

Clark, W.C. and Dickson, N.M. (2003). Sustainability science: The emerging research program. *Proceedings of the National Academy of Sciences*, 100(14): 8059–8061. https://doi.org/10.1073/pnas.1231333100

Davison, A.M., De Koning, K., Taubert, F. and Schakel, J.-K. (2025). Automated near real-time monitoring in ecology: Status quo and ways forward. *Ecological Informatics*, 89: 103157. https://doi.org/10.1016/j.ecoinf.2025.103157

Demissie, W.A., Sebastiani, L. and Rossetto, R. (2026). Integration of artificial intelligence and remote sensing for crop yield prediction and crop growth parameter estimation in Mediterranean agroecosystems: Methodologies, emerging technologies, research gaps, and future directions. *European Journal of Agronomy*, 173: 127894. https://doi.org/10.1016/j.eja.2025.127894

Drogkoula, M., Kokkinos, K. and Samaras, N. (2023). A Comprehensive Survey of Machine Learning Methodologies with Emphasis in Water Resources Management. *Applied Sciences*, 13(22): 12147. https://doi.org/10.3390/app132212147

Espíndola, R.P., Picanço, M.M., De Andrade, L.P. and Ebecken, N.F. F. (2025). Applications of Machine Learning Methods in Sustainable Forest Management. *Climate*, 13(8): 159. https://doi.org/10.3390/cli13080159

Gaur, S. and Singh, R. (2023). A Comprehensive Review on Land Use/Land Cover (LULC) Change Modeling for Urban Development: Current Status and Future Prospects. *Sustainability*, 15(2): 903. https://doi.org/10.3390/su15020903

Grant, M.J. and Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information & Libraries Journal*, 26(2): 91–108. https://doi.org/10.1111/j.1471-1842.2009.00848.x

Grønstad, A. (2026). Scoping reviews vis-à-vis systematic reviews and integrative reviews in management research: Guidelines for good practice. *Review of Managerial Science*, 20(1): 221–250. https://doi.org/10.1007/s11846-025-00901-1

Hasan, Md. K., Alam, Md. A., Roy, S., Dutta, A., Jawad, Md. T. and Das, S. (2021). Missing value imputation affects the performance of machine learning: A review and analysis of the literature (2010–2021). *Informatics in Medicine Unlocked*, 27: 100799. https://doi.org/10.1016/j.imu.2021.100799

Hognogi, G.-G., Meltzer, M., Alexandrescu, F. and Ştefănescu, L. (2023). The role of citizen science mobile apps in facilitating a contemporary digital agora. *Humanities and Social Sciences Communications*, 10(1): 863. https://doi.org/10.1057/s41599-023-02358-7

Ikuemonisan, F.E., Kayode, Y.O., Ogunjo, S.T., Odubote, O.B. and Okedeyi, S.A. (2025). Modeling heatwave trends from land cover dynamics using satellite observations and machine learning in Ibadan, Nigeria. *Discover Geoscience*, 3(1): 254. https://doi.org/10.1007/s44288-025-00361-w

Janiesch, C., Zschech, P. and Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, 31(3): 685–695. https://doi.org/10.1007/s12525-021-00475-2

Jarocińska, A., Kopeć, D. and Kycko, M. (2024). Comparison of dimensionality reduction methods on hyperspectral images for the identification of heathlands and mires. *Scientific Reports*, 14(1): 27662. https://doi.org/10.1038/s41598-024-79209-1

Jemeļjanova, M., Kmoch, A. and Uuemaa, E. (2024). Adapting machine learning for environmental spatial data—A review. *Ecological Informatics*, 81: 102634. https://doi.org/10.1016/j.ecoinf.2024.102634

Khanam, M.M., Uddin, M.K. and Kazi, J.U. (2025). Advances in machine learning for the detection and characterization of microplastics in the environment. *Frontiers in Environmental Science*, 13: 1573579. https://doi.org/10.3389/fenvs.2025.1573579

Koldasbayeva, D., Tregubova, P., Gasanov, M., Zaytsev, A., Petrovskaia, A. and Burnaev, E. (2024). Challenges in data-driven geospatial modeling for environmental research and practice. *Nature Communications*, 15(1): 10700. https://doi.org/10.1038/s41467-024-55240-8

Kumar, S., Imen, S., Sridharan, V.K., Gupta, A., McDonald, W., Ramirez-Avila, J.J., Abdul-Aziz, O.I., Talchabhadel, R., Gao, H., Quinn, N.W.T., Weiss, W.J., Poulose, T., Palmate, S.S., Lee, C.M. and Baskaran, L. (2024). Perceived barriers and advances in integrating earth observations with water resources modeling. *Remote Sensing Applications: Society and Environment*, 33: 101119. https://doi.org/10.1016/j.rsase.2023.101119

Li, H., Cui, J., Zhang, X., Han, Y. and Cao, L. (2022). Dimensionality Reduction and Classification of Hyperspectral Remote Sensing Image Feature Extraction. *Remote Sensing*, 14(18): 4579. https://doi.org/10.3390/rs14184579

Li, S., Yang, X., Cui, P., Sun, Y. and Song, B. (2024). Machine-Learning-Algorithm-Based Prediction of Land Use/Land Cover and Land Surface Temperature Changes to Characterize the Surface Urban Heat Island Phenomena over Harbin, China. *Land*, 13(8): 1164.

https://doi.org/10.3390/land13081164

Li, Y., Chen, B., Yang, S., Jiao, Z., Zhang, M., Yang, Y. and Gao, Y. (2025). Advances in environmental pollutant detection techniques: Enhancing public health monitoring and risk assessment. *Environment International*, 197: 109365. https://doi.org/10.1016/j.envint.2025.109365

Liu, Y. (2025). Applicability of synthetic aperture radar in remote sensing flood area calculation. *Systems and Soft Computing*, 7: 200253. https://doi.org/10.1016/j.sasc.2025.200253

Loukika, K.N., Keesara, V.R. and Sridhar, V. (2021). Analysis of Land Use and Land Cover Using Machine Learning Algorithms on Google Earth Engine for Munneru River Basin, India. *Sustainability*, 13(24): 13758. https://doi.org/10.3390/su132413758

Maurya, N.K. and Pandey, M.K. (2026). Application of machine learning in forest monitoring: Recent progress and future challenges. In *Advanced Geospatial and Ground Based Techniques in Forest Monitoring* (pp. 273–299). Elsevier. https://doi.org/10.1016/B978-0-443-18949-4.00006-3

Mohammed, S., Budach, L., Feuerpfeil, M., Ihde, N., Nathansen, A., Noack, N., Patzlaff, H., Naumann, F. and Harmouch, H. (2025). The effects of data quality on machine learning performance on tabular data. *Information Systems*, 132: 102549. https://doi.org/10.1016/j.is.2025.102549

Montavon, G., Kauffmann, J., Samek, W. and Müller, K.-R. (2022). Explaining the Predictions of Unsupervised Learning Models. In A. Holzinger, R. Goebel, R. Fong, T. Moon, K.-R. Müller, and W. Samek (Eds.), *xxAI - Beyond Explainable AI* (Vol. 13200, pp. 117–138). Springer International Publishing. https://doi.org/10.1007/978-3-031-04083-2_7

Moradi, S., Moradi, H. and Rezvani, A. (2025). Integrating remote sensing and species distribution modeling to assess the impacts of flood disturbance on the species habitat in the central Zagros. *Discover Applied Sciences*, 7(9): 935. https://doi.org/10.1007/s42452-025-07582-1

Nakhle, P., Stamos, I., Proietti, P. and Siragusa, A. (2024). Environmental monitoring in European regions using the sustainable development goals (SDG) framework. *Environmental and Sustainability Indicators*, 21: 100332. https://doi.org/10.1016/j.indic.2023.100332

Nguyen, H.A.T., Sophea, T., Gheewala, S.H., Rattanakom, R., Areerob, T. and Prueksakorn, K. (2021). Integrating remote sensing and machine learning into environmental monitoring and assessment of land use change. *Sustainable Production and Consumption*, 27: 1239–1254. https://doi.org/10.1016/j.spc.2021.02.025

Obaido, G., Mienye, I.D., Egbelowo, O.F., Emmanuel, I.D., Ogunleye, A., Ogbuokiri, B., Mienye, P. and Aruleba, K. (2024). Supervised machine learning in drug discovery and development: Algorithms, applications, challenges, and prospects. *Machine Learning with Applications*, 17: 100576. https://doi.org/10.1016/j.mlwa.2024.100576

Ojo, O.S. and Ogunjo, S.T. (2022). Machine learning models for prediction of rainfall over Nigeria. *Scientific African*, 16: e01246. https://doi.org/10.1016/j.sciaf.2022.e01246

Ojwang, G.O., Ogutu, J.O., Said, M.Y., Ojwala, M.A., Kifugo, S.C., Verones, F., Graae, B.J., Buitenwerf, R. and Olff, H. (2024). An integrated hierarchical classification and machine learning approach for mapping land use and land cover in complex social-ecological systems. *Frontiers in Remote Sensing*, 4: 1188635. https://doi.org/10.3389/frsen.2023.1188635

Olawade, D.B., Wada, O.Z., Ige, A.O., Egbewole, B.I., Olojo, A. and Oladapo, B.I. (2024). Artificial intelligence in environmental monitoring: Advancements, challenges, and future directions. *Hygiene and Environmental Health Advances*, 12: 100114. https://doi.org/10.1016/j.heha.2024.100114

Ortenzi, S., Cencetti, C., Mincu, F.-I., Neculau, G., Chendeş, V., Ciabatta, L., Massari, C. and Di Matteo, L. (2024). Comparing Satellite Soil Moisture Products Using *in-situ* Observations over an Instrumented Experimental Basin in Romania. *Remote Sensing*, 16(17): 3283. https://doi.org/10.3390/rs16173283

Pause, M., Schweitzer, C., Rosenthal, M., Keuck, V., Bumberger, J., Dietrich, P., Heurich, M., Jung, A. and Lausch, A. (2016). In Situ/Remote Sensing Integration to Assess Forest Health—A Review. *Remote Sensing*, 8(6): 471. https://doi.org/10.3390/rs8060471

Pichler, M. and Hartig, F. (2023). Machine learning and deep learning—A review for ecologists. *Methods in Ecology and Evolution*, 14(4): 994–1016. https://doi.org/10.1111/2041-210X.14061

Roy, S.D., Khekare, G., Chhajed, S. and Victor, A.S. (2025). Integrating classification, regression, and time series models to assess biochar safety, optimize pollutant removal, and predict environmental impacts. *Frontiers in Soil Science*, *5*, 1661097. https://doi.org/10.3389/fsoil.2025.1661097

Russo, S., Lürig, M., Hao, W., Matthews, B. and Villez, K. (2020). Active learning for anomaly detection in environmental data. *Environmental Modelling and Software*, 134: 104869. https://doi.org/10.1016/j.envsoft.2020.104869

Shen, Z., Miao, J., Wang, J., Zhao, D., Tang, A. and Zhen, J. (2023). Evaluating Feature Selection Methods and Machine Learning Algorithms for Mapping

Mangrove Forests Using Optical and Synthetic Aperture Radar Data. *Remote Sensing*, 15(23): 5621. https://doi.org/10.3390/rs15235621

Song, B., Zhao, S., Dang, L., Wang, H. and Xu, L. (2025). A survey on learning from data with label noise via deep neural networks. *Systems Science & Control Engineering*, 13(1): 2488120. https://doi.org/10.1080/21642583.2025.2488120

Sun, C., Wu, J., Chen, H. and Du, C. (2022). SemiSANet: A Semi-Supervised High-Resolution Remote Sensing Image Change Detection Model Using Siamese Networks with Graph Attention. *Remote Sensing*, 14(12): 2801. https://doi.org/10.3390/rs14122801

Tahir, Z., Haseeb, M., Mahmood, S.A., Batool, S., Abdullah-Al-Wadud, M., Ullah, S. and Tariq, A. (2025). Predicting land use and land cover changes for sustainable land management using CA-Markov modelling and GIS techniques. *Scientific Reports*, 15(1): 3271. https://doi.org/10.1038/s41598-025-87796-w

Valencia-Arias, A., Vásquez Coronado, M.H., Clavo Medina, M.L., Morales Garcia, M.J., Requejo Arias, A.R. and Saavedra Tirado, J.I. (2025). Research trends in the application of machine learning in sustainability practices based on a bibliometric analysis. *Sustainable Futures*, 10: 100987. https://doi.org/10.1016/j.sftr.2025.100987

Wang, J. and Biljecki, F. (2022). Unsupervised machine learning in urban studies: A systematic review of applications. *Cities*, 129: 103925. https://doi.org/10.1016/j.cities.2022.103925

Wang, L.A.Y., Wang, G., Xue, B., Wu, J., Wang, Y. and Wang, Q. (2025). Building a global early-warning system for environmental risks with remote sensing. *Fundamental Research*, S2667325825006181. https://doi.org/10.1016/j.fmre.2025.11.002

Wang, Z., Akber, M.A. and Aziz, A.A. (2025). Application of Remote Sensing and Geographic Information Systems for Monitoring and Managing Chili Crops: A Systematic Review. *Remote Sensing*, 17(16): 2827. https://doi.org/10.3390/rs17162827

Wu, B., Zhang, Y., Wang, Y., He, Y., Wang, J., Wu, Y., Lin, X. and Wu, S. (2024). Mitigation of urban heat island in China (2000–2020) through vegetation-induced cooling. *Sustainable Cities and Society*, 112 105599. https://doi.org/10.1016/j.scs.2024.105599

Yang, R., Hu, J., Li, Z., Mu, J., Yu, T., Xia, J., Li, X., Dasgupta, A. and Xiong, H. (2024). Interpretable machine learning for weather and climate prediction: A review. *Atmospheric Environment*, 338: 120797. https://doi.org/10.1016/j.atmosenv.2024.120797

Zafar, Z., Zubair, M., Zha, Y., Fahd, S. and Ahmad Nadeem, A. (2024). Performance assessment of machine learning algorithms for mapping of land use/land cover using remote sensing data. *The Egyptian Journal of Remote Sensing and Space Sciences*, 27(2): 216–226. https://doi.org/10.1016/j.ejrs.2024.03.003

Zare, M., Moameri, M., Ghorbani, A., Sahragard, H.P., Mostafazadeh, R., Dadjou, F. and Biswas, A. (2024). Modeling habitat suitability of Dorema ammoniacum D Don. In the rangelands of central Iran. *Scientific Reports*, 14(1): 16185. https://doi.org/10.1038/s41598-024-61073-8

Zhang, H., Liu, Y., Zhang, C. and Li, N. (2025). Machine Learning Methods for Weather Forecasting: A Survey. *Atmosphere*, 16(1): 82. https://doi.org/10.3390/atmos16010082

Zhong, S., Zhang, K., Bagheri, M., Burken, J.G., Gu, A., Li, B., Ma, X., Marrone, B.L., Ren, Z.J., Schrier, J., Shi, W., Tan, H., Wang, T., Wang, X., Wong, B.M., Xiao, X., Yu, X., Zhu, J.-J. and Zhang, H. (2021). Machine Learning: New Ideas and Tools in Environmental Science and Engineering. *Environmental Science and Technology*, acs.est.1c01339. https://doi.org/10.1021/acs.est.1c01339

Zhu, H., Zhou, Q. and Krisp, J.M. (2025). Exploring machine learning approaches for precipitation downscaling. *Geo-Spatial Information Science*, 28(6): 2673–2689. https://doi.org/10.1080/10095020.2025.2477547

Zhu, M., Wang, J., Yang, X., Zhang, Y., Zhang, L., Ren, H., Wu, B. and Ye, L. (2022). A review of the application of machine learning in water quality evaluation. *Eco-Environment and Health*, 1(2): 107–116. https://doi.org/10.1016/j.eehl.2022.06.001

Zou, S., Ju, H. and Zhang, J. (2025). Water Quality Management in the Age of AI: Applications, Challenges, and Prospects. *Water*, 17(11): 1641. https://doi.org/10.3390/w17111641

Zuccotto, M., Castellini, A., Torre, D.L., Mola, L. and Farinelli, A. (2024). Reinforcement learning applications in environmental sustainability: A review. *Artificial Intelligence Review*, 57(4): 88. https://doi.org/10.1007/s10462-024-10706-5