water

MDPI

Review

Recent Progress on Surface Water Quality Models Utilizing Machine Learning Techniques

Mengjie He [1], Qin Qian [2,*], Xinyu Liu [1], Jing Zhang [3] and James Curry [1]

[1] Department of Industry and Systems Engineering, Lamar University, Beaumont, TX 77710, USA; mhe@lamar.edu (M.H.); xliu1@lamar.edu (X.L.); jcurry@lamar.edu (J.C.)
[2] Department of Civil and Environmental Engineering, Lamar University, Beaumont, TX 77710, USA
[3] Department of Computer Science, Lamar University, Beaumont, TX 77710, USA; jzhang9@lamar.edu
* Correspondence: qqian1@lamar.edu

Abstract: Surface waterbodies are heavily exposed to pollutants caused by natural disasters and human activities. Empowering sensor technologies in water quality monitoring, sufficient measurements have become available to develop machine learning (ML) models. Numerous ML models have quickly been adopted to predict water quality indicators in various surface waterbodies. This paper reviews 78 recent articles from 2022 to October 2024, categorizing water quality models utilizing ML into three groups: Point-to-Point (P2P), which estimates the current target value based on other measurements at the same time point; Sequence-to-Point (S2P), which utilizes previous time series data to predict the target value at one time point ahead; and Sequence-to-Sequence (S2S), which uses previous time series data to forecast sequential target values in the future. The ML models used in each group are classified and compared according to water quality indicators, data availability, and model performance. Widely used strategies for improving performance, including feature engineering, hyperparameter tuning, and transfer learning, are recognized and described to enhance model effectiveness. The interpretability limitations of ML applications are discussed. This review provides a perspective on emerging ML for surface water quality models.

Keywords: surface water management; machine learning (ML); water quality model; model selection; model improvement

check for updates

Citation: He, M.; Qian, Q.; Liu, X.; Zhang, J.; Curry, J. Recent Progress on Surface Water Quality Models Utilizing Machine Learning Techniques. *Water* **2024**, *16*, 3616. https://doi.org/10.3390/w16243616

Academic Editors: Tianxiang Wang and Alex Neumann

Received: 8 November 2024
Revised: 11 December 2024
Accepted: 13 December 2024
Published: 15 December 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

Surface water quality deterioration has become a global problem due to climate changes and human activities, including extreme weather (e.g., droughts, heat), natural disasters (e.g., hurricanes, floodings), land erosion, urban runoff, agricultural and livestock wastewater, and unmanaged domestic sewage [1–4]. It is urgent to develop sustainable water management strategies to preserve water resources. Traditional water quality models such as EFDC [5], WASP [6], MIKE 3 [7], CE-QUAL-W2 (V3) [8], CH3D [9], and Delft3D [10] simulate the spatial and temporal variations in water quality based on governing equations. These partial differential equations determine hydrodynamics and water quality by physical mechanisms coupled with biogeochemical processes, requiring costly field and lab measurements for determining empirical coefficients [11,12].

In recent years, machine learning (ML) applications to water resource management have garnered significant interest [13,14]. ML techniques have been applied in hydrological models to discuss the transport of sediment and its correlation to water quality due to rains and runoffs [15,16]. On the other hand, ML models are utilized to explore the potential relationships of water quality measurements by directly training with water quality data, which are the analyzed targets in this study. The water quality data used to train ML models generally come from two sources. One is the discrete field/lab sample data collected from multiple monitoring stations over several decades with relatively low sampling frequencies (days/weeks/months). The other is real-time data at a fixed interval of 15 to 60 min from on-site automated recording equipment enabled by advancements in smart sensing, wireless communication, etc. Vast quantities of sensor data make it possible for ML models to explore complex relationships among measurements [17–19]. Therefore, an increasing number of water quality models have been developed using ML for timely and cost-effective water management decision making.

Generally, ML models can be divided into two categories, traditional ML models and deep learning (DL) models (Figure 1), which are introduced in [20,21] with more details. The traditional ML models have demonstrated competitive performance in many studies [12,22–31]. They include kernel-based models like Support Vector Machines (SVMs) [24], example-based models like K-Nearest Neighbor (KNN) [25], Bayesian models like Naïve Bayes (NB) [26], tree-based models including Decision Trees (DTs) [27] and Random forests (RFs) [28], boosting models including the Gradient Boosting Machine (GBM) [29] and Extreme Gradient Boosting (XGBoost) [30], shallow Artificial Neural Networks (ANNs) like Multilayer Perceptron Neural Networks (MLPNNs) [31], and stacking models integrated with several ML models. On the other hand, the DL model, as an evolution of ANNs with complex and deep neural network structures, has been classified as a separate category [32,33]. Among the presented DL models, Recurrent Neural Networks (RNNs), Gated Recurrent Units (GRUs), Long Short-Term Memory (LSTM), and Convolutional Neural Networks (CNNs) are the most commonly adopted for surface water quality predictions [21].

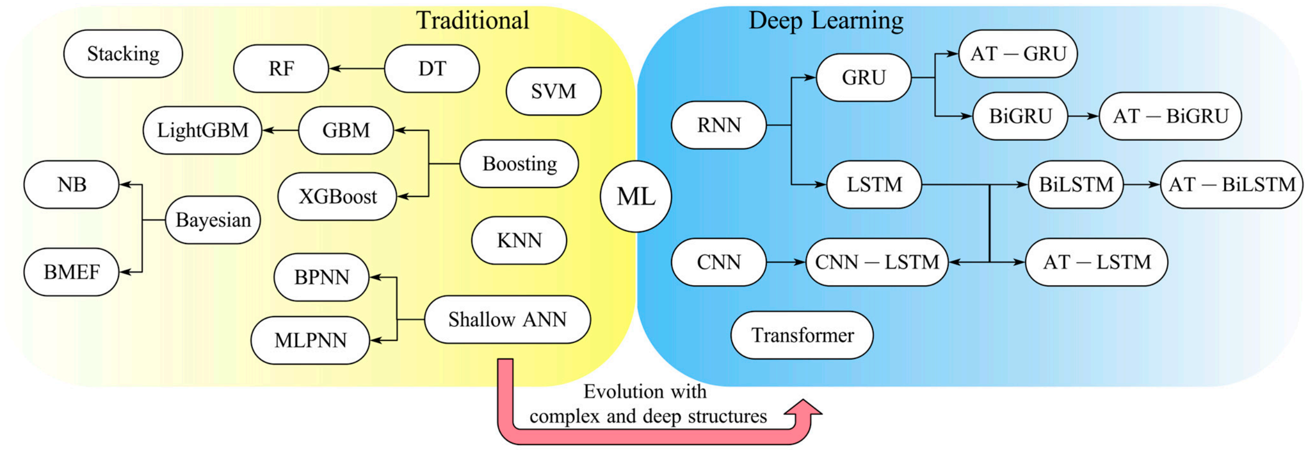

Figure 1. Summary of the main traditional models and deep learning models in this review.

It is well known that water quality monitoring data are highly dependent on local hydro-meteorological, environmental, and social conditions [34] and vary with the data-collecting tools and measuring methods used. The limited reusability of the best ML model from one region in another has been recognized [35]. The non-stationarity and noise of the datasets decrease the model performance in ML applications [36]. Meanwhile, employing different statistical methods for model evaluation leads to inconsistent reporting across datasets and regions [37]. Choosing a suitable ML model for a specific dataset becomes an emerging question since the model performance varies due to data heterogeneity and inconsistent model evaluation methods.

Two review articles [21,38] on ML applications in surface water resources were published before 2022. The detailed procedures of developing ML models such as SVM, RF, and ANN for water quality prediction were introduced in [38], including data acquisition and data cleaning, as well as model selection, training, and evaluation. The performance of ML models depends on data splitting, input features, and model architecture [38]. DL models such as RNN, GRU, LSTM, and CNN applied for water quality prediction have achieved high performance for large datasets and multiple input features [21]. Both reviews [21,38] concluded the limitations of data size and quality, as well as the lack of a universal model for different domains. The challenges of hyperparameter tuning and model comparison have been discussed in [21] to provide perspectives on self-adjusting strategies for hyperparameters and standards for fairly comparing models between different datasets in future works.

However, no model classification makes it difficult to select the suitable model for specific goals. To systematically review recent advancements in water quality models and assist in model selection, this paper classifies the 78 recent articles published from 2022 to date (October 2024) into the following three categories:

- Point-to-Point (P2P) models are developed to estimate current water quality indicators depending on other measurements monitored at the same time point. They aim to discover relationships between the target and other measurements, finding an alternative way to estimate the target water quality indicators.
- Sequence-to-Point (S2P) models consider water quality variations with time, utilizing previous time series data to predict future water quality indicators at only one time point ahead.
- Sequence-to-Sequence (S2S) models also utilize previous time series data to forecast multiple target values as an output sequence for a future period, extending beyond just one time point.

This paper aims to summarize water quality indicators, data availability, and model performance for each group while providing suggestions to facilitate model selection, evaluation, and performance improvement for water quality prediction. The rest of the paper is organized as follows. Section 2 introduces the sources and profiles of the reviewed articles. Section 3 classifies the ML water quality models as P2P, S2P, or S2S models, and analyzes the ML models in each group. Section 4 describes the existing strategies used to improve model performance, including feature engineering, hyperparameter optimization, and transfer learning methods. Section 5 summarizes and compares the common statistical metrics of model evaluation. Finally, Sections 6 and 7 discuss the limitations of ML applications and their possible solutions and draw conclusions.

2. Source and Profile of Reviewed Articles

This study searched for journal articles on Google Scholar with the keywords "machine learning", "water quality", and "estimate/predict/forecast". This review covered 78 journal articles published from 2022 to date (October 2024), including but not limited to Water, Sustainability, Science of The Total Environment, Journal of Environmental Management, Journal of Hydrology, Environmental Science and Pollution Research, Ecological Informatics, Ecological Indicators, Environmental Pollution, Process Safety and Environmental Protection, and Multimedia Tools and Applications. These articles applied traditional ML and DL techniques to develop water quality models for rivers, lakes, reservoirs, canals, estuaries, and coastal areas. The water quality indicators used in those papers include but are not limited to water quality index/class (WQI/WQC), dissolved oxygen (DO), water temperature (T_w), pH, turbidity, conductivity (EC), chlorophyll-a (Chl-a), ammonia–nitrogen (NH_3-N), total phosphorus (TP), total nitrogen (TN), total dissolved solids (TDS), suspended solids (SS), biochemical oxygen demand (BOD), chemical oxygen demand (COD), permanganate index (COD_{Mn}), and the oxidation–reduction potential (ORP).

3. Analysis of Machine Learning Models for Water Quality Prediction

Among the reviewed articles, P2P models represent 57.8% of 78 articles. However, the proportion declined considerably from 64% in 2022 to 49% after 2022. S2P and S2S models account for 32.5% and 9.6%, respectively. The popularity of S2P models increased from 30% in 2022 to 36% after 2022 and that of S2S models jumped from 6% in 2022 to 15% after 2022. Although most previous studies focus on traditional ML models in the field of water resource management [39], surface water quality models utilizing DL increased from 50% in 2022 to 64% after 2022. This demonstrates that the benefits and importance of advanced DL models have been recognized, as shown by the increasing demand for S2P and S2S models in the field of water quality management. The increasing number of applications with DL is also related to advancements in software availability, including the expansion of the programming libraries PyTorch (https://pytorch.org/, accessed on 10 December 2024) and TensorFlow (https://www.tensorflow.org/, accessed on 10 December 2024) as well as hardware accessibility, such as Graphics Processing Units (GPUs) and platforms like Google Colab (https://colab.google/, accessed on 10 December 2024). Programming libraries greatly simplify the complex modeling processes of DL while GPUs accelerate model training processes with powerful parallel computing capabilities.

3.1. ML Applications on P2P Models

The dataset for the P2P model is mostly derived from weekly, monthly, or seasonal water monitoring data with small data sizes (ranging from 30 to 3700), as shown in Table 1. A total of 57.1% of the articles chose WQI/WQC as their target model for the holistic assessment of water quality. The calculation of WQI/WQC relies on many physical, chemical, and biological measurements, making this assessment inconvenient and costly [40]. P2P models are developed as alternatives to estimate WQI/WQC with limited measurements. The top four water quality indicators used in P2P models are DO (16.3%), EC (16.3%), TDS (14.3%), and TP (12.2%). In terms of the popularity of the ML models (Figure 2), RF has the highest proportion and has been used in 22.7% of the articles, followed by MLPNN and SVM, accounting for 17.3% and 16.4%, respectively.

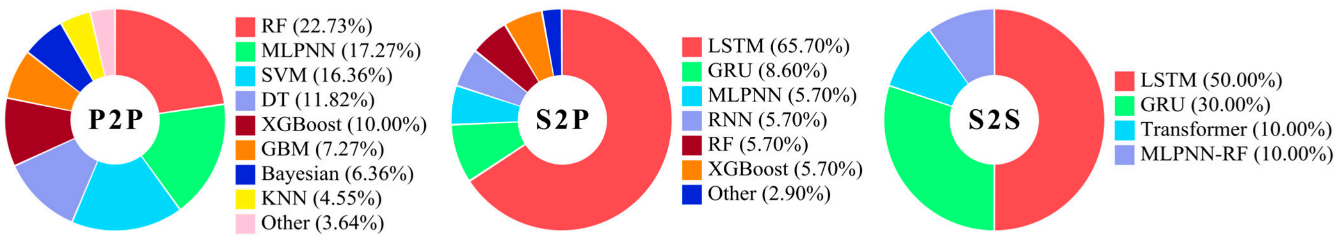

Figure 2. Proportions of different ML models applied on P2P, S2P, and S2S models.

1. The RF, ensembled by many DTs, is often used to develop both regression and classification P2P models. It is adaptable to different-sized datasets in many domains by adjusting the number of DTs and is less likely to suffer from overfitting problems [41–43]. It has been applied to estimate the monthly EC and TDS in the Indus River, Pakistan, with the highest coefficient of determination (R^2) = 0.93 [44] and superior to multi-expression programming (MEP) [45]. It also resulted in the smallest root mean square error (RMSE) in a WQI estimation from 11 water quality measurements (salinity, T_w, DO, NH_3-N, Chl-a, etc.) in Cork Harbour, Ireland [46].
2. The MLPNN is the basic ANN, commonly comprising three neural layers. It gains popularity due to its robustness and capability to capture nonlinear patterns in data [47]. Meanwhile, it is flexible to adjust the structure of hidden layers to obtain a better performance in the WQI regression model as compared to KNN, DT, SVR, Nonlinear Autoregressive Network (NARNet), and even LSTM in some cases [42,48]. The Back-Propagation Neural Network (BPNN) is utilized to predict one of the water quality indicators, namely WQC, pH, DO, COD, NH_3-N, and TP, with four other measurements in Luoyang River, China, and it is superior to SVM and LSTM by combining an AEABC optimizer constructed with an adaptive evolution strategy and artificial bee colony algorithm [49].

3. The SVM has also been widely utilized for the classification and regression of water quality indicators [24]. Its advantage is in its exploration of the potential relationships of inputs and outputs by mapping the data into a high-dimensional or even infinite-dimensional feature space. It demonstrates the best accuracy (96.35%) for the WQC model on DO, BOD, COD, SS, pH, and NH_3-N, superior to the MLPNN and DT [50]. Meanwhile, it achieved the highest R^2 (0.95) in a monthly WQI estimation in Gulshan Lake, Bangladesh [51].
4. The DT is known for its interpretable flow chart structure [47], but it may not be appropriate for time series problems and it is prone to overfit the training dataset [52]. It is built with the bagging technique for monthly EC and TDS based on T_w, pH, bicarbonates (HCO_3^-), chloride (Cl^-), sulfate (SO_4^{2-}), calcium (Ca^{2-}), magnesium (Mg^{2+}), and sodium (Na^+) in the Upper Indus River [53]. The bagging technique significantly decreases the Mean absolute error (MAE) of the DT from 16.82 to 9.04, outperforming the RF, SVM, and MLPNN.
5. XGBoost and GBM, which are ensemble meta-algorithm models, improve the performance of several weaker models by reducing bias and variance [30]. XGBoost results in the best performance (accuracy = 1) when estimating WQC in Cork Harbour, Ireland [54]. In addition, XGBoost and LightGBM are applied to predict daily mean pH values collected from USGS stations and outperform linear regression (LR), RF, and MLPNN, while the error of LightGBM is slightly lower than that of XGBoost [55]. Moreover, the classifier developed with GBM achieves the highest accuracy among RF, XGBoost, and Adaptive Boosting (AdaBoost) [48].
6. The Bayesian and KNN are often used as the benchmark models for performance comparison [46,48]. However, the WQC model developed by Naïve Bayes (NB) correctly predicted 64 out of 68 cases in Vojvodina Province, Serbia [26]. A variant of the Bayesian, Bayesian Maximum Entropy-based Fusion (BMEF), has the best performance among RF, SVR, and MLPNN estimating Chl-a, DO, T_w, EC, and ORP with the latitude, longitude, and depth of sample locations [56]. In addition, the KNN also achieved the second-best results (accuracy = 94%) in a WQI classification, outperforming SVM, NB, DT, and ANN [54].
7. Other models such as Gaussian Process Regression (GPR) [57], Classification and Regression Tree (CART), Chi-squared Automatic Interaction Detector (CHAID) [58], reduced error pruning tree (REPT) [59], Gene Expression Programming (GEP) [44], and Stochastic Gradient Descent (SGD) [60] are adopted in a few articles. In addition, the stacking models combining several ML models achieved favorable results by taking advantage of each model and avoiding its shortcomings. The SG classification model integrates SVM, RF, Adaboost, and GBM to avoid erroneous results in each model [61]. The stacking model combining GBM, ANN, and the Generalized Linear Model (GLM) provides reliable WQI estimations even with limited data. Only the DL models such as DNN and CNN can achieve a similar performance [62]. The stacking model for WQI prediction consisting of RF, NB, and the LR meta-classifier obtains the highest accuracy among LR, NB, MLPNN, and RF [63].

Table 1. Recent ML applications on P2P water quality models for different indicators.

Target	Study Area	Dataset	Input Feature	ML Model	Article
WQI/WQC	Luoyang River, China	390 weekly data between 2014 and 2021	pH, DO, COD, NH_3-N, TP	BPNN	[49]
	M River, China	300 $^{\Delta}$ monthly data in October 2011–August 2020	pH, T_w, DO, TN, TP, NH_3-N, Nitrate Nitrogen (NO_3-N), COD_{Mn}, Cl^-, SO_4^{2-}, Fluoride (F^-), Manganese (Mn), *F. coli*, Iron (Fe)	RF	[43]
	Western Bug River, Ukraine	977 monthly data on 1 January 2016–31 December 2021 from 8 stations	BOD, SS, DO, SO_4^{2-}, NH_4^+, Cl^-, Phosphate (PO_4^{3-}), NO_2-N, NO_3-N	MLPNN	[64]
WQI/WQC	La Buong River, Vietnam	200 $^{\Delta}$ bimonthly data in 2010–2017 from 4 stations	T_w, pH, COD, BOD, DO, Turbidity, TSS, Coliform, Ammonium Nitrogen (NH_4-N), Phosphate	RF, MLPNN, XGBoost *, GBM, LightGBM, AdaBoost, CNN	[47]
	Bhavani river, India	320 seasonal data on 1 January 2016–31 December 2020 from 11 stations	Alkalinity, Hardness, Phosphate, Potassium (K^-), Coliform, EC, BOD, DO, NO_3-N, Cl^-, SO_4^{2-}, Na^+, F^-, and so on (27 features)	RF, MLPNN *, SVM, DT, NB	[65]
	Saf-Saf River Basin, Algeria	70 seasonal data in wet and dry seasons from 35 sites	T_w, pH, BOD, COD, DO, Oxygen Saturation (OS), EC, TDS, SS, Turbidity, Nitrite Nitrogen (NO_2-N), NO_3-N, NH_4-N, Cl^-, PO_4^{3-}	RF	[66]
	Langat River Basin, Malaysia	560 seasonal data in January 2012–December 2016 from 14 stations	DO, BOD, COD, SS, pH, NH_3-N	MLPNN, SVM *, DT	[50]
	Nainital Lake, Uttarakhand	530 daily data in 2018–2019	pH, TSS, Turbidity	RF *, SVM *, MLPNN, DT, SGD *, KNN, GBM	[60]
	Almanor Lake, USA	108 $^{\Delta}$ monthly data in 2014–2016	TDS, T_w, pH, Turbidity, TP, NO_3-N, BOD, Coliforms, DO	MLPNN *, GBM *	[67]
	Gulshan Lake, Bangladesh	108 monthly data in 2016 from 9 sites	pH, SS, EC, TDS, Turbidity, DO, COD, Cl^-, Alkalinity	RF, SVM *, GBM *, AdaBoost	[51]
	Erhai Lake, China	1056 monthly data in January 2016–December 2021 from 11 sites	pH, T_w, DO, Transparency (Tran), TN, TP, NH_3-N, COD, COD_{Mn}, BOD, Chl-a	XGBoost	[68]
	Loktak Lake, India	30 seasonal data in the post-monsoon season from 30 sites	pH, T_w, EC, TDS, Turbidity, DO, BOD, COD, NO_3-N	CNN, DNN, GBM-GLM-NN *	[62]
	Loktak Lake, India	60 $^{\Delta}$ seasonal data after the monsoon season from 60 sites	NO_3-N, COD, BOD, DO, Turbidity, T_w, pH, EC, TDS	RF *, GBM, DNN	[69]
	Cork Harbour, Ireland	100 $^{\Delta}$ seasonal data in 2020 from 29 sites	Salinity, T_w, pH, Tran, DO, BOD, Total Organic Nitrogen (TON), NH_3-N, Molybdate Reactive Phosphorus (MRP), Chl-a	RF *, DT, KNN, XGBoost *, ExT, SVM	[46]
	Cork Harbour, Ireland	100 $^{\Delta}$ seasonal data in October 2019–September 2020 from 29 stations	Salinity, T_w, pH, Tran, DO, BOD, TON, NH_3-N, dissolved inorganic nitrogen (DIN), MRP, Chl-a	XGBoost	[70]
	Cork Harbour, Ireland	100 $^{\Delta}$ seasonal data in 2019 from 29 stations	T_w, DO, pH, TON, NH_3-N, MRP, BOD, Tran, Chl-a, DIN	MLPNN, SVM, DT, XGBoost *, NB, KNN	[54]

Table 1. *Cont.*

Target	Study Area	Dataset	Input Feature	ML Model	Article
WQI/WQC	Tongi Canal, Bangladesh	90 seasonal data in June 2022 from 30 sites	T_w, DO, EC, Lead (Pb), Cadmium (Cd), Copper (Cu), Fe	GPR	[57]
DO	Luoyang River, China	390 weekly data in 2014–2021	WQC, pH, COD, NH_3-N, TP	BPNN	[49]
	Bullfrog River, USA	200 $^{\Delta}$ monthly data in 1998–2016	Land Use/Land Cover, Meteorological, Soil Animal Statistics, Antecedent Conditions, Water Quality (i.e., pH, Turbidity)	RF, MLPNN, SVM, XGBoost *, RF-XGBoost *	[71]
	Adige River, Italy	3700 $^{\Delta}$ seasonal data in 1994–2014 from 45 sites	Arsenic (As), EC, T_w, DO, Cl^-, SO_4^{2-}	RF, DNN *	[72]
	Lake Erie, USA	300 $^{\Delta}$ biweekly data in 2006–2020	EC, TDS, pH, PO_4^{3-}, NH_3-N	CART *, CHAID	[58]
	Feitsui Reservoir, China	348 monthly data in 29 years from 12 sites	T_w, BOD, Fe, TOC	MLPNN	[73]
	Wadi Dayqah Dam, Oman	34 $^{\Delta}$ sampled data on 15 January 2023 from 34 sites	EC, ORP, pH, T_w, Chl-a	RF, MLPNN, SVR, BMEF *	[56]
EC	Indus River, Pakistan	372 monthly data in 1975–2005	T_w, pH, HCO_3^-, Cl^-, SO_4^{2-}, Na^+, Ca^{2+}, Mg^{2+}	RF, MLPNN, DT *, SVM	[53]
	Indus River, Pakistan	360 monthly data in 1975–2005	T_w, pH, HCO_3^-, Cl^-, SO_4^{2-}, Na^+, Ca^{2+}, Mg^{2+}	MEP	[74]
	Indus River, Pakistan	321 monthly data in 1975–2005	pH, HCO_3^-, Cl^-, SO_4^{2-}, Na^+, Ca^{2+}, Mg^{2+}	RF *, MLPNN, GEP	[44]
	Wadi Dayqah Dam, Oman	34 $^{\Delta}$ sampled data on 15 January 2023 from 34 sites	ORP, pH, T_w, DO, Chl-a	RF, MLPNN, SVR, BMEF *	[56]
	Adige River, Italy	3700 $^{\Delta}$ seasonal data in 1994–2014 from 45 sites	Arsenic (As), EC, T_w, DO, Cl^-, SO_4^{2-}	RF, DNN *	[72]
TDS	Indus River, Pakistan	321 monthly data in 1975–2005	pH, HCO_3^-, Cl^-, SO_4^{2-}, Na^+, Ca^{2+}, Mg^{2+}	RF *, MLPNN, GEP	[44]
	Karoon River, Iran	600 $^{\Delta}$ monthly data in 1968–2020	Cl^-, SO_4^{2-}, Na^+, Ca^{2+}, Mg^{2+}	RF *, REPT	[59]
	Indus River, Pakistan	360 monthly data in 1975–2005	T_w, pH, HCO_3^-, Cl^-, SO_4^{2-}, Na^+, Ca^{2+}, Mg^{2+}	MEP	[74]
	Indus River, Pakistan	360 monthly data in 1975–2005	T_w, pH, HCO_3^-, Cl^-, SO_4^{2-}, Na^+, Ca^{2+}, Mg^{2+}	RF *, MEP	[45]

Table 1. *Cont.*

Target	Study Area	Dataset	Input Feature	ML Model	Article
TP	Euiam Lake, Korea	546 weekly data in March 2007–November 2020	T_w, EC, pH, COD, BOD, SS, TN, Oxygen Demand (OD), and so on (40 features)	RF, XGBoost, DNN, LSTM *	[75]
	Feitsui Reservoir, China	300 $^{\Delta}$ monthly data in 1986–2014	NH_3-N, NO_3-N, TDS, Hardness, Mg^{2+}	RF, MLPNN *, SVM	[76]
pH	Luoyang River, China	390 weekly data in 2014–2021	WQC, DO, COD, NH_3-N, TP	BPNN	[49]
	Water System in Georgia, USA	26,085 daily data in 2016–2018 from 37 stations	Min, mean, max of DO, T_w, SC, min and max of pH	RF, MLPNN, XGBoost, LightGBM *	[55]
T_w	Wadi Dayqah Dam, Oman	34 $^{\Delta}$ sampled data on 15 January 2023 from 34 sites	EC, ORP, T_w, DO, Chl-a	RF, MLPNN, SVR, BMEF *	[56]
	Adige River, Italy	3700 $^{\Delta}$ seasonal data in 1994–2014 from 45 sites	Arsenic (As), EC, T_w, DO, Cl^-, SO_4^{2-}	RF, DNN *	[72]
Chl-a	Wadi Dayqah Dam, Oman	34 $^{\Delta}$ sampled data on 15 January 2023 from 34 sites	EC, ORP, pH, T_w, DO	RF, MLPNN, SVR, BMEF *	[56]
	Geum River, Korea	126 $^{\Delta}$ monthly data in October 2017–March 2021 from 3 stations	T_w, EC, DO, Chl-a, pH, Turbidity	XGBoost	[77]
Bacteria	Rejeka Bay, Croatia	1800 $^{\Delta}$ biweekly data in 2009–2020 bathing season from 15 sites	Salinity, Global Horizontal Irradiance (GHI), Surface Pressure, T_{air}, Water Level, Relative Humidity (RH), and so on (33 features)	RF, Catboost *	[78]

Notes: * The best-performing model. $^{\Delta}$ Estimated number according to intervals and time windows.

3.2. ML Applications on S2P Models

Datasets with daily and hourly intervals and medium to large data sizes (ranging from 300 to 39,752) are often considered for S2P model development, as shown in Table 2. The S2P model takes time series data as inputs. DO, NH_3-N, TP, TN, and pH are the most commonly used water quality indicators, employed in 62%, 38%, 27%, 19%, and 19% of the relevant articles, respectively. As shown in Figure 2, although MLPNN, RNN, RF, and XGBoost are considered, the most widely adopted DL model out of the S2P models is the LSTM-based model (65.7%), followed by the GRU-based model (8.6%).

1. The LSTM, as an improved RNN, possesses inner loops within hidden layers to memorize previous data and transfer them to the current output. Three special gates are designed to solve the problems of either an exploding or vanishing gradient over 10 time steps [64]. The LSTM is adopted to predict the current DO, NH_3-N, TP, and TN from the previous water quality measurements of DO, NH_3-N, TP, TN, T_w, turbidity, EC, and COD_{Mn} [65]. The value of R^2 = 0.88 represents the highest performance compared with SVM, MLP, RF, and XGBoost. The LSTM is combined with a CNN layer (CNN-LSTM) to extract and filter the main features of inputs to predict COD with a 4 h interval dataset consisting of COD, TN, TP, COD_{Mn}, NH_3-N, and pH (Figure 3). The MAE and RMSE of the CNN-LSTM are 29.1% and 20.7% lower than those of LSTM [66]. The Bi-directional LSTM (BiLSTM) model is an enhanced version of LSTM that uses both forward and backward hidden layers to capture the past and future information at specific time steps (Figure 4). The BiLSTM is utilized to predict the BOD and COD in the next month based on the previous 13-month data of T_w, DO, pH, COD, BOD, total coliform, *F. coli*, EC, and total Kjeldahl nitrogen (TKN) [67]. Its performance is significantly superior to CNN-LSTM and LSTM, with its mean square error (MSE), RMSE, MAE, and Mean Average Percentage Error (MAPE) equal to 0.11, 0.11, 0.12, and 18.22% for COD, and 0.015, 0.12, 0.12, and 20.32% for BOD, respectively. The Attention-based BiLSTM (AT-BiSLTM) is developed to predict the hourly DO with previous T_w, pH, DO, and EC values and obtained the best performance with ranges of RMSE = [0.10–0.16], MAE = [0.057–0.12], and R^2 = [0.93–0.94] for four stations in the Yellow River, China [68]. The AT mechanism emulates the allocation of human attention to enhance the neural network to variably focus on distinct parts of inputs [69]. Thus, it becomes a pivotal tool for augmenting the expressiveness and generalization of the model.

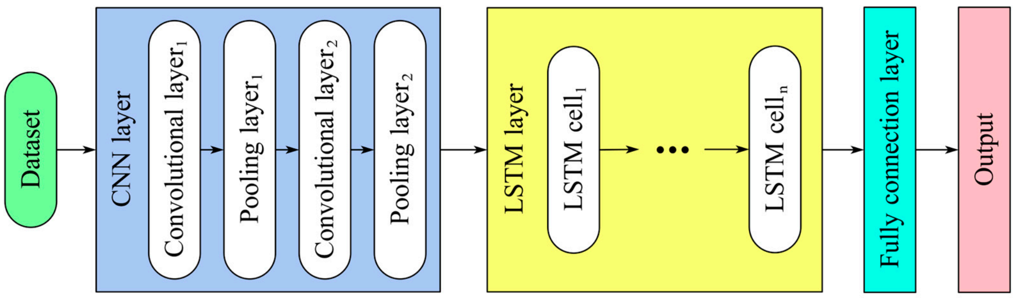

Figure 3. The architecture of the CNN-LSTM model.

2. The GRU is a variant of RNN with only two gated units. It is simpler and can train data faster than the LSTM [70]. The GRU is adopted to predict hourly DO levels in Han River, Korea, with 16 measurements including water quality, hydrological, and meteorological data, as shown in Table 2. It outperforms ANN and LSTM with the best R^2 = 0.98 and RMSE = 0.14 [71]. A bi-directional GRU with an AT layer (AT-BiGRU) is utilized to address the temporal dependencies of time series data in the prediction of multivariate water quality indicators [72]. A graph learning layer and an auto-regression module are considered in the AT-BiGRU to capture the hidden relationships of input features and to generate the final prediction, respectively. The

combined AT-BiGRU model outperforms LSTM, RNN-LSTM, and CNN-LSTM with all four datasets [72].

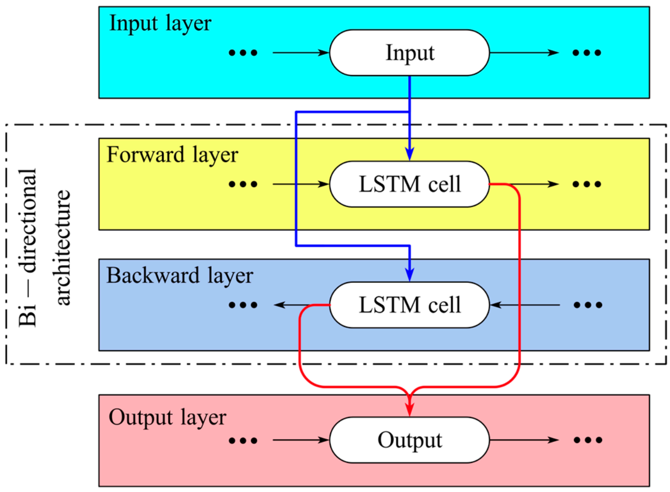

Figure 4. The bi-directional architecture of the BiLSTM model.

Table 2. Recent ML applications on S2P water quality models for different indicators.

Target	Study Area	Dataset	Input Feature	ML Model	Article
DO	Yellow River, China	39,026 hourly data in December 2017–December 2019 from 4 stations	T_w, pH, DO, EC	RNN, LSTM, BiLSTM, AT-BiLSTM *, GRU	[68]
	Burnett River, Australia	39,752 hourly data in January 2015–January 2020	T_w, DO, pH, EC, Chl-a, Turbidity	LSTM, AT-LSTM *	[73]
DO	Lijiang River, China	30,000 $^{\Delta}$ hourly data in April 2016–December 2019 from 15 stations	EC, pH, Turbidity, Discharge (Q), T_w, DO, COD_{Mn}, NH_3-N, TP, TN, Precipitation (P), Land Use, Vegetation, Slope	LSTM	[74]
	Han River, Korea	9870 hourly data in 2017–2019	EC, pH, Turbidity, T_w, SS, DO, TOC, TP, TN, Q, Water Level, Tide Level, P, Cumulative P, Humidity, Sunlight	Classifier: RF *, DT, XGBoost; Regressor: MLPNN, LSTM, GRU *	[71]
	Pearl River, China	5058 data every 4 h on 8 November 2020–28 February 2023	NH_3-N, COD_{Mn}, T_w, pH, DO, TP, TN, EC, Turbidity	SVM, MLPNN, LSTM *	[65]
	Ganjiang River, China	738 data every 4 honn 1 October 2021–31 January 2022	T_w, pH, DO, NH_3-N, COD_{Mn}, TP, TN, Turbidity, EC	LSTM	[75]
	Li River and Liu River, China	865 daily data on 19 April 2021–31 August 2023	DO, T_{air}, T_w, pH, COD_{Mn}, NH_3-N, P, Q, RH, Evaporation, Hours of Sunshine, Wind Velocity	AT-GRU	[76]
	Liao River, China	2900 $^{\Delta}$ daily data on 1 January 2013–31 December 2020	DO	RF, SVM *	[77]
	Surma River, Bangladesh	1000 $^{\Delta}$ daily data in January 2017–December 2019	Humidity, T_w, T_{air}, P, TDS, pH, Turbidity, DO	MARS	[78]
	Yellow River, China	365 weekly data in 2010–2016	DO	Deep Extreme Learning Machine (DELM)	[79]
	Tai Lake, China	5095 data every 4 h on 1 November 2020–28 February 2023 from 6 stations	T_w, T_{air}, Sea Level Pressure, Ground Pressure, AT, Surface T, Dew Point T, DO	AT-CNN-BiLSTM	[80]
	Poyang Lake, China	882 daily data on 1 October 2017–29 February 2020	DO, NH_3-N, pH	BiLSTM	[81]
	Taihu Lake, China	35,000 $^{\Delta}$ hourly data in October 2017–October 2021	DO, EC, pH, Turbidity	AT-BiGRU	[72]

Table 2. *Cont.*

Target	Study Area	Dataset	Input Feature	ML Model	Article
NH_3-N	Lijiang River, China	30,000 $^{\Delta}$ hourly data in April 2016–December 2019 from 15 stations	EC, pH, Turbidity, Discharge (Q), T_w, DO, COD_{Mn}, NH_3-N, TP, TN, Precipitation (P), Land Use, Vegetation, Slope	LSTM	[74]
	Pearl River, China	5058 data every 4 h on 8 November 2020–28 February 2023	NH_3-N, COD_{Mn}, T_w, pH, DO, TP, TN, EC, Turbidity	SVM, MLPNN, LSTM *	[65]
	Yellow River, China	300 weekly data in January 2012–September 2017	pH, DO, COD_{Mn}, NH_3-N	VARMA-BiLSTM	[82]
	Poyang Lake, China	882 daily data on 1 October 2017–29 February 2020	DO, NH_3-N, pH	BiLSTM	[81]
	Xili Reservoir, China	943 daily data on 1 January 2019–31 July 2021	TN, NH_3-N, TP, pH, DO, EC, Turbidity, T_w, COD	LSTM	[83]
TP	Lijiang River, China	30,000 $^{\Delta}$ hourly data in April 2016–December 2019 from 15 stations	EC, pH, Turbidity, Discharge (Q), T_w, DO, COD_{Mn}, NH_3-N, TP, TN, Precipitation (P), Land Use, Vegetation, Slope	LSTM	[74]
TP	Pearl River, China	5058 data every 4 h on 8 November 2020–28 February 2023	NH_3-N, COD_{Mn}, T_w, pH, DO, TP, TN, EC, Turbidity	SVM, MLPNN, LSTM *	[65]
	Jinjiang River, China	443 weekly data on 7 January 2013–21 June 2021	DO, NH_3-N, COD_{Mn}, TP	LSTM	[84]
	Haihe River Basin, China	4000 $^{\Delta}$ monthly data in January 2010–September 2014 from 76 stations	BOD, DO, NH_3-N, TP, pH, COD_{Mn}, COD_{Cr}	LSTM	[85]
	Poyang Lake, China	1096 daily data on 1 January 2017–1 January 2020	DO, EC, Turbidity, NH_3-N, pH, T_w, P, Water Level	LSTM	[86]
	Xili Reservoir, China	943 daily data on 1 January 2019–31 July 2021	TN, NH_3-N, TP, pH, DO, EC, Turbidity, T_w, COD	LSTM	[83]
TN	Pearl River, China	5058 data every 4 h on 8 November 2020–28 February 2023	NH_3-N, COD_{Mn}, T_w, pH, DO, TP, TN, EC, Turbidity	SVM, MLPNN, LSTM *	[65]
	Juhe River, China	1086 data every 4 h on 18 August 2018–18 March 2019	T_w, pH, DO, EC, Turbidity, COD_{Mn}, TP	CNN-LSTM	[87]
	Poyang Lake, China	1096 daily data on 1 January 2017–1 January 2020	DO, EC, Turbidity, NH_3-N, pH, T_w, P, Water Level	LSTM	[86]
	Xili Reservoir, China	943 daily data on 1 January 2019–31 July 2021	TN, NH_3-N, TP, pH, DO, EC, Turbidity, T_w, COD	LSTM	[83]
pH	Poyang Lake, China	882 daily data on 1 October 2017–29 February 2020	DO, NH_3-N, pH	BiLSTM	[81]
	Xili Reservoir, China	943 daily data on 1 January 2019–31 July 2021	TN, NH_3-N, TP, pH, DO, EC, Turbidity, T_w, COD	LSTM	[83]
COD_{Mn}	Lijiang River, China	30,000 $^{\Delta}$ hourly data in April 2016–December 2019 from 15 stations	EC, pH, Turbidity, Discharge (Q), T_w, DO, COD_{Mn}, NH_3-N, TP, TN, Precipitation (P), Land Use, Vegetation, Slope	LSTM	[74]
	Jinjiang River, China	443 weekly data on 7 January 2013–21 June 2021	DO, NH_3-N, COD_{Mn}, TP	LSTM	[84]
	Haihe River Basin, China	4000 $^{\Delta}$ monthly data in January 2010–September 2014 from 76 stations	BOD, DO, NH_3-N, TP, pH, COD_{Mn}, COD_{Cr}	LSTM	[85]
COD	Yellow River, China	1000 $^{\Delta}$ data every 4 h in January 2018–December 2020	COD, TN, TP, NH_3-N, COD_{Mn}, pH	CNN-LSTM	[66]
Turbidity	Hong Kong's Coasts, China	18,000 $^{\Delta}$ monthly data on 1 January 1997–1 December 2016 from 76 sites	NH_3-N, pH, SS, P, TP, DO, Mean T_{air}, Pressure	MLPNN, LSTM-RNN *	[88]
Salinity	Salton Sea (Lake), USA	300 $^{\Delta}$ seasonal data in 2003–2008 from 14 sites	T_w, EC, DO	LSTM	[89]

Notes: * The best-performing model. $^{\Delta}$ Estimated number according to intervals and time windows.

3.3. ML Applications on S2S Models

S2S models are often developed with weekly, daily, and hourly datasets and with data sizes ranging from 730 to 70,000, as shown in Table 3. DO and NH_3-N are the most commonly considered targets, followed by T_w, pH, COD_{Mn}, and WQI. According to the proportions shown in Figure 2, 90% of S2S models are implemented with advanced DL models.

1. LSTM-based models are frequently selected for S2S models because of their strong long-term memory capabilities. The LSTM model forecasts the following 5-day WQI with the previous 30-day NH_3-N, BOD, COD, DO, pH, and SS in Klang River, Malaysia. The MAE = 1.03, RMSE = 1.24, mean bias error (MBE) = 0.14, and Kling–Gupta efficiency (KGE) = 0.68 are better than those of the GRU [90]. The daily NH_3-N in Haihe River, China, is predicted by combining the LSTM with a dual-stage AT, namely the DA-LSTM [91]. It forecasts daily NH_3-N within the following 1 day and 7 days, resulting in RMSE = 0.054, MAE = 0.038, and Nash–Sutcliffe Efficiency coefficient (NSE) = 0.99 for 1 day, and RMSE = 0.29, MAE = 0.19, and NSE = 0.94 for 7 days, respectively. The Encoder–Decoder (ED) structure combined with the AT-BiLSTM (Figure 5) can be employed to effectively forecast multi-step time series values [92]. It is applied to forecast hourly NH_3-N and DO within the following 1 to 5 days with two datasets from rivers in the Beijing–Tianjin–Hebei region, China, and the Merced River, USA, respectively. The result of the NH_3-N model has MAPE = 15.64%, RMSE = 0.04, MAE = 0.03, and R^2 = 0.96. The larger dataset of DO improves the model's performance compared with that of the NH_3-N model and obtains a notably lower value of MAPE = 4.02%, with RMSE = 0.04, MAE = 0.03, and R^2 = 0.94 [92].

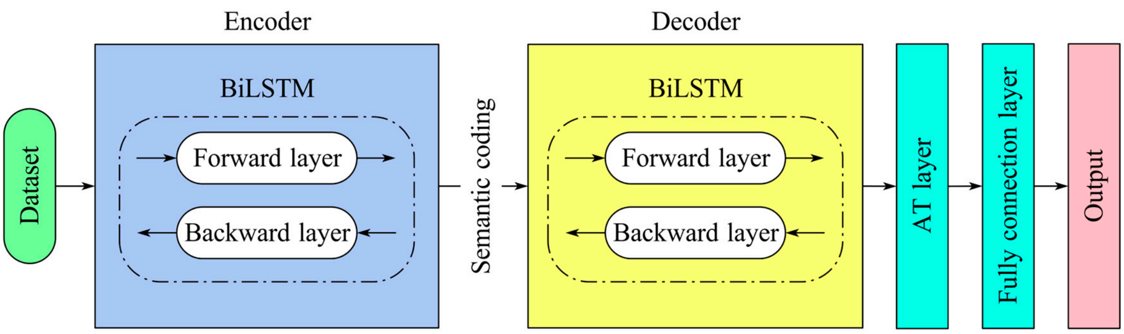

Figure 5. The architecture of the AT-BiLSTM model with ED structure.

2. GRU-based models achieve the second highest proportion of use for S2S models (Figure 2) because they are easier than LSTM-based models to implement with acceptable performance. The BiGRU model results in a reasonable NSE in the range of 0.83–0.98 for 7-day DO and NH_3-N forecasting at three stations in Poyang Lake, China [93]. An AT-GRU model is applied to the multi-step-ahead prediction of daily DO with two datasets from Li River and Liu River, China. Notably higher performances are demonstrated with RMSE = [0.084–0.097], MAE = [0.069–0.088], and NSE = [0.93–0.97] for three steps ahead, and RMSE = [0.21–0.28], MAE = [0.17–0.27], and NSE = [0.43–0.81] for five steps ahead, respectively [76].
3. Transformer is developed with the ED structure and is widely used for natural language processing tasks such as machine translation [69]. It has been verified in forecasting runoff to capture time-dependent patterns [94]. The Transformer is utilized to forecast weekly pH, DO, NH_3-N, and COD_{Mn} in the next 4 weeks with the data from the previous 52 weeks (12 months) [95]. It obtains a lower MSE and MAE compared to MLPNN, CNN, LSTM, Bi-LSTM, and AT-CNN-LSTM, with its MSE and MAE equal to 0.86 and 0.72, 0.64 and 0.61, 1.02 and 0.77, and 0.90 and 0.68 for pH, DO, NH_3-N, and COD_{Mn}, respectively.

Table 3. Recent ML applications on S2S water quality models for different indicators.

Target	Study Area	Dataset	Input Feature	ML Model	Article
DO	Burnett River, Australia	39,752 hourly data in January 2015–January 2020	T_w, DO, pH, EC, Chl-a, Turbidity	LSTM, AT-LSTM *	[73]
	Merced River, USA	70,000 hourly data in September 2018–December 2021	DO, month, day, hour	AT-BiLSTM-ED	[92]
DO	Li River and Liu River, China	865 daily data on 19 April 2021–31 August 2023	DO, T_{air}, T_w, pH, COD_{Mn}, NH_3-N, P, Evaporation, Hours of Sunshine, Q, RH, Wind velocity	AT-GRU	[76]
	Poyang Lake, China	900 $^{\Delta}$ daily data on 1 October 2017–30 April 2020	DO	BiGRU	[93]
	Major rivers and lakes in China	13,200 weekly data in 2013–2016 from 120 stations	pH, DO, NH_3-N, COD_{Mn}	Transformer	[95]
NH_3-N	Haihe River, China	4000 $^{\Delta}$ daily data in 2004–2015	NH_3–N, pH, DO, Wind direction speed	AT-LSTM	[91]
	Rivers in the Beijing-Tianjin-Hebei region, China	7100 seasonal data in May 2012–August 2020	NH_3-N	AT-BiLSTM-ED	[92]
	Poyang Lake, China	900 $^{\Delta}$ daily data on 1 October 2017–30 April 2020	NH_3-N	BiGRU	[93]
	Major rivers and lakes in China	13,200 weekly data in 2013–2016 from 120 stations	pH, DO, NH_3-N, COD_{Mn}	Transformer	[95]
T_w	8 lakes in Poland	10,000 $^{\Delta}$ daily data in 1987–2016	Lake surface T_w, T_{air}	MLPNN-RF	[96]
pH COD_{Mn}	Major rivers and lakes in China	13,200 weekly data in 2013–2016 from 120 stations	pH, DO, NH_3-N, COD_{Mn}	Transformer	[95]
WQI	Klang River, Malaysia	730 daily data on 1 January 2019–31 December 2020	NH_3-N, BOD, COD, DO, pH, SS	LSTM *, GRU	[90]

Notes: * The best-performing model. $^{\Delta}$ Estimated number according to intervals and time windows.

4. Analysis of Strategies to Improve Model Performance

ML water quality models often struggle to capture the full relationships of the interconnecting water quality measurements in dynamic water systems due to data noise and non-stationarity [97]. Model performance is heavily dependent on hyperparameter selection [98]. Some DL models may suffer from insufficient training data in many instances [99]. To improve their performance and exploit their potential capabilities, the following three strategies are reported in the literature.

4.1. Feature Engineering

Robust feature engineering methods are helpful in processing time series water quality data when exploring relationships [80]. Two mainstream algorithms adopted are Discrete Wavelet Transform (DWT) [100,101] and Mode Decomposition (MD) [102]. The main processes are (1) decomposing time series data into several signals according to different algorithms, (2) replacing original data with signals as inputs to train models, and (3) generating final target predictions by integrating outputs. Removing the short-term random disturbance noise from the complex variation in the original time series data makes ML models learn useful and important information more efficiently.

DWT can decompose time series data into numerous scale-specific time series with a simpler computation and development process [100,101]. The DO model developed with ANN-LSTM and DWT achieves a perfect performance with error indices close to 0 [84]. In addition, maximal overlap DWT (MODWT) has an enhanced ability to extract additional information and is applicable to any sample size [78,103]. The MODWT combined with boundary correction is utilized to improve the NSE of the DA-LSTM from 0.76 to 0.94 in 7-day NH_3-N forecasting [91]. In addition, Synchro-squeezed Wavelet Transform (SWT) is applied to improve the DO model performance with R^2 = 0.97, showing greater noise reduction effectiveness than several MD approaches [79].

MD methods segregate time series data into a series of intrinsic mode functions (IMFs) with distinct central frequencies, extracting features pertinent to each mode [102]. Higher correlations of IMF1, IMF2, and IMF3 with features than those of the original data are reported [80]. Variational Mode Decomposition (VMD) as a widely used algorithm has shown excellent performance in decomposing unstable time series data with high complexity and strong nonlinearity into relatively stable subsequences [104]. It is used to improve the R^2 of the RF from 0.83 to 0.96 in TDS predictions [59]. VMD is utilized for DO models, decreasing the RMSE of the LSTM to 0.29 in Ganjiang River and the RMSE of the GRU to 0.05 in Li River [75,76]. In addition, VMD effectively separates high-frequency noises in the WQI model and obtains better forecasting results than predictions with the Empirical Mode Decomposition (EMD) algorithm [90].

Moreover, Complete Ensemble EMD with Adaptive Noise (CEEMDAN) is proposed to improve the decomposition efficiency by eliminating partial noise amplitudes [105]. It is developed in the LSTM for TN, TP, NH_3-N, and pH predictions, significantly increasing the value of R^2 from 0.83, 0.71, 0.72, and 0.52 to 0.96, 0.95, 0.93, and 0.87 [83]. The improved CEEMDAN is applied for NH_3-H, DO, and pH models to overcome the limitations of mode confusion, reconstruction error, and pseudo-components in EMD, EEMD, and original CEEMDAN [81,106]. Furthermore, Empirical Wavelet Transform (EWT) is more effective in real-time data decomposition-based modeling as it consists of both the mathematical basis of DWT and the adaptability of EMD [107,108]. A BiGRU employing EWT obtains up to a 0.98 NSE for its DO predictions [93].

4.2. Hyperparameter Tuning Methods

The performance of the ML model is affected by some hyperparameters, which must be configured before the training process [21,39]. Hyperparameter tuning determines the set of hyperparameters that yields an optimal model on a given dataset. Some meta-heuristic intelligence optimization algorithms are used to find the best combination of hyperparameters for ML water quality models [72,77,79,80,86,92,109]. Particle Swarm Optimization (PSO), inspired by the activities of bird swarms, can find optimal solutions through cooperation and information sharing with low design costs [110]. It is applied to find the best hyperparameter combinations for the DL model and enhance the model's performance [72]. The Genetic Simulated annealing-based PSO (GSPSO) is utilized to greatly reduce the tedious process of manual parameter setting, achieving the lowest MAPE of 4.02% for DO forecasting (the second lowest MAPE was 10.35%) [92].

The Sparrow Search Algorithm (SSA) is a new type of swarm intelligence optimization algorithm with high accuracy, superior to PSO in some articles [111]. The SWT-DELM model is developed for DO prediction with the SSA to obtain the best R^2 = 0.97 [79]. The Chaos Sparrow Search Algorithm (CSSA) is implemented to automatically calibrate the hyperparameters of the LSTM in TN and TP predictions, performing better than SSA and PSO [86]. The improved SSA (IMSSA), depending on Tent chaotic sequences, outperforms Gray Wolf Optimization (GWO) and the Genetic Algorithm (GA) in terms of convergence accuracy and merit-seeking ability for improvement in the DO model in Liao River, China [77].

The Whale Optimization Algorithm (WOA) has the advantages of fast optimization speed, strong global convergence, and few parameters [112]. It significantly enhances model performance in DO predictions, increasing the R^2 from 0.90 to 0.98 [80]. Moreover, WOA based on the Self-adapting parameter adjustment and Mix mutation strategy (SMWOA) reduces the RMSE from 6.18 to 3.93 in the WQI prediction, showing better performance than a hybrid PSO combined with the Genetic Algorithm (HPSOGA) [109].

4.3. Transfer Learning Methods

Transfer learning is an effective method to solve the insufficient data issue by transferring knowledge across similar domains, features, and distributions [113]. For example, the pre-training DO model based on BiLSTM is developed with a large dataset from the Lake Taihu and uses transfer learning to fine-tune with the small target dataset from another lake near Yixing, China [114]. The results show a drastic increase in the R^2 from 0.38 to 0.79.

The recurrent transfer learning method solves that the pre-trained model is prone to overfitting to the source domain data and reducing the performance in the fine-tuning stage [95]. The Transformer with recurrent transfer learning (TLT) obtains MAE = 0.60, lower than with transfer learning (MAE = 0.81). The transfer learning method has been applied cross-domain to predict DO, pH, NH_3-N, and COD_{Mn} in the Haihe River and lakes in China [95]. The first TLT model built for the Haihe River is pre-trained based on the dataset in lakes, and the second TLT model built for the dataset in lakes is transferred from the Haihe River dataset. The performances of both TLT models pre-trained by cross-domain datasets are better than those of the models without transfer learning [95].

5. Analysis of Metrics for Model Evaluation

More than 30 metrics are applied to evaluate model performance in the reviewed articles, as shown in Table A1. Choosing the correct metrics to produce favorable results is important for model performance evaluation [39]. To conveniently compare model performance, the most used metrics for regression and classification models are summarized and discussed as follows.

1. Metrics for water quality regression models, which can be grouped into three major categories: (1) standard regression, including R^2 and correlation coefficient (r) [115]; (2) dimensionless metrics, such as NSE [116]; (3) error indices, including MSE, RMSE, MAE, and MAPE [116]. The equations for the metrics are presented in Table A2, and the percentages of these metrics used for the reviewed P2P, S2P, and S2S models are illustrated in Figure 2. Other metrics with proportions of less than 5% are not included in this study. R^2 is very informative in regression analysis evaluation [117] and is more often selected (51%) than NSE (22%) and r (13%) (Figure 6). The error indices of RMSE (76%) are the most popular in water quality model evaluation (Figure 6). The reason is that water resource management pays more attention to abnormal and extreme events and RMSE is more sensitive to outliers than MAE [118]. MAE obtains the second highest proportion (61%) and is most frequently used to evaluate S2S models because it better describes the capability of predictive values to follow the trend in observations. MAPE is the most widely used for unit-free error indices due to the advantages of scale independence and interpretability, but it may overestimate or underestimate errors above or below the measurements, respectively [119].
2. Metrics for water quality classification models, which are relatively unified compared to the metrics for regression models. Accuracy is applied to evaluate every ML water quality classification model. Precision, F1-Score, and Sensitivity (also known as Recall and True Positive Rate (TPR)) are used in 92% of the reviewed articles. Accuracy, Precision, and F1-score are sensitive to imbalanced data [120,121], while Sensitivity is suitable for evaluating model performance with both balanced and imbalanced data [122]. Using these metrics can comprehensively evaluate water quality models with imbalanced class distribution [50,123]. Moreover, Specificity (also known as True Negative Rate (TNR)) is selected in 25% of articles to help in evaluating the correctness

of predictions for each class for imbalanced datasets [25,54,71]. Equations of these metrics are presented in Table A3.

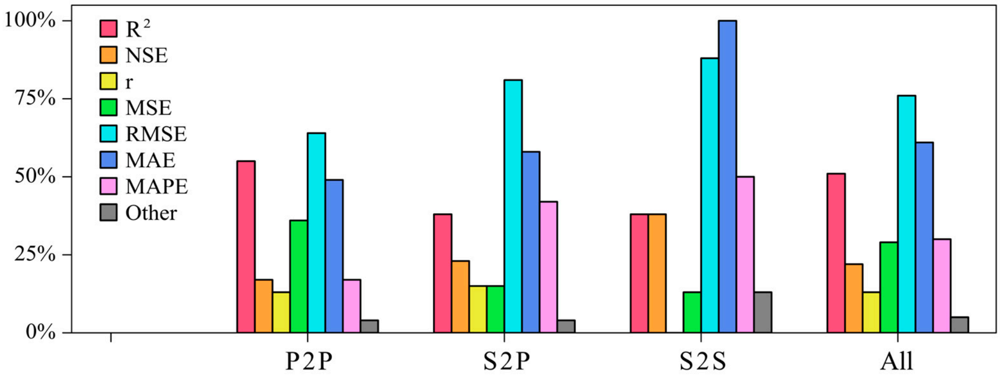

Figure 6. Proportion of metrics used in ML water quality regression models.

6. Discussion

Although ML models provide alternatives in water resource management, the interpretability limitation is still notable. In RF, SVM, ANNs, and DL, individual model parameters have a limited relationship to the physical process and only take meaning in context with other model parameters. What if analysis could be conducted on the fit model to study the physical and biogeochemical processes that impact water quality [124]?

On the other hand, some statistical algorithms have been adopted to interpret the importance of model input features by computing the contribution of each feature to the outputs. Shapley additive explanations (SHAP) is one of the common explanatory frameworks based on the game's theoretically optimal Shapley values [125] and has been used to interpret ML water quality models [126,127]. In addition, LIME [128], Integrated Gradients [129], and DeepLift [130] are useful explanation techniques proposed to interpret ML black-boxed structures. Sensitivity analysis is also a powerful tool to assess the impact of each input feature on the output and figure out the most sensitive inputs [44].

In addition, comparing the models with different combinations of input features is feasible to find out the key features [131,132]. WQI models are developed with different combinations of DO, NH_3-N, TP, TN, T_w, and pH in the Erhai Lake, China, while DO, NH_3-N, TP, and TN are reported more important to build the minimal WQI model [131]. The best DO model in the Fei-Tsui reservoir, China, is developed with T_w, BOD, and Fe, while TOC does not need to be considered [132].

Moreover, the DT has the advantage of an interpretable structure similar to the flow chart. The RF based on DTs also provides feature importance through the computation of the mean decrease in impurity [41]. The GEP [133] can generate the explicit mathematical expression utilizing some of the mathematical functions (i.e., $+$, $-$, $\times$, $\div$, $\ln(x)$, e^x, x^2, x^3, sin, cos). It is more convenient to interpret water quality models by utilizing these ML models.

Although feature importance analysis can provide information to explain the ML water quality model to some extent, it is recommended to apply governing equations under local environmental conditions to understand the relationships among input features. For example, DO and NH_3-N show high importance levels in lakes and rivers in India [48], but Mn, Fe, and *F. coli* are the top important features in M River, China, due to the local mining industry and the leak in the sewer system [43]. These differences in local environments also cause performance degradation, affecting the reusability of ML models [134]. This demonstrates that research targets, domains, and data availability are crucial factors when selecting appropriate models.

Furthermore, the reviewed articles primarily focus on either spatial (i.e., P2P model at multiple sites) or temporal dependencies (i.e., S2P/S2S model at one site), rarely considering the spatiotemporal consistency. On the contrary, traditional process-driven water quality models involving hydrological and geo-spatiotemporal information can build two-dimensional and three-dimensional models for the study areas. Therefore, the combination of process-driven models and machine learning techniques will be one of the future research directions through which to take advantage of traditional and advanced techniques.

7. Conclusions

A total of 78 articles on ML water quality modeling for surface waterbodies were reviewed and classified as P2P, S2P, and S2S models. The widely used ML models were discussed to assist in model selection based on water quality indicators, data availability, and model performance. Meanwhile, the benefits of adopting certain strategies to improve model performance were recognized. Finally, the commonly used metrics were summarized and discussed for comprehensive model evaluation. The main conclusions are the following:

1. Traditional ML models such as RF, MLPNN, and SVM were common for P2P models, while the LSTM-based and GRU-based models were capable of handling time series data for S2P models. More complex DL models including AT-GRU, AT-LSTM, and Transformer were frequently utilized for S2S models to learn interconnected relationships from longer time sequences. These findings could help researchers in the community select the most suitable ML models for specific local databases depending on the water quality indicator, data availability, and model performance, rather than by simply choosing the most advanced ones.
2. Adopting certain strategies to improve ML water quality model performance is explored. Firstly, feature engineering approaches, such as DWT, VMD, and EMD, can decompose unstable time series data into relatively stable subsequences, reducing the noise and complexity of the original datasets. Meta-heuristic intelligence optimization algorithms such as PSO, SSA, WOA, and variants can help in finding the best combination of hyperparameters. Finally, transfer learning methods can solve underfitting problems due to insufficient data, showing the positive effects of knowledge transferring between datasets.
3. The metrics of R^2, NSE, RMSE, MAE, and MAPE were most recommended for regression models, while Accuracy, Precision, F1-score, Sensitivity, and Specificity were suggested for classification models to evaluate their performance.
4. This review provides a perspective on the emerging ML models for water quality prediction. S2S water quality models are rising in number due to rapid DL model development, as well as their high performance and forecasting capability.

Author Contributions: M.H.: Conceptualization, investigation, formal analysis, data curation, writing—original draft, review and editing, and visualization. Q.Q.: Writing—review and editing, supervision, project administration, and funding acquisition. X.L.: Writing—review and editing, and supervision. J.Z.: writing—review and editing, and supervision. J.C.: writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the US Department of Energy, Office of Science's, Biological and Environmental Research Program under Award Number DE-SC0023216, as well as the Center for Resiliency (231016), Lamar University.

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

Appendix A

Table A1. Performance of the best ML water quality models in the reviewed articles.

Target	Model Type	ML Model	Results of Model Performance	Article
WQI/WQC	P2P	RF	R^2 = 0.97, MSE = 1.6, MAE = 0.95, MAPE = 1.35%	[43]
	P2P	RF	RMSE = 2.9, MSE = 7.8, MAE = 2.2	[46]
	P2P	RF	MAE = 0.02, MSE = 0.02, RMSE = 0.12	[60]
	P2P	RF	RMSE = 5.17, MAE = 4.71, r = 0.82	[135]
	P2P	RF	R^2 = 0.9	[136]
	P2P	RF	RMSE = 2.9, MSE = 7.8, MAE = 2.2	[46]
	P2P	MLPNN	Accuracy = 85.11, Precision = 89.01, Sensitivity = 90.92, F1 = 90.91	[137]
	P2P	MLPNN	MAPE = 9.6%; R^2 = 0.96	[138]
	P2P	MLPNN	MAE = 1.91, RMSE = 2.43, MSE = 7.10, R^2 = 0.73; Accuracy = 81, Precision = 63, Recall = 61, F1 = 59	[139]
	P2P	BPNN	MSE = 0.077, MAPE = 2.78	[49]
	P2P	SVM	R^2 = 0.95, RMSE = 4.93, MAE = 4.37; Accuracy = 100, Precision = 100, Sensitivity = 100, F1 = 100	[51]
	P2P	SVM	Accuracy = 96.35, Precision = 91.97, Sensitivity = 84.89, F1 = 87.98	[50]
	P2P	RF, SVM	Accuracy = 98.7, Precision = 98.8, Sensitivity = 98.7, F1 = 98.8	[60]
	P2P	XGBoost	R^2 = 0.82, MSE = 3.43, MAE = 1.46	[131]
	P2P	XGBoost	R^2 = 0.99, RMSE = 0.107	[47]
	P2P	XGBoost	R^2 = 0.91–0.97, RMSE = 3.1–4.4	[140]
	P2P	XGBoost	Accuracy = 100, Precision = 99, Sensitivity = 100, F1 = 99	[54]
	P2P	GBM	MSE = 3.66, RMSE = 4.27, R^2 = 3.90	[137]
	P2P	GBM-GLM-NN	MSE = 25.77, RMSE = 5.07, MAE = 3.5, R^2 = 0.98	[62]
	P2P	GPR	R^2 = 1.0, NSE = 1.0	[57]
	S2S	LSTM	MAPE = 1.92%, MAE = 1.03, RMSE = 1.24	[90]
DO	P2P	MLPNN	r = 0.99, R^2 = 0.98, RMSE = 0.14	[132]
	P2P	BPNN	MSE = 0.275, MAPE = 5.03	[49]
	P2P	BMEF	NSE = 0.95, R^2 = 0.94, RMSE = 0.24, MAE = 0.27	[56]
	P2P	XGBoost, RF-XGBoost	NSE = 0.94–0.96	[141]
	P2P	CART	214 cases differ by less than 2 units from the observed data	[58]
	P2P	DNN	NSE = 0.60, RMSE = 0.78	[142]
	S2P	RF	ACC 0.99, Sensitivity = 0.99, Specificity = 1	[71]
	S2P	SVM	RMSE = 0.96, MAE = 0.60, MAPE = 5.06%, R^2 = 0.99	[77]
	S2P	LSTM	R^2 = 0.88, MSE = 3.36, RMSE = 1.83, NSE = 0.88	[65]
	S2P	LSTM	RMSE = 0.26, MAE = 0.23	[74]
	S2P	LSTM	MSE = 0.003, RMSE = 0.05, MAE 0.04, MAPE = 0.4%, R^2 = 0.98	[75]
	S2P	AT-LSTM	RMSE = 0.13, MAE = 0.09, R^2 = 0.95	[73]
	S2P	BiLSTM	MAE = 0.02–0.14, RMSE = 0.03–0.16, MAPE = 0.25–1.76%	[81]
	S2P	AT-BiLSTM	RMSE = 0.10–0.16, MAE = 0.06–0.12, R^2 = 0.93–0.97	[68]
	S2P	AT-CNN-BiLSTM	MAPE = 0.02–0.03, MAE = 0.15–0.17, RMSE = 0.18–0.23, R^2 = 0.99–0.97	[80]
	S2P	GRU	RMSE = 0.14–0.51, R^2 = 0.94–0.98	[71]

Table A1. *Cont.*

Target	Model Type	ML Model	Results of Model Performance	Article
DO	S2P	AT-GRU	RMSE = 0.05, MAE = 0.04, NSE = 0.98	[76]
	S2P	AT-BiGRU	Root Relative Squared Error (RSE) = 0.058–0.070, r = 0.85–0.91	[72]
	S2P	MARS	r = 0.98, MAE = 0.09, RMSE = 0.12	[78]
	S2P	DELM	MAE = 0.13–0.18, MAPE = 1.45–2.58%, RMSE = 0.17–0.26, R^2 = 0.97–0.99, NSE = 0.97–0.99	[79]
	S2S	AT-LSTM	RMSE = 0.20–0.41, MAE = 0.15–0.31, R^2 = 0.89–0.54	[73]
	S2S	AT-BiLSTM	MAPE = 4.02, MAE = 0.03, RMSE = 0.04, R^2 = 0.94	[92]
	S2S	AT-GRU	RMSE = 0.084–0.28, MAE = 0.069–0.27, NSE = 0.43–0.97	[76]
	S2S	BiGRU	MAE = 0.02–0.09, RMSE = 0.03–0.12, MAPE = 0.29–1.25%, NSE = 0.94–0.98	[93]
	S2S	Transformer	MSE = 0.48, MAE = 0.50	[95]
EC	P2P	RF	R^2 = 0.93, RMSE = 3.52, NSE = 0.93, MAE = 2.5	[44]
	P2P	DT	R^2 = 0.93, RMSE = 18.3, MAE = 13.6	[53]
	P2P	BMEF	NSE = 0.91, R^2 = 0.91, RMSE = 0.66, MAE = 0.5	[56]
	P2P	DNN	NSE = 0.76, RMSE = 33.08	[142]
	P2P	MEP	R^2 = 0.97, MAPE = 6.12, MAE = 11.22, RMSE = 16.43	[143]
TP	P2P	MLPNN	RMSE = 1.20, MAE = 0.86, R^2 = 0.98, MSE = 1.44, r = 0.99	[144]
	P2P	XGBoost, RF-XGBoost	NSE = 0.8–0.95	[141]
	P2P	LSTM	r = 0.92, R^2 = 0.85, MSE = 0.006, MAE = 0.05, RMSE = 0.07, MAPE = 24.57%	[145]
	S2P	LSTM	R^2 = 0.77, MSE = 5.68, RMSE = 2.33, NSE = 0.76	[65]
	S2P	LSTM	RMSE = 0.005, MAE = 0.004	[74]
	S2P	LSTM	R^2 = 0.95, RMSE = 0.003, MAE = 0.002	[83]
	S2P	LSTM	RMSE = 0.03, MAPE = 0.24%, MAE = 0.01	[84]
	S2P	LSTM	NSE = 0.71	[85]
	S2P	LSTM	MAE = 0.005, MAPE = 10.83%, RMSE = 0.006, NSE = 0.92	[86]
TDS	P2P	RF	R^2 = 0.93, RMSE = 3.1, NSE = 0.91, MAE = 5.10	[44]
	P2P	RF	R^2 = 0.99, RMSE = 6.99, MAE = 5.45, MAPE = 3.34%	[45]
	P2P	RF	R^2 = 0.96, RMSE = 74.42, NSE = 0.96, MAE = 57.33	[59]
	P2P	MEP	R^2 = 0.97, MAPE = 5.54%, MAE = 8.27, RMSE = 11.36	[143]
pH	P2P	BPNN	MSE = 0.014, MAPE = 1.2%	[49]
	P2P	LightGBM	RMSE = 10.84–14.42, MAPE = 9.22–14.54%	[55]
	S2P	LSTM	R^2 = 0.87, RMSE = 0.20, MAE = 0.12, MRE = 0.02	[83]
	S2P	BiLSTM	MAE = 0.07–0.18, RMSE = 0.09–0.23, MAPE = 1.01–2.55	[81]
	S2S	Transformer	MSE = 0.70, MAE = 0.62	[95]
T_w	P2P	BMEF	NSE = 0.96, R^2 = 0.93, RMSE = 2.2, MAE = 0.75	[56]
	P2P	DNN	NSE = 0.89, RMSE = 1.35	[142]
	S2S	MLPNN-RF	R^2 = 0.86–0.99, RMSE = 0.36–2.67, MAE = 0.23–2.17, MAPE = 3.21–147.22%	[96]
Chl-a	P2P	BMEF	NSE = 0.91, R^2 = 0.98, RMSE = 0.3, MAE = 0.25	[56]
	P2P	XGBoost	RMSE = 1.87, NSE = 0.60	[146]
Bacteria	P2P	Catboost	R^2 = 0.69–0.71, RMS = 0.38–0.43	[147]
TN	S2P	LSTM	R^2 = 0.75, MSE = 5.75, RMSE = 2.35, NSE = 0.75	[65]
	S2P	LSTM	R^2 = 0.96, RMSE = 0.08, MAE = 0.05, MRE = 0.07	[83]
	S2P	LSTM	MAE = 0.13, MAPE = 8.09%, RMSE = 0.15, NSE = 0.92	[86]
	S2P	CNN-LSTM	MAE = 0.39, RMSE = 0.47, MAPE = 3.1%, R^2 = 0.84	[87]

Table A1. *Cont.*

Target	Model Type	ML Model	Results of Model Performance	Article
NH_3-N	S2P	LSTM	RMSE = 0.017 MAE = 0.012	[74]
	S2P	LSTM	R^2 = 0.83, MSE = 5.61, RMSE = 2.33, NSE = 0.83	[65]
	S2P	LSTM	R^2 = 0.93, RMSE = 0.002, MAE = 0.002	[83]
	S2P	BiLSTM	MAE = 0.01–0.06, RMSE = 0.02–0.07, MAPE = 3.18–10.63%	[81]
	S2P	VARMA-BiLSTM	RMSE = 0.001, MAPE = 0.008%, NSE = 1	[82]
	S2S	AT-LSTM	MAE = 0.04–0.19, RMSE = 0.05–0.29, NSE = 0.99–0.94	[91]
	S2S	AT-BiLSTM	MAPE = 15.64, MAE = 0.03, RMSE = 0.04, R^2 = 0.96	[92]
	S2S	BiGRU	MAE = 0.008–0.043, RMSE = 0.015–0.062, MAPE = 6.27–9.79%, NSE = 0.83–0.94	[93]
COD_{Mn}	S2P	LSTM	RMSE = 0.09, MAE = 0.05	[74]
	S2P	LSTM	RMSE = 0.02, MAPE = 0.02%, MAE = 0.01	[84]
	S2P	LSTM	NSE = 0.84	[85]
	S2S	Transformer	MSE = 0.83, MAE = 0.69	[95]
COD	S2P	CNN-LSTM	MAE = 0.22–0.29, RMSE = 0.29–0.38	[66]
Turbidity	S2P	LSTM-RNN	R^2 = 0.88, r = 0.94, MAE = 0.21, RMSE = 0.29	[88]
Salinity	S2P	LSTM	RMSE = 0.24, MAPE = 94.8%	[89]

Table A2. Common metrics for ML water quality regression model evaluation.

Metrics	Abbreviation	Equation
Coefficient of determination	R^2	$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - \overline{y})^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}$
Nash–Sutcliffe efficiency coefficient	NSE	$NSE = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \overline{y})^2}$
Correlation coefficient	r	$r = \frac{\sum_{i=1}^{n}(y_i - \overline{y})(\hat{y}_i - \overline{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2 \cdot \sum_{i=1}^{n}(\hat{y}_i - \overline{\hat{y}})^2}}$
Mean squared error	MSE	$MAE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$
Root mean squared error	RMSE	$RMAE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$
Mean absolute error	MAE	$MAE = \frac{1}{n}\sum_{i=1}^{n}\lvert y_i - \hat{y}_i \rvert$
Mean absolute percentage error	MAPE	$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left\lvert \frac{y_i - \hat{y}_i}{y_i} \right\rvert \times 100\%$

Notes: y_i, $\overline{y}$, $\hat{y}_i$, and $\overline{\hat{y}}$ denote the observed value, the mean of the observed value, the predicted value, and the mean of the predicted value, respectively. n is the total amount of values.

Table A3. Commonly used metrics for ML water quality classification model evaluation.

Metrics	Equation
Accuracy	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$
Precision	$Precision = \frac{TP}{TP+FP}$
F1-Score	$F1\text{-}Score = \frac{2 \cdot Precision \cdot Sensitivity}{Sensitivity + Precision}$
True Positive Rate (TPR)/Sensitivity/Recall	$TPR = \frac{TP}{TP+FN}$
True Negative Rate (TNR)/Specificity	$TNR = \frac{TN}{TN+FP}$

Notes: TP/TN is true positive/negative, indicating that the true class is classified into a true/false category; FP/FN is false positive, representing that the false class is classified into a false/true category.

References

1. Giri, S. Water Quality Prospective in Twenty First Century: Status of Water Quality in Major River Basins, Contemporary Strategies and Impediments: A Review. *Environ. Pollut.* **2021**, *271*, 116332. [CrossRef] [PubMed]
2. Akhtar, N.; Syakir Ishak, M.I.; Bhawani, S.A.; Umar, K. Various Natural and Anthropogenic Factors Responsible for Water Quality Degradation: A Review. *Water* **2021**, *13*, 2660. [CrossRef]

3. Sur, I.M.; Moldovan, A.; Micle, V.; Polyak, E.T. Assessment of Surface Water Quality in the Baia Mare Area, Romania. *Water* **2022**, *14*, 3118. [CrossRef]
4. Malistani, H.A.; Jawadi, H.A.; Sidle, R.C.; Khawary, M.; Khan, A.A. Water Resources and Water Quality Assessment, Central Bamyan, Afghanistan. *Water* **2022**, *14*, 3060. [CrossRef]
5. DSI LLC EFDC+ Theory, Version 11. Available online: https://www.eemodelingsystem.com/wp-content/Download/Documentation/EFDC_Theory_Document_Ver_11.pdf (accessed on 1 November 2024).
6. Wool, T.; Davie, S.; Rodriguez, H. Development of Three-Dimensional Hydrodynamic and Water Quality Models to Support Total Maximum Daily Load Decision Process for the Neuse River Estuary, North Carolina. *J. Water Resour. Plan. Manag.-ASCE* **2003**, *129*, 295–306. [CrossRef]
7. DHI Products | MIKE Powered by DHI. Available online: https://www.dhigroup.com/technologies/mikepoweredbydhi/products (accessed on 18 November 2024).
8. Wells, S.; Berger, C. CE-QUAL-W2 V3 Model Information. Available online: https://cee.pdx.edu/w2/download_model_information.html (accessed on 18 November 2024).
9. Sheng, Y.P. CH3D, CH3D-IMS Integrated Modeling System, & Next Generation CH3D and CH3D-IMS. Available online: https://aces.coastal.ufl.edu/CH3D/ (accessed on 18 November 2024).
10. Delft3D Home-Delft3D-Oss.Deltares.Nl. Available online: https://oss.deltares.nl/web/delft3d (accessed on 18 November 2024).
11. Khullar, S.; Singh, N. Machine Learning Techniques in River Water Quality Modelling: A Research Travelogue. *Water Supply* **2020**, *21*, 1–13. [CrossRef]
12. Yang, Y.; Shang, X.; Chen, Z.; Mei, K.; Wang, Z.; Dahlgren, R.A.; Zhang, M.; Ji, X. A Support Vector Regression Model to Predict Nitrate-Nitrogen Isotopic Composition Using Hydro-Chemical Variables. *J. Environ. Manag.* **2021**, *290*, 112674. [CrossRef]
13. Jayaraman, P.; Nagarajan, K.K.; Partheeban, P.; Krishnamurthy, V. Critical Review on Water Quality Analysis Using IoT and Machine Learning Models. *Int. J. Inf. Manag. Data Insights* **2024**, *4*, 100210. [CrossRef]
14. Razavi, S.; Hannah, D.M.; Elshorbagy, A.; Kumar, S.; Marshall, L.; Solomatine, D.P.; Dezfuli, A.; Sadegh, M.; Famiglietti, J. Coevolution of Machine Learning and Process-based Modelling to Revolutionize Earth and Environmental Sciences: A Perspective. *Hydrol. Process.* **2022**, *36*, e14596. [CrossRef]
15. Valyrakis, M.; Diplas, P.; Dancey, C.L. Prediction of Coarse Particle Movement with Adaptive Neuro-Fuzzy Inference Systems. *Hydrol. Process.* **2011**, *25*, 3513–3524. [CrossRef]
16. Shen, C.; Chen, X.; Laloy, E. Editorial: Broadening the Use of Machine Learning in Hydrology. *Front. Water* **2021**, *3*, 681023. [CrossRef]
17. Kumar, A.; Boehm, M.; Yang, J. Data Management in Machine Learning: Challenges, Techniques, and Systems. In Proceedings of the 2017 ACM International Conference on Management of Data, Chicago, IL, USA, 14–19 May 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 1717–1722.
18. Quinn, C.; Shabestari, A.Z.; Misic, T.; Gilani, S.; Litoiu, M.; McArthur, J.J. Building Automation System-BIM Integration Using a Linked Data Structure. *Autom. Constr.* **2020**, *118*, 103257. [CrossRef]
19. Himeur, Y.; Elnour, M.; Fadli, F.; Meskin, N.; Petri, I.; Rezgui, Y.; Bensaali, F.; Amira, A. AI-Big Data Analytics for Building Automation and Management Systems: A Survey, Actual Challenges and Future Perspectives. *Artif. Intell. Rev.* **2023**, *56*, 4929–5021. [CrossRef] [PubMed]
20. Huang, R.; Ma, C.; Ma, J.; Huangfu, X.; He, Q. Machine Learning in Natural and Engineered Water Systems. *Water Res.* **2021**, *205*, 117666. [CrossRef]
21. Wai, K.P.; Chia, M.Y.; Koo, C.H.; Huang, Y.F.; Chong, W.C. Applications of Deep Learning in Water Quality Management: A State-of-the-Art Review. *J. Hydrol.* **2022**, *613*, 128332. [CrossRef]
22. 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]
23. Ayesha Jasmin, S.; Ramesh, P.; Tanveer, M. An Intelligent Framework for Prediction and Forecasting of Dissolved Oxygen Level and Biofloc Amount in a Shrimp Culture System Using Machine Learning Techniques. *Expert Syst. Appl.* **2022**, *199*, 117160. [CrossRef]
24. Haghiabi, A.H.; Nasrolahi, A.H.; Parsaie, A. Water Quality Prediction Using Machine Learning Methods. *Water Qual. Res. J.* **2018**, *53*, 3–13. [CrossRef]
25. Malek, N.H.A.; Wan Yaacob, W.F.; Md Nasir, S.A.; Shaadan, N. Prediction of Water Quality Classification of the Kelantan River Basin, Malaysia, Using Machine Learning Techniques. *Water* **2022**, *14*, 1067. [CrossRef]
26. Ilić, M.; Srdjević, Z.; Srdjević, B. Water Quality Prediction Based on Naïve Bayes Algorithm. *Water Sci. Technol.* **2022**, *85*, 1027–1039. [CrossRef]
27. Geetha, A.; Nasira, G.M. Data Mining for Meteorological Applications: Decision Trees for Modeling Rainfall Prediction. In Proceedings of the 2014 IEEE International Conference on Computational Intelligence and Computing Research, Coimbatore, India, 18–20 December 2014; pp. 1–4.
28. 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] [PubMed]

29. Prakash, R.; Tharun, V.P.; Renuga Devi, S. A Comparative Study of Various Classification Techniques to Determine Water Quality. In Proceedings of the 2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT), Coimbatore, India, 20–21 April 2018; pp. 1501–1506.
30. Wu, T.; Zhang, W.; Jiao, X.; Guo, W.; Hamoud, Y.A. Comparison of Five Boosting-Based Models for Estimating Daily Reference Evapotranspiration with Limited Meteorological Variables. *PLoS ONE* **2020**, *15*, e0235324. [CrossRef] [PubMed]
31. Han, K.; Wang, Y. A Review of Artificial Neural Network Techniques for Environmental Issues Prediction. *J. Therm. Anal. Calorim.* **2021**, *145*, 2191–2207. [CrossRef]
32. Ly, Q.V.; Nguyen, X.C.; Lê, N.C.; Truong, T.-D.; Hoang, T.-H.T.; Park, T.J.; Maqbool, T.; Pyo, J.; Cho, K.H.; Lee, K.-S.; et al. Application of Machine Learning for Eutrophication Analysis and Algal Bloom Prediction in an Urban River: A 10-Year Study of the Han River, South Korea. *Sci. Total Environ.* **2021**, *797*, 149040. [CrossRef]
33. Dogo, E.M.; Nwulu, N.I.; Twala, B.; Aigbavboa, C. A Survey of Machine Learning Methods Applied to Anomaly Detection on Drinking-Water Quality Data. *Urban Water J.* **2019**, *16*, 235–248. [CrossRef]
34. Syeed, M.M.M.; Hossain, M.S.; Karim, M.R.; Uddin, M.F.; Hasan, M.; Khan, R.H. Surface Water Quality Profiling Using the Water Quality Index, Pollution Index and Statistical Methods: A Critical Review. *Environ. Sustain. Indic.* **2023**, *18*, 100247. [CrossRef]
35. Hutson, M. Artificial Intelligence Faces Reproducibility Crisis. *Science* **2018**, *359*, 725–726. [CrossRef]
36. Gupta, S.; Gupta, A. Dealing with Noise Problem in Machine Learning Data-Sets: A Systematic Review. *Procedia Comput. Sci.* **2019**, *161*, 466–474. [CrossRef]
37. Lowe, M.; Qin, R.; Mao, X. A Review on Machine Learning, Artificial Intelligence, and Smart Technology in Water Treatment and Monitoring. *Water* **2022**, *14*, 1384. [CrossRef]
38. Zhu, M.; Wang, J.; Yang, X.; Zhang, Y.; Zhang, L.; Ren, H.; Wu, B.; Ye, L. A Review of the Application of Machine Learning in Water Quality Evaluation. *Eco-Environ. Health* **2022**, *1*, 107–116. [CrossRef]
39. Ghobadi, F.; Kang, D. Application of Machine Learning in Water Resources Management: A Systematic Literature Review. *Water* **2023**, *15*, 620. [CrossRef]
40. Asadollah, S.B.H.S.; Sharafati, A.; Motta, D.; Yaseen, Z.M. River Water Quality Index Prediction and Uncertainty Analysis: A Comparative Study of Machine Learning Models. *J. Environ. Chem. Eng.* **2021**, *9*, 104599. [CrossRef]
41. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [CrossRef]
42. Rahu, M.A.; Chandio, A.F.; Aurangzeb, K.; Karim, S.; Alhussein, M.; Anwar, M.S. Toward Design of Internet of Things and Machine Learning-Enabled Frameworks for Analysis and Prediction of Water Quality. *IEEE Access* **2023**, *11*, 101055–101086. [CrossRef]
43. Zhou, Y.; Wang, X.; Li, W.; Zhou, S.; Jiang, L. Water Quality Evaluation and Pollution Source Apportionment of Surface Water in a Major City in Southeast China Using Multi-Statistical Analyses and Machine Learning Models. *IJERPH* **2023**, *20*, 881. [CrossRef]
44. Alqahtani, A.; Shah, M.I.; Aldrees, A.; Javed, M.F. Comparative Assessment of Individual and Ensemble Machine Learning Models for Efficient Analysis of River Water Quality. *Sustainability* **2022**, *14*, 1183. [CrossRef]
45. Aldrees, A.; Javed, M.F.; Bakheit Taha, A.T.; Mustafa Mohamed, A.; Jasiński, M.; Gono, M. Evolutionary and Ensemble Machine Learning Predictive Models for Evaluation of Water Quality. *J. Hydrol. Reg. Stud.* **2023**, *46*, 101331. [CrossRef]
46. Uddin, M.G.; Nash, S.; Mahammad 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]
47. Khoi, D.N.; Quan, N.T.; Linh, D.Q.; Nhi, P.T.T.; Thuy, N.T.D. Using Machine Learning Models for Predicting the Water Quality Index in the La Buong River, Vietnam. *Water* **2022**, *14*, 1552. [CrossRef]
48. Shams, M.Y.; Elshewey, A.M.; El-kenawy, E.-S.M.; Ibrahim, A.; Talaat, F.M.; Tarek, Z. Water Quality Prediction Using Machine Learning Models Based on Grid Search Method. *Multimedia Tools Appl.* **2023**, *83*, 35307–35334. [CrossRef]
49. Chen, L.; Wu, T.; Wang, Z.; Lin, X.; Cai, Y. A Novel Hybrid BPNN Model Based on Adaptive Evolutionary Artificial Bee Colony Algorithm for Water Quality Index Prediction. *Ecol. Indic.* **2023**, *146*, 109882. [CrossRef]
50. Shamsuddin, I.I.S.; Othman, Z.; Sani, N.S. Water Quality Index Classification Based on Machine Learning: A Case from the Langat River Basin Model. *Water* **2022**, *14*, 2939. [CrossRef]
51. Islam Khan, M.S.; Islam, N.; Uddin, J.; Islam, S.; Nasir, M.K. Water Quality Prediction and Classification Based on Principal Component Regression and Gradient Boosting Classifier Approach. *J. King Saud Univ.-Comput. Inf. Sci.* **2022**, *34*, 4773–4781. [CrossRef]
52. Tso, G.K.F.; Yau, K.K.W. Predicting Electricity Energy Consumption: A Comparison of Regression Analysis, Decision Tree and Neural Networks. *Energy* **2007**, *32*, 1761–1768. [CrossRef]
53. Aldrees, A.; Awan, H.H.; Javed, M.F.; Mohamed, A.M. Prediction of Water Quality Indexes with Ensemble Learners: Bagging and Boosting. *Process Saf. Environ. Prot.* **2022**, *168*, 344–361. [CrossRef]
54. Uddin, M.G.; Nash, S.; Rahman, A.; Olbert, A.I. Performance Analysis of the Water Quality Index Model for Predicting Water State Using Machine Learning Techniques. *Process Saf. Environ. Prot.* **2023**, *169*, 808–828. [CrossRef]
55. Li, Y.; Mao, S.; Yuan, Y.; Wang, Z.; Kang, Y.; Yao, Y. Beyond Tides and Time: Machine Learning's Triumph in Water Quality Forecasting. *AJAMS* **2023**, *11*, 89–97. [CrossRef]
56. Zamani, M.G.; Nikoo, M.R.; Niknazar, F.; Al-Rawas, G.; Al-Wardy, M.; Gandomi, A.H. A Multi-Model Data Fusion Methodology for Reservoir Water Quality Based on Machine Learning Algorithms and Bayesian Maximum Entropy. *J. Clean. Prod.* **2023**, *416*, 137885. [CrossRef]

57. Sajib, A.M.; Diganta, M.T.M.; Moniruzzaman, M.; Rahman, A.; Dabrowski, T.; Uddin, M.G.; Olbert, A.I. Assessing Water Quality of an Ecologically Critical Urban Canal Incorporating Machine Learning Approaches. *Ecol. Inform.* **2024**, *80*, 102514. [CrossRef]
58. Gorgan-Mohammadi, F.; Rajaee, T.; Zounemat-Kermani, M. Decision Tree Models in Predicting Water Quality Parameters of Dissolved Oxygen and Phosphorus in Lake Water. *Sustain. Water Resour. Manag.* **2023**, *9*, 1. [CrossRef]
59. Panahi, J.; Mastouri, R.; Shabanlou, S. Insights into Enhanced Machine Learning Techniques for Surface Water Quantity and Quality Prediction Based on Data Pre-Processing Algorithms. *J. Hydroinform.* **2022**, *24*, 875–897. [CrossRef]
60. Koranga, M.; Pant, P.; Kumar, T.; Pant, D.; Bhatt, A.K.; Pant, R.P. Efficient Water Quality Prediction Models Based on Machine Learning Algorithms for Nainital Lake, Uttarakhand. *Mater. Today Proc.* **2022**, *57*, 1706–1712. [CrossRef]
61. Yan, T.; Zhou, A.; Shen, S.-L. Prediction of Long-Term Water Quality Using Machine Learning Enhanced by Bayesian Optimisation. *Environ. Pollut.* **2023**, *318*, 120870. [CrossRef] [PubMed]
62. Talukdar, S.; Shahfahad; Ahmed, S.; Naikoo, M.W.; Rahman, A.; Mallik, S.; Ningthoujam, S.; Bera, S.; Ramana, G.V. Predicting Lake Water Quality Index with Sensitivity-Uncertainty Analysis Using Deep Learning Algorithms. *J. Clean. Prod.* **2023**, *406*, 136885. [CrossRef]
63. Dritsas, E.; Trigka, M. Efficient Data-Driven Machine Learning Models for Water Quality Prediction. *Computation* **2023**, *11*, 16. [CrossRef]
64. Abba, S.I.; Linh, N.T.T.; Abdullahi, J.; Ali, S.I.A.; Pham, Q.B.; Abdulkadir, R.A.; Costache, R.; Nam, V.T.; Anh, D.T. Hybrid Machine Learning Ensemble Techniques for Modeling Dissolved Oxygen Concentration. *IEEE Access* **2020**, *8*, 157218–157237. [CrossRef]
65. Wang, X.; Li, Y.; Qiao, Q.; Tavares, A.; Liang, Y. Water Quality Prediction Based on Machine Learning and Comprehensive Weighting Methods. *Entropy* **2023**, *25*, 1186. [CrossRef]
66. Wu, X.; Zhang, Q.; Wen, F.; Qi, Y. A Water Quality Prediction Model Based on Multi-Task Deep Learning: A Case Study of the Yellow River, China. *Water* **2022**, *14*, 3408. [CrossRef]
67. Khullar, S.; Singh, N. Water Quality Assessment of a River Using Deep Learning Bi-LSTM Methodology: Forecasting and Validation. *Environ. Sci. Pollut. Res.* **2022**, *29*, 12875–12889. [CrossRef]
68. Zhang, Q.; Wang, R.; Qi, Y.; Wen, F. A Watershed Water Quality Prediction Model Based on Attention Mechanism and Bi-LSTM. *Environ. Sci. Pollut. Res.* **2022**, *29*, 75664–75680. [CrossRef]
69. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Newry, UK, 2017; Volume 30.
70. Zhu, S.; Hrnjica, B.; Ptak, M.; Choiński, A.; Sivakumar, B. Forecasting of Water Level in Multiple Temperate Lakes Using Machine Learning Models. *J. Hydrol.* **2020**, *585*, 124819. [CrossRef]
71. Lee, J.H.; Lee, J.Y.; Lee, M.H.; Lee, M.Y.; Kim, Y.W.; Hyung, J.S.; Kim, K.B.; Cha, Y.K.; Koo, J.Y. Development of a Short-Term Water Quality Prediction Model for Urban Rivers Using Real-Time Water Quality Data. *Water Supply* **2022**, *22*, 4082–4097. [CrossRef]
72. Ni, Q.; Cao, X.; Tan, C.; Peng, W.; Kang, X. An Improved Graph Convolutional Network with Feature and Temporal Attention for Multivariate Water Quality Prediction. *Environ. Sci. Pollut. Res.* **2022**, *30*, 11516–11529. [CrossRef] [PubMed]
73. Chen, H.; Yang, J.; Fu, X.; Zheng, Q.; Song, X.; Fu, Z.; Wang, J.; Liang, Y.; Yin, H.; Liu, Z.; et al. Water Quality Prediction Based on LSTM and Attention Mechanism: A Case Study of the Burnett River, Australia. *Sustainability* **2022**, *14*, 13231. [CrossRef]
74. Wan, H.; Xu, R.; Zhang, M.; Cai, Y.; Li, J.; Shen, X. A Novel Model for Water Quality Prediction Caused by Non-Point Sources Pollution Based on Deep Learning and Feature Extraction Methods. *J. Hydrol.* **2022**, *612*, 128081. [CrossRef]
75. Wang, Z.; Wang, Q.; Wu, T. A Novel Hybrid Model for Water Quality Prediction Based on VMD and IGOA Optimized for LSTM. *Front. Environ. Sci. Eng.* **2023**, *17*, 88. [CrossRef]
76. Wang, Z.; Wang, Q.; Liu, Z.; Wu, T. A Deep Learning Interpretable Model for River Dissolved Oxygen Multi-Step and Interval Prediction Based on Multi-Source Data Fusion. *J. Hydrol.* **2024**, *629*, 130637. [CrossRef]
77. Hu, Y.; Lyu, L.; Wang, N.; Zhou, X.; Fang, M. Application of Machine Learning Model Optimized by Improved Sparrow Search Algorithm in Water Quality Index Time Series Prediction. *Multimedia Tools Appl.* **2024**, *83*, 16097–16120. [CrossRef]
78. Ahmed, A.A.M.; Jui, S.J.J.; Chowdhury, M.A.I.; Ahmed, O.; Sutradha, A. The Development of Dissolved Oxygen Forecast Model Using Hybrid Machine Learning Algorithm with Hydro-Meteorological Variables. *Environ. Sci. Pollut. Res.* **2023**, *30*, 7851–7873. [CrossRef]
79. Song, C.; Yao, L. Application of Artificial Intelligence Based on Synchrosqueezed Wavelet Transform and Improved Deep Extreme Learning Machine in Water Quality Prediction. *Environ. Sci. Pollut. Res.* **2022**, *29*, 38066–38082. [CrossRef]
80. Tan, R.; Wang, Z.; Wu, T.; Wu, J. A Data-Driven Model for Water Quality Prediction in Tai Lake, China, Using Secondary Modal Decomposition with Multidimensional External Features. *J. Hydrol. Reg. Stud.* **2023**, *47*, 101435. [CrossRef]
81. Kim, J.; Yu, J.; Kang, C.; Ryang, G.; Wei, Y.; Wang, X. A Novel Hybrid Water Quality Forecast Model Based on Real-Time Data Decomposition and Error Correction. *Process Saf. Environ. Prot.* **2022**, *162*, 553–565. [CrossRef]
82. Zheng, Z.; Ding, H.; Weng, Z.; Wang, L. Research on a Multiparameter Water Quality Prediction Method Based on a Hybrid Model. *Ecol. Inform.* **2023**, *76*, 102125. [CrossRef]
83. Zhang, L.; Jiang, Z.; He, S.; Duan, J.; Wang, P.; Zhou, T. Study on Water Quality Prediction of Urban Reservoir by Coupled CEEMDAN Decomposition and LSTM Neural Network Model. *Water Resour Manag.* **2022**, *36*, 3715–3735. [CrossRef]
84. Wu, J.; Wang, Z. A Hybrid Model for Water Quality Prediction Based on an Artificial Neural Network, Wavelet Transform, and Long Short-Term Memory. *Water* **2022**, *14*, 610. [CrossRef]

85. Li, Q.; Yang, Y.; Yang, L.; Wang, Y. Comparative Analysis of Water Quality Prediction Performance Based on LSTM in the Haihe River Basin, China. *Environ. Sci. Pollut. Res.* **2023**, *30*, 7498–7509. [CrossRef]
86. He, M.; Wu, S.; Huang, B.; Kang, C.; Gui, F. Prediction of Total Nitrogen and Phosphorus in Surface Water by Deep Learning Methods Based on Multi-Scale Feature Extraction. *Water* **2022**, *14*, 1643. [CrossRef]
87. Yan, J.; Gao, Q.; Yu, Y.; Chen, L.; Xu, Z.; Chen, J. Combining Knowledge Graph with Deep Adversarial Network for Water Quality Prediction. *Environ. Sci. Pollut. Res.* **2022**, *30*, 10360–10376. [CrossRef]
88. 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]
89. 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]
90. Wai, K.P.; Koo, C.H.; Huang, Y.F.; Chong, W.C. Decomposed Intrinsic Mode Functions and Deep Learning Algorithms for Water Quality Index Forecasting. *Neural Comput. Appl.* **2024**, *36*, 13223–13242. [CrossRef]
91. Li, Y.; Li, R. Predicting Ammonia Nitrogen in Surface Water by a New Attention-Based Deep Learning Hybrid Model. *Environ. Res.* **2023**, *216*, 114723. [CrossRef] [PubMed]
92. Bi, J.; Zhang, L.; Yuan, H.; Zhang, J. Multi-Indicator Water Quality Prediction with Attention-Assisted Bidirectional LSTM and Encoder-Decoder. *Inf. Sci.* **2023**, *625*, 65–80. [CrossRef]
93. Yu, J.-W.; Kim, J.-S.; Li, X.; Jong, Y.-C.; Kim, K.-H.; Ryang, G.-I. Water Quality Forecasting Based on Data Decomposition, Fuzzy Clustering and Deep Learning Neural Network. *Environ. Pollut.* **2022**, *303*, 119136. [CrossRef]
94. Yin, H.; Guo, Z.; Zhang, X.; Chen, J.; Zhang, Y. RR-Former: Rainfall-Runoff Modeling Based on Transformer. *J. Hydrol.* **2022**, *609*, 127781. [CrossRef]
95. Peng, L.; Wu, H.; Gao, M.; Yi, H.; Xiong, Q.; Yang, L.; Cheng, S. TLT: Recurrent Fine-Tuning Transfer Learning for Water Quality Long-Term Prediction. *Water Res.* **2022**, *225*, 119171. [CrossRef]
96. Di Nunno, F.; Zhu, S.; Ptak, M.; Sojka, M.; Granata, F. A Stacked Machine Learning Model for Multi-Step Ahead Prediction of Lake Surface Water Temperature. *Sci. Total Environ.* **2023**, *890*, 164323. [CrossRef]
97. Pyo, J.; Pachepsky, Y.; Kim, S.; Abbas, A.; Kim, M.; Kwon, Y.S.; Ligaray, M.; Cho, K.H. Long Short-Term Memory Models of Water Quality in Inland Water Environments. *Water Res. X* **2023**, *21*, 100207. [CrossRef]
98. Haggerty, R.; Sun, J.; Yu, H.; Li, Y. Application of Machine Learning in Groundwater Quality Modeling—A Comprehensive Review. *Water Res.* **2023**, *233*, 119745. [CrossRef]
99. Espejo-Garcia, B.; Mylonas, N.; Athanasakos, L.; Fountas, S.; Vasilakoglou, I. Towards Weeds Identification Assistance through Transfer Learning. *Comput. Electron. Agric.* **2020**, *171*, 105306. [CrossRef]
100. Shi, G.; Qin, C.; Tao, J.; Liu, C. A VMD-EWT-LSTM-Based Multi-Step Prediction Approach for Shield Tunneling Machine Cutterhead Torque. *Knowl.-Based Syst.* **2021**, *228*, 107213. [CrossRef]
101. Jamei, M.; Ahmadianfar, I.; Chu, X.; Yaseen, Z.M. Prediction of Surface Water Total Dissolved Solids Using Hybridized Wavelet-Multigene Genetic Programming: New Approach. *J. Hydrol.* **2020**, *589*, 125335. [CrossRef]
102. Dragomiretskiy, K.; Zosso, D. Variational Mode Decomposition. *IEEE Trans. Signal Process.* **2014**, *62*, 531–544. [CrossRef]
103. Quilty, J.; Adamowski, J. Addressing the Incorrect Usage of Wavelet-Based Hydrological and Water Resources Forecasting Models for Real-World Applications with Best Practices and a New Forecasting Framework. *J. Hydrol.* **2018**, *563*, 336–353. [CrossRef]
104. Li, Y.; Wang, S.; Wei, Y.; Zhu, Q. A New Hybrid VMD-ICSS-BiGRU Approach for Gold Futures Price Forecasting and Algorithmic Trading. *IEEE Trans. Comput. Soc. Syst.* **2021**, *8*, 1357–1368. [CrossRef]
105. Wu, Z.; Huang, N.E. A Study of the Characteristics of White Noise Using the Empirical Mode Decomposition Method. *Proc. R. Soc. Lond. A* **2004**, *460*, 1597–1611. [CrossRef]
106. Colominas, M.A.; Schlotthauer, G.; Torres, M.E. Improved Complete Ensemble EMD: A Suitable Tool for Biomedical Signal Processing. *Biomed. Signal Process. Control* **2014**, *14*, 19–29. [CrossRef]
107. Gilles, J. Empirical Wavelet Transform. *IEEE Trans. Signal Process.* **2013**, *61*, 3999–4010. [CrossRef]
108. Kim, J.; Wang, X.; Kang, C.; Yu, J.; Li, P. Forecasting Air Pollutant Concentration Using a Novel Spatiotemporal Deep Learning Model Based on Clustering, Feature Selection and Empirical Wavelet Transform. *Sci. Total Environ.* **2021**, *801*, 149654. [CrossRef]
109. Chia, S.L.; Chia, M.Y.; Koo, C.H.; Huang, Y.F. Integration of Advanced Optimization Algorithms into Least-Square Support Vector Machine (LSSVM) for Water Quality Index Prediction. *Water Supply* **2022**, *22*, 1951–1963. [CrossRef]
110. Li, P.; Liu, X.; Chen, H.; Li, B.; Ma, T.; Jiang, W. Optimization of Three-Dimensional Magnetic Field in Vacuum Interrupter Using Particle Swarm Optimization Algorithm. *IEEE Trans. Appl. Supercond.* **2021**, *31*, 1–4. [CrossRef]
111. Xue, J.; Shen, B. A Novel Swarm Intelligence Optimization Approach: Sparrow Search Algorithm. *Syst. Sci. Control Eng.* **2020**, *8*, 22–34. [CrossRef]
112. Mirjalili, S.; Lewis, A. The Whale Optimization Algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [CrossRef]
113. Lumini, A.; Nanni, L. Deep Learning and Transfer Learning Features for Plankton Classification. *Ecol. Inform.* **2019**, *51*, 33–43. [CrossRef]
114. Zhu, N.; Ji, X.; Tan, J.; Jiang, Y.; Guo, Y. Prediction of Dissolved Oxygen Concentration in Aquatic Systems Based on Transfer Learning. *Comput. Electron. Agric.* **2021**, *180*, 105888. [CrossRef]
115. Moriasi, D.N.; Arnold, J.G.; Van Liew, M.W.; Bingner, R.L.; Harmel, R.D.; Veith, T.L. Veith Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations. *Trans. ASABE* **2007**, *50*, 885–900. [CrossRef]

116. Legates, D.R.; McCabe, G.J. Evaluating the Use of "Goodness-of-fit" Measures in Hydrologic and Hydroclimatic Model Validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]
117. Chicco, D.; Warrens, M.J.; Jurman, G. The Coefficient of Determination R-Squared Is More Informative than SMAPE, MAE, MAPE, MSE and RMSE in Regression Analysis Evaluation. *PeerJ Comput. Sci.* **2021**, *7*, e623. [CrossRef]
118. Willmott, C.J.; Matsuura, K. On the Use of Dimensioned Measures of Error to Evaluate the Performance of Spatial Interpolators. *Int. J. Geogr. Inf. Sci.* **2006**, *20*, 89–102. [CrossRef]
119. Goodwin, P.; Lawton, R. On the Asymmetry of the Symmetric MAPE. *Int. J. Forecast.* **1999**, *15*, 405–408. [CrossRef]
120. Fawcett, T. An Introduction to ROC Analysis. *Pattern Recognit. Lett.* **2006**, *27*, 861–874. [CrossRef]
121. Tharwat, A. Classification Assessment Methods. *Appl. Comput. Inform.* **2021**, *17*, 168–192. [CrossRef]
122. Garcia, V.; Mollineda, R.A.; Sanchez, J.S. Theoretical Analysis of a Performance Measure for Imbalanced Data. In Proceedings of the 2010 20th International Conference on Pattern Recognition, Istanbul, Turkey, 23–26 August 2010; pp. 617–620.
123. Islam, N.; Irshad, K. Artificial Ecosystem Optimization with Deep Learning Enabled Water Quality Prediction and Classification Model. *Chemosphere* **2022**, *309*, 136615. [CrossRef] [PubMed]
124. Wexler, J.; Pushkarna, M.; Bolukbasi, T.; Wattenberg, M.; Viegas, F.; Wilson, J. The What-If Tool: Interactive Probing of Machine Learning Models. *IEEE Trans. Visual. Comput. Graph.* **2019**, *26*, 56–65. [CrossRef]
125. Lundberg, S.M.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. *arXiv* **2017**. [CrossRef]
126. Lee, D.; Kim, M.; Lee, B.; Chae, S.; Kwon, S.; Kang, S. Integrated Explainable Deep Learning Prediction of Harmful Algal Blooms. *Technol. Forecast. Soc. Chang.* **2022**, *185*, 122046. [CrossRef]
127. Zheng, H.; Liu, Y.; Wan, W.; Zhao, J.; Xie, G. Large-Scale Prediction of Stream Water Quality Using an Interpretable Deep Learning Approach. *J. Environ. Manag.* **2023**, *331*, 117309. [CrossRef]
128. Ribeiro, M.T.; Singh, S.; Guestrin, C. "Why Should I Trust You?": Explaining the Predictions of Any Classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016.
129. Sundararajan, M.; Taly, A.; Yan, Q. Axiomatic Attribution for Deep Networks. In Proceedings of the Proceedings of the 34th International Conference on Machine Learning, Sydney, NSW, Australia, 6–11 August 2017; pp. 3319–3328.
130. Shrikumar, A.; Greenside, P.; Kundaje, A. Learning Important Features Through Propagating Activation Differences. *arXiv* **2017**, arXiv:1704.02685. [CrossRef]
131. Xu, B.; Zhou, T.; Kuang, C.; Wang, S.; Liao, C.; Liu, J.; Guo, C. Water Quality Assessment in a Large Plateau Lake in China from 2014 to 2021 with Machine Learning Models: Implications for Future Water Quality Management. *Sci. Total Environ.* **2024**, *946*, 174212. [CrossRef]
132. 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]
133. Kisi, O. Modeling of Dissolved Oxygen in River Water Using Artificial Intelligence Techniques. *J. Environ. Inform.* **2013**, *22*, 92–101. [CrossRef]
134. Rawat, P.; Bajaj, M.; Sharma, V.; Vats, S. A Comprehensive Analysis of the Effectiveness of Machine Learning Algorithms for Predicting Water Quality. In Proceedings of the 2023 International Conference on Innovative Data Communication Technologies and Application (ICIDCA), Uttarakhand, India, 14–16 March 2023; pp. 1108–1114.
135. Sakaa, B.; Elbeltagi, A.; Boudibi, S.; Chaffaï, H.; Islam, A.R.M.T.; Kulimushi, L.C.; Choudhari, P.; Hani, A.; Brouziyne, Y.; Wong, Y.J. Water Quality Index Modeling Using Random Forest and Improved SMO Algorithm for Support Vector Machine in Saf-Saf River Basin. *Environ. Sci. Pollut. Res.* **2022**, *29*, 48491–48508. [CrossRef] [PubMed]
136. Talukdar, S.; Shahfahad; Bera, S.; Naikoo, M.W.; Ramana, G.V.; Mallik, S.; Kumar, P.A.; Rahman, A. Optimisation and Interpretation of Machine and Deep Learning Models for Improved Water Quality Management in Lake Loktak. *J. Environ. Manag.* **2024**, *351*, 119866. [CrossRef] [PubMed]
137. Azrour, M.; Mabrouki, J.; Fattah, G.; Guezzaz, A.; Aziz, F. Machine Learning Algorithms for Efficient Water Quality Prediction. *Model. Earth Syst. Environ.* **2022**, *8*, 2793–2801. [CrossRef]
138. Trach, R.; Trach, Y.; Kiersnowska, A.; Markiewicz, A.; Lendo-Siwicka, M.; Rusakov, K. A Study of Assessment and Prediction of Water Quality Index Using Fuzzy Logic and ANN Models. *Sustainability* **2022**, *14*, 5656. [CrossRef]
139. Nair, J.P.; Vijaya, M.S. River Water Quality Prediction and Index Classification Using Machine Learning. *J. Phys. Conf. Ser.* **2022**, *2325*, 012011. [CrossRef]
140. 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]
141. Adedeji, I.C.; Ahmadisharaf, E.; Sun, Y. Predicting In-Stream Water Quality Constituents at the Watershed Scale Using Machine Learning. *J. Contam. Hydrol.* **2022**, *251*, 104078. [CrossRef]
142. Zanoni, M.G.; Majone, B.; Bellin, A. A Catchment-Scale Model of River Water Quality by Machine Learning. *Sci. Total Environ.* **2022**, *838*, 156377. [CrossRef]
143. Aldrees, A.; Khan, M.A.; Tariq, M.A.U.R.; Mustafa Mohamed, A.; Ng, A.W.M.; Bakheit Taha, A.T. Multi-Expression Programming (MEP): Water Quality Assessment Using Water Quality Indices. *Water* **2022**, *14*, 947. [CrossRef]

144. Latif, S.D.; Birima, A.H.; Ahmed, A.N.; Hatem, D.M.; Al-Ansari, N.; Fai, C.M.; El-Shafie, A. Development of Prediction Model for Phosphate in Reservoir Water System Based Machine Learning Algorithms. *Ain Shams Eng. J.* **2022**, *13*, 101523. [CrossRef]
145. Lee, H.W.; Kim, M.; Son, H.W.; Min, B.; Choi, J.H. Machine-Learning-Based Water Quality Management of River with Serial Impoundments in the Republic of Korea. *J. Hydrol. Reg. Stud.* **2022**, *41*, 101069. [CrossRef]
146. Park, J.; Lee, W.H.; Kim, K.T.; Park, C.Y.; Lee, S.; Heo, T.-Y. Interpretation of Ensemble Learning to Predict Water Quality Using Explainable Artificial Intelligence. *Sci. Total Environ.* **2022**, *832*, 155070. [CrossRef] [PubMed]
147. 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]

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.