

Review

A Comprehensive Review of Machine Learning for Water Quality Prediction over the Past Five Years

Xiaohui Yan ^{1,2,3,*} , Tianqi Zhang ³, Wenying Du ² , Qingjia Meng ^{1,*}, Xinghan Xu ³ and Xiang Zhao ¹

- ¹ State Environmental Protection Key Laboratory of Estuarine and Coastal Environment, Chinese Research Academy of Environmental Sciences, Beijing 100012, China; zhaoxiang1312@foxmail.com
- ² National Engineering Research Center for Geographic Information System, China University of Geosciences, Wuhan 430074, China; duwenying@cug.edu.cn
- ³ Department of Hydraulic Engineering, Dalian University of Technology, Dalian 116024, China; zhangtianqi_adela@163.com (T.Z.); xuxh2023@dlut.edu.cn (X.X.)
- * Correspondence: mengqj@craes.org.cn (Q.M.); yanxh@dlut.edu.cn (X.Y.)

Abstract: Water quality prediction, a well-established field with broad implications across various sectors, is thoroughly examined in this comprehensive review. Through an exhaustive analysis of over 170 studies conducted in the last five years, we focus on the application of machine learning for predicting water quality. The review begins by presenting the latest methodologies for acquiring water quality data. Categorizing machine learning-based predictions for water quality into two primary segments—indicator prediction and water quality index prediction—further distinguishes between single-indicator and multi-indicator predictions. A meticulous examination of each method's technical details follows. This article explores current cutting-edge research trends in machine learning algorithms, providing a technical perspective on their application in water quality prediction. It investigates the utilization of algorithms in predicting water quality and concludes by highlighting significant challenges and future research directions. Emphasis is placed on key areas such as hydrodynamic water quality coupling, effective data processing and acquisition, and mitigating model uncertainty. The paper provides a detailed perspective on the present state of application and the principal characteristics of emerging technologies in water quality prediction.

Keywords: machine learning; water quality prediction; water quality index; remote sensing; coastal area



Citation: Yan, X.; Zhang, T.; Du, W.; Meng, Q.; Xu, X.; Zhao, X.

A Comprehensive Review of Machine Learning for Water Quality Prediction over the Past Five Years. *J. Mar. Sci. Eng.* **2024**, *12*, 159. <https://doi.org/10.3390/jmse12010159>

Academic Editor: Ryan J.K. Dunn

Received: 24 November 2023

Revised: 24 December 2023

Accepted: 11 January 2024

Published: 13 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the industrialization and urbanization of coastal areas have experienced increasing population pressures [1–4]. A significant volume of wastewater generated by local residents is often discharged into the sea after undergoing rudimentary water treatment [5–7]. The discharge of sewage into the receiving water body will significantly increase turbidity and organic and inorganic substances, thereby changing the living environment of marine organisms [8]. Discrepancies between sewage treatment standards and marine water quality standards, particularly concerning specific key pollution values in certain countries, have led to the degradation of seawater quality due to the discharge of untreated sewage [9]. Furthermore, the growing industrial sector and population necessitate additional land for infrastructure development [10–12], making coastal reclamation [13–15] a prominent topic. This activity, in turn, has instigated alterations in marine hydrodynamic factors in coastal regions [16,17]. Reports illustrate the adverse effects of coastal erosion, hurricanes, typhoons [18,19], and coastal flooding [20]. These occurrences not only have environmental consequences but also impact aquaculture development [21,22] and bathing areas [23,24]. To address these challenges, considerable efforts have been directed towards predicting water quality and hydrodynamic movements in coastal regions.

Numerical models, which derive results from rules and data [25], have been a frequently employed method for predicting changes in water quality [8,26–28] and hydro-

dynamic movement [29], with a variety of models available. Despite their capability to generate accurate simulations, their usage is significantly constrained by limitations. Firstly, accuracy heavily depends on the selection of model parameters [30,31], posing a challenge for beginners lacking a solid grasp of the underlying theories. Secondly, in an effort to minimize complexity, numerous assumptions are integrated into numerical models, often tailored to specific situations, rendering them less adaptable for direct application in different environments. Furthermore, long-term and large-scale predictions have proven to be exceedingly time consuming [32,33] and demand substantial computational space, making it challenging to achieve real-time and prompt results during unexpected situations. In response to these challenges, scientists have been actively seeking solutions to address these inherent shortcomings.

In recent years, advancements in computer science have enabled integrating machine learning into numerical simulations, offering high-speed and precise predictions [34,35]. Initially, the application of machine learning in coastal simulation faced limitations due to the challenge of obtaining oceanic data. However, recent progress in satellite remote sensing and unmanned aerial vehicle observation has alleviated these data constraints [36–38]. For instance, Nagur Cherukuru et al. [39] developed a semi-analytical remote sensing model, facilitating the retrieval of suspended sediment and dissolved organic carbon in coastal waters. This breakthrough enables the exploration of potential correlations between water quality metrics and satellite imagery. The substantial increase in available data, a fundamental aspect for training machine learning models [40,41], has played a pivotal role. Machine learning, as a field in computer science, seeks implicit relationships between input and output values, facilitating the rapid discovery of connections and the establishment of criteria for prediction without restrictive assumptions. The expanded dataset has significantly enhanced the effectiveness of machine learning applications.

This review primarily focuses on presenting the latest applications of machine learning technologies in the identification and prediction of water quality. Additionally, we provide an overview of the most recent methods for acquiring water quality data. Various cases of water quality prediction, including chlorophyll-a, salinity, dissolved oxygen, and water quality index prediction, are examined. Furthermore, the paper explores water quality predictions through the coupling of hydrodynamics.

The subsequent sections of the paper are organized as follows: Section 2 delves into data acquisition techniques, providing insights into the methods employed. Section 3 comprehensively discusses the most recent advancements in water quality prediction through machine learning. Finally, the conclusions drawn and directions for future work are explored in Section 4.

2. Acquiring Water Quality Data

The identification and collection of water pollution data constitute pivotal steps in understanding the status of water quality [42,43]. Seawater quality is predominantly characterized by its chemical, physical, and biological properties. Various water parameters [5], including physical parameters (such as water temperature, total suspended solids, turbidity, and total dissolved solids), chemical parameters (such as chemical and biochemical oxygen demand, and dissolved oxygen) [44–46], biological parameters (such as *Escherichia Coli* and enterococci levels [47]), can be collected by water quality inspectors.

Remote sensing, particularly via satellites, offers broader spatial coverage [48,49] and requires less time compared to traditional field measurements. Leveraging the specific reflection wavelength characteristics of objects, remote sensing enables the extraction of data from images. For instance, water with a higher algae content exhibits the reflectance at wavelengths of 550 nm [50]. Researchers have explored the potential correlation between water quality and satellite images, deriving water quality parameters from these images through empirical formulas [51,52]. Remote sensing algorithms can convert spectral reflectance into chlorophyll concentration, and widely used sensors such as SeaWiFS [53–57],

MODIS [53,57,58], CZCS, MERIS [57,59], and OLCI [60] facilitate this process. Figure 1 illustrates the steps in obtaining water quality data through remote sensing.

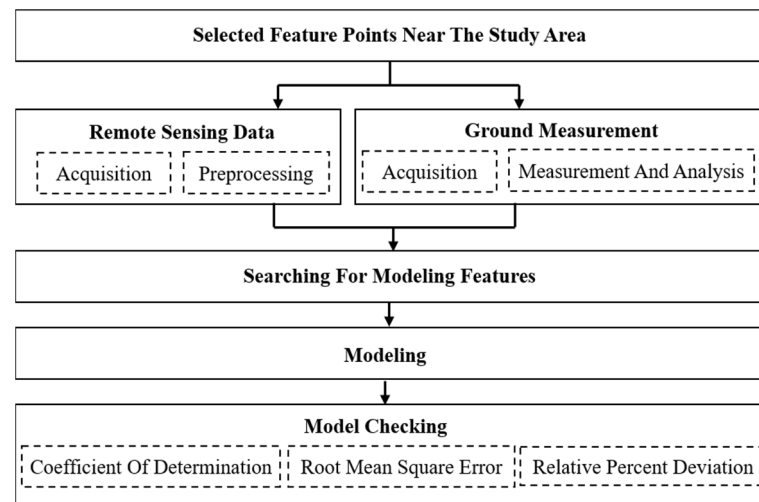


Figure 1. Steps for inverting water quality data from remote sensing data. The establishment of spectral reference libraries is a prerequisite for extracting data from remote sensing monitoring [61].

In recent years, substantial strides have been conducted on leveraging optical features to obtain water quality information [62–64]. Advancing water quality research necessitates more precise data [65,66]. Consequently, methods that integrate optical features with water quality parameters have been developed. Demonstrating feasibility, the use of remote sensing to measure physical and chemical parameters in marine water has been validated, such as turbidity, suspended sediment, dissolved organic carbon, chemical oxygen demand (COD), ammonia nitrogen ($\text{NH}_3\text{-N}$), dissolved oxygen (DO), Secchi disc depth, and total suspended solids. Vaibhav Garg et al. [67] have shown that it is possible to detect turbidity using Sentinel-2 multispectral remote sensing data through red and near-infrared wavelengths. Cherukuru et al. [39] have successfully developed a new semi-analytical remote sensing inversion model for retrieving suspended sediment and dissolved organic carbon in coastal waters. High-performance inversion results were achieved for four water quality parameters: chemical oxygen demand (COD), turbidity, ammonia nitrogen ($\text{NH}_3\text{-N}$), and dissolved oxygen (DO), indicating the potential application value of near-surface remote sensing in inland, coastal, and various water bodies [68]. Yuan Fong Su et al. [69] established univariate and multivariate water quality evaluation models for retrieving sea surface reflectance using SPOT remote sensing images, applying them to SPOT multispectral images to generate distribution maps for three water quality variables: Secchi disc depth, turbidity, and total suspended solids. The study demonstrated the feasibility of utilizing satellite remote sensing images for coastal water quality mapping.

Machine learning has excelled in normalizing the difference in chlorophyll, turbidity, and the salinity index, successfully classifying water quality in Sentinel-2 images. The Classification and Regression Tree method has accurately identified macroscopic bloom locations with over 98% accuracy [70]. Figure 2 shows the steps in building a machine learning model.

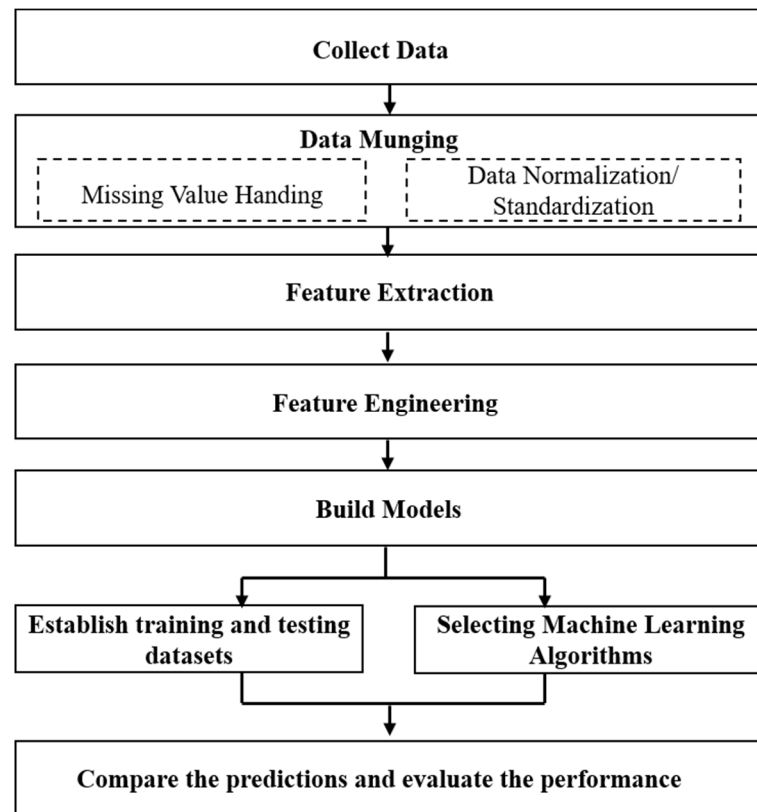


Figure 2. Steps for building a model through machine learning. (1) Data collection: Identify the specific water quality parameters you want to predict. Gather a comprehensive dataset. (2) Data mugging: Clean the dataset by handling missing values, outliers, and irrelevant features. (3) Feature extraction: Identify relevant features that contribute to the prediction task. (4) Feature engineering: Transforming raw data into better representations of the essential features of a problem. (5) Model selection: Choose appropriate machine learning algorithms. (6) Model training: Train the selected model using the training dataset. (7) Model evaluation: Assess the model’s performance on the testing dataset using appropriate evaluation metrics.

It is crucial to emphasize that water quality indicators lacking optical mechanisms, which rely on mathematical models or intermittent methods, cannot be directly measured through remote sensing. Based on a high correlation between non-optical active parameters and optical active parameters, in a study by Hanyu Li et al. [71], Landsat 5/8 remote sensing images and measured total nitrogen (TN) and total phosphorus (TP) were utilized to investigate the modeling effects of machine learning methods. The results confirmed that machine learning algorithms are well suited for inverting non-optical activity parameters in coastal water bodies.

3. Utilization of Machine Learning in Water Quality Prediction

Changes in water quality in coastal areas are influenced by both natural and anthropogenic factors [72–76]. Pollutants in the marine environment are not only from sewage discharge but also from natural activities, whereby water bodies transport pollutants through inherent circulation processes [77,78], including rainfall, runoff [79], seawater intrusion, and tidal intrusion, ultimately merging into the ocean. Water quality pollution encompasses a broad spectrum of sources and complex causes, rendering accurate results challenging through mechanistic analysis. Nonetheless, given the extensive impact and threats posed to industrial production, human lives, and the ecological environment, real-time identification and prediction of water quality pollution are imperative.

3.1. Single Water Quality Prediction Using Machine Learning

Machine learning is well suited for predicting water quality as it can identify the factors causing changes and reveal potential complex relationships between variables [80,81] and their predicted outcomes. Machine learning models have found extensive applications across various fields [82–85]. For instance, a neural network-based algorithm has been utilized to monitor turbidity in the marine environment [86]. Jun Ma et al. [87] employed a combination of Deep Matrix Factorization and Deep Neural Network to accurately predict BOD values. Furthermore, Decision Forest, Decision Jungle, and Boosted Decision Tree achieved accuracy scores exceeding 99% for predicting *Escherichia Coli* and enterococci levels [47].

3.1.1. Prediction of Chlorophyll-a

Chlorophyll is a crucial water quality parameter commonly employed to assess biomass [88,89]. Elevated chlorophyll values indicate eutrophication within a water body [90,91], often associated with increased pollutant input, diminished dissolved oxygen, and the emergence of toxic cyanobacteria blooms. Two primary sources provide chlorophyll data: on-site measurements, which possess inherent limitations discussed earlier, and numerical simulation technology. Despite numerous attempts to predict chlorophyll content through numerical simulations, challenges arise due to the complex interplay of physical and mixing processes involving other physicochemical parameters related to water quality, as well as external factors such as light and temperature.

Given the uncertainties associated with marine biochemical parameters, numerical simulations have not yielded entirely satisfactory results in simulating phytoplankton biomass. Growing efforts and promising results are emerging in the realm of data-driven modeling for water quality [92–94]. For instance, Xin Yu et al. [95] employed Visible Infrared Imaging Radiometer Suite satellite data from 2011 to 2018 and to train a machine learning-based model and utilized a data interpolation empirical orthogonal function. Driven by external forcing including river discharge, nutrient loadings, solar radiation, wind, and air temperature, the data-driven model achieved an average root mean square error of 1.85 $\mu\text{g/L}$ for the entire bay with overall satisfactory performance.

In a study by Yong Sung Kwon et al. [96], machine learning techniques utilizing bands from 1 to 4 obtained from Landsat-8 Operational Land Imager satellite images demonstrated satisfactory performance, highlighting the effectiveness of combining remote sensing and machine learning for estimating chlorophyll-a concentration. Huanmei Yao et al. [97] utilized the Gradient Boosting Decision Tree model to estimate Chl-a concentrations, combining Landsat 8 OLI satellite data with a nominal 30 m spatial resolution from the United States Geological Survey with field measurements. The Gradient Boosting Decision Tree model exhibited a higher accuracy (MAE = 0.998 $\mu\text{g/L}$, MAPE = 19.413%, and RMSE = 1.626 $\mu\text{g/L}$) compared to different physics models.

Acknowledging the efficacy of machine learning as a prediction method, scientists continually explore optimal machine learning-based models for prediction. Hae-Ran Kim et al. [98] determined that the integrated learning method Extreme Gradient Boosting, combined with the single model Support Vector Regression, achieved superior results compared to six other machine learning algorithms (Regression Tree, Support Vector Regression, Bagging, Random Forest, Gradient Boosting Machine, and Extreme Gradient Boosting) in predicting Chl-a concentration. In another study, Diego Gómez et al. [99] evaluated the performance of Random Forest, Support Vector Machine, Artificial Neural Network, and Deep Neural Network algorithms. ANN demonstrated better performance under specific conditions, but when considering more factors, the other three methods were preferable. The most successful outcome was achieved by the Random Forest algorithm without using any feature selection techniques, yielding $R^2 = 0.92$ and $\text{RMSE} = 0.82 \text{ mg/m}^3$. Mohebzadeh and Lee [100] employed three machine learning techniques (Support Vector Regression, Random Forest Regression, and Long Short-Term Memory) as a downscaling approach, concluding that second degree multiple polynomial

regression and Support Vector Regression–Radial Basis Function could produce high-resolution Chl-a maps. Junan Lin et al. [101] used four machine learning models (Support Vector Regression, Random Forest Regression, Wavelet Analysis—Back Propagation Neural Network, and Wavelet Analysis—Long Short-Term Memory) to predict chlorophyll-a in coastal waters, successfully forecasting algal blooms, with Wavelet Analysis—Back Propagation Neural Network and Wavelet Analysis—Long Short-Term Memory outperforming the others. Jie Niu et al. [102] incorporated particulate organic carbon and particulate inorganic carbon as predictive factors in machine learning and deep learning models to estimate Chl-a concentration along the coast. The results showed that the Gaussian process regression model outperformed the deep learning model in terms of stability and robustness. Due to differences and varying data characteristics, a universally superior method cannot be determined; the appropriate approach should be selected based on data characteristics and local conditions.

Although the mentioned machine learning algorithms can yield highly satisfactory results, certain studies have noted limitations in algorithms like Support Vector Machines [103], attributing this to their low-computational efficiency due to the nonlinear relationship between variables and outputs.

In addition to commonly used traditional machine learning methods, increasingly specialized machine learning algorithms are being developed. Hua Su et al. [104] introduced LightGBM, surpassing traditional methods and OLCI Chl-a products. Nima Pahlevan et al. [105] presented the mixed density network (MDN) simulation applicable to MSI and OLCI data.

3.1.2. Prediction of Salinity

Salinity is a pivotal factor influencing the physical, chemical, and biological processes in the ocean [106–108]. The salt content in seawater directly impacts its density [109,110], consequently influencing circulation and stratification. Salinity distribution plays a crucial role in the growth and reproduction of marine organisms. To gain a deeper understanding of salinity distribution, a method has been devised, relying on historical data to train machine learning models and extract developmental trends. Guillou et al. [111] delved into machine learning algorithms to simulate the nonlinear and intricate relationship between salinity and input parameters (such as tide-induced free-surface elevation, river discharges, and wind velocity). Priyanka Chawla et al. [112] developed an effective regression and machine learning model for predicting water quality salinity and forecasting future salinity levels based on historical records. Lal and Datta [113] created independent models (Artificial Neural Network, Gaussian Process, and Support Vector Regression) to construct both homogeneous and heterogeneous models capable of predicting salinity concentrations. The results highlight the superiority of the heterogeneous model over all independent models and numerical salt transport models.

3.1.3. Prediction of Dissolved Oxygen

Dissolved oxygen, representing molecular oxygen in the water environment [114,115], is influenced by factors such as atmospheric pressure, water temperature, and salt content. A decrease in atmospheric pressure, an increase in water temperature, or higher salt content can lead to a decrease in dissolved oxygen. The dissolved oxygen level in water results from the comprehensive effects of water quality and environmental conditions. Typically ranging between 5–10 mg/L in source water, it approaches saturation in natural water surfaces. Excessive reproduction of algae can cause dissolved oxygen to become supersaturated. In cases of pollution by organic and inorganic reducing substances, dissolved oxygen may decrease or approach zero, activating anaerobic bacteria and deteriorating water quality. Water hypoxia poses a severe environmental challenge in coastal areas worldwide [116], leading to significant economic losses when dissolved oxygen falls below a critical threshold, causing mass deaths of aquatic organisms [116,117].

The successful application of past cases has proven the possibility of machine learning for predicting dissolved oxygen. Eric Ariel L. Salas et al. [118] trained two machine learning algorithms, Random Forest and Support Vector Machine, to predict spatiotemporal variations in dissolved oxygen concentrations using spectral predictors derived from Sentinel-2 images, yielding accurate results. Manuel Valera et al. [119] compared the performance of Random Forest and Support Vector Machine, with Random Forest consistently performing slightly better, given its ease of tuning and training.

Despite the feasibility of machine learning-based dissolved oxygen prediction, acquiring water quality data in challenging environments remains a hurdle. Seongsik Park et al. [120] introduced redox potential as a preferred input variable, using machine learning to predict dissolved oxygen—a cost-effective method.

3.1.4. Prediction of Multiple Water Quality Parameters

Compared with the prediction of specific water quality parameters, methods for predicting multiple water quality parameters can achieve better results. For instance, Yuan-Fong Su et al. [69] discovered that a multivariate model considering the wavelength-dependent comprehensive impact of various seawater components on sea surface reflectance outperformed the univariate model. Identifying the model with multiple parameters which predicts the best results has been widely studied. Yong Hoon Kim et al. [121] used three machine learning methods (Random Forest, Cubist, Support Vector Regression) to predict chlorophyll-a and suspended particulate matter indicators in coastal environments, with Support Vector Regression showing superiority. Shang Tian et al. [122] compared four machine learning algorithms (Extreme Gradient Boost, Support Vector Regression, Random Forest, and Artificial Neural Network) in retrieving chlorophyll-a, dissolved oxygen, and ammonia nitrogen from inland reservoirs, with Extreme Gradient Boost showing superior performance. Patricia Jimeno-Sáez et al. [123] utilized machine learning (Multi-layer Neural Networks and Support Vector Regression) to predict chlorophyll-a levels based on target dataset information from nine different water quality parameters, demonstrating satisfactory results with the Support Vector Regression model outperforming them all. Xiaotong Zhu et al. [124] estimated chlorophyll-a, turbidity, and dissolved oxygen in the Shenzhen Bay area using an ensemble machine learning model based on Sentinel-2 satellite remote sensing images, yielding satisfactory performance. Nguyen et al. [125] applied three machine learning methods (Decision Tree, Random Forest, Gradient Augmented Regression, and Ada Augmented Regression) based on Sentinel-2 images to establish a seawater quality parameter model, with Random Forest producing the best results. Shengyue Chen et al. [126] trained Random Forest, Support Vector Machine, and Backpropagation Neural Network models using water temperature, hydrogen ion concentration, conductivity, dissolved oxygen, and turbidity as water quality datasets to estimate total phosphorus, total nitrogen, and ammonia nitrogen in small-scale coastal basins, with the Random Forest model outperforming the Support Vector Machine and Backpropagation Neural Network models in estimation performance. Table 1 lists the simulation parameters mentioned above and the best model in the simulation. Researchers select models based on the simulation area and simulation parameters.

Table 1. Comparison table of existing research performance.

Author	Algorithms	Predicted Parameters	Best Algorithms
Yong Hoon Kim et al. [121]	Random Forest, Cubist, Support Vector Regression	Chlorophyll-a and suspended particulate matter indicators	Support Vector Regression
Shang Tian et al. [122]	Extreme Gradient Boost, Support Vector Regression, Random Forest, and Artificial Neural Network	Chlorophyll-a, dissolved oxygen, and ammonia nitrogen	Extreme Gradient Boost
Patricia Jimeno-Sáez et al. [123]	Multi-layer Neural Networks and Support Vector Regression	Chlorophyll-a (based on target dataset of nine different water quality parameters)	Support Vector Regression

Table 1. Cont.

Author	Algorithms	Predicted Parameters	Best Algorithms
Xiaotong Zhu [124]	ensemble machine learning model (Extreme Gradient Boosting, Support vector regression, Multi-Layer Perception, and mixture density networks)	Chlorophyll-a, turbidity, and dissolved oxygen	Ensemble machine learning model
Nguyen et al. [125]	Decision Tree, Random Forest, Gradient Augmented Regression, and Ada Augmented Regression	Total suspended solids, chlorophyll-a, chemical oxygen demand, and dissolved oxygen	Random Forest
Shengyue Chen et al. [126]	Random Forest, Support Vector Machine, and Backpropagation Neural Network models	Total phosphorus, total nitrogen, and ammonia nitrogen	Random Forest

3.2. Prediction of Coastal Water Quality Index Using Machine Learning

The water quality index, providing an overall assessment of a water body at a specific location and moment [127–129], is a method developed for analyzing the water quality of marine systems [130]. Its primary advantage lies in the use of simple mathematical functions that can convert multifaceted information into a straightforward numerical expression, conveying the environmental status to the public. This method is relatively easy for non-professionals to comprehend. Typically, this technique comprises four crucial elements [131,132]: (i) indicator selection process [133]; (ii) sub-index process; (iii) indicator weighting process; and (iv) aggregation function. The calculation is as follows:

$$WQI = \sum W_i Q_i \quad (1)$$

where Q_i is the water quality of a single parameter, W_i is the weight of the corresponding water quality parameter.

Sangeeta Pati et al. [134] developed a water quality index employing cluster analysis to categorize data into three water quality characteristics. Discriminant analysis was then applied to generate discriminant functions, effectively measuring multiple parameters of Indian coastal waters. The results demonstrated the utility of the WQI method in handling complex nutrient data and identifying pollution sources. Values of water quality parameters corresponding to the water quality index value are shown in Figure 3.

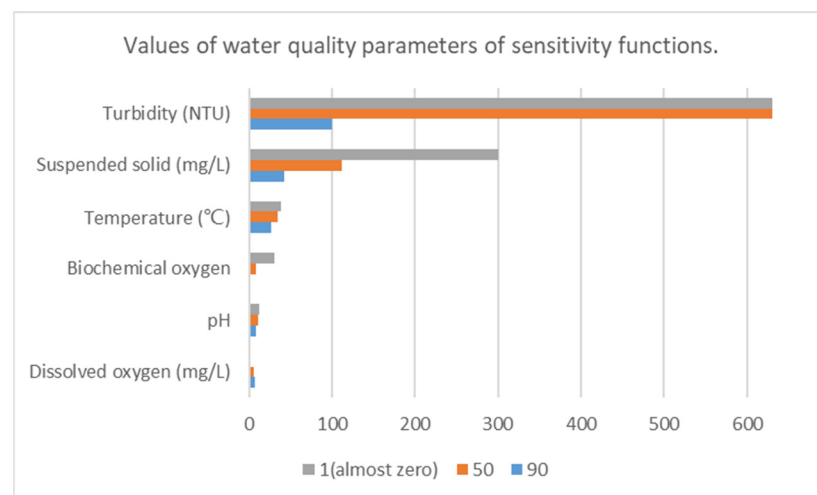


Figure 3. Values of water quality parameters corresponding to water quality index value (data from [135]).

Currently, the uncertainty of the WQI model has been revealed by several studies [136–140]. The uncertainty of the water quality index model mainly arises from the indicator selection process and the indicator weighting process. Problems occur due to inappropriate sub-indexes, parameter weightings, inappropriate aggregation functions,

or an overestimation of the WQI index that does not reflect the real information of water quality. To address these issues, some research has attempted to solve the uncertainty and provide accurate predictions of the water quality index [138,141–145]. Several studies have applied various machine learning algorithms such as Extreme Gradient Boosting, Support Vector Machine, Random Forest, and Decision Tree algorithms [146] to compare the algorithm performance in predicting WQIs correctly.

Michelle C. Tanega et al. [147] applied machine learning classification algorithms such as Random Forest, Decision Tree, and Support Vector to calculate the water quality index and water quality classification of Lake Taal in the Philippines. The results showed that the accuracy of Random Forest was the highest at 95.0%, followed by decision trees with the same accuracy. The accuracy of support vector machines was 93.33%. Md Galal Uddin et al. [137] proposed an improved Water Quality Index (WQI) model for predicting coastal water quality, which is more objective and data driven, and is less susceptible to masking and fuzzy errors. This model uses the machine learning algorithm Extreme Gradient Boosting to rank water quality indicators based on their impact and importance. Zohreh Sheikh Khozani et al. [148] attempted to use intelligent models with different functions to predict the water quality index. Using three different machine learning techniques, including Multi-layer Perceptron, Convolutional Neural Network, and Short-term Memory, to perform WQI prediction. All three models have shown good performance in predicting WQI, effectively shortening the calculation time and reducing errors in the derivation process of sub-indicators. Guize Liu et al. [149] proposed a prediction system based on the Support Vector Machine and Particle Swarm Optimization algorithm. The results show that the maximum error of the water pollution index prediction model for sample prediction is 2.41%, the average error is 1.24%, and the root mean square error is 5.36×10^{-4} , with a correlation coefficient of 0.91 squared. The SVM-PSO algorithm has good sewage prediction ability. Md Galal Uddin et al. [134] conducted in-depth research on indicator selection techniques to reduce significant uncertainty in evaluation. Analyzing the effects of 18 different FS technologies, constructing 15 water quality indicator combinations, and testing the performance of the model using nine machine learning algorithms. The results indicate that the Random Forest algorithm can effectively select key water quality indicators. The Deep Neural Network algorithm predicts a subset more accurately.

These methods have achieved relatively effective results but have not fundamentally improved WQI. Md Galal Uddin et al. [136] used machine learning techniques to improve the newly developed Weighted Quadratic Mean-WQI model architecture to reduce model uncertainty. They used eight widely used machine learning algorithms, Decision Tree, Extra Tree, Extreme Gradient Boosting, Random Forest, Support Vector Machine, K-Nearest Neighbors, Linear Regression, and Gaussian Naïve Bayes, to reduce the uncertainty of the WQI model and improve the model architecture in modeling coastal WQIs.

3.3. Prediction of Water Quality through Coupling Hydrodynamics and Water Quality

The primary distinction between nearshore waters and natural rivers is that the upstream of natural river reservoirs is nearly unaffected by downstream water bodies, while nearshore water bodies are not only influenced by shore discharge but also by ocean water bodies due to tides and storms [150,151]. Pollutants may be mixed and washed by seawater, making the changes in water quality more complex. Therefore, it is necessary to combine models with hydrodynamic factors to predict changes in ocean water quality.

Currently, hydrodynamic factor prediction has made outstanding progress [152–154], and simulation of hydrodynamic processes can yield relatively accurate results. Kai Fei et al. [155] established a numerical simulation model for storm surges coupled with hydrology and hydrodynamics, successfully calculating the relative contributions of each driving factor to water level using Extreme Gradient Boosting. Shamshirband et al. [156] proposed a nested grid numerical model that utilizes water depth and surface wind field data for wave height modeling. The combination of hydrodynamic models and machine learning can improve analysis reliability and computational efficiency [157,158]. For example, his-

torical data or simulated data from hydrodynamic models can be used to train machine learning models to predict wave heights [159,160], hurricane storm surge hazards [161,162], floods [163], erosion [164], and water level characteristics of storm surges [155]. Notably, a nearshore wave and hydrodynamic prediction model by Wei and Davison [165] based on Convolutional Neural Networks can accurately predict the propagation and fragmentation of waves on nearshore slopes, including detailed wave peak bending and separation. Riaz et al. [166] achieved the prediction of near-bottom hydrodynamic conditions in rapidly changing water flows. This detailed simulation provides a reference for future analyses of water quality changes.

Based on the results of hydrodynamic prediction, it is possible to predict the water quality status of coastal areas through hydrodynamic coupling hydrology. Hung Vuong Pham et al. [167] has designed a coastal risk assessment model based on Bayesian networks, predicting the water quality parameters of seawater. This method proposes a multi-model chain approach, integrating regional and global climate models with machine learning and satellite images. It combines ocean fluid dynamics, wave fields, and coastline extraction models to predict suspended particulate matter. Importantly, this method has the potential to integrate climate and risk data, enabling more accurate predictions in the future. Bayesian [168] is a probability-based method that combines prior information about unknown parameters with sample information, using the Bayesian formula to obtain posterior information. Based on the posterior information, unknown parameters can be inferred, yielding good results in small samples. Therefore, it can be used in situations where coastal data are relatively insufficient. Cebe and Balas [169] studied the prediction of nitrite, nitrate, and dissolved oxygen concentrations using water quality coupled three-dimensional hydrodynamic methods. The hydrodynamic model simulates wind-driven nearshore water flow, cycling of nitrogen, phosphorus, and oxygen in ecological sub-models, as well as dominant aquatic organisms such as phytoplankton, zooplankton, and planktonic bacteria.

4. Discussion

This paper presents a comprehensive exploration of the challenges and advancements in predicting water quality in coastal areas, with an emphasis on the integration of machine learning into water quality parameter simulations. The identified challenges, including the impact of urbanization, industrialization, coastal reclamation, and environmental events such as erosion, hurricanes, and flooding, underscore the importance of accurate predictive models for maintaining and safeguarding coastal ecosystems.

This review rightly emphasizes the significant role of machine learning in overcoming the limitations of numerical models. Traditional models face challenges related to parameter selection, adaptability, and computational efficiency. The integration of machine learning, enabled by recent progress in satellite remote sensing and unmanned aerial vehicle observation, offers a promising solution. The ability to process large datasets and extract meaningful correlations between water quality metrics and satellite imagery represents a substantial leap forward. The review of water quality data acquisition methods highlights the pivotal role of remote sensing, particularly satellite technology. The exploration of optical features and the establishment of spectral reference libraries demonstrate innovative approaches to derive water quality information. Machine learning models trained on optical features prove effective in predicting water quality parameters, overcoming challenges associated with non-optical indicators. The presented steps for inverting water quality data from remote sensing maps provide a clear framework for researchers and practitioners.

This paper also provides a comprehensive overview of machine learning applications in predicting specific water quality parameters such as chlorophyll-a, salinity, and dissolved oxygen. We provide a thorough examination of various machine learning algorithms' performance in predicting these parameters. Noteworthy is the recognition that the choice of the appropriate algorithm depends on data characteristics and local conditions. The inclusion of studies comparing different algorithms for specific parameters adds valuable insights for researchers seeking optimal prediction models.

Different machine learning algorithms can be selected based on the simulated water quality parameters. The Classification and Regression Tree method can accurately identify macroscopic bloom locations [71]. Decision Forest, Decision Jungle, and Boosted Decision Tree can be used to predict *Escherichia Coli* and enterococci levels [47]. Extreme Gradient Boosting, combined with the single model Support Vector Regression and Long Short-Term Memory Long Short-Term Memory is better when estimating Chl-a concentrations. Artificial Neural Network, Gaussian Process, and Support Vector Regression can be used in predicting salinity concentrations. Random Forest performing slightly better than Support Vector Machine in the prediction of dissolved oxygen. Support Vector Regression, Extreme Gradient Boost, and Random Forest can be use in predicting multiple water quality parameters, and ensemble machine learning model is also a good choice. Random Forests, Extreme Gradient Boosting, Multi-layer Perceptron, Convolutional Neural Network, and Short-term Memory are all models which have shown good performance in predicting WQI, while the Random Forest algorithm can effectively select key water quality indicators. The Deep Neural Network algorithm predicts a subset more accurately.

The water quality index (WQI) is a crucial tool for assessing overall water quality, and this paper addresses the uncertainties associated with traditional WQI models. Machine learning algorithms, including Random Forest, Support Vector Machine, and Decision Tree, have been successfully applied to enhance WQI prediction. However, we point out that while these methods achieve effective results, they fall short of fundamentally improving WQI. Our discussion on ongoing efforts to reduce model uncertainty and improve architecture adds depth to the exploration of WQI prediction.

The integration of hydrodynamics with water quality prediction is a key focus of the current paper. The challenges posed by nearshore waters, influenced by both shore discharge and oceanic forces, necessitate a holistic approach. The success of hydrodynamic prediction models, coupled with machine learning, is evident in accurately simulating storm surges, wave heights, and other dynamic factors. Our discussion on the potential of Bayesian networks in coastal risk assessment, integrating climate and risk data, presents a forward-looking perspective on predicting water quality in coastal regions.

The presented research underscores the transformative impact of machine learning on water quality prediction in coastal areas. Our review on the limitations of current models, the need for diverse datasets, and the consideration of evolving environmental conditions points to avenues for future research.

5. Conclusions

This review provides a comprehensive overview of the recent advancements in machine learning applied to water quality prediction. Despite an extensive survey and comparison of existing literature, establishing a singular best-performing machine learning approach proves challenging. The efficacy of machine learning models tends to vary significantly across different parameters and regions. A promising avenue for further exploration involves a deeper analysis of water quality parameter characteristics, aiming to propose more universally applicable methodologies. Generalizing results for the intricate prediction of coastal water quality remains challenging when solely relying on machine learning models trained on data devoid of consideration for physical and chemical processes. Models with regional characteristics may lack the capacity to predict processes involving key factors beyond the training dataset. To enhance the prediction accuracy of machine learning, the following aspects can be addressed: (a) diversifying data sources and increasing data volume, where remote sensing satellite maps can serve as a reference for inverting optical characteristic parameters, necessitating further research for rapid and effective data acquisition on non-optical water quality parameters in coastal areas; (b) addressing missing data through interpolation methods such as univariate input, k-nearest neighbors' input, and multiple-input denoising techniques. This paper offers an exploration of the complexities and advancements in predicting water quality in coastal regions, providing a valuable resource for researchers, practitioners, and policymakers

involved in environmental management and conservation. The integration of machine learning into numerical simulations emerges as a promising paradigm shift, offering more accurate and adaptable predictive models for safeguarding the delicate balance of coastal ecosystems.

Author Contributions: Investigation, T.Z.; writing—original draft preparation, T.Z.; writing—review and editing, X.Y.; supervision, Q.M.; project administration, X.Y., Q.M., X.X., X.Z. and W.D. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Open Research Fund of State Environmental Protection Key Laboratory of Estuarine and Coastal Environment, Chinese Research Academy of Environmental Sciences (HKHA2022012) and the Fundamental Research Funds for the Central Public-interest Scientific Institution (2022YSKY-03), the Open Fund of National Engineering Research Center for Geographic Information System, University of Geosciences, Wuhan 430074, China (Grant No. 2022KFJJ03), and the National Natural Science Foundation of China (Grant No. 52309079).

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ye, J. Stayers in China's "hollowed-out" villages: A counter narrative on massive rural–urban migration. *Popul. Space Place* **2018**, *24*, e2128. [\[CrossRef\]](#)
- Thinh, H.B. Migration and Education in Vietnam: Opportunities and Challenges. *Adv. Sci. Lett.* **2017**, *23*, 2166–2168. [\[CrossRef\]](#)
- Wang, Y.; Bai, J. Experience, Lessons of India's Urbanization and Its Warnings to China. In *Proceedings of the 20th International Symposium on Advancement of Construction Management and Real Estate*; Wu, Y., Zheng, S., Luo, J., Wang, W., Mo, Z., Shan, L., Eds.; Springer: Singapore, 2017; pp. 489–496.
- Meng, G.; Guo, Z.; Li, J. The dynamic linkage among urbanisation, industrialisation and carbon emissions in China: Insights from spatiotemporal effect. *Sci. Total Environ.* **2021**, *760*, 144042. [\[CrossRef\]](#)
- Misbari, S.; Hashim, M. Water quality changes using GIS-based approach at seagrass meadows along the Straits of Johor. In *Proceedings of the 3rd Symposium on Industrial Science and Technology (SISTEC2021)*, Pahang, Malaysia, 25–26 August 2021; p. 050002.
- Johnson, D.C.; Enriquez, C.E.; Pepper, I.L.; Davis, T.L.; Gerba, C.P.; Rose, J.B. Survival of *Giardia*, *Cryptosporidium*, poliovirus and *Salmonella* in marine waters. *Water Sci. Technol.* **1997**, *35*, 261–268. [\[CrossRef\]](#)
- Wang, C.; Guo, Z.; Li, Q.; Fang, J. Study on layout optimization of sewage outfalls: A case study of wastewater treatment plants in Xiamen. *Sci. Rep.* **2021**, *11*, 18326. [\[CrossRef\]](#)
- Rohmana, Q.A.; Fischer, A.M.; Cumming, J.; Blackwell, B.D.; Gemmill, J. Increased Transparency and Resource Prioritization for the Management of Pollutants from Wastewater Treatment Plants: A National Perspective from Australia. *Front. Mar. Sci.* **2020**, *7*, 564598. [\[CrossRef\]](#)
- Yang, M.; Zhang, X. Comparative Developmental Toxicity of New Aromatic Halogenated DBPs in a Chlorinated Saline Sewage Effluent to the Marine Polychaete *Platynereis dumerilii*. *Environ. Sci. Technol.* **2013**, *47*, 10868–10876. [\[CrossRef\]](#)
- Li, T.; Long, H.; Liu, Y.; Tu, S. Multi-scale analysis of rural housing land transition under China's rapid urbanization: The case of Bohai Rim. *Habitat Int.* **2015**, *48*, 227–238. [\[CrossRef\]](#)
- Esbah, H. Land Use Trends During Rapid Urbanization of the City of Aydin, Turkey. *Environ. Manag.* **2007**, *39*, 443–459. [\[CrossRef\]](#)
- Wu, Y.; Luo, J.; Zhang, X.; Skitmore, M. Urban growth dilemmas and solutions in China: Looking forward to 2030. *Habitat Int.* **2016**, *56*, 42–51. [\[CrossRef\]](#)
- Tian, B.; Wu, W.; Yang, Z.; Zhou, Y. Drivers, trends, and potential impacts of long-term coastal reclamation in China from 1985 to 2010. *Estuar. Coast. Shelf Sci.* **2016**, *170*, 83–90. [\[CrossRef\]](#)
- McKinstry, M.C.; Anderson, S.H. Evaluation of wetland creation and waterfowl use in conjunction with abandoned mine lands in northeast Wyoming. *Wetlands* **1994**, *14*, 284–292. [\[CrossRef\]](#)
- Chen, S.; Chen, L.; Liu, Q.; Li, X.; Tan, Q. Remote sensing and GIS-based integrated analysis of coastal changes and their environmental impacts in Lingding Bay, Pearl River Estuary, South China. *Ocean Coast. Manag.* **2005**, *48*, 65–83. [\[CrossRef\]](#)
- Chen, L.; Ren, C.; Zhang, B.; Li, L.; Wang, Z.; Song, K. Spatiotemporal Dynamics of Coastal Wetlands and Reclamation in the Yangtze Estuary during Past 50 Years (1960s–2015). *Chin. Geogr. Sci.* **2018**, *28*, 386–399. [\[CrossRef\]](#)
- Talke, S.A.; Jay, D.A. Changing Tides: The Role of Natural and Anthropogenic Factors. *Annu. Rev. Mar. Sci.* **2020**, *12*, 121–151. [\[CrossRef\]](#)
- Shen, Y.; Jia, H.; Li, C.; Tang, J. Numerical simulation of saltwater intrusion and storm surge effects of reclamation in Pearl River Estuary, China. *Appl. Ocean Res.* **2018**, *79*, 101–112. [\[CrossRef\]](#)
- Pan, Z.; Liu, H. Impact of human projects on storm surge in the Yangtze Estuary. *Ocean Eng.* **2020**, *196*, 106792. [\[CrossRef\]](#)
- Lee, C.; Hwang, S.; Do, K.; Son, S. Increasing flood risk due to river runoff in the estuarine area during a storm landfall. *Estuar. Coast. Shelf Sci.* **2019**, *221*, 104–118. [\[CrossRef\]](#)

21. Van Wesenbeeck, B.K.; Balke, T.; Van Eijk, P.; Tonneijck, F.; Siry, H.Y.; Rudianto, M.E.; Winterwerp, J.C. Aquaculture induced erosion of tropical coastlines throws coastal communities back into poverty. *Ocean Coast. Manag.* **2015**, *116*, 466–469. [\[CrossRef\]](#)
22. Parvin, S.; Sakib, M.H.; Islam, M.L.; Brown, C.L.; Islam, M.S.; Mahmud, Y. Coastal aquaculture in Bangladesh: Sundarbans's role against climate change. *Mar. Pollut. Bull.* **2023**, *194*, 115431. [\[CrossRef\]](#)
23. Isla, F.I. From touristic villages to coastal cities: The costs of the big step in Buenos Aires. *Ocean Coast. Manag.* **2013**, *77*, 59–65. [\[CrossRef\]](#)
24. Greco, M.; Martino, G.; Guariglia, A.; Trivigno, L.; Sansanelli, V.; Losurdo, A.; Mussuto, G. Integrated SDSS for Environmental Risk Analysis in Sustainable Coastal Area Planning. In *Computational Science and Its Applications—ICCSA 2018*; Gervasi, O., Murgante, B., Misra, S., Stankova, E., Torre, C.M., Rocha, A.M.A.C., Taniar, D., Apduhan, B.O., Tarantino, E., Ryu, Y., Eds.; Springer International Publishing: Cham, Switzerland, 2018; pp. 671–684.
25. Kim, T.; Lee, W.-D. Review on Applications of Machine Learning in Coastal and Ocean Engineering. *J. Ocean Eng. Technol.* **2022**, *36*, 194–210. [\[CrossRef\]](#)
26. Kang, M.; Tian, Y.; Zhang, H.; Wan, C. Effect of hydrodynamic conditions on the water quality in urban landscape water. *Water Supply* **2022**, *22*, 309–320. [\[CrossRef\]](#)
27. Lee, M.E.; Seo, I.W. Analysis of pollutant transport in the Han River with tidal current using a 2D finite element model. *J. Hydro-Environ. Res.* **2007**, *1*, 30–42. [\[CrossRef\]](#)
28. Li, D.; Wang, J.; Dong, Z.; Lai, X. Research and Application of 1-D and 2-D Coupling Water Environment Numerical Model for Taihu Basin. In Proceedings of the 2009 3rd International Conference on Bioinformatics and Biomedical Engineering (iCBBE), Beijing, China, 11–13 June 2009; pp. 1–4.
29. Krapesch, G.; Tritthart, M.; Habersack, H. A model-based analysis of meander restoration. *River Res. Appl.* **2009**, *25*, 593–606. [\[CrossRef\]](#)
30. Chen, K.; Chen, H.; Zhou, C.; Huang, Y.; Qi, X.; Shen, R.; Liu, F.; Zuo, M.; Zou, X.; Wang, J.; et al. Comparative analysis of surface water quality prediction performance and identification of key water parameters using different machine learning models based on big data. *Water Res.* **2020**, *171*, 115454. [\[CrossRef\]](#)
31. Maier, H.R.; Dandy, G.C. Modelling cyanobacteria (blue-green algae) in the River Murray using artificial neural networks. *Math. Comput. Simul.* **1997**, *43*, 377–386. [\[CrossRef\]](#)
32. Loewenthal, R.E.; Morrison, I.; Wentzel, M.C. Control of corrosion and aggression in drinking water systems. *Water Sci. Technol.* **2004**, *49*, 9–18. [\[CrossRef\]](#)
33. McKay, P.; Blain, C.A. An automated approach to extracting river bank locations from aerial imagery using image texture: Automated river bank extraction from imagery. *River Res. Appl.* **2014**, *30*, 1048–1055. [\[CrossRef\]](#)
34. Tang, X.; Huang, M. Simulation of chlorophyll-a concentration in Donghu Lake based on GA-ELM and multiple water quality indexes. In Proceedings of the International Conference on Algorithms, High Performance Computing, and Artificial Intelligence (AHPICAI 2021), Sanya, China, 21 December 2021; p. 26.
35. Su, H.; Zou, R.; Zhang, X.; Liang, Z.; Ye, R.; Liu, Y. Exploring the type and strength of nonlinearity in water quality responses to nutrient loading reduction in shallow eutrophic water bodies: Insights from a large number of numerical simulations. *J. Environ. Manag.* **2022**, *313*, 115000. [\[CrossRef\]](#)
36. Chen, J.; Chen, S.; Fu, R.; Li, D.; Jiang, H.; Wang, C.; Peng, Y.; Jia, K.; Hicks, B.J. Remote Sensing Big Data for Water Environment Monitoring: Current Status, Challenges, and Future Prospects. *Earth's Future* **2022**, *10*, e2021EF002289. [\[CrossRef\]](#)
37. Latif, S.D.; Alyaa Binti Hazrin, N.; Hoon Koo, C.; Lin Ng, J.; Chaplot, B.; Feng Huang, Y.; El-Shafie, A.; Ahmed, A.N. Assessing rainfall prediction models: Exploring the advantages of machine learning and remote sensing approaches. *Alex. Eng. J.* **2023**, *82*, 16–25. [\[CrossRef\]](#)
38. Holloway, J.; Mengersen, K. Statistical Machine Learning Methods and Remote Sensing for Sustainable Development Goals: A Review. *Remote Sens.* **2018**, *10*, 1365. [\[CrossRef\]](#)
39. Cherukuru, N.; Martin, P.; Sanwlani, N.; Mujahid, A.; Müller, M. A Semi-Analytical Optical Remote Sensing Model to Estimate Suspended Sediment and Dissolved Organic Carbon in Tropical Coastal Waters Influenced by Peatland-Draining River Discharges off Sarawak, Borneo. *Remote Sens.* **2020**, *13*, 99. [\[CrossRef\]](#)
40. Qiu, J.; Sun, Y. A Research on Machine Learning Methods for Big Data Processing. In Proceedings of the 4th International Conference on Information Technology and Management Innovation, Shenzhen, China, 15 October 2015.
41. Mogha, G.; Ahlawat, K.; Singh, A.P. Performance Analysis of Machine Learning Techniques on Big Data Using Apache Spark. In *Data Science and Analytics*; Panda, B., Sharma, S., Roy, N.R., Eds.; Springer: Singapore, 2018; pp. 17–26.
42. Koelmans, A.A.; Mohamed Nor, N.H.; Hermesen, E.; Kooi, M.; Mintenig, S.M.; De France, J. Microplastics in freshwaters and drinking water: Critical review and assessment of data quality. *Water Res.* **2019**, *155*, 410–422. [\[CrossRef\]](#)
43. Moroni, D.; Pieri, G.; Salvetti, O.; Tampucci, M.; Domenici, C.; Tonacci, A. Sensorized buoy for oil spill early detection. *Methods Oceanogr.* **2016**, *17*, 221–231. [\[CrossRef\]](#)
44. Siyang, S.; Kerdcharoen, T. Development of unmanned surface vehicle for smart water quality inspector. In Proceedings of the 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, Thailand, 28 June–1 July 2016; pp. 1–5.

45. Adhipramana, M.; Mardiaty, R.; Mulyana, E. Remotely Operated Vehicle (ROV) Robot for Monitoring Quality of Water Based on IoT. In Proceedings of the 2020 6th International Conference on Wireless and Telematics (ICWT), Yogyakarta, Indonesia, 3–4 September 2020; pp. 1–7.
46. Balbuena, J.; Quiroz, D.; Song, R.; Bucknall, R.; Cuellar, F. Design and Implementation of an USV for Large Bodies of Fresh Waters at the Highlands of Peru. In Proceedings of the OCEANS 2017—Anchorage Conference, Anchorage, AK, USA, 18–21 September 2017.
47. Grbčić, L.; Družeta, S.; Mauša, G.; Lipić, T.; Lušić, D.V.; Alvir, M.; Lučin, I.; Sikirica, A.; Davidović, D.; Travaš, V.; et al. Coastal water quality prediction based on machine learning with feature interpretation and spatio-temporal analysis. *Environ. Model. Softw.* **2022**, *155*, 105458. [\[CrossRef\]](#)
48. Lin, H.; Li, S.; Xing, J.; Yang, J.; Wang, Q.; Dong, L.; Zeng, X. Fusing Retrievals of High Resolution Aerosol Optical Depth from Landsat-8 and Sentinel-2 Observations over Urban Areas. *Remote Sens.* **2021**, *13*, 4140. [\[CrossRef\]](#)
49. Cui, T.; Li, F.; Wei, Y.; Yang, X.; Xiao, Y.; Chen, X.; Liu, R.; Ma, Y.; Zhang, J. Super-resolution optical mapping of floating macroalgae from geostationary orbit. *Appl. Opt.* **2020**, *59*, C70–C77. [\[CrossRef\]](#)
50. Lamote, M.; Darko, E.; Schoefs, B.; Lemoine, Y. Assembly of the photosynthetic apparatus in embryos from *Fucus serratus* L. *Photosynth. Res.* **2003**, *77*, 45–52. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Osińska-Skotak, K. Influence of atmospheric correction on determination of lake water quality parameters based on CHRIS/PROBA images. In Proceedings of the 25th EARSeL Symposium, Porto, Portugal, 6 June–11 June 2005.
52. Hsu, P.-H. Using SPOT Images for Monitoring Water Quality of Reservoir. *Sens. Mater.* **2016**, *1*, 455–462. [\[CrossRef\]](#)
53. LiQin, Q.; Lei, G.; MingXia, H. The Global Availabilities of SeaWiFS, MODIS and Merged Chlorophyll-a Data. *J. Ocean. Univ. China* **2006**, *36*, 321–326.
54. Pahlevan, N.; Sarkar, S.; Franz, B.A.; Balasubramanian, S.V.; He, J. Sentinel-2 MultiSpectral Instrument (MSI) data processing for aquatic science applications: Demonstrations and validations. *Remote Sens. Environ.* **2017**, *201*, 47–56. [\[CrossRef\]](#)
55. Sathyendranath, S.; Brewin, R.J.W.; Brockmann, C.; Brotas, V.; Calton, B.; Chuprin, A.; Cipollini, P.; Couto, A.B.; Dingle, J.; Doerffer, R.; et al. An Ocean-Colour Time Series for Use in Climate Studies: The Experience of the Ocean-Colour Climate Change Initiative (OC-CCI). *Sensors* **2019**, *19*, 4285. [\[CrossRef\]](#)
56. Seegers, B.N.; Stumpf, R.P.; Schaeffer, B.A.; Loftin, K.A.; Werdell, P.J. Performance metrics for the assessment of satellite data products: An ocean color case study. *Opt. Express* **2018**, *26*, 7404–7422. [\[CrossRef\]](#)
57. O'Reilly, J.E.; Werdell, P.J. Chlorophyll algorithms for ocean color sensors—OC4, OC5 & OC6. *Remote Sens. Environ.* **2019**, *229*, 32–47. [\[CrossRef\]](#)
58. Groom, S.; Sathyendranath, S.; Ban, Y.; Bernard, S.; Brewin, R.; Brotas, V.; Brockmann, C.; Chauhan, P.; Choi, J.K.; Chuprin, A.; et al. Satellite Ocean Colour: Current Status and Future Perspective. *Front. Mar. Sci.* **2019**, *6*, 485. [\[CrossRef\]](#)
59. Blondeau-Patissier, D.; Gower, J.F.R.; Dekker, A.G.; Phinn, S.R.; Brando, V.E. A review of ocean color remote sensing methods and statistical techniques for the detection, mapping and analysis of phytoplankton blooms in coastal and open oceans. *Prog. Oceanogr.* **2014**, *123*, 123–144. [\[CrossRef\]](#)
60. Cao, Z.; Ma, R.; Duan, H.; Pahlevan, N.; Melack, J.; Shen, M.; Xue, K. A machine learning approach to estimate chlorophyll-a from Landsat-8 measurements in inland lakes. *Remote Sens. Environ.* **2020**, *248*, 111974. [\[CrossRef\]](#)
61. Pardeshi, S.; Gandre, P.; Poojari, N.; Pansare, S.; Alte, B. Water Quality Analysis from Satellite Images. In Proceedings of the 2023 International Conference on Data Science and Network Security (ICDSNS), Tiptur, India, 28–29 July 2023; pp. 1–6.
62. Matthews, M.W. A current review of empirical procedures of remote sensing in inland and near-coastal transitional waters. *Int. J. Remote Sens.* **2011**, *32*, 6855–6899. [\[CrossRef\]](#)
63. Odermatt, D.; Gitelson, A.; Brando, V.E.; Schaepman, M. Review of constituent retrieval in optically deep and complex waters from satellite imagery. *Remote Sens. Environ.* **2012**, *118*, 116–126. [\[CrossRef\]](#)
64. Werdell, P.J.; McKinnis, L.I.W.; Boss, E.; Ackleson, S.G.; Craig, S.E.; Gregg, W.W.; Lee, Z.; Maritorena, S.; Roesler, C.S.; Rousseaux, C.S.; et al. An overview of approaches and challenges for retrieving marine inherent optical properties from ocean color remote sensing. *Prog. Oceanogr.* **2018**, *160*, 186–212. [\[CrossRef\]](#)
65. Garver, S.A.; Siegel, D.A. Inherent optical property inversion of ocean color spectra and its biogeochemical interpretation: 1. Time series from the Sargasso Sea. *J. Geophys. Res. Ocean.* **1997**, *102*, 18607–18625. [\[CrossRef\]](#)
66. Werdell, P.J.; Franz, B.A.; Bailey, S.W.; Feldman, G.C.; Boss, E.; Brando, V.E.; Dowell, M.; Hirata, T.; Lavender, S.J.; Lee, Z.; et al. Generalized ocean color inversion model for retrieving marine inherent optical properties. *Appl. Opt.* **2013**, *52*, 2019–2037. [\[CrossRef\]](#)
67. Garg, V.; Aggarwal, S.P.; Chauhan, P. Changes in turbidity along Ganga River using Sentinel-2 satellite data during lockdown associated with COVID-19. *Geomat. Nat. Hazards Risk* **2020**, *11*, 1175–1195. [\[CrossRef\]](#)
68. Zhao, Y.; Yu, T.; Hu, B.; Zhang, Z.; Liu, Y.; Liu, X.; Liu, H.; Liu, J.; Wang, X.; Song, S. Retrieval of Water Quality Parameters Based on Near-Surface Remote Sensing and Machine Learning Algorithm. *Remote Sens.* **2022**, *14*, 5305. [\[CrossRef\]](#)
69. Su, Y.-F.; Liou, J.-J.; Hou, J.-C.; Hung, W.-C.; Hsu, S.-M.; Lien, Y.-T.; Su, M.D.; Cheng, K.S.; Wang, Y.-F. A Multivariate Model for Coastal Water Quality Mapping Using Satellite Remote Sensing Images. *Sensors* **2008**, *8*, 6321–6339. [\[CrossRef\]](#)
70. Medina-López, E.; Navarro, G.; Santos-Echeandía, J.; Bernárdez, P.; Caballero, I. Machine Learning for Detection of Macroalgal Blooms in the Mar Menor Coastal Lagoon Using Sentinel-2. *Remote Sens.* **2023**, *15*, 1208. [\[CrossRef\]](#)

71. Li, H.; Zhang, G.; Zhu, Y.; Kaufmann, H.; Xu, G. Inversion and Driving Force Analysis of Nutrient Concentrations in the Ecosystem of the Shenzhen-Hong Kong Bay Area. *Remote Sens.* **2022**, *14*, 3694. [\[CrossRef\]](#)
72. Deng, X. Influence of water body area on water quality in the southern Jiangsu Plain, eastern China. *J. Clean. Prod.* **2020**, *254*, 120136. [\[CrossRef\]](#)
73. Yin, Z.; Duan, R.; Li, P.; Li, W. Water quality characteristics and health risk assessment of main water supply reservoirs in Taizhou City, East China. *Hum. Ecol. Risk Assess. Int. J.* **2021**, *27*, 2142–2160. [\[CrossRef\]](#)
74. Liu, C.; Zhang, F.; Wang, X.; Chan, N.W.; Rahman, H.A.; Yang, S.; Tan, M.L. Assessing the factors influencing water quality using environment water quality index and partial least squares structural equation model in the Ebinur Lake Watershed, Xinjiang, China. *Environ. Sci. Pollut. Res.* **2022**, *29*, 29033–29048. [\[CrossRef\]](#)
75. Zhao, Y.; Yang, Q.; Yuan, P. Study on the Influence of Land Use on Water Environment Quality in Riverside Zone Based on GIS. *Appl. Sci.* **2020**, *10*, 1262. [\[CrossRef\]](#)
76. Bian, J.M.; Ma, H.Y.; Sun, X.Q. Analysis on Water Quality Change and Influence Factors in the Yitong River Basin. *Appl. Mech. Mater.* **2011**, *71–78*, 2970–2973. [\[CrossRef\]](#)
77. Shen, W.; Jin, Y.; Cong, P.; Li, G. Dynamic Coupling Model of Water Environment of Urban Water Network in Pearl River Delta Driven by Typhoon Rain Events. *Water* **2023**, *15*, 1084. [\[CrossRef\]](#)
78. Losno, R.; Colin, J.L.; Spokes, L.; Jickells, T.; Schulz, M.; Rebers, A.; Leermakers, M.; Meuleman, C.; Baeyens, W. Non-rain deposition significantly modifies rain samples at a coastal site. *Atmos. Environ.* **1998**, *32*, 3445–3455. [\[CrossRef\]](#)
79. Li, J.; Ma, M.; Li, Y.; Zhang, Z. Influence analysis of different design conditions on urban runoff and nonpoint source pollution. *Water Environ. Res.* **2019**, *91*, 1546–1557. [\[CrossRef\]](#)
80. Li, L.; Rong, S.; Wang, R.; Yu, S. Recent advances in artificial intelligence and machine learning for nonlinear relationship analysis and process control in drinking water treatment: A review. *Chem. Eng. J.* **2021**, *405*, 126673. [\[CrossRef\]](#)
81. Ahmed, A.N.; Othman, F.B.; Afan, H.A.; Ibrahim, R.K.; Fai, C.M.; Hossain, M.S.; Ehteram, M.; Elshafie, A. Machine learning methods for better water quality prediction. *J. Hydrol.* **2019**, *578*, 124084. [\[CrossRef\]](#)
82. Ibrahim, H.; Yaseen, Z.M.; Scholz, M.; Ali, M.; Gad, M.; Elsayed, S.; Khadr, M.; Hussein, H.; Ibrahim, H.H.; Eid, M.H.; et al. Evaluation and Prediction of Groundwater Quality for Irrigation Using an Integrated Water Quality Indices, Machine Learning Models and GIS Approaches: A Representative Case Study. *Water* **2023**, *15*, 694. [\[CrossRef\]](#)
83. Hanoon, M.S.; Ahmed, A.N.; Fai, C.M.; Birima, A.H.; Razzaq, A.; Sherif, M.; Sefelnasr, A.; El-Shafie, A. Application of Artificial Intelligence Models for Modeling Water Quality in Groundwater: Comprehensive Review, Evaluation and Future Trends. *Water Air Soil Pollut.* **2021**, *232*, 411. [\[CrossRef\]](#)
84. Allawi, M.F.; Hussain, I.R.; Salman, M.I.; El-Shafie, A. Monthly inflow forecasting utilizing advanced artificial intelligence methods: A case study of Haditha Dam in Iraq. *Stoch. Environ. Res. Risk Assess.* **2021**, *35*, 2391–2410. [\[CrossRef\]](#)
85. Tikhamarine, Y.; Souag-Gamane, D.; Ahmed, A.N.; Sammen, S.S.; Kisi, O.; Huang, Y.F.; El-Shafie, A. Rainfall-runoff modelling using improved machine learning methods: Harris hawks optimizer vs. particle swarm optimization. *J. Hydrol.* **2020**, *589*, 125133. [\[CrossRef\]](#)
86. Kumar, L.; Afzal, M.S.; Ahmad, A. Prediction of water turbidity in a marine environment using machine learning: A case study of Hong Kong. *Reg. Stud. Mar. Sci.* **2022**, *52*, 102260. [\[CrossRef\]](#)
87. Ma, J.; Ding, Y.; Cheng, J.C.P.; Jiang, F.; Xu, Z. Soft detection of 5-day BOD with sparse matrix in city harbor water using deep learning techniques. *Water Res.* **2020**, *170*, 115350. [\[CrossRef\]](#)
88. Vörös, L.; Padisák, J. Phytoplankton biomass and chlorophyll-a in some shallow lakes in central Europe. *Hydrobiologia* **1991**, *215*, 111–119. [\[CrossRef\]](#)
89. Lionard, M.; Muylaert, K.; Tackx, M.; Vyverman, W. Evaluation of the performance of HPLC–CHEMTAX analysis for determining phytoplankton biomass and composition in a turbid estuary (Schelde, Belgium). *Estuar. Coast. Shelf Sci.* **2008**, *76*, 809–817. [\[CrossRef\]](#)
90. Brito, A.C.; Newton, A.; Tett, P.; Fernandes, T.F. Changes in the yield of microphytobenthic chlorophyll from nutrients: Considering denitrification. *Ecol. Indic.* **2012**, *19*, 226–230. [\[CrossRef\]](#)
91. Kim, K.B.; Jung, M.-K.; Tsang, Y.F.; Kwon, H.-H. Stochastic modeling of chlorophyll-a for probabilistic assessment and monitoring of algae blooms in the Lower Nakdong River, South Korea. *J. Hazard. Mater.* **2020**, *400*, 123066. [\[CrossRef\]](#) [\[PubMed\]](#)
92. Shen, J.; Qin, Q.; Wang, Y.; Sisson, M. A data-driven modeling approach for simulating algal blooms in the tidal freshwater of James River in response to riverine nutrient loading. *Ecol. Model.* **2019**, *398*, 44–54. [\[CrossRef\]](#)
93. Yu, X.; Shen, J.; Du, J. A Machine-Learning-Based Model for Water Quality in Coastal Waters, Taking Dissolved Oxygen and Hypoxia in Chesapeake Bay as an Example. *Water Resour. Res.* **2020**, *56*, e2020WR027227. [\[CrossRef\]](#)
94. Uddin, M.G.; Nash, S.; Rahman, A.; Dabrowski, T.; Olbert, A.I. Data-driven modelling for assessing trophic status in marine ecosystems using machine learning approaches. *Environ. Res.* **2023**, *242*, 117755. [\[CrossRef\]](#) [\[PubMed\]](#)
95. Yu, X.; Shen, J.; Zheng, G.; Du, J. Chlorophyll-a in Chesapeake Bay based on VIIRS satellite data: Spatiotemporal variability and prediction with machine learning. *Ocean Model.* **2022**, *180*, 102119. [\[CrossRef\]](#)
96. Kwon, Y.S.; Baek, S.H.; Lim, Y.K.; Pyo, J.; Ligaray, M.; Park, Y.; Cho, K.H. Monitoring Coastal Chlorophyll-a Concentrations in Coastal Areas Using Machine Learning Models. *Water* **2018**, *10*, 1020. [\[CrossRef\]](#)
97. Yao, H.; Huang, Y.; Wei, Y.; Zhong, W.; Wen, K. Retrieval of Chlorophyll-a Concentrations in the Coastal Waters of the Beibu Gulf in Guangxi Using a Gradient-Boosting Decision Tree Model. *Appl. Sci.* **2021**, *11*, 7855. [\[CrossRef\]](#)

98. Kim, H.-R.; Soh, H.Y.; Kwak, M.-T.; Han, S.-H. Machine Learning and Multiple Imputation Approach to Predict Chlorophyll-a Concentration in the Coastal Zone of Korea. *Water* **2022**, *14*, 1862. [\[CrossRef\]](#)
99. Gómez, D.; Salvador, P.; Sanz, J.; Casanova, J.L. A new approach to monitor water quality in the Menor sea (Spain) using satellite data and machine learning methods. *Environ. Pollut.* **2021**, *286*, 117489. [\[CrossRef\]](#)
100. Mohebzadeh, H.; Lee, T. Spatial downscaling of MODIS Chlorophyll-a with machine learning techniques over the west coast of the Yellow Sea in South Korea. *J. Oceanogr.* **2021**, *77*, 103–122. [\[CrossRef\]](#)
101. Lin, J.; Liu, Q.; Song, Y.; Liu, J.; Yin, Y.; Hall, N.S. Temporal Prediction of Coastal Water Quality Based on Environmental Factors with Machine Learning. *J. Mar. Sci. Eng.* **2023**, *11*, 1608. [\[CrossRef\]](#)
102. Niu, J.; Feng, Z.; He, M.; Xie, M.; Lv, Y.; Zhang, J.; Sun, L.; Liu, Q.; Hu, B.X. Incorporating marine particulate carbon into machine learning for accurate estimation of coastal chlorophyll-a. *Mar. Pollut. Bull.* **2023**, *192*, 115089. [\[CrossRef\]](#)
103. Deng, T.; Chau, K.-W.; Duan, H.-F. Machine learning based marine water quality prediction for coastal hydro-environment management. *J. Environ. Manag.* **2021**, *284*, 112051. [\[CrossRef\]](#)
104. Su, H.; Lu, X.; Chen, Z.; Zhang, H.; Lu, W.; Wu, W. Estimating Coastal Chlorophyll-A Concentration from Time-Series OLCI Data Based on Machine Learning. *Remote Sens.* **2021**, *13*, 576. [\[CrossRef\]](#)
105. Pahlevan, N.; Smith, B.; Schalles, J.; Binding, C.; Cao, Z.; Ma, R.; Alikas, K.; Kangro, K.; Gurlin, D.; Hà, N.; et al. Seamless retrievals of chlorophyll-a from Sentinel-2 (MSI) and Sentinel-3 (OLCI) in inland and coastal waters: A machine-learning approach. *Remote Sens. Environ.* **2020**, *240*, 111604. [\[CrossRef\]](#)
106. Woody, C.; Shih, E.; Miller, J.; Royer, T.; Atkinson, L.P.; Moody, R.S. Measurements of Salinity in the Coastal Ocean: A Review of Requirements and Technologies. *Mar. Technol. Soc. J.* **2000**, *34*, 26–33. [\[CrossRef\]](#)
107. Luo, L.; Li, S.; Wang, D. Hypoxia in the Pearl River Estuary, the South China Sea, in July 1999. *Aquat. Ecosyst. Health Manag.* **2009**, *12*, 418–428. [\[CrossRef\]](#)
108. He, H.; Chen, Y.; Li, X.; Cheng, Y.; Yang, C.; Zeng, G. Influence of salinity on microorganisms in activated sludge processes: A review. *Int. Biodeterior. Biodegrad.* **2017**, *119*, 520–527. [\[CrossRef\]](#)
109. Johnson, G.C.; Schmidt, S.; Lyman, J.M. Relative contributions of temperature and salinity to seasonal mixed layer density changes and horizontal density gradients. *J. Geophys. Res. Oceans* **2012**, *117*, 2011JC007651. [\[CrossRef\]](#)
110. Schmidt, H.; Seitz, S.; Hassel, E.; Wolf, H. The density–salinity relation of standard seawater. *Ocean Sci.* **2018**, *14*, 15–40. [\[CrossRef\]](#)
111. Guillou, N.; Chapalain, G.; Petton, S. Predicting sea surface salinity in a tidal estuary with machine learning. *Oceanologia* **2023**, *65*, 318–332. [\[CrossRef\]](#)
112. Chawla, P.; Cao, X.; Fu, Y.; Hu, C.; Wang, M.; Wang, S.; Gao, J.Z. Water quality prediction of salton sea using machine learning and big data techniques. *Int. J. Environ. Anal. Chem.* **2023**, *103*, 6835–6858. [\[CrossRef\]](#)
113. Lal, A.; Datta, B. Performance Evaluation of Homogeneous and Heterogeneous Ensemble Models for Groundwater Salinity Predictions: A Regional-Scale Comparison Study. *Water Air Soil Pollut.* **2020**, *231*, 320. [\[CrossRef\]](#)
114. Zaitsev, N.K.; Dvorkin, V.I.; Melnikov, P.V.; Kozhukhova, A.E. A Dissolved Oxygen Analyzer with an Optical Sensor. *J. Anal. Chem.* **2018**, *73*, 102–108. [\[CrossRef\]](#)
115. Ziyad Sami, B.F.; Latif, S.D.; Ahmed, A.N.; Chow, M.F.; Murti, M.A.; Suhendi, A.; Ziyad Sami, B.H.; Wong, J.K.; Birima, A.H.; El-Shafie, A. Machine learning algorithm as a sustainable tool for dissolved oxygen prediction: A case study of Feitsui Reservoir, Taiwan. *Sci. Rep.* **2022**, *12*, 3649. [\[CrossRef\]](#)
116. Chan, F.; Barth, J.A.; Lubchenco, J.; Kirincich, A.; Weeks, H.; Peterson, W.T.; Menge, B.A. Emergence of Anoxia in the California Current Large Marine Ecosystem. *Science* **2008**, *319*, 920. [\[CrossRef\]](#)
117. Grantham, B.A.; Chan, F.; Nielsen, K.J.; Fox, D.S.; Barth, J.A.; Huyer, A.; Lubchenco, J.; Menge, B.A. Upwelling-driven nearshore hypoxia signals ecosystem and oceanographic changes in the northeast Pacific. *Nature* **2004**, *429*, 749–754. [\[CrossRef\]](#) [\[PubMed\]](#)
118. Salas, E.A.L.; Kumaran, S.S.; Partee, E.B.; Willis, L.P.; Mitchell, K. Potential of mapping dissolved oxygen in the Little Miami River using Sentinel-2 images and machine learning algorithms. *Remote Sens. Appl. Soc. Environ.* **2022**, *26*, 100759. [\[CrossRef\]](#)
119. Valera, M.; Walter, R.K.; Bailey, B.A.; Castillo, J.E. Machine Learning Based Predictions of Dissolved Oxygen in a Small Coastal Embayment. *J. Mar. Sci. Eng.* **2020**, *8*, 1007. [\[CrossRef\]](#)
120. Park, S.; Kim, K.; Hibino, T.; Sakai, Y.; Furukawa, T.; Kim, K. An Antifouling Redox Sensor with a Flexible Carbon Fiber Electrode for Machine Learning-Based Dissolved Oxygen Prediction in Severely Eutrophic Waters. *Water* **2023**, *15*, 2467. [\[CrossRef\]](#)
121. Kim, Y.H.; Im, J.; Ha, H.K.; Choi, J.-K.; Ha, S. Machine learning approaches to coastal water quality monitoring using GOCI satellite data. *GIScience Remote Sens.* **2014**, *51*, 158–174. [\[CrossRef\]](#)
122. Tian, S.; Guo, H.; Xu, W.; Zhu, X.; Wang, B.; Zeng, Q.; Mai, Y.; Huang, J.J. Remote sensing retrieval of inland water quality parameters using Sentinel-2 and multiple machine learning algorithms. *Environ. Sci. Pollut. Res.* **2022**, *30*, 18617–18630. [\[CrossRef\]](#)
123. Jimeno-Sáez, P.; Senent-Aparicio, J.; Cecilia, J.M.; Pérez-Sánchez, J. Using Machine-Learning Algorithms for Eutrophication Modeling: Case Study of Mar Menor Lagoon (Spain). *Int. J. Environ. Res. Public Health* **2020**, *17*, 1189. [\[CrossRef\]](#) [\[PubMed\]](#)
124. Zhu, X.; Guo, H.; Huang, J.J.; Tian, S.; Xu, W.; Mai, Y. An ensemble machine learning model for water quality estimation in coastal area based on remote sensing imagery. *J. Environ. Manag.* **2022**, *323*, 116187. [\[CrossRef\]](#) [\[PubMed\]](#)
125. Quang, N.H.; Dinh, N.T.; Dien, N.T.; Son, L.T. Calibration of Sentinel-2 Surface Reflectance for Water Quality Modelling in Binh Dinh's Coastal Zone of Vietnam. *Sustainability* **2023**, *15*, 1410. [\[CrossRef\]](#)
126. Chen, S.; Zhang, Z.; Lin, J.; Huang, J. Machine learning-based estimation of riverine nutrient concentrations and associated uncertainties caused by sampling frequencies. *PLoS ONE* **2022**, *17*, e0271458. [\[CrossRef\]](#) [\[PubMed\]](#)

127. Radu, V.-M.; Ionescu, P.; Deak, G.; Diacu, E.; Ivanov, A.A.; Zamfir, S.; Marcus, M.-I. Overall assessment of surface water quality in the Lower Danube River. *Environ. Monit. Assess.* **2020**, *192*, 135. [\[CrossRef\]](#)
128. Said, A.; Stevens, D.K.; Sehlke, G. An Innovative Index for Evaluating Water Quality in Streams. *Environ. Manag.* **2004**, *34*, 406–414. [\[CrossRef\]](#) [\[PubMed\]](#)
129. Qu, X.; Chen, Y.; Liu, H.; Xia, W.; Lu, Y.; Gang, D.-D.; Lin, L.-S. A holistic assessment of water quality condition and spatiotemporal patterns in impounded lakes along the eastern route of China's South-to-North water diversion project. *Water Res.* **2020**, *185*, 116275. [\[CrossRef\]](#)
130. Uddin, M.G.; Nash, S.; Rahman, A.; Olbert, A.I. A sophisticated model for rating water quality. *Sci. Total Environ.* **2023**, *868*, 161614. [\[CrossRef\]](#)
131. Akhtar, N.; Ishak, M.I.S.; Ahmad, M.I.; Umar, K.; Md Yusuff, M.S.; Anees, M.T.; Qadir, A.; Ali Almanasir, Y.K. Modification of the Water Quality Index (WQI) Process for Simple Calculation Using the Multi-Criteria Decision-Making (MCDM) Method: A Review. *Water* **2021**, *13*, 905. [\[CrossRef\]](#)
132. Uddin, M.G.; Nash, S.; Rahman, A.; Olbert, A.I. A novel approach for estimating and predicting uncertainty in water quality index model using machine learning approaches. *Water Res.* **2023**, *229*, 119422. [\[CrossRef\]](#)
133. Uddin, M.G.; Nash, S.; Rahman, A.; Olbert, A.I. Assessing optimization techniques for improving water quality model. *J. Clean. Prod.* **2023**, *385*, 135671. [\[CrossRef\]](#)
134. Pati, S.; Dash, M.K.; Mukherjee, C.K.; Dash, B.; Pokhrel, S. Assessment of water quality using multivariate statistical techniques in the coastal region of Visakhapatnam, India. *Environ. Monit. Assess.* **2014**, *186*, 6385–6402. [\[CrossRef\]](#)
135. Gupta, A.K.; Gupta, S.K.; Patil, R.S. A Comparison of Water Quality Indices for Coastal Water. *J. Environ. Sci. Health Part A* **2003**, *38*, 2711–2725. [\[CrossRef\]](#) [\[PubMed\]](#)
136. Uddin, M.G.; Nash, S.; Mohammad Diganta, M.T.; Rahman, A.; Olbert, A.I. Robust machine learning algorithms for predicting coastal water quality index. *J. Environ. Manag.* **2022**, *321*, 115923. [\[CrossRef\]](#) [\[PubMed\]](#)
137. Uddin, M.G.; Nash, S.; Rahman, A.; Olbert, A.I. A comprehensive method for improvement of water quality index (WQI) models for coastal water quality assessment. *Water Res.* **2022**, *219*, 118532. [\[CrossRef\]](#) [\[PubMed\]](#)
138. Rezaie-Balf, M.; Attar, N.F.; Mohammadzadeh, A.; Murti, M.A.; Ahmed, A.N.; Fai, C.M.; Nabipour, N.; Alaghmand, S.; El-Shafie, A. Physicochemical parameters data assimilation for efficient improvement of water quality index prediction: Comparative assessment of a noise suppression hybridization approach. *J. Clean. Prod.* **2020**, *271*, 122576. [\[CrossRef\]](#)
139. Uddin, M.G.; Nash, S.; Olbert, A.I. A review of water quality index models and their use for assessing surface water quality. *Ecol. Indic.* **2021**, *122*, 107218. [\[CrossRef\]](#)
140. Sutadian, A.D.; Muttill, N.; Yilmaz, A.G.; Perera, B.J.C. Development of river water quality indices—A review. *Environ. Monit. Assess.* **2016**, *188*, 58. [\[CrossRef\]](#)
141. Babbar, R.; Babbar, S. Predicting river water quality index using data mining techniques. *Environ. Earth Sci.* **2017**, *76*, 504. [\[CrossRef\]](#)
142. Bui, D.T.; Khosravi, K.; Tiefenbacher, J.; Nguyen, H.; Kazakis, N. Improving prediction of water quality indices using novel hybrid machine-learning algorithms. *Sci. Total Environ.* **2020**, *721*, 137612. [\[CrossRef\]](#)
143. Gao, Y.; Qian, H.; Ren, W.; Wang, H.; Liu, F.; Yang, F. Hydrogeochemical characterization and quality assessment of groundwater based on integrated-weight water quality index in a concentrated urban area. *J. Clean. Prod.* **2020**, *260*, 121006. [\[CrossRef\]](#)
144. Kouadri, S.; Elbeltagi, A.; Islam, A.R.M.T.; Kateb, S. Performance of machine learning methods in predicting water quality index based on irregular data set: Application on Illizi region (Algerian southeast). *Appl. Water Sci.* **2021**, *11*, 190. [\[CrossRef\]](#)
145. Wang, X.; Zhang, F.; Ding, J. Evaluation of water quality based on a machine learning algorithm and water quality index for the Ebinur Lake Watershed, China. *Sci. Rep.* **2017**, *7*, 12858. [\[CrossRef\]](#)
146. Ho, J.Y.; Afan, H.A.; El-Shafie, A.H.; Koting, S.B.; Mohd, N.S.; Jaafar, W.Z.B.; Sai, H.L.; Malek, M.A.; Ahmed, A.N.; Mohtar, W.H.M.W.; et al. Towards a time and cost effective approach to water quality index class prediction. *J. Hydrol.* **2019**, *575*, 148–165. [\[CrossRef\]](#)
147. Tanega, M.C.; Fajardo, A.; Limbago, J.S. Analysis of Water Quality for Taal Lake Using Machine Learning Classification Algorithm. In Proceedings of the 2023 20th International Joint Conference on Computer Science and Software Engineering (JCSSE), Phitsanulok, Thailand, 28 June–1 July 2023; pp. 397–402.
148. Sheikh Khozani, Z.; Iranmehr, M.; Wan Mohtar, W.H.M. Improving Water Quality Index prediction for water resources management plans in Malaysia: Application of machine learning techniques. *Geocarto Int.* **2022**, *37*, 10058–10075. [\[CrossRef\]](#)
149. Liu, G.; Ye, J.; Chen, Y.; Yang, X.; Gu, Y. Analysis of Water Pollution Causes and Control Countermeasures in Liaohe Estuary via Support Vector Machine Particle Swarm Optimization under Deep Learning. *Comput. Model. Eng. Sci.* **2022**, *130*, 315–329. [\[CrossRef\]](#)
150. McKenzie, T.; Dulai, H.; Fuleky, P. Traditional and novel time-series approaches reveal submarine groundwater discharge dynamics under baseline and extreme event conditions. *Sci. Rep.* **2021**, *11*, 22570. [\[CrossRef\]](#)
151. Gündoğdu, S.; Ayat, B.; Aydoğan, B.; Çevik, C.; Karaca, S. Hydrometeorological assessments of the transport of microplastic pellets in the Eastern Mediterranean. *Sci. Total Environ.* **2022**, *823*, 153676. [\[CrossRef\]](#)
152. Patryniak, K.; Collu, M.; Coraddu, A. Rigid body dynamic response of a floating offshore wind turbine to waves: Identification of the instantaneous centre of rotation through analytical and numerical analyses. *Renew. Energy* **2023**, *218*, 119378. [\[CrossRef\]](#)

153. Sanada, Y.; Kim, D.-H.; Sadat-Hosseini, H.; Stern, F.; Hossain, M.A.; Wu, P.-C.; Toda, Y.; Otzen, J.; Simonsen, C.; Abdel-Maksoud, M.; et al. Assessment of EFD and CFD capability for KRISO Container Ship added power in head and oblique waves. *Ocean Eng.* **2022**, *243*, 110224. [\[CrossRef\]](#)
154. Ehteram, M.; Ahmed, A.N.; Latif, S.D.; Huang, Y.F.; Alizamir, M.; Kisi, O.; Mert, C.; El-Shafie, A. Design of a hybrid ANN multi-objective whale algorithm for suspended sediment load prediction. *Environ. Sci. Pollut. Res.* **2021**, *28*, 1596–1611. [\[CrossRef\]](#) [\[PubMed\]](#)
155. Fei, K.; Du, H.; Gao, L. The contribution of typhoon local and remote forcings to storm surge along the Makou-Dahengqin tidal reach of Pearl River Estuary. *Sci. Total Environ.* **2023**, *899*, 165592. [\[CrossRef\]](#) [\[PubMed\]](#)
156. Shamshirband, S.; Mosavi, A.; Rabczuk, T.; Nabipour, N.; Chau, K. Prediction of significant wave height; comparison between nested grid numerical model, and machine learning models of artificial neural networks, extreme learning and support vector machines. *Eng. Appl. Comput. Fluid Mech.* **2020**, *14*, 805–817. [\[CrossRef\]](#)
157. Ayyad, M.; Hajj, M.R.; Marsooli, R. Machine learning-based assessment of storm surge in the New York metropolitan area. *Sci. Rep.* **2022**, *12*, 19215. [\[CrossRef\]](#) [\[PubMed\]](#)
158. Xu, G.; Ji, C.; Xu, Y.; Yu, E.; Cao, Z.; Wu, Q.; Lin, P.; Wang, J. Machine learning in coastal bridge hydrodynamics: A state-of-the-art review. *Appl. Ocean Res.* **2023**, *134*, 103511. [\[CrossRef\]](#)
159. Ali, M.; Prasad, R. Significant wave height forecasting via an extreme learning machine model integrated with improved complete ensemble empirical mode decomposition. *Renew. Sustain. Energy Rev.* **2019**, *104*, 281–295. [\[CrossRef\]](#)
160. Franklin, G.L.; Torres-Freyermuth, A. On the runup parameterisation for reef-lined coasts. *Ocean Model.* **2022**, *169*, 101929. [\[CrossRef\]](#)
161. Ayyad, M.; Hajj, M.R.; Marsooli, R. Climate change impact on hurricane storm surge hazards in New York/New Jersey Coastlines using machine-learning. *NPJ Clim. Atmos. Sci.* **2023**, *6*, 88. [\[CrossRef\]](#)
162. Sampurno, J.; Vallaes, V.; Ardianto, R.; Hanert, E. Integrated hydrodynamic and machine learning models for compound flooding prediction in a data-scarce estuarine delta. *Nonlinear Process. Geophys.* **2022**, *29*, 301–315. [\[CrossRef\]](#)
163. Huang, P.-C. An effective alternative for predicting coastal floodplain inundation by considering rainfall, storm surge, and downstream topographic characteristics. *J. Hydrol.* **2022**, *607*, 127544. [\[CrossRef\]](#)
164. Simmons, J.A.; Splinter, K.D. A multi-model ensemble approach to coastal storm erosion prediction. *Environ. Model. Softw.* **2022**, *150*, 105356. [\[CrossRef\]](#)
165. Wei, Z.; Davison, A. A convolutional neural network based model to predict nearshore waves and hydrodynamics. *Coast. Eng.* **2022**, *171*, 104044. [\[CrossRef\]](#)
166. Riaz, M.Z.B.; Iqbal, U.; Yang, S.-Q.; Sivakumar, M.; Enever, K.; Khalil, U.; Ji, R.; Miguntanna, N.S. SedimentNet—A 1D-CNN machine learning model for prediction of hydrodynamic forces in rapidly varied flows. *Neural Comput. Appl.* **2022**, *35*, 9145–9166. [\[CrossRef\]](#)
167. Pham, H.V.; Dal Barco, M.K.; Cadau, M.; Harris, R.; Furlan, E.; Torresan, S.; Rubinetti, S.; Zanchettin, D.; Rubino, A.; Kuznetsov, I.; et al. Multi-model chain for climate change scenario analysis to support coastal erosion and water quality risk management for the Metropolitan city of Venice. *Sci. Total Environ.* **2023**, *904*, 166310. [\[CrossRef\]](#)
168. Kim, J.; Park, J. Bayesian structural equation modeling for coastal management: The case of the Saemangeum coast of Korea for water quality improvements. *Ocean Coast. Manag.* **2017**, *136*, 120–132. [\[CrossRef\]](#)
169. Cebe, K.; Balas, L. Water quality modelling in kaş bay. *Appl. Math. Model.* **2016**, *40*, 1887–1913. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.