Environmental Science and Pollution Research (2024) 31:26415–26431
https://doi.org/10.1007/s11356-024-33058-7

REVIEW ARTICLE

Research progress in water quality prediction based on deep learning technology: a review

Wenhao Li[1,2] · Yin Zhao[1] · Yining Zhu[2,3] · Zhongtian Dong[3] · Fenghe Wang[2,3] · Fengliang Huang[1,2]

Received: 27 November 2023 / Accepted: 20 March 2024 / Published online: 27 March 2024

Abstract
Water, an invaluable and non-renewable resource, plays an indispensable role in human survival and societal development. Accurate forecasting of water quality involves early identification of future pollutant concentrations and water quality indices, enabling evidence-based decision-making and targeted environmental interventions. The emergence of advanced computational technologies, particularly deep learning, has garnered considerable interest among researchers for applications in water quality prediction because of its robust data analytics capabilities. This article comprehensively reviews the deployment of deep learning methodologies in water quality forecasting, encompassing single-model and mixed-model approaches. Additionally, we delineate optimization strategies, data fusion techniques, and other factors influencing the efficacy of deep learning-based water quality prediction models, because understanding and mastering these factors are crucial for accurate water quality prediction. Although challenges such as data scarcity, long-term prediction accuracy, and limited deployments of large-scale models persist, future research aims to address these limitations by refining prediction algorithms, leveraging high-dimensional datasets, evaluating model performance, and broadening large-scale model application. These efforts contribute to precise water resource management and environmental conservation.

Keywords Deep learning · Water quality prediction · Hybrid model · Optimization algorithm · Data decomposition algorithm · Neural network

Introduction

Water, an essential resource for human survival, is inherently vulnerable and non-renewable. Rapid industrialization and urbanization have caused ecological and environmental destruction and a considerable upsurge in water pollution (Tirkolaee et al. 2018). Both human activities and natural processes, such as rock weathering, erosion, and climate change, impact water quality (Lyu et al. 2020). The persistent presence of pollution and deteriorating water environments pose serious threats to human health (Vörösmarty et al. 2010). Water quality prediction plays a crucial role in addressing specific environmental challenges, such as effective management and pollution reduction. It enables early detection, warning, and water pollution treatment, ensuring the safe use of water. First, water quality prediction aids water resource managers in understanding the current status and trends in water pollution. This insight enables the implementation of targeted management measures, optimization of water supply strategies, and preservation of water resources through judicious usage. Second, it assists in monitoring and controlling pollution sources, promptly detecting and responding to potential pollution incidents. This, in turn, helps reduce water pollution and preserve the health of water ecosystems. Last, the early detection and resolution of potential water pollution issues contribute to ensuring personal safety and well-being.

Water quality prediction methods can be categorized into traditional and machine learning methods. Traditional

Responsible Editor: Marcus Schulz

Fengliang Huang
huangfengliang@njnu.edu.cn

1 School of Electrical and Automation Engineering, Nanjing Normal University, Nanjing, China

2 Jiangsu Province Engineering Research Center of Environmental Risk Prevention and Emergency Response Technology, School of Environment, Nanjing 210023, China

3 Key Laboratory for Soft Chemistry and Functional Materials of Ministry of Education, Nanjing University of Science and Technology, Nanjing 210094, Jiangsu, China

methods typically consist of physical and statistical models. Physical models utilize mathematical equations to describe water movement, transport, and transformation processes (Magar et al. 2017, Post et al. 2018, Rong et al. 2019, Wool et al. 2020, Zamani et al. 2018), such as the Soil and Water Assessment Tool, Hydrological Simulation Program-FORTRAN, and MIKE System Hydrological European. When applied to water quality prediction, these models are typically build upon a foundation of understanding physical processes and factors with parameters possessing rigorous physical explanations. However, challenges such as difficulty with parameter calibration, modeling structure complexity, parameter uncertainty, and high computational costs restrict their utility in river basin water quality prediction. Additionally, these models are often challenging to calibrate and require a high level of professional expertise to achieve accurate results (Liu and Tong 2011, Moshtaghi et al. 2018; Wan et al. 2021; Zhou et al. 2021). Statistical models are based on probability theory and mathematical statistical methods. While some processes cannot be derived through theoretical analyses, the functional relationship between variables can be obtained through methods such as multiple regression and principal component analysis (PCA) using experimental data. Statistical models, unlike physical models, require only historical data for water quality prediction, making them simpler and more effective (Guo et al. 2020; Shi et al. 2019). However, the simplistic nature of statistical models means they typically assume a normal and linear distribution in the correlation between water quality and explanatory variables, and linearizing nonlinear relationships between water quality changes and influencing factors can reduce the model's accuracy (Avila et al. 2018; Yang et al. 2017).

Deep learning (LeCun et al. 2015) is a machine learning branch that utilizes neural network models. These methods involve learning complex feature representations and pattern recognitions through multi-layer neural networks. These multi-layer neural networks offer enhanced expressive power and adaptability, with wide-ranging applications in big data analytics, including but not limited to computer vision (Badrinarayanan et al. 2015), image classification (Rawat and Wang 2017), speech recognition (Zhang et al. 2020), time series prediction (Sezer et al. 2019), natural language processing (Otter et al. 2020), and anomaly detection (Malki et al. 2022). In the context of water quality prediction, deep learning technology has significantly improved both prediction accuracy and model effectiveness. It enables not only water quality concentration predictions but also classifications based on water quality levels (Islam and Irshad 2022). In the water quality prediction field, deep learning technology and traditional machine learning methods exhibit notable distinctions. Deep learning methods employ neural network models that autonomously learn intricate features for feature extraction and generalization. Conversely, traditional machine learning methods utilize simpler models that necessitate manual feature design and selection. Consequently, deep learning tends to achieve higher accuracy. Despite this advantage, deep learning models are often considered 'black boxes' because of their lower interpretability compared with traditional machine learning models. Nonetheless, researchers are striving to enhance deep learning model interpretability. Therefore, advancements in deep learning techniques are crucial for water quality prediction technologies and governing strategies for water resource management and environmental preservation.

This study aims to furnish a comprehensive review of deep learning algorithms applied to water quality prediction. The screened literature covers a variety of pollutants and a wide range of algorithms used for prediction. Literature screening for research studies on deep learning technology in water quality prediction involves three stages: search, collation, and analysis. During the search stage, the search engines used include Web of Science, Google Scholar, SpringerLink, and ScienceDirect; the specific keywords used include water quality prediction, artificial intelligence, deep learning, meta-heuristic method, data decomposition, groundwater quality, and surface water quality. Additional articles were identified using the cross-citation method by reviewing the selected articles' references, resulting in a total of 253 articles. The collation stage involved a comprehensive evaluation of articles based on title, keywords, abstracts, etc., leading to the selection of 87 highly relevant articles. Finally, in the analysis stage, each article was carefully reviewed to examine the advancements in deep learning technology for water quality prediction. "The application of deep learning technology in water quality prediction" of this article provides an overview of existing deep learning technology applied to water quality prediction, as shown in Fig. 1, categorized into single and hybrid deep learning models. "Model performance evaluation indicators and factors affecting deep learning technologies for water quality prediction" discusses the evaluation indicators for model performance and the factors influencing water quality prediction using deep learning technology. "The limitations of current deep learning techniques in water quality prediction" explores deep learning technology limitations in water quality prediction. Finally, this article concludes by outlining prospective research avenues and opportunities in this burgeoning field.

The application of deep learning technology in water quality prediction

A systematic analysis of the 87 highly correlated articles was conducted to examine the deep learning prediction of water quality based on the number of deep learning models

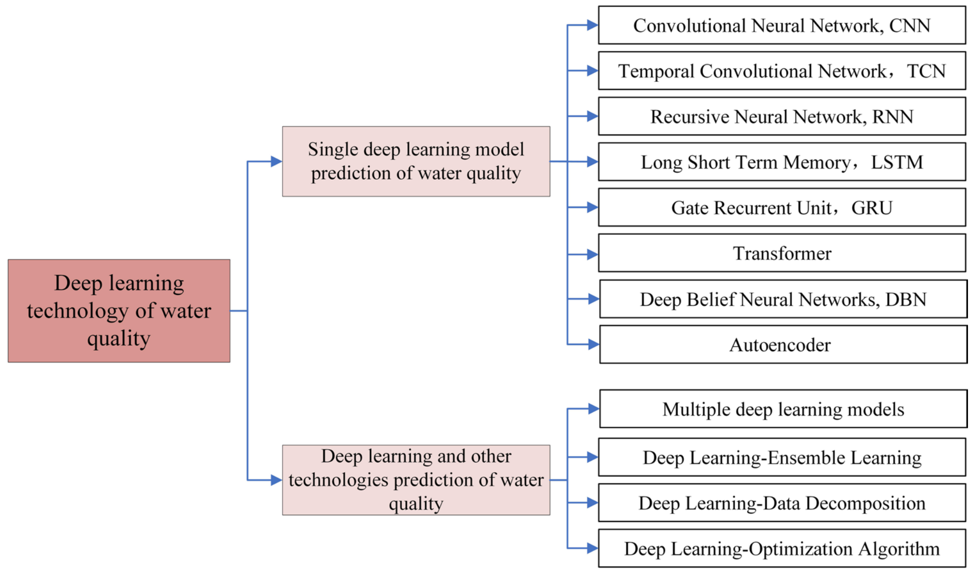


Fig. 1 Classification of water quality prediction methods based on deep learning technology

utilized. This analysis revealed two main categories: single-model predictions of water quality and hybrid-model predictions of water quality, as illustrated in Fig. 1. In single-model predictions, a deep learning model is employed to forecast water quality. In contrast, to enhance prediction accuracy, hybrid-model predictions involve integrating various deep learning models or combining deep learning models with other techniques, such as traditional machine learning, data decomposition methods, and optimization algorithms.

General steps in water quality prediction based on deep learning technology

The process of predicting water quality using deep learning is schematically represented in Fig. 2 and encompasses the following steps:

1. Data gollection: gather relevant data on water quality indicators and environmental factors from sensors, monitoring stations, and historical records.
2. Data preprocessing: clean data by addressing missing values and outliers and standardizing the features.
3. Feature dimension reduction: apply methods such as PCA to reduce feature dimensionality.
4. Model selection and training: choose an appropriate deep learning model for water quality prediction and train and optimize the model using the preprocessed dataset and evaluate its performance using a test dataset.
5. Model optimization: fine-tune the model parameters or improve the model structure based on the evaluation results to enhance its performance further.
6. Real-time prediction and monitoring: utilize the optimized model to predict real-time water quality, include monitoring data into the model to obtain water quality prediction results, monitor the water quality in real-time, and take necessary actions accordingly.

Single deep learning model predictions of water quality

Convolutional neural network (CNN)

Inspired by biology, CNNs have been successfully applied to tasks such as image recognition, object detection, and text processing (Banan et al. 2020; Kumar et al. 2021). A CNN consists of three main components: the convolution layer, the pooling layer, and the fully connected layer. The convolution layer applies a series of filters that slide over the input image to perform dot products based on input data. The pooling layer reduces the dimensionality of the resulting matrix, and the fully connected layer compresses the extracted features to produce the final result.

CNNs can be utilized for time series prediction by constructing an end-to-end model. For instance, water quality time series data can serve as the model input, with the predicted water quality as the output. Pyo et al. (2020) employed a CNN by inputting synthetic nutrient, environmental, and atmospheric grid unit data to predict the cyanobacteria concentration in water. Ta andWei (2018) proposed a simplified reverse understanding CNN model for predicting dissolved oxygen, which showed a faster convergence speed and better prediction stability compared with a back propagation neural network (BPNN). Additionally, CNN can effectively reduce data dimension and extract spatial features. When CNNs are used for predictions in this context, they are commonly incorporated into mixed models, as explained in "Hybrid-model predictions of water quality."

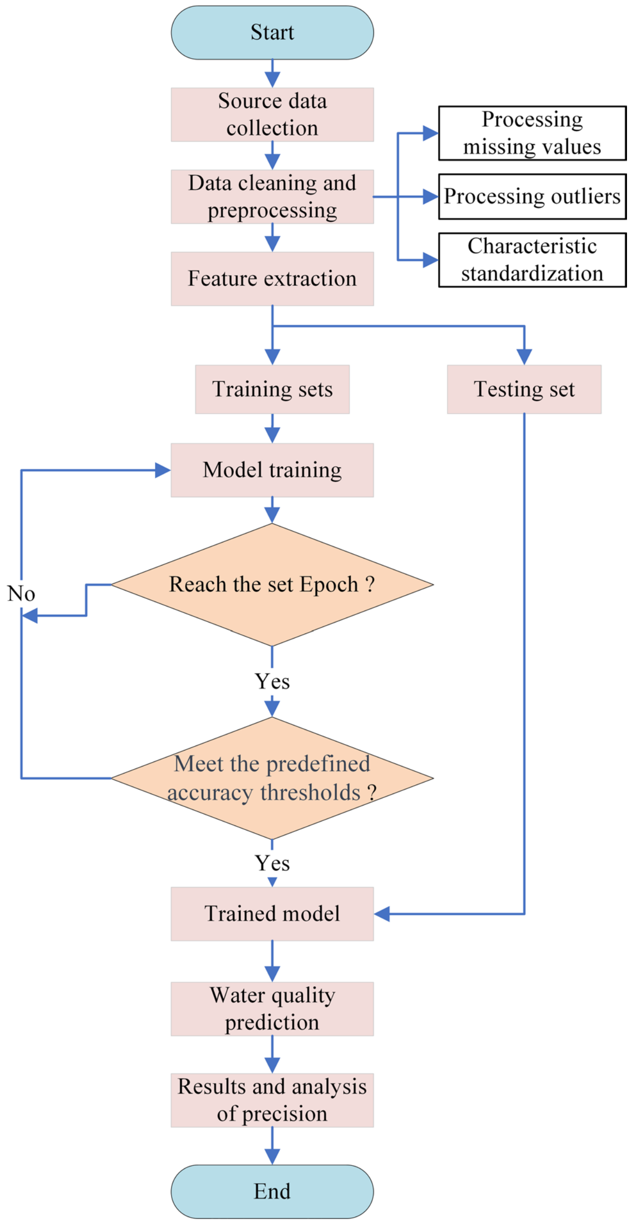


Fig. 2 General procedure for utilizing deep learning technologies in water quality prediction

CNNs are capable of effectively extracting both spatial and temporal features for water quality prediction while effectively reducing data feature dimensionality. Despite their relatively simple architecture and training ease, CNNs possess certain limitations. Specifically, they exhibit restricted capabilities in processing extended sequences and offer limited model interpretability. Currently, CNNs are primarily used as data processing models to extract various features from water quality data to enhance prediction accuracy (Fan et al. 2020, Habib and Qureshi 2020, Xue et al. 2021).

Temporal convolutional network (TCN)

TCNs (Bai et al. 2018) represent a specialized sequence modeling architecture that leverages CNNs to discern patterns and features within sequence data. Unlike traditional CNNs, TCNs incorporate a dilation factor, allowing the convolution kernel to skip a certain number of inputs. This expands the receptive field and captures long-distance dependencies in the sequence effectively. Upon gradually increasing the expansion factor layer by layer, TCNs can model temporal patterns at various scales and consider multiple time scales simultaneously.

Zhang and Li (2022) proposed a multi-input, multi-output end-to-end prediction model called MIMO-TCN based on a TCN. ConvNeXt was utilized to extract features from input data, and a TCN was employed to enhance the prediction accuracy using extracted feature data. To address the issue of gradient disappearance with an increasing number of network layers, the model incorporates skip connections between its modules.

TCNs offer several advantages, including the ability to consider time series correlations, process indefinite long time series data, and enable parallel computing. However, TCNs also possess some limitations, such as a lack of consideration for spatial correlation, data imbalance issues, and limited explanatory power. Despite these shortcomings, TCNs show promising application prospects, particularly in time series prediction tasks that prioritize computational efficiency, modeling long-term dependence, and achieving better results with fewer parameters (Zhang et al. 2019c).

Recurrent neural network (RNN)

RNNs are a type of neural network capable of processing data with a time series structure (Schmidhuber 2015). Unlike traditional feedforward neural networks (FFNNs), RNN neurons receive activation values from the hidden layer in the previous time step as well as input from the current time step. This allows RNNs to update their internal states dynamically and to consider both past inputs and states when computing the current output, making them particularly apt for sequence modeling and forecasting.

Mohamed et al. (2012) implemented deep learning techniques using RNN algorithms. Xiang and Demir (2020) utilized an RNN and sequence learning to develop a neural runoff model and enhanced flow prediction by integrating water level data from an upstream flowmeter. Zhang et al. (2019b) proposed a water quality prediction model based on kernel PCA (kPCA) and an RNN for predicting trends in dissolved oxygen. kPCA was employed to reduce noise in the original dataset while retaining relevant information. The RNN was used to leverage past information for future trend prediction. Compared with FFNNs, support vector regression (SVR), and general regression neural networks, this model demonstrated higher prediction accuracy.

RNNs possess the theoretical capability to handle sequences of any length. However, they are susceptible to issues such as gradient disappearance or explosion when dealing with long sequences. Additionally, the sequential nature of RNN computations can negatively impact the model's training speed and efficiency, especially when working with large-scale data. Furthermore, the transmission of information in RNNs occurs step by step through time steps, leading to subpar modeling performance in long-term relationships. These limitations restrict its practical applicability (Chen et al. 2018, Hochreiter and Schmidhuber 1997).

Long short-term memory network (LSTM)

LSTMs (Hochreiter and Schmidhuber 1997) are an enhanced model derived from RNNs. They incorporate three essential gating mechanisms, namely, the input gate, forgetting gate, and output gate, aiding in effective information retention and omission. This improves the model's long-term memory capacity and its ability to capture time series patterns (Gers et al. 2000, 2003). Building upon the LSTM, Bi-LSTM (Graves and Schmidhuber 2005) introduces a reverse LSTM network to consider both the forward and backward information in a sequence.

LSTM and bidirectional LSTM (Bi-LSTM) models have gained significant popularity in the water quality prediction field. These models have proven to be effective tools for predicting various aspects of water quality, such as water quality indexes (Saroja et al. 2023), drinking water quality (Liu et al. 2019a), and river algal blooms (Lee et al. 2018). Several studies have demonstrated that LSTM models outperform other models, such as support vector machines (SVMs) and artificial neural networks (ANNs) in water quality prediction (Essam et al. 2022; Yang et al. 2023a). Additionally, an improved LSTM model has shown promise as a practical method for early warnings of water pollution risks (Guan et al. 2022). Furthermore, Bi-LSTMs have also been widely adopted (Khullar and Singh 2022), as they provide more accurate water quality prediction results. Overall, the research literature indicates that LSTM and Bi-LSTM models possess significant application value and high prediction accuracy in water quality prediction.

LSTMs effectively address the issues of gradient disappearance and explosion through their gating structure. They excel at preserving long-term memory and exhibit a robust modeling capability for handling lengthy sequences and intricate temporal dependencies. However, it is important to note that LSTMs incur a higher computational cost compared with ordinary RNNs, as they necessitate more parameters and computing resources. Additionally, when handling shorter sequence data, they may become overly complex and prone to overfitting (Chen et al. 2018; Zhou 2020).

Gated recurrent unit (GRU)

GRUs (Chen et al. 2018) are an RNN structure that improves upon LSTMs. By simplifying the LSTM gating mechanism, GRUs reduce the number of parameters and improve model efficiency. Unlike LSTMs, GRUs only comprise two gating units: the reset and the update gates. The reset gate controls the degree of reset for the hidden state at the previous time step, while the update gate determines how the current input updates the hidden state. By reducing the number of gating units, GRUs reduce computational complexity and enhance their ability to handle long-term dependencies (Chung et al. 2014).

Researchers have successfully integrated GRUs with other models or employed different data processing methods to improve prediction accuracy. For instance, Liu et al. (2020) introduced a deep learning network called bidirectional stacked simple recurrent units. Additionally, GRU has the advantage of faster convergence compared with LSTM and higher training efficiency (Cheng et al. 2020).

GRUs offer advantages such as fewer parameters and increased computational efficiency compared with LSTMs. GRU utilizes a memory unit similar to LSTM, enabling it to learn both long-term and short-term dependencies. However, GRU's modeling ability in very long sequence data remains constrained. In certain tasks, GRU demonstrates comparable or slightly better performance than LSTM. Additionally, the reduced parameter count in GRUs makes them less prone to overfitting (Gruber and Jockisch 2020, Yan et al. 2021).

Transformer

Transformer (Vaswani et al. 2017) is a neural network architecture that addresses the issue of gradient disappearance or explosion in traditional RNNs when handling long sequences. It comprises an encoder and a decoder, which are composed of multiple identical layers. Each layer consists of two sub-layers: a multi-head self-attention and an FFNN. The encoder encodes the input sequence, while the decoder generates the output sequence.

Currently, there are limited Transformer applications in water quality prediction. Yao et al. (2022) conducted a study in the Chaohu area and employed various deep learning models, including RNN, LSTM, multi-layer perceptron (MLP), and transformer-based models, to predict a long-term comprehensive water quality index. The results demonstrated that all selected models performed well in the study area. However, as the length of the prediction sequence increased, the performance of informer, a transformer-based model, was notably better. Particularly, informer showed significant advantages in long-term water quality prediction, offering effective modern tools for water quality monitoring and management.

The transformer model is a promising tool for water quality prediction, especially for large-scale predictions. It offers parallel computing capabilities for processing long sequence data, resulting in high efficiency. However, the transformer model applied to water quality prediction requires a substantial amount of high-quality training data. Additionally, this model has numerous parameters and high computational complexity, which limits its application research range. Nevertheless, the research prospects for the transformer model in water quality prediction are extensive.

Deep belief neural network (DBN)

DBNs (Mohamed et al. 2012) are a probabilistic generation model consisting of a series of constrained Boltzmann machine elements. They serve as a tool for unsupervised learning, similar to an autoencoder, and can also be used for supervised learning and classification purposes. These models comprise multiple hidden layers interconnected by weights. The DBN training process involves a greedy layer-by-layer approach. Initially, each restricted Boltzmann machine (RBM) is trained to obtain the weight parameters for each layer. Subsequently, the entire DBN is established by connecting these layers. By combining unsupervised pre-training and supervised fine-tuning, the model's expressive capability can be enhanced, making it adaptable to the target task.

Yan et al. (2020) proposed a water quality prediction model called PSO-DBN-LSSVR, which combines the particle swarm optimization (PSO) algorithm and the least squares SVR machine. This model demonstrates improved accuracy and robustness in predicting water quality parameters compared with traditional neural networks and model combination methods. In order to address the complex relationship between variables in wastewater treatment processes, Niu et al. (2020) introduced a GA-DBN method that utilizes genetic algorithms (GAs) to reduce dimensionality and simplify network structure. Comparing GA-DBN with traditional DBN and back propagation neural network models, it achieves higher accuracy in predicting variables in complex wastewater treatment processes and improves prediction accuracy.

DBNs are seldom utilized as standalone approaches in water quality modeling because of their relatively lower prediction accuracy compared with other deep learning models. Instead, DBNs are commonly combined with other optimization algorithms or data processing methods to evaluate and showcase the effectiveness of these algorithms. Additionally, DBNs can be used as a benchmark model to highlight the superior prediction accuracy of other models (Niu et al. 2020; Ren et al. 2020).

Autoencoder

The autoencoder is an unsupervised learning algorithm used to learn high-dimensional representation and extract feature data (Zhao et al. 2019). It is composed of two parts: the encoder and the decoder. The encoder combines input data into a low-dimensional coding representation, which is then decoded into an output that resembles the original input data. Through training, the encoder can learn meaningful features from the input data and map these features back to the original data using the decoder.

Autoencoders are unsuitable for water quality prediction. They are commonly employed for reducing data dimensionality or enhancing prediction accuracy when combined with other models. For instance, Kayalvizhi et al. (2023) developed a denoising autoencoder (DAE) model by combining an autoencoder with an LSTM. They used the LSTM as both an encoder and decoder to predict the nitrate and chloride levels in groundwater.

While autoencoders may not be directly applicable to water quality prediction, they can serve as a valuable auxiliary tool in such tasks. Autoencoders can be utilized for data preprocessing and feature extraction, ultimately enhancing water quality prediction model performance. The potential for autoencoders in data preprocessing is vast and holds promising application prospects.

Hybrid-model predictions of water quality

Our comparative analysis of recent literature studies on deep learning for water quality prediction revealed a significant increase in research focusing on mixed model predictions. In contrast, applying single-model predictions is on a downward trend. Typically, researchers combine multiple deep learning models or integrate deep learning with traditional machine learning algorithms, data decomposition algorithms, optimization algorithms, etc., to leverage their respective strengths in capturing complex relationships and patterns within data.

Fusion of multiple deep learning models to predict water quality

The fusion of multiple deep learning models leverages the unique characteristics of different methods, such as CNN, RNN, LSTM, TCN, and attention, to address the challenges posed by complex time series problems arising from spatial and temporal variations in datasets. By combining these deep learning methods, hybrid models can achieve improved prediction results. This approach is particularly effective in handling large quantities of time series data and can adapt well to diverse data structures.

Using the timing processing ability of the original RNN and the ability of attention to weight or focus on different input positions (Geng et al. 2022; Liu et al. 2019b), more accurate predictions can be achieved. LSTM can be improved by combining it with attention (Chen et al. 2022). Upon combining the spatiotemporal feature extraction ability of CNN with the timing processing ability of LSTM, the prediction accuracy and training speed for water quality can be improved (Prasad et al. 2022). LSTM-TCN (Li et al. 2022) outperforms LSTM in capturing characteristics from historical data, while MPA-RNN (Geng et al. 2022) improves prediction accuracy compared with RNN. There are various applications of CNN and RNN (including LSTM and GRU) after fusing with Attention (Mei et al. 2022; Yang et al. 2023b, c, 2021). These models primarily utilize CNN to extract features, RNNs and their improved models to capture long-term dependencies, and the attention mechanism to dynamically adjust the model's focus. The prediction accuracy of these fusion models is superior to that of single models (LSTM, GRU, etc.) and simple hybrid models (such as CNN-LSTM and LSTM-attention).

The fusion of different deep learning models in time series prediction is an efficient method that combines their individual advantages. This approach, known as the hybrid prediction model, demonstrates improved prediction accuracy and stability compared with single deep learning methods. By integrating multiple deep learning models, this approach offers novel ideas and methods for addressing time series prediction problems.

Fusion of deep learning and traditional machine learning to predict water quality

Traditional machine learning methods refer to using statistical theory and algorithms to construct models for solving machine learning problems. These methods include linear regression (LR), random forest (RF), SVM, PCA, SVR, MLP, and other algorithms. While some traditional machine learning methods can be used alone for water quality prediction (such as LR, RF, SVM, MLP, etc.), their prediction accuracy is generally lower compared with deep learning methods. Therefore, they are often used as benchmark models to compare the prediction accuracy of deep learning models. Additionally, traditional machine learning algorithms are utilized for data processing.

Juan et al. (2022) utilized RF to interpolate missing data and then fed these processed data into an RNN with an attention mechanism for the multi-step prediction of dissolved oxygen. The findings demonstrate that RF can compensate for a lack of dissolved oxygen monitoring data, contribute to creating high-quality water quality monitoring datasets, and enhance the model's prediction accuracy. Similarly, Shan et al. (2022) proposed a hybrid deep learning architecture called XG-LSTM, which comprises an XGBoost module and two parallel LSTM models. XGBoost is employed to process variables and predict algal cell density and microcystin concentration in the Three Gorges Reservoir. The results indicate that the XG-LSTM model outperforms other models in terms of prediction accuracy, and the ensemble learning approach exhibits advantages in handling noise and missing data in water quality datasets. The utilization of various algorithms in combination enhances the model performance, accelerates convergence speed, and improves prediction accuracy for water quality prediction challenges. Moreover, integrating deep learning models with ensemble techniques effectively addresses complex temporal and spatial dependencies, allowing for powerful expression capabilities. This approach enables the model to learn intricate patterns and features from data, ultimately reducing prediction errors and enhancing prediction accuracy (Zamani et al. 2023).

Traditional machine learning methods possess certain advantages in handling raw and noisy data. However, in terms of improving water quality prediction accuracy, their effectiveness is limited compared with deep learning methods. It is important to carefully consider the characteristics and suitable scenarios for both methods.

Fusion of deep learning and data decomposition algorithms to predict water quality

Deep learning models often encounter complex data when performing time series prediction. This complexity can greatly reduce the model's prediction efficiency and render simple predictors unreliable. To address this issue, data decomposition methods have been implemented in data processing. These methods aim to handle larger and more complex data sequences. In recent years, the significance of data decomposition methods in time series prediction has grown, leading to their widespread use in signal decomposition and noise reduction to enhance prediction accuracy.

The general steps for predicting water quality using deep learning and data decomposition are outlined in Fig. 3. First, data for the time series prediction are collected and organized, ensuring a time sequence and performing necessary preprocessing. Next, data decomposition algorithms, such as EMD, EEMD, and VMD, are employed to decompose the original time series data into different components such as trend, periodicity, and seasonality. Subsequently, deep learning models such as RNN or CNN are utilized to train and learn from the decomposed data. Finally, the trained

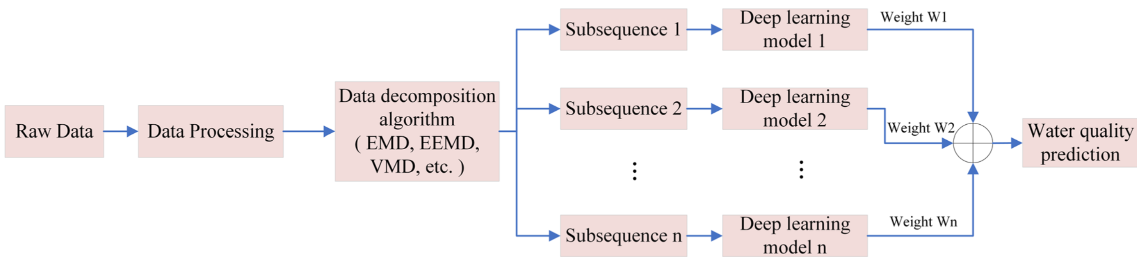


Fig. 3 General steps of deep learning and data decomposition to predict water quality

model is employed for timing prediction, and the results are adjusted and optimized as required.

The commonly used data decomposition methods are empirical mode decomposition (EMD), ensemble empirical mode decomposition (EEMD), and variational mode decomposition (VMD). These methods have distinct advantages, such as identifying vibration modes, suppressing modal aliasing, and reducing data smoothness. Researchers have applied these decomposition methods to process original data in order to reduce noise. These denoised data are subsequently fed into a deep learning model to enhance prediction accuracy (He et al. 2022; Wang et al. 2023c; Zhang et al. 2021, 2023). For instance, Zhang et al. (2021) proposed an EEMD-LSTM model, which combines EEMD and LSTM networks. By establishing an LSTM sub-model for each sub-sequence and aggregating the prediction results, better prediction accuracy was achieved compared with CNN, LSTM, and EEMD-CNN. Another example is the VMD-LSTM model proposed by He et al. (2022) for water quality data denoising and prediction. VMD was utilized to denoise water quality data, while LSTM/GRU was employed for prediction, resulting in improved prediction performance. Additionally, the secondary decomposition method, which employs two data decomposition methods, can further enhance the deep learning model (Dong and Zhang 2021). Furthermore, combining the data decomposition method with other technologies, such as the two-level attention mechanism or optimization algorithm, can also improve the model's prediction capability (Li and Li 2023, Song et al. 2021).

Therefore, choosing an appropriate data decomposition method is crucial for enhancing the water quality prediction accuracy attained using deep learning models. However, it is necessary to fully consider the needs of practical applications and data characteristics and select the appropriate data decomposition method.

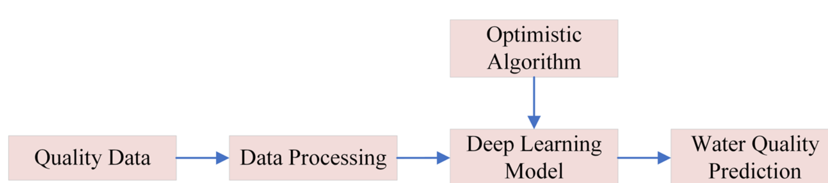
Fusion of deep learning and optimization algorithms to predict water quality

Optimization algorithms are a practical method for improving prediction model performance. They effectively enhance the efficiency and accuracy of deep learning models when dealing with complex data. By utilizing optimization algorithms, we can efficiently search for the model's optimal parameter set, optimize the feature engineering process, and enhance the stability and accuracy of the learning algorithm. These technologies are extensively applied in deep learning, enabling models to better adapt to and learn complex data relationships.

The integration of deep learning and optimization algorithms for prediction involves several steps, as shown in Fig. 4. First, time series data need to be prepared, including collection, collation, and preprocessing. Then, relevant features are extracted from time series data to enable the deep learning model to comprehend the patterns and relationships in these data. Subsequently, the deep learning model is designed and trained, with careful selection of the appropriate network structure and optimization algorithm to maximize the time series prediction accuracy. Additionally, the model parameters are further optimized using optimization algorithms such as the chaotic sparrow search algorithm (CSSA) to enhance the prediction performance. Finally, the trained model is utilized for timing prediction, with the option to adjust and optimize the results as necessary.

He et al. (2022) utilized the CSSA to determine the optimal hyperparameters for an LSTM model. Yang and Liu (2022) employed the VMD and wavelet threshold joint denoising methods to eliminate mixed noise in water quality

Fig. 4 General steps of deep learning and optimization algorithms for predicting water quality



time series and enhanced the whale optimization algorithm to identify the optimal hyperparameters for GRU. Wang et al. (2023c) developed an optimized LSTM prediction model using VMD and an improved grasshopper optimization algorithm. Furthermore, PSO (Zhang et al. 2023), adaptive hybrid mutation particle swarm optimization (Liu et al. 2021), pathfinder optimization algorithm (Guo et al. 2022), and GA (Niu et al. 2020) have also been utilized to optimize deep learning model performance, with successful applications in water quality monitoring and wastewater treatment. Other optimization algorithms, such as the gray wolf optimizer algorithm (Yang et al. 2020) and modified teaching–learning-based optimization algorithm (Larijani and Dehghani 2023), could be considered for enhancing water quality prediction models in future research. These optimization techniques enable us to enhance prediction models' accuracy and reliability, resulting in more precise data predictions. In some cases, optimization algorithms can not only improve the prediction accuracy but also reduce the model's run time (Farsi et al. 2020).

In recent years, researchers have focused on studying hybrid models to enhance water quality prediction accuracy. These models typically combine multiple deep learning models or integrate deep learning with traditional machine learning, data decomposition algorithms, and optimization algorithms. The fusion of multiple deep learning models or deep learning with traditional machine learning leverages the strengths of different methods to address complex time series challenges arising from spatial and temporal dataset changes. This approach is adaptable to diverse data structures and can efficiently handle large quantities of time series data. Integrating deep learning with data decomposition algorithms involves utilizing these algorithms to reduce or break down original data using various techniques, extracting the most relevant features to enhance prediction accuracy. Similarly, integrating deep learning with optimization algorithms focuses on leveraging optimization algorithms to effectively search for optimal model hyperparameters, optimize feature engineering, and enhance the stability and accuracy of learning algorithms. Some researchers have also explored Transformer-based methods, improved models (Yao et al. 2022), and satellite remote sensing data (Wang et al. 2023b) for water quality prediction, although this research area requires further exploration and improvement.

There are alternative methods for water quality prediction aside from deep learning. Given the need to acquire substantial data to ensure accurate predictions, researchers have explored virtual sample generation to enhance prediction accuracy (El Bilali et al. 2022). This involves creating virtual samples to expand datasets and improve the model's generalization ability. Transfer learning is another approach where researchers pre-train the model in a source domain and then optimize it for the target domain to boost prediction accuracy (Cao et al. 2022; Chen et al. 2023). To address the interpretability issue of deep learning, some researchers combine physical models with deep learning models to achieve a balance between physical and data-driven approaches to enhance prediction accuracy and interpretability (Dong et al. 2023).

Comparison of different deep learning methods

In summary, different deep learning methods usually have different roles in water quality prediction, and their characteristics and disadvantages also differ, as shown in Table 1.

The advantage of hybrid models is that their prediction accuracy is usually higher than that of single models. However, this improvement comes at the expense of increasing the number of model parameters, which makes hyperparameter tuning more difficult and increases the computational cost.

There are various types of deep learning models, each with its own set of advantages and disadvantages. It is essential to choose the model that best fits the characteristics of the original data and the model itself. Once a deep learning model is selected, determining the appropriate number of layers in the neural network and other parameters, such as the size of the convolution kernel, is crucial, often conducted using empirical methods. Conducting multiple experiments, comparing prediction results, and fine-tuning parameters or utilizing meta-heuristic algorithms to determine optimal values is common practice. Moreover, before making predictions, data preprocessing techniques such as handling missing values and standardizing data are employed to minimize the impact of raw data on prediction accuracy.

Data sources for predicting water quality can encompass national water quality platforms, sensor networks, and collaborations with relevant companies to gather water quality data detected by these entities. Preprocessing data is crucial and involves outlier detection, missing value interpolation, and standardization. Outlier detection assists in removing the influence of anomalous data on the model while missing value interpolation maintains data integrity. Standardization harmonizes features of varying scales into a consistent standard scale, mitigating unit constraints and differences in initial data magnitudes. This process aims to enhance model training and prediction outcomes.

Model performance evaluation indicators and factors affecting deep learning technologies for water quality prediction

Model performance evaluation indicators

Performance metrics are statistical measures that assist developers in assessing and fine-tuning prediction

Table 1 The characteristics and disadvantages of deep learning technologies in water quality prediction

Deep learning method	Feature	Inferiority
CNN	• Suitable for both prediction and feature extraction • Spatial feature extraction ability • Capable of automated feature representation learning	• The ability to capture long-term dependence is weak • Data demand is large, and data quality requirements are high
TCN	• An improved CNN model for time series prediction • Extracting time features is more effective in capturing long-term dependencies than traditional RNNs • Enables parallel computing for high efficiency	• Lack of spatial correlation considerations • Data demand is large, and data quality requirements are high
RNN	• Handling temporal dependencies • Automatic feature extraction • Accommodates input sequences of varying lengths	• Gradient vanishing / explosion problem • Step-by-step calculation, low efficiency • Prediction error accumulation
LSTM	• An improved model of RNN to capture long-term dependencies • The gated mechanism is introduced, which can effectively alleviate the gradient disappearance/explosion problem, and the computational efficiency is also improved	• When dealing with short sequence data, it may be too complex and easy to over-fit • When facing a very long sequence, it may still face the problem of gradient disappearance or gradient explosion
GRU	• Improved LSTM model • Fewer parameters, higher computational efficiency	• Modeling complex sequence data may be slightly worse • Facing a very long sequence, it may still face the problem of gradient disappearance or gradient explosion
Transformer	• Has strong modeling ability • Supports parallel computing • The self-attention mechanism is used to model the context information in the sequence data	• Compared with traditional models such as RNNs, the requirements for data volume will be higher • Require a lot of computing resources when training and reasoning
DBN	• Versatile for both supervised and unsupervised learning • Consists of multiple RBMs, is capable of learning and representing complex patterns and relationships in input data • The forward propagation of DBNs can perform parallel computing and accelerate the process of training and reasoning	• Requires a large amount of training data and sufficient sample diversity to obtain better generalization ability • Large consumption of computing resources
Autoencoder	• Unsupervised learning, which does not require labeled training data, automatically learns feature representations from unlabeled data • Demonstrates robust modeling capabilities	• Generally not used directly for water quality prediction • Requires integration with supervised learning techniques, such as LSTM
CNN-LSTM	• Use CNN for feature extraction • Capture long-term dependencies with LSTM	• Compared with CNN and LSTM, the model's prediction parameters increase, and the calculation quantity increases • Large data demand and high data quality requirements
LSTM-TCN	• Better capture of historical data features	• A large number of parameters, difficult-to-optimize hyperparameters, and high calculation cost • Large data demand and high data quality requirements

performance on various platforms. Moreover, performance accuracy and effectiveness are translated into comprehensible and quantifiable formats. In the literature pertaining to water quality prediction, the primary model evaluation indicators are mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), and the coefficient of determination (R^2) (Irwan et al. 2023).

MAE

An objective quantitative evaluation of the model's prediction error. It measures the gap between the model's predicted and true values. A smaller MAE indicates a closer prediction result to the true value and a better prediction effect. MAE primarily focuses on the error size rather than the error distribution.

$$\mathrm{MAE} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left| \mathrm{y}_{\mathrm{pred}} - \mathrm{y}_{\mathrm{true}} \right|$$

MSE

An objective quantitative measure used to evaluate the model's prediction error. It calculates the square sum of the difference between the true and predicted values. A smaller MSE indicates that the predicted result is closer to the true result, indicating better model performance. However, MSE only focuses on quantifying the error and does not consider the error distribution.

$$\mathrm{MSE} = \frac{1}{\mathrm{n}} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left(\mathrm{y}_{\mathrm{pred}} - \mathrm{y}_{\mathrm{true}} \right)^2$$

RMSE

RMSE has the same effect and significance as MSE. The main difference is that RMSE places a higher penalty on samples with large errors, making it more sensitive to outliers.

$$\mathrm{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_{\mathrm{pred}} - y_{\mathrm{true}} \right)^2}$$

R.^{2}

R^2 is an objective quantitative measure that evaluates the degree of model fitting. It indicates how well a model fits data, with values ranging from 0 to 1. A value closer to 1 implies a better degree of fitting.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_{\mathrm{true},i} - y_{\mathrm{pred},i} \right)^2}{\sum_{i=1}^{n} \left(y_{\mathrm{true},i} - \overline{y_{\mathrm{true}}} \right)^2}$$

In the above four formulas, y_{true} is the true value, y_{pred} is the predicted value, $\overline{y_{\mathrm{true}}}$ is the average value of the true value, and n is the number of samples.

Factors affecting deep learning technologies in water quality prediction

Water quality indicators are categorized into biological, chemical, and physical indicators (Tchobanoglous and Schroeder 1985). Biological indicators encompass fecal coliforms and algae, while chemical indicators include dissolved oxygen, chemical oxygen demand, and ammonia nitrogen. Physical indicators consist of pH, temperature, and turbidity (Wu et al. 2014).

Various factors, such as climate change, geological terrain, soil type, hydrological characteristics, land use, and management, influence water quality (Lintern et al. 2018, Liu et al. 2017, Shi et al. 2017, Wilhm and Dorris 1968). These factors interact in intricate ways, resulting in multiple forms of pollution that significantly impact water quality. Therefore, when utilizing deep learning methods to predict water quality and construct water quality datasets, it is essential to collect different data types. The reasons for the diverse factors affecting water quality are as follows.

Water area

The impact of various water types on water quality varies, and each type exhibits spatial non-stationarity. Different water types exhibit different relationship models with water quality, attributable to variations in purification processes in rivers and lakes (Deng 2020).

Land use/cover change

River water quality is influenced by land use, with the extent of this effect depending on the specific river area and the spatial scale used to measure land use (Wang et al. 2023a). City and cultivated land showed a negative correlation with water quality, while forest land and water bodies exhibited a positive correlation with water quality (Zhang et al. 2019a).

Natural factors

Natural factors include rainfall, topography, and hydrogeology and can significantly influence water flow, oxygen content, and pollutant concentration, consequently impacting water quality. For instance, rainfall can alter water flow and velocity. At the same time, the terrain's fluctuation and slope direction can affect water flow velocity and direction, ultimately influencing water mixing and circulation.

Internal factors for water bodies

Water temperature, pH value, conductivity, turbidity, color, and redox potential play a significant role in determining water quality. For instance, a low water temperature can slow or impede certain chemical reactions, while a high water temperature can accelerate reaction rates. The pH value influences element dissolution in water, while the redox potential reflects the water's redox properties, which in turn affect the presence and exchange of oxygen and oxygen compounds as well as biological and chemical reactions. These indicators collectively contribute to overall water quality.

Human activities

Large-scale industrial, agricultural, and urban activities can significantly contribute to poor water quality. These activities involve the discharge of wastewater, pollutants, and other substances into water bodies, which disrupt the ecological balance and impair the regulation and self-purification capabilities of water.

Biological factors

Algae, bacteria, and plankton also play a role in influencing water quality. Algae, through photosynthesis, impact gas concentrations and proportions in water. Bacteria, through decomposing organic matter, affect chemical indicators and overall water quality. Excessive bacterial proliferation can lead to water body eutrophication. Plankton influences the nutritional status, color, turbidity, and oxygen concentration in water.

The limitations of current deep learning techniques in water quality prediction

Constraints of raw data availability

Deep learning models require a large quantity of data to achieve accurate prediction results. However, data collection constraints often restrict many studies to small-scale datasets. In the water quality prediction field, current deep learning techniques primarily utilize single-dimension raw data for modeling and prediction without fully considering other factors that may impact water quality, such as land use, forest coverage, and population. However, these additional factors play a significant role in water quality prediction. To obtain more precise and comprehensive prediction results, it is crucial to expand the dataset size and incorporate these influential factors. This approach can enhance the accuracy and practicality of water quality models to better support decision-making in water environment management.

Failure of data processing

Data processing methods, including wavelet transform, can be susceptible to errors resulting from data preprocessing and processing (Du et al. 2017, Quilty and Adamowski 2018). As the utilization of deep learning techniques for water quality management and prediction becomes more extensive, it becomes crucial to comprehend the errors and limitations of these models, particularly in relation to data selection and processing.

Challenges of long-term prediction

In the water quality prediction field, long-term prediction poses a significant challenge. Unlike short-term and medium-term forecasts, long-term forecasts involve intricate and adaptable spatiotemporal correlations, as well as increased uncertainties, resulting in reduced prediction accuracy. This is due to the diminishing impact of historical data on future predictions and the presence of multiple ambiguous features. Additionally, long-term prediction is influenced by uncertain factors such as the difficulty in accurately predicting weather changes, water mobility, human activities, and a lack of sufficient historical data to establish precise models.

Poor interpretability of models

Deep learning models have faced challenges due to the black box problem since their inception. These models' intricate structures and multiple parameters obscure the understanding of their operational mechanisms. While deep learning models produce more accurate prediction results compared with traditional models, the rationale and methodology behind parameter selection remain unclear.

Directions for future research

Model selection for optimal prediction

In order to achieve the best prediction results, it is crucial to select the appropriate model based on the characteristics and requirements of the task. However, the process of choosing the most suitable model still requires further research and exploration. Additionally, when dealing with prediction problems involving noise, it is important to consider whether data noise can impact the model's performance and quality. Therefore, it is necessary to conduct further studies on data processing methods that can effectively reduce noise in data and enhance prediction robustness.

Construction of high-dimensional datasets

In order to enhance the accuracy of water quality prediction, researchers should compile a comprehensive dataset that includes various indicators such as water quality variables, meteorological variables, population data, and forest coverage. By constructing datasets that incorporate these indicators, researchers can employ different analysis methods to investigate correlations between these indicators. This approach will facilitate a deeper understanding of water quality issues and their influencing factors and provide a scientific foundation for relevant departments and decision-makers to develop more effective water quality management strategies. By improving the accuracy and efficiency of water quality prediction, we can take prompt and precise measures to address water quality problems and ensure the sustainable development of human life and the ecological environment.

Enhancing long-term prediction accuracy

Current studies largely concentrate on short-term water quality predictions, which perform inadequately for long-term forecasting. Only a few studies have successfully achieved long-term predictions based on water quality principles. Analyzing the importance of features in long-term prediction and improving prediction accuracy are important issues. Factors such as optimizing feature selection, feature engineering, data processing, and considering model complexity play crucial roles in improving the accuracy of long-term water quality prediction. These improvements will provide a more reliable basis for decision-making in water quality monitoring and management, ultimately leading to more accurate and sustainable water quality protection and management.

Advancements in large model prediction

The complexity and multidimensionality of water quality prediction have prompted researchers to focus on large-scale models. These models can better capture the complex relationship among water quality indicators, thereby enhancing prediction accuracy. Additionally, increasing the depth and breadth of the network in these large models enhances their expressive power. Leveraging distributed computing and parallelization technology accelerates the training process, further improving efficiency. Although research on large-scale models in water quality prediction remains in the early stages, advancements in computing resources and algorithms are expected to drive further research in this area. This will yield enhanced accuracy and reliability of water quality prediction, providing crucial support for macro-level water quality control.

Bridging academic research and practical application

The majority of the literature focuses on assessing the viability of deep learning techniques in predicting water quality, with the goal of enhancing the accuracy and accessibility of such predictions. However, these studies often fail to provide explicit guidelines on effectively connecting academic research with industry and government management. Consequently, it is pivotal for scholars to actively engage in the practical application of deep learning technologies to facilitate precise water quality assessment and management.

Conclusion

In recent years, deep learning technology has been widely used in water quality prediction, yielding positive results. Its effectiveness lies in handling high-dimensional feature representation and nonlinear relationships within water quality data. Through multi-layer nonlinear processing units, more intricate structures can be constructed to better model data. Deep learning excels in processing large-scale data efficiently, utilizing batch processing and parallel computing to handle massive and high-dimensional water quality data to support effective water quality prediction. Furthermore, deep learning models can autonomously learn and be iteratively optimized to enhance prediction accuracy over time. However, the success of deep learning technology hinges on high-quality datasets. Original water quality data often includes missing values, outliers, and noise, which impact prediction accuracy. Despite outperforming physical and statistical models in prediction accuracy, deep learning models are criticized for their lack of interpretability, often referred to as a 'black box.' Moreover, training deep learning models demands significant computing resources, necessitating high hardware and computing power requirements.

To further advancements in water quality prediction research and application, it is crucial to integrate various technologies including machine learning, data mining, cloud computing, multi-source data fusion, and deep reinforcement learning. Data mining plays a key role in uncovering underlying rules and relationships, offering valuable insights for water quality prediction. Cloud computing and distributed platforms provide the necessary computational power for handling large-scale water quality data, while multi-source data fusion enhances monitoring accuracy and temporal resolution. Deep reinforcement learning optimizes decision-making processes for water quality treatment, thereby enhancing overall efficacy. Furthermore, exploring the interpretability of deep learning models in water quality prediction enhances model credibility and practicality. Simplifying algorithms and computational requirements, along with promoting understanding of deep learning methods through educational resources, can greatly support the widespread application of deep learning in water quality prediction.

Interdisciplinary collaboration is essential for advancing water quality prediction research. Environmental scientists can utilize sensor networks and remote monitoring technology developed by embedded engineers to access up-to-date water quality data in real time. By leveraging their domain knowledge, environmental scientists can identify key water quality characteristics and factors, creating datasets for water quality prediction. Artificial intelligence researchers can then use these datasets to develop feature importance analysis methods and prediction models. Subsequently, environmental scientists can analyze the prediction results using their expertise in water systems and provide feedback to AI researchers. This collaborative effort leads to the innovation and optimization of water quality prediction models, ultimately enabling real-time monitoring and efficient water quality prediction.

The application of deep learning technology in water quality prediction can enhance prediction accuracy and monitoring quality to provide strong support for water environment protection and management. Further research and exploration of deep learning technology in water environment protection can contribute to promoting water environment improvement and sustainable development.

Author contribution Wenhao Li: investigation, data curation, visualization, writing—original draft preparation, and funding acquisition. Yin Zhao: data curation and writing—original draft preparation. Yining Zhu: data curation. Zhongtian Dong: writing—review and editing. Fenghe Wang: conceptualization, funding acquisition, project administration, and writing—review and editing. Fengliang Huang: conceptualization, funding acquisition, project administration, and writing—review and editing. All authors read and approved the manuscript.

Funding This work was supported by the Yangtze River ecological Environment protection and restoration joint research project II (2022-LHYJ-02–0502-05) and Postgraduate Research & Practice Innovation Program of Jiangsu Province (SJCX23_0570).

Declarations

Ethical Approval The authors declare that the manuscript has not been published previously.

Consent to participate All authors voluntarily participated in this research study.

Consent to publish All authors consent to the publication of the manuscript.

Competing interests The authors declare no competing interests.

References

Avila RG, Horn B, Moriarty EM, Hodson R, Moltchanova E, Joem J (2018) Evaluating statistical model performance in water quality prediction. J Environ Manag 206:910–919. https://doi.org/10.1016/j.jenvman.2017.11.049

Badrinarayanan V, Kendall A, Cipolla RJIToPA, Intelligence M (2015) SegNet: a deep convolutional encoder-decoder architecture for image segmentation. 39:2481–2495

Bai S, Kolter JZ, Koltun VJA (2018) An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. abs/1803.01271

Banan A, Nasiri A, Taheri-Garavand A (2020) Deep learning-based appearance features extraction for automated carp species identification. Aquacult Eng 89:102053. https://doi.org/10.1016/j.aquaeng.2020.102053

Cao H, Xie X, Shi J, Jiang G, Wang Y-xJEs, technology (2022) Siamese network-based transfer learning model to predict geogenic contaminated groundwaters

Chen HL, Yang JB, Fu XH, Zheng QX, Song XY, Fu ZD, Wang JC, Liang YQ, Yin HL, Liu ZM, Jiang J, Wang H, Yang XX (2022) Water quality prediction based on LSTM and attention mechanism: a case study of the Burnett River, Australia. Sustainability 14. https://doi.org/10.3390/su142013231

Chen S, Huang J, Wang P, Tang X, Zhang ZJWr (2023) A coupled model to improve river water quality prediction towards addressing non-stationarity and data limitation. 248:120895

Chen Y, Cheng Q, Cheng Y, Yang H, Yu H (2018) Applications of Recurrent neural networks in environmental factor forecasting: a review. Neural Comput 30:2855–2881. https://doi.org/10.1162/neco_a_01134

Cheng TY, Harrou F, Kadri F, Sun Y, Leiknes T (2020) Forecasting of wastewater treatment plant key features using deep learning-based models: a case study. Ieee Access 8:184475–184485. https://doi.org/10.1109/Access.2020.3030820

Chung J, Gülçehre Ç, Cho K, Bengio YJA (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. abs/1412.3555

Deng XJ (2020) Influence of water body area on water quality in the southern Jiangsu Plain, eastern China. J Clean Prod 254:120136. https://doi.org/10.1016/j.jclepro.2020.120136

Dong L, Zhang J (2021) Predicting polycyclic aromatic hydrocarbons in surface water by a multiscale feature extraction-based deep learning approach. Sci Total Environ 799:149509. https://doi.org/10.1016/j.scitotenv.2021.149509

Dong W, Zhang Y, Zhang L, Ma W, Luo L (2023) What will the water quality of the Yangtze River be in the future?. Sci Total Environ 857. https://doi.org/10.1016/j.scitotenv.2022.159714

Du KC, Zhao Y, Lei JQ (2017) The incorrect usage of singular spectral analysis and discrete wavelet transform in hybrid models to predict hydrological time series. J Hydrol 552:44–51. https://doi.org/10.1016/j.jhydrol.2017.06.019

El Bilali A, Lamane H, Taleb A, Nafii AJJoCP (2022) A framework based on multivariate distribution-based virtual sample generation and DNN for predicting water quality with small data

Essam Y, Huang YF, Birima AH, Ahmed AN, El-Shafie A (2022) Predicting suspended sediment load in Peninsular Malaysia using support vector machine and deep learning algorithms. Sci Rep 12:302. https://doi.org/10.1038/s41598-021-04419-w

Fan W, Zhang Z JnICoAiCT, Information Science, Communications (2020) a CNN-SVR hybrid prediction model for wastewater index measurement. 90–94

Farsi M, Hosahalli D, Manjunatha BR, Gad I, Gad I, Atlam E-S, Atlam E-S, Ahmed A, Elmarhomy G, Elmarhoumy M, Ghoneim OAJaej (2020) Parallel genetic algorithms for optimizing the SARIMA model for better forecasting of the NCDC weather data

Geng JX, Yang CH, Li YG, Lan LJ, Luo QW (2022) MPA-RNN: a novel attention-based recurrent neural networks for total nitrogen prediction. IEEE Trans Industr Inf 18:6516–6525. https://doi.org/10.1109/Tii.2022.3161990

Gers FA, Schmidhuber J, Cummins F (2000) Learning to forget: continual prediction with LSTM. Neural Comput 12:2451–2471. https://doi.org/10.1162/089976600300015015

Gers FA, Schraudolph NN, Schmidhuber J (2003) Learning precise timing with LSTM recurrent networks. J Mach Learn Res 3:115–143. https://doi.org/10.1162/153244303768966139

Graves A, Schmidhuber J (2005) Framewise phoneme classification with bidirectional LSTM and other neural network architectures. Neural Netw 18:602–610. https://doi.org/10.1016/j.neunet.2005.06.042

Gruber N, Jockisch A (2020) Are GRU cells more specific and LSTM cells more sensitive in motive classification of text? Front Artif Intell 3:40. https://doi.org/10.3389/frai.2020.00040

Guan G, Wang Y, Yang L, Yue J, Li Q, Lin J, Liu Q (2022) Water-quality assessment and pollution-risk early-warning system based on web crawler technology and LSTM. Int J Environ Res Public Health 19. https://doi.org/10.3390/ijerph191811818

Guo D, Lintern A, Webb JA, Ryu D, Bende-Michl U, Liu S, Western AWJH, Sciences ES (2020) A data-based predictive model for spatiotemporal variability in stream water quality

Guo JJ, Dong JQ, Zhou B, Zhao XH, Liu SY, Han QY, Wu HL, Xu LQ, Hassan SG (2022) A hybrid model for the prediction of dissolved oxygen in seabass farming. Comput Electron Agric 198:106971. https://doi.org/10.1016/j.compag.2022.106971

Habib G, Qureshi S (2020) Optimization and acceleration of convolutional neural networks: a survey. Journal of King Saud University - Computer and Information Sciences 34:4244–4268. https://doi.org/10.1016/j.jksuci.2020.10.004

He M, Wu SF, Huang BB, Kang CX, Gui FL (2022) Prediction of total nitrogen and phosphorus in surface water by deep learning methods based on multi-scale feature extraction. Water 14. https://doi.org/10.3390/w14101643

Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Comput 9:1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Irwan D, Ali M, Ahmed AN, Jacky G, Nurhakim A, Han MCP, AlDahoul N, El-Shafie A (2023) Predicting water quality with artificial intelligence: a review of methods and applications. Arch Comput Methods Eng 30:4633–4652. https://doi.org/10.1007/s11831-023-09947-4

Islam N, Irshad K (2022) Artificial ecosystem optimization with deep learning enabled water quality prediction and classification model. Chemosphere 309:136615. https://doi.org/10.1016/j.chemosphere.2022.136615

Juan H, Li MB, Xu XG, Hao Z, Yang BE, Jiang JM, Bing S (2022) Multi-step prediction of dissolved oxygen in river based on random forest missing value imputation and attention mechanism coupled with recurrent neural network. Water Supply 22:5480–5493. https://doi.org/10.2166/ws.2022.154

Kayalvizhi S, Jiavana KFK, Suganthi K, Malarvizhi S (2023) Prediction of ground water quality in western regions of Tamil Nadu using deep auto encoders. Urban Climate 49. https://doi.org/10.1016/j.uclim.2023.101458

Khullar S, Singh N (2022) Water quality assessment of a river using deep learning Bi-LSTM methodology: forecasting and validation. Environ Sci Pollut Res Int 29:12875–12889. https://doi.org/10.1007/s11356-021-13875-w

Kumar KK, Kumar MD, Samsonu C, Krishna KVJMTP (2021) Role of convolutional neural networks for any real time image classification, recognition and analysis

Larijani A, Dehghani FJF (2023) An efficient optimization approach for designing machine models based on combined algorithm

LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521:436–444. https://doi.org/10.1038/nature14539

Lee S, Lee DJIJoER, Health P (2018) Improved prediction of harmful algal blooms in four major South Korea's rivers using deep learning models. 15

Li W, Wei Y, An D, Jiao Y, Wei Q (2022) LSTM-TCN: dissolved oxygen prediction in aquaculture, based on combined model of long short-term memory network and temporal convolutional network. Environ Sci Pollut Res Int 29:39545–39556. https://doi.org/10.1007/s11356-022-18914-8

Li Y, Li R (2023) Predicting ammonia nitrogen in surface water by a new attention-based deep learning hybrid model. Environ Res 216:114723. https://doi.org/10.1016/j.envres.2022.114723

Lintern A, Webb JA, Ryu D, Liu S, Bende-Michl U, Waters D, Leahy P, Wilson P, Western AW (2018) Key factors influencing differences in stream water quality across space. Wiley Interdisciplinary Reviews-Water 5. https://doi.org/10.1002/wat2.1260

Liu J, Zhang X, Wu B, Pan G, Xu J, Wu S (2017) Spatial scale and seasonal dependence of land use impacts on riverine water quality in the Huai River basin, China. Environ Sci Pollut Res Int 24:20995–21010. https://doi.org/10.1007/s11356-017-9733-7

Liu JT, Yu C, Hu ZH, Zhao YC, Bai Y, Xie MS, Luo J (2020) Accurate prediction scheme of water quality in smart mariculture with deep Bi-S-SRU learning network. Ieee Access 8:24784–24798. https://doi.org/10.1109/Access.2020.2971253

Liu P, Wang J, Sangaiah AK, Xie Y, Yin XC (2019a) Analysis and prediction of water quality using LSTM deep neural networks in IoT environment. Sustainability 11. https://doi.org/10.3390/su11072058

Liu X, Shi QM, Liu Z, Yuan J (2021) Using LSTM neural network based on improved PSO and attention mechanism for predicting the effluent COD in a wastewater treatment plant. Ieee Access 9:146082–146096. https://doi.org/10.1109/Access.2021.3123225

Liu YQ, Zhang Q, Song LH, Chen YY (2019b) Attention-based recurrent neural networks for accurate short-term and long-term dissolved oxygen prediction. Computers and Electronics in Agriculture 165. https://doi.org/10.1016/j.compag.2019.104964

Liu Z, Tong S (2011) Using HSPF to model the hydrologic and water quality impacts of riparian land-use change in a small watershed. J Environ Inform. https://doi.org/10.3808/jei.201100182

Lyu H-M, Shen S, Zhou AJJoCP (2020) The development of IFN-SPA: a new risk assessment method of urban water quality and its application in Shanghai. 124542

Magar MR, Khatry SBJAJoW, Environment, Pollution (2017) Vollenweider model for temporal eutrophication characteristics of Nagdaha Lake, Nepal. 14, 29-39

Malki A, Atlam E, Gad IJAEJ (2022) Machine learning approach of detecting anomalies and forecasting time-series of IoT devices

Mei P, Li M, Zhang Q, Li GL, Song L (2022) Prediction model of drinking water source quality with potential industrial-agricultural pollution based on CNN-GRU-Attention. J Hydrol 610. https://doi.org/10.1016/j.jhydrol.2022.127934

Mohamed AR, Dahl GE, Hinton G (2012) Acoustic modeling using deep belief networks. IEEE Trans Audio Speech Lang Process 20:14–22. https://doi.org/10.1109/Tasl.2011.2109382

Moshtaghi B, Niksokhan MH, Ghazban F, Dalilsafaee SJI, Drainage (2018) Assessing the impacts of climate change on the quantity and quality of agricultural runoff (case study: Golgol River Basin). 67, 17 - 28

Niu GQ, Yi XH, Chen C, Li XY, Han DH, Yan B, Huang MZ, Ying GG (2020) A novel effluent quality predicting model based on genetic-deep belief network algorithm for cleaner production in a full-scale paper-making wastewater treatment. J Clean Prod 265:121787. https://doi.org/10.1016/j.jclepro.2020.121787

Otter D, Medina JR, Kalita JKJIToNN, Systems L (2020) A survey of the usages of deep learning for natural language processing. 32, 604-624

Post CJ, Cope MP, Gerard PD, Masto NM, Vine JR, Stiglitz RY, Hallstrom JO, Newman JC, Mikhailova EA (2018) Monitoring spatial and temporal variation of dissolved oxygen and water temperature in the Savannah River using a sensor network. Environ Monit Assess 190:272. https://doi.org/10.1007/s10661-018-6646-y

Prasad DVV, Venkataramana LY, Kumar PS, Prasannamedha G, Harshana S, Srividya SJ, Harrinei K, Indraganti S (2022) Analysis and prediction of water quality using deep learning and auto deep learning techniques. Sci Total Environ 821:153311. https://doi.org/10.1016/j.scitotenv.2022.153311

Pyo J, Park LJ, Pachepsky Y, Baek SS, Kim K, Cho KH (2020) Using convolutional neural network for predicting cyanobacteria concentrations in river water. Water Res 186:116349. https://doi.org/10.1016/j.watres.2020.116349

Quilty J, Adamowski J (2018) 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 563:336–353. https://doi.org/10.1016/j.jhydrol.2018.05.003

Rawat W, Wang Z (2017) Deep convolutional neural networks for image classification: a comprehensive review. Neural Comput 29:2352–2449. https://doi.org/10.1162/NECO_a_00990

Ren Q, Wang XY, Li WS, Wei YG, An D (2020) Research of dissolved oxygen prediction in recirculating aquaculture systems based on deep belief network. Aquacult Eng 90:102085. https://doi.org/10.1016/j.aquaeng.2020.102085

Rong QQ, Cai YP, Su MR, Yue WC, Dang Z, Yang ZF (2019) Identification of the optimal agricultural structure and population size in a reservoir watershed based on the water ecological carrying capacity under uncertainty. J Clean Prod 234:340–352. https://doi.org/10.1016/j.jclepro.2019.06.179

Saroja, Haseena, Dharshini S (2023): Deep learning approach for prediction and classification of potable water. Anal Sci 39, 1179-1189.https://doi.org/10.1007/s44211-023-00328-2

Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117. https://doi.org/10.1016/j.neunet.2014.09.003

Sezer OB, Gudelek MU, Ozbayoglu AMJA (2019) Financial time series forecasting with deep learning : a systematic literature review: 2005–2019. abs/1911.13288

Shan K, Ouyang T, Wang XX, Yang H, Zhou BT, Wu ZX, Shang MS (2022) Temporal prediction of algal parameters in Three Gorges Reservoir based on highly time-resolved monitoring and long short-term memory network. J Hydrol 605. https://doi.org/10.1016/j.jhydrol.2021.127304

Shi B, Bach PM, Lintern A, Zhang K, Coleman RA, Metzeling L, Mccarthy DT, Deletic AJJoem (2019) Understanding spatiotemporal variability of in-stream water quality in urban environments - a case study of Melbourne, Australia. 246, 203-213

Shi P, Zhang Y, Li ZB, Li P, Xu GC (2017) Influence of land use and land cover patterns on seasonal water quality at multi-spatial scales. CATENA 151:182–190. https://doi.org/10.1016/j.catena.2016.12.017

Song CG, Yao LH, Hua CY, Ni QH (2021) A novel hybrid model for water quality prediction based on synchrosqueezed wavelet transform technique and improved long short-term memory. J Hydrol 603:126879. https://doi.org/10.1016/j.jhydrol.2021.126879

Ta XX, Wei YG (2018) Research on a dissolved oxygen prediction method for recirculating aquaculture systems based on a convolution neural network. Comput Electron Agric 145:302–310. https://doi.org/10.1016/j.compag.2017.12.037

Tchobanoglous G, Schroeder ED (1985) Water quality : characteristics, modeling, modification

Tirkolaee EB, Hosseinabadi AAR, Soltani M, Sangaiah AK, Wang J (2018) A hybrid genetic algorithm for multi-trip green capacitated arc routing problem in the scope of urban services. Sustainability 10:1366. https://doi.org/10.3390/su10051366

Vaswani A, Shazeer NM, Parmar N, Uszkoreit J, Jones L, Gomez AN, Kaiser L, Polosukhin I (2017) Attention is all you need, neural information processing systems

Vörösmarty CJ, McIntyre PB, Gessner MO, Dudgeon D, Prusevich AA, Green P, Glidden SJ, Bunn SE, Sullivan CA, Liermann CR, Davies PMJN (2010) Global threats to human water security and river biodiversity. 468, 334-334

Wan H, Mao Y, Cai Y, Li R, Feng J, Yang HJEwC (2021) An SPH-based mass transfer model for simulating hydraulic characteristics and mass transfer process of dammed rivers. 38, 3169 - 3184

Wang H, Xiong X, Wang K, Li X, Hu H, Li Q, Yin H, Wu C (2023a) The effects of land use on water quality of alpine rivers: a case study in Qilian Mountain. China Sci Total Environ 875:162696. https://doi.org/10.1016/j.scitotenv.2023.162696

Wang L, Li W, Wang X, Xu JJPCS (2023b) Remote sensing image analysis and prediction based on improved Pix2Pix model for water environment protection of smart cities. 9

Wang ZC, Wang QY, Wu TH (2023c) A novel hybrid model for water quality prediction based on VMD and IGOA optimized for LSTM. Front Environ Sci Eng 17. https://doi.org/10.1007/s11783-023-1688-y

Wilhm JL, Dorris TCJB (1968) Biological parameters for water quality criteria. 18, 477-481

Wool T, Ambrose RB, Martin JL, Comer ARJJW (2020) WASP 8: the next generation in the 50-year evolution of USEPA's water quality model. 12, 1398 - 1398

Wu WY, Dandy GC, Maier HR (2014) Protocol for developing ANN models and its application to the assessment of the quality of the ANN model development process in drinking water quality modelling. Environ Model Softw 54:108–127. https://doi.org/10.1016/j.envsoft.2013.12.016

Xiang ZR, Demir I (2020) Distributed long-term hourly streamflow predictions using deep learning - a case study for state of Iowa. Environ Model Softw 131:104761. https://doi.org/10.1016/j.envsoft.2020.104761

Xue Y, Zhu L, Zou B, Wen Y-M, Long Y, Zhou S-lJW (2021) Research on inversion mechanism of chlorophyll—a concentration in water bodies using a convolutional neural network model

Yan JZ, Gao Y, Yu YC, Xu HX, Xu ZB (2020) A prediction model based on deep belief network and least squares SVR applied to cross-section water quality. Water 12:1929. https://doi.org/10.3390/w12071929

Yan JZ, Liu JX, Yu YC, Xu HX (2021) Water quality prediction in the Luan River based on 1-DRCNN and BiGRU hybrid neural network model. Water 13. https://doi.org/10.3390/w13091273

Yang B, Xiao Z, Meng Q, Yuan Y, Wang W, Wang H, Wang Y, Feng X (2023a) Deep learning-based prediction of effluent quality of a

constructed wetland. Environ Sci Ecotechnol 13:100207. https://doi.org/10.1016/j.ese.2022.100207
Yang H, Wang Z, Song KJEwC (2020) A new hybrid grey wolf optimizer-feature weighted-multiple kernel-support vector regression technique to predict TBM performance. 38, 2469 - 2485
Yang H, Liu S (2022) Water quality prediction in sea cucumber farming based on a GRU neural network optimized by an improved whale optimization algorithm. PeerJ Comput Sci 8:e1000. https://doi.org/10.7717/peerj-cs.1000
Yang JC, Jia LL, Guo ZW, Shen Y, Li XW, Mou ZP, Yu KP, Lin JCW (2023b) Prediction and control of water quality in recirculating aquaculture system based on hybrid neural network. Eng Appl Artif Intell 121:106002. https://doi.org/10.1016/j.engappai.2023.106002
Yang W, Liu W, Gao Q (2023c) Prediction of dissolved oxygen concentration in aquaculture based on attention mechanism and combined neural network. Math Biosci Eng 20:998–1017. https://doi.org/10.3934/mbe.2023046
Yang X, Liu Q, Luo X, Zheng ZJSR (2017) Spatial regression and prediction of water quality in a watershed with complex pollution sources. 7
Yang Y, Xiong Q, Wu C, Zou Q, Yu Y, Yi H, Gao M (2021) A study on water quality prediction by a hybrid CNN-LSTM model with attention mechanism. Environ Sci Pollut Res Int 28:55129–55139. https://doi.org/10.1007/s11356-021-14687-8
Yao S, Zhang Y, Wang P, Xu Z, Wang Y, Zhang YJAS (2022) Long-term water quality prediction using integrated water quality indices and advanced deep learning models: a case study of Chaohu Lake, China, 2019–2022
Zamani B, Koch M, Hodges BR, Fakheri-Fard AJJoAWE, Research (2018) Pre-impoundment assessment of the limnological processes and eutrophication in a reservoir using three-dimensional modeling: Abolabbas reservoir, Iran. 6, 48 - 61
Zamani MG, Nikoo MR, Jahanshahi S, Barzegar R, Meydani AJES, Research P (2023) Forecasting water quality variable using deep learning and weighted averaging ensemble models. 30, 124316-124340
Zhang DY, Chang RK, Wang HS, Wang Y, Wang H, Chen SQ (2021) Predicting water quality based on EEMD and LSTM networks. Proceedings of the 33rd Chinese Control and Decision Conference (Ccdc 2021), 2372–2377. https://doi.org/10.1109/Ccdc52312.2021.9602800
Zhang J, Li SY, Dong RZ, Jiang CS, Ni MF (2019a) Influences of land use metrics at multi-spatial scales on seasonal water quality: a case study of river systems in the Three Gorges Reservoir Area, China. J Clean Prod 206:76–85. https://doi.org/10.1016/j.jclepro.2018.09.179
Zhang JH, Yin Z, Chen P, Nichele S (2020) Emotion recognition using multi-modal data and machine learning techniques: a tutorial and review. Information Fusion 59:103–126. https://doi.org/10.1016/j.inffus.2020.01.011
Zhang X, Li D (2022) Multi-input multi-output temporal convolutional network for predicting the long-term water quality of ocean ranches. Environ Sci Pollut Res 30:7914–7929
Zhang X, Chen X, Zheng G, Cao G (2023) Improved prediction of chlorophyll-a concentrations in reservoirs by GRU neural network based on particle swarm algorithm optimized variational modal decomposition. Environ Res 221:115259. https://doi.org/10.1016/j.envres.2023.115259
Zhang Y, Fitch P, Thorburn PJJW (2019b) Predicting the trend of dissolved oxygen based on the kPCA-RNN model
Zhang Y, Thorburn PJ, Fitch P (2019c) Multi-task temporal convolutional network for predicting water quality sensor data, International Conference on Neural Information Processing
Zhao R, Yan RQ, Chen ZH, Mao KZ, Wang P, Gao RX (2019) Deep learning and its applications to machine health monitoring. Mech Syst Signal Process 115:213–237. https://doi.org/10.1016/j.ymssp.2018.05.050
Zhou W, Zhu Z, Xie Y-f, Cai YJJoH (2021) Impacts of rainfall spatial and temporal variabilities on runoff quality and quantity at the watershed scale
Zhou YL (2020) Real-time probabilistic forecasting of river water quality under data missing situation: deep learning plus post-processing techniques. J Hydrol 589:125164. https://doi.org/10.1016/j.jhydrol.2020.125164

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.