

Article

Prediction of Typical Power Plant Circulating Cooling Tower Blowdown Water Quality Based on Explicable Integrated Machine Learning

Yongjie Wan ¹, Xing Tian ¹, Hanhua He ¹, Peng Tong ², Ruiying Gao ¹, Xiaohui Ji ¹, Shaojie Li ¹, Shan Luo ¹, Wei Li ¹ and Zhenguo Chen ^{3,*}

¹ China Energy Engineering Group Guangdong Electric Power Design Institute Co., Ltd., Guangzhou 510700, China; wanyongjie@gedi.com.cn (Y.W.); tianxing@gedi.com.cn (X.T.); hehanhua@gedi.com.cn (H.H.); gaoruiying@gedi.com.cn (R.G.); jixiaohui@gedi.com.cn (X.J.); lishaojie@gedi.com.cn (S.L.); luoshan@gedi.com.cn (S.L.); g3426@gedi.com.cn (W.L.)

² Huaneng Dongguan Gas Turbine Thermoelectricity Co., Ltd., Dongguan 523593, China; hndggcb2@163.com

³ School of Environment, South China Normal University, Guangzhou 510006, China

* Correspondence: zhenguo.chen@m.scnu.edu.cn

Abstract: This paper establishes an explicable integrated machine learning model for predicting the discharge water quality in a circulating cooling water system of a power plant. The performance differences between three deep learning models, a Temporal Convolutional Network (TCN), Long Short-Term Memory (LSTM), and a Convolutional Neural Network (CNN), and traditional machine learning models, namely eXtreme Gradient Boosting (XGboost) and Support Vector Machine (SVM), were evaluated and compared. The TCN model has high fitting accuracy and low error in predicting ammonia nitrogen, nitrate nitrogen, total nitrogen, chemical oxygen demand (COD), and total phosphorus in the effluent of a circulating cooling tower. Compared to other traditional machine learning models, the TCN has a larger R^2 (maximum 0.911) and lower Root Mean Square Error (RMSE, minimum 0.158) and Mean Absolute Error (MAE, minimum 0.118), indicating the TCN has better feature extraction and fitting performance. Although the TCN takes additional time, it is generally less than 1 s, enabling the real-time prediction of drainage water quality. The main water quality indices have the greatest causal inference relationship with those of makeup water, followed by the concentration ratio, indicating that concentrations of ammonia nitrogen, nitrate nitrogen, total nitrogen, and COD have a more decisive impact. Shapley Additive Explanations (SHAP) analysis further reveals that the concentration ratio has a weaker decisive impact on circulating cooling water drainage quality. The results of this study facilitate the optimization of industrial water resource management and offer a feasible technical pathway for water resource utilization in power plants.

Keywords: machine learning; circulating cooling water system; water quality prediction; causal inference; SHAP



Academic Editors: Guocheng Zhu and Andrew S. Hursthouse

Received: 2 April 2025

Revised: 8 May 2025

Accepted: 9 May 2025

Published: 17 June 2025

Citation: Wan, Y.; Tian, X.; He, H.; Tong, P.; Gao, R.; Ji, X.; Li, S.; Luo, S.; Li, W.; Chen, Z. Prediction of Typical Power Plant Circulating Cooling Tower Blowdown Water Quality Based on Explicable Integrated Machine Learning. *Processes* **2025**, *13*, 1917. <https://doi.org/10.3390/pr13061917>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In China, the management of cooling water used in energy production processes has become increasingly important, especially in regions with limited water resources. In some chemical production processes, circulating cooling water accounts for 70% to 80% of total water usage, making it a primary target for water conservation and technological improvement. Power plants, being among the most water-intensive industrial sectors, use cooling processes that contribute significantly to their overall water consumption [1,2].

These systems often rely on treated reclaimed water—such as effluent from sewage treatment plants—as intake water, which promotes wastewater reuse and sustainable development [3].

However, the recirculating cooling water systems in power plants face unique challenges. As water continuously evaporates in cooling towers, salt and contaminant concentrations rise, requiring blowdown to maintain water chemistry balance. Increasing the concentration ratio has been adopted as a water-saving strategy, yet it also leads to more complex discharge water quality due to rising impurity concentrations and variable reclaimed water quality. Compounding these challenges, power plants in China generally exhibit low informatization levels, lack real-time monitoring at key water flow nodes, and rely on scattered or delayed data, which hinders rapid anomaly detection and informed decision-making [4].

To predict discharge water quality and guide process control, researchers have adopted mechanistic and data-driven modeling approaches [5]. While mechanistic models attempt to describe the underlying physicochemical processes, they are often limited by their complexity, reliance on numerous poorly understood parameters, and lack of scalability [6]. In contrast, data-driven machine learning models such as Support Vector Machine (SVM, imported via Python's SVM library as `from sklearn.svm import SVR`), XGBoost, Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM), and Temporal Convolutional Networks (TCNs) have shown promise in capturing nonlinear relationships and temporal patterns in water quality data [7–11]. However, most existing studies focus on a single modeling approach without systematically comparing multiple deep learning and traditional machine learning models under consistent datasets. Furthermore, these models are often treated as “black boxes”, lacking interpretability, which limits their practical utility in guiding operational decisions in industrial contexts [12–14].

To address these gaps, this study presents a comprehensive framework for predicting the discharge water quality in power plant circulating cooling systems. Specifically, we systematically compare the performance of multiple deep learning models (TCN, LSTM, CNN) and machine learning models (XGBoost, SVM) using the same industrial dataset to determine the most effective predictive model, demonstrate the superior ability of the TCN model to capture long- and short-term dependencies in time series water quality data, and apply interpretable AI techniques such as SHAP (Shapley Additive Explanations) and causal inference to analyze the contribution of input variables and clarify causal relationships between inflow characteristics and discharge water quality. This integrated approach not only improves prediction accuracy but also enhances model transparency and interpretability, providing practical guidance for intelligent water quality management. The findings offer a novel technical pathway for optimizing cooling water reuse efficiency and mitigating environmental risks in power plant operations.

2. Materials and Methods

2.1. Sample Data from Paper Industrial Wastewater Treatment Plant

In a simulation experiment, samples from the influent and effluent (i.e., make-up water and discharge water) in a cooling circulation tower of a power plant in Guangdong Province were collected from October 2024 to December 2024. The collected samples were frozen and sent to the laboratory for an analysis of major water quality indicators, including pH, ammonia nitrogen, nitrate nitrogen, total nitrogen, COD, total phosphorus, and Cl^- concentration multiple ratio (concentration ratio), after being filtered through 0.45 μm syringe filters. All water quality indicators were measured according to standard methods [15].

The data used in this study included the concentrations of influent ammonia nitrogen, nitrate nitrogen, total nitrogen, total phosphorus, and concentration ratio as inputs and the concentrations of effluent ammonia nitrogen, nitrate nitrogen, total nitrogen, COD, and total phosphorus as outputs (shown in Table S1). The dataset was split into training and test sets in an 8:2 ratio, where the number of samples in the training set $n = 100$ and the number of samples in the test set $n = 26$. Due to the small sample size of the data, the test set also served as the validation set.

2.2. Data Preprocessing

Before feeding the datasets into a model, normalization needed to be performed on both the training and testing datasets to accelerate the convergence speed of the model and eliminate the influence of dimensions. After the removal of outliers and nulls, normalization involves using specific algorithms to keep the data within a certain range. The method used in this study was normalization by a linear function. The expression is as follows [16]:

$$Y = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (1)$$

In the formula, Y is the normalized data, x represents the original value of the corresponding variable, and $x_{\max}$ and $x_{\min}$ are the maximum and minimum values of the corresponding variable.

2.3. TCN Neural Network

A TCN is a type of deep learning model used for processing sequence data. The TCN model extracts features from sequence data through a series of one-dimensional convolutional layers. Compared to RNNs, a TCN improves training efficiency through parallel computation and addresses the vanishing gradient problem in long sequence dependencies. As the water quality data are a type of time series data, the TCN model was employed to construct a prediction model for cooling water discharge. The causal convolution in the TCN model ensures the temporal order of the input sequence, dilated convolution enlarges the receptive field, residual blocks enhance the model's predictive accuracy, and additionally, the model incorporates Dropout layers to prevent overfitting.

A TCN has two core features: causal convolution and dilated convolution [17]. Causal convolution ensures that the output depends only on current and previous inputs in time, thus maintaining sequence consistency of the time series. Dilated convolution introduces gaps in the convolution computation, allowing the receptive field of the convolutional kernel to expand without increasing the number of parameters or computational complexity, thereby capturing longer sequence dependencies.

The basic formula of the TCN model can be described as follows:

Causal Convolution [17]:

$$y[t] = \sum_{k=0}^{K-1} \omega[k] \cdot x[t - k] \quad (2)$$

where y represents the convolution output, x is the input signal at time step $t - k$, ω is the convolution kernel parameter, and K is the size of the convolution kernel.

Dilated Convolution [17]:

$$y[t] = \sum_{k=0}^{K-1} \omega[k] \cdot x[t - d \cdot k] \quad (3)$$

Here, d is the dilation rate, which can control the receptive field size by adjusting d .

A TCN alleviates the gradient vanishing problem by stacking multiple layers of this convolutional structure combined with ResNet-style skip connections.

2.4. Construction of TCN Prediction Model

A TCN has a strong feature extraction capability, which can integrate original features into high-dimensional abstract features, thereby more efficiently uncovering specific information content. When constructing a TCN for a water quality prediction model, it is necessary to design the specific network architecture. The core mechanisms of a TCN include causal convolution and dilated convolution, which together form the foundation of the model. Causal convolution ensures that the model's output depends only on the current and past inputs, thereby ensuring that the causal relationship of the time series is maintained. Dilated convolution introduces gaps in the convolution computation, allowing the model to capture longer-range temporal dependencies without increasing the number of parameters.

According to the results of hyperparameter optimization for the model, the final hyperparameter settings for the TCN model were as follows: the number of TCN layers was 1, the number of convolutional kernels was 64, the kernel size was 3, and the dilation factors d were 1, 2, 4, 8, and 16, with other parameters set to their default values. The Adam optimizer was set with a learning rate of 0.02, the loss function was set to MSE, the number of iterations was set to 300, and the batch size was set to 32.

As for the output layer of the TCN, the final prediction output was obtained through a series of computations and transformations. After processing through multiple layers of causal convolutions and dilated convolutions, the network extracted deep features from the input data. These features, after passing through the fully connected layer, were transformed into output predictions. Supposing the convolution output feature of the model's last layer is h , then the prediction output $\hat{y}$ was obtained through a linear transformation. The specific formula is as follows [17]:

$$\hat{y} = W \cdot h + b \quad (4)$$

Here, W is the weight matrix of the fully connected layer, and b is the bias vector. This linear transformation maps the multidimensional convolutional feature maps to the target output dimension through a weighted sum.

2.5. Predictive Performance Evaluation Metrics

To evaluate the accuracy of point predictions, three evaluation metrics were used: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and the correlation coefficient (R^2). The formulas for their calculation are shown below. n is the number of test sample prediction data, y' is the point prediction value obtained after training, y is the actual measurement value, y'_i is the point prediction value obtained after training for the i -th output, and y_i is the actual measurement value for the i -th output. The smaller the values of RMSE and MAE, the closer the predicted results are to the actual situation, indicating better prediction performance. R^2 reflects the goodness of fit, and the closer it is to 1, the stronger the ability to fit the data is.

The calculation method for RMSE is as follows [18]:

$$\text{RMSE} = \sqrt{\left[\sum_{i=1}^n \frac{(y - y')^2}{n} \right]} \quad (5)$$

The calculation method for R^2 is as follows [18]:

$$R^2 = \frac{(n \sum_{i=1}^n y'_i y_i - \sum_{i=1}^n y'_i \sum_{i=1}^n y_i)^2}{\left[n \sum_{i=1}^n y_i'^2 - (\sum_{i=1}^n y'_i)^2 \right] \left[n \sum_{i=1}^n y_i^2 - (\sum_{i=1}^n y_i)^2 \right]} \quad (6)$$

The calculation method for MAE is as follows [18]:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |y_i - y'_i| \quad (7)$$

To evaluate the model's running speed, an additional time (TT) evaluation metric is added, which is calculated by measuring the time taken for the model to make predictions.

2.6. Hyperparameter Setting

In this study, the hyperparameters of all models were systematically optimized through grid search to ensure the best performance configuration for each model. For the TCN, the parameters adjusted including the number of filters (16, 32, 64), kernel size (2, 3, 5), dilations (1, 2, 4, 8, 16), learning rate (0.001, 0.01, 0.1), batch size (16, 32, 64), and number of training epochs (100 to 500). For LSTM, the optimized parameters included the number of hidden units (16, 32, 64), learning rate (0.001, 0.01, 0.1), batch size (16, 32, 64), and number of training epochs (100 to 500). For the CNN, the number of convolutional layers (1, 2, 3), number of filters (32, 64, 128), learning rate (0.001, 0.01, 0.1), batch size (16, 32, 64), and number of training epochs (100 to 500) were all tested. For the SVM, the adjusted parameters contained the C value (0.1, 1, 10, 100), kernel function (linear, polynomial, RBF), and gamma value (0.001, 0.01, 0.1). Finally, for the XGBoost model, we explored the learning rate (eta) (0.01, 0.1, 0.2), maximum depth (3, 5, 7), number of trees (n_estimators) (50, 100, 200), and proportion of features used by each tree (colsample_bytree) (0.3, 0.7, 1.0).

2.7. Causal Inference Model (EconML)

The EconML method estimates personalized causal responses from observational or experimental data by incorporating machine learning steps into explicable causal models. EconML, based on the potential outcome framework, was applied to analyze the impact of the treatment variable T (i.e., features) on the outcome variable Y (i.e., water quality) after controlling for a set of confounding variables X [19,20]. Causal inference was implemented in Python software using the EconML library (econml-0.15.1). In causal inference analysis, numerical values typically represent the magnitude of the impact that the treatment variable (as feature T) has on the outcome variable (such as water quality Y).

2.8. Shapley Additive Explanations (SHAP)

SHAP was employed to quantify the effects of various water supply factors on effluent water quality and calculate the marginal effect of single factors on effluent water quality. The Generalized Additive Model (GAM) method was used to fit the SHAP values with the feature values.

Generalized Additive Models (GAMs) are a class of models extending linear models, where the relationship between predictor variables and the response variable can be captured through nonlinear functions. Compared to linear models, a GAM offers flexibility and interpretability, allowing it to fit complex relationships without easily leading to overfitting. A Generalized Additive Model can be represented as follows [21]:

$$g(E(Y)) = \alpha + \sum_{i=1}^p f_i(X_i) \quad (8)$$

where $g(\cdot)$ is the link function, used to map the expected value $E(Y)$ to a linear scale. α is the intercept term. $f_i(X_i)$ is a univariate smooth function, intended to capture the nonlinear relationship between the response variable Y and the predictor X_i . p is the number of predictor variables.

The SHAP method is a classic interpretability framework [22] that is based on the Shapley value from game theory as a fair distribution method used to measure the contributions of each participant in cooperative games. In SHAP, Shapley values are used to calculate the contribution of each feature to the prediction result in a machine learning model [23], thereby enhancing the interpretability of “black box models” and clarifying their decision-making processes.

The SHAP method was used to obtain the SHAP value of each feature, standardize the feature variables, remove the outliers that are more than three times the standard deviation after standardization, and fit the relationship between the feature variables and the SHAP values by using ExpectileGAM in the pygam library, in which the parameter of the ExpectileGAM was set to 0.5, to explain the effect of a feature on the target variable (here, “total phosphorus in effluent”) by using a Generalized Additive Model (GAM) to predict the effect of a feature on the target variable.

2.9. Environment

The deep learning architecture Keras (Tensorflow backend) was applied to build the environment required for the simulation experiment in this study. The specific environment parameters were as follows: CPU, AMD Ryzen 7 7735H with Radeon Graphics 3.20 GHz; GPU, NVIDIA GeForce RTX 4060 Laptop GPU; 16 G computer memory; operating system, Win11 64-bit; Keras version, 2.3.1; Tensorflow version, 2.6.0; programming language, Python3.9.21.

3. Results and Discussion

3.1. Results of the TCN Model

Figure 1 shows the comparison results between predicted values using the TCN model and actual values, using R^2 , RMSE, MAE, and TT as evaluation metrics. From Figure 1, it can be seen that the R^2 values for training set prediction results of ammonia nitrogen, nitrate nitrogen, total nitrogen, COD, and total phosphorus from the outlet of the power plant’s circulating cooling tower using the TCN model were 0.778, 0.881, 0.938, 0.691, and 0.887, respectively. The RMSE values were 0.348, 3.758, 3.583, 12.825, and 0.128, and the MAE values were 0.215, 2.771, 2.698, 9.692, and 0.096, respectively. The R^2 values for testing set prediction results of ammonia nitrogen, nitrate nitrogen, total nitrogen, COD, and total phosphorus from the outlet of the power plant’s circulating cooling tower using the TCN model were 0.720, 0.910, 0.909, 0.756, and 0.911, respectively. The RMSE values were 0.478, 4.841, 3.432, 13.471, and 0.158, and the MAE values were 0.325, 3.739, 2.553, 9.979, and 0.118, respectively. These results demonstrate that the TCN model has high fitting accuracy and low error, accurately reflecting the variation patterns of water quality, indicating that the TCN model can effectively identify the relationship between influent parameters and effluent water quality [24], and efficiently extract the data features of influent parameters.

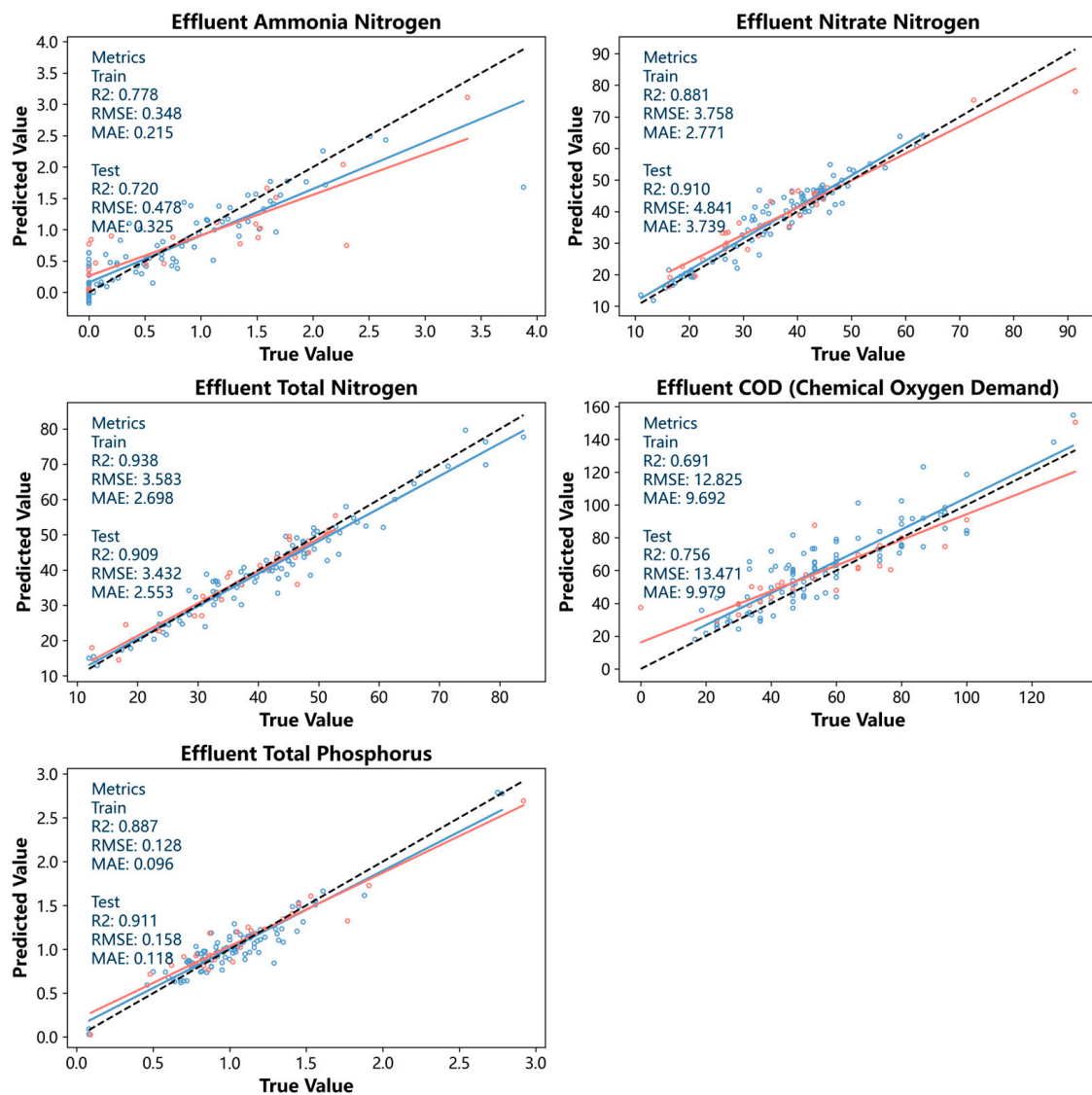


Figure 1. Prediction results of TCN model.

3.2. Comparison of Different Models

To further demonstrate the advantages of the TCN model in achieving real-time, efficient, and accurate prediction of circulating water effluent quality, the results of the TCN model were compared with those of two common deep learning models, a CNN and LSTM, as well as traditional machine learning models, XGBoost and SVM. As shown in Figure 2, overall, the three deep-learning-based models, the TCN, LSTM, and the CNN, exhibited a larger R^2 and smaller RMSE and MAE compared to traditional machine learning models, but had a greater TT. These results indicate that deep learning models have a greater advantage in fitting data due to their self-learning and self-organizing characteristics [25]. In a comparison of the TCN with LSTM and the CNN, the TCN demonstrates a larger R^2 and lower RMSE and MAE for model fitting metrics, indicating that the TCN has better feature extraction and fitting performance. In terms of TT comparison, deep learning models like the CNN, LSTM, and the TCN require more time compared to traditional machine learning models like SVM and XGBoost. However, all TT values were less than 1 s, enabling real-time prediction [26]. Simulation results showed that the TCN model can detect changes in water quality in a timely manner and provide real-time warnings to power plants, thus ensuring more accurate and faster water management. In the case of key water quality indicators such as COD and total phosphorus, there was a time lag

in the monitoring of these parameters using standard methods, which does not allow for early warning in the event of a major water quality incident. Using the TCN model prediction method mentioned in this study could realize the real-time prediction of key water quality, timely feedback on water quality, and real-time control. However, there are some potential limitations of this approach. For example, TCN modeling is highly dependent on data quality, and noisy or incomplete data may affect prediction accuracy. Meanwhile, when applied to different wastewater treatment processes and datasets of wastewater treatment processes, the TCN model needs to be re-trained and is less portable. In addition, the complexity of TCN models may affect the scalability of other systems. In particular, when integrating different water management systems, the model needs to be adapted to different environments and data structures.

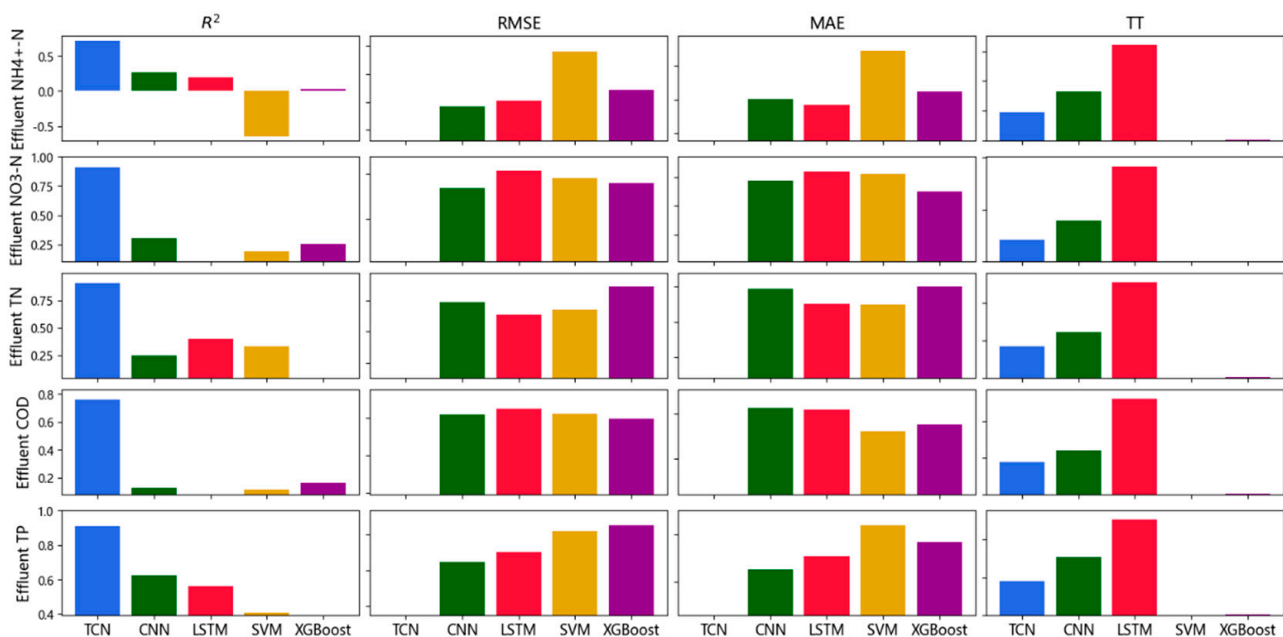


Figure 2. Comparison of evaluation indices of five models.

3.3. Causal Inference

Figure 3 presents the causal inference results for five effluent variables. For effluent ammonia nitrogen, nitrate nitrogen, and total nitrogen concentrations, the causal inference results for influent ammonia nitrogen, nitrate nitrogen, and total nitrogen were relatively large, followed by the concentration ratio. In terms of effluent COD, in addition to influent COD, influent ammonia nitrogen, nitrate nitrogen, and total nitrogen also showed large causal inference results, followed by the concentration ratio. For effluent parameters, the load of corresponding influent parameters played a crucial role. Regarding COD, it was associated with influent ammonia nitrogen, nitrate nitrogen, and total nitrogen, potentially due to two factors: (1) Ammonia nitrogen and nitrate nitrogen in water can be transformed through microbial processes under specific conditions. Nitrification and denitrification are not only nitrogen transformations but may also affect the degradation of organic matter. Under aerobic conditions, nitrification consumes oxygen, which may lead to oxygen depletion that impacts the oxidation process of organic matter [27]. (2) Ammonia nitrogen and total nitrogen serve as nutrients for microbial growth. A sufficient nitrogen source can promote microbial reproduction and activity, thereby enhancing the ability to degrade organic matter, which in turn influences the removal efficiency of COD [27]. For the results of the five effluent parameters, the causal inference results for the pH value were relatively low, potentially due to the small variation of pH, thus resulting in a minor

impact. To further analyze the effect of influent parameters on effluent parameters, SHAP analysis was conducted on the pH and concentration ratio.

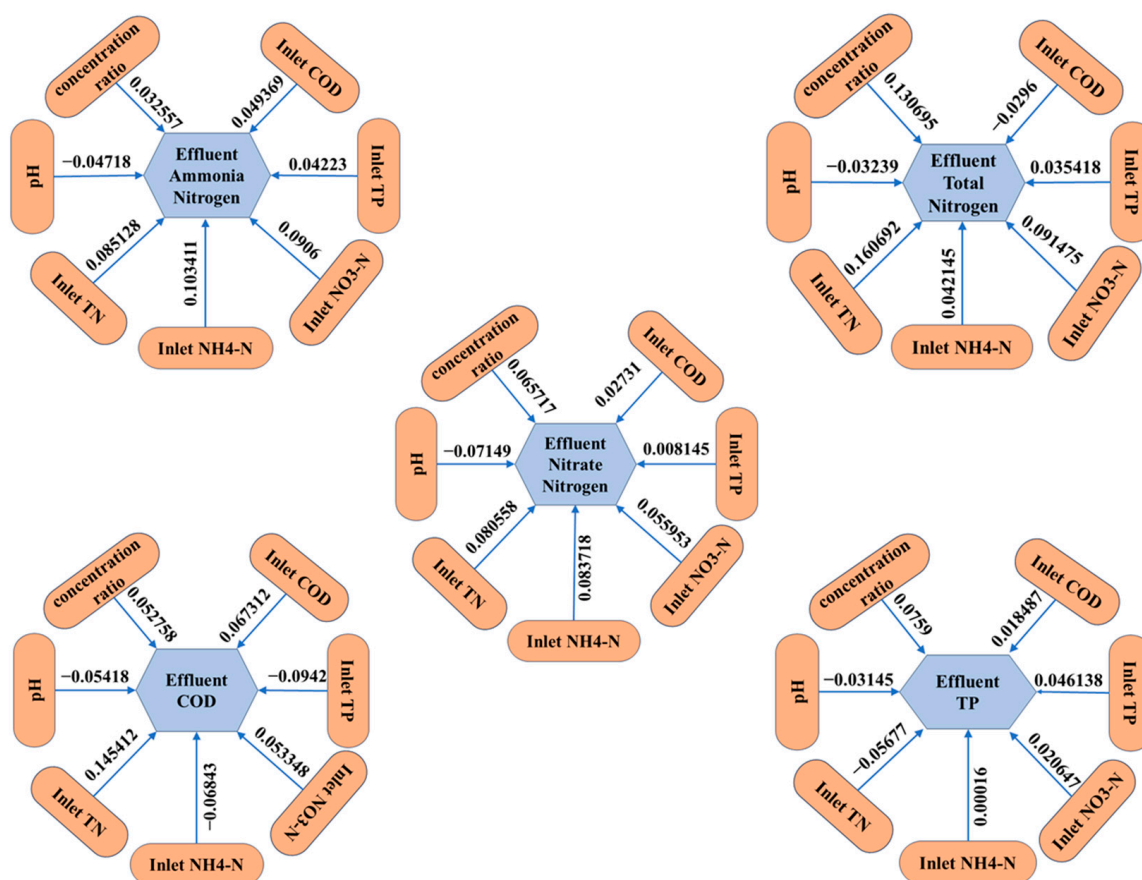


Figure 3. Causal inference results of water quality indicators.

3.4. SHAP Analysis

Figure 4 shows the SHAP results of five effluent indicators for influent pH. As could be seen from the SHAP plot, the SHAP value showed a gradually decreasing trend for influent pH as the pH value increases. When the pH value was higher than 7, the alkalinity of the water was enhanced and then influenced the concentrations of ammonia nitrogen, total nitrogen, nitrate nitrogen, COD, and total phosphorus in the effluent of the recirculating cooling tower. Firstly, ammonia nitrogen can be more easily converted into free ammonia (NH₃) under alkaline conditions [28], making it more volatile, thereby reducing the total nitrogen concentration. Meanwhile, the nitrification process may be inhibited at a higher pH condition (pH > 6) since the activity of nitrifying bacteria can weaken under increased alkalinity [29], thus affecting the conversion of nitrate nitrogen. Additionally, high-pH conditions can promote chemical reactions of some organic matter, such as saponification, accelerating the degradation of COD and causing changes in its concentration [30]. As for total phosphorus, alkaline conditions increase the likelihood of phosphates reacting with metal ions (such as calcium and magnesium) to form insoluble salts, thereby reducing the total phosphorus concentration [31].

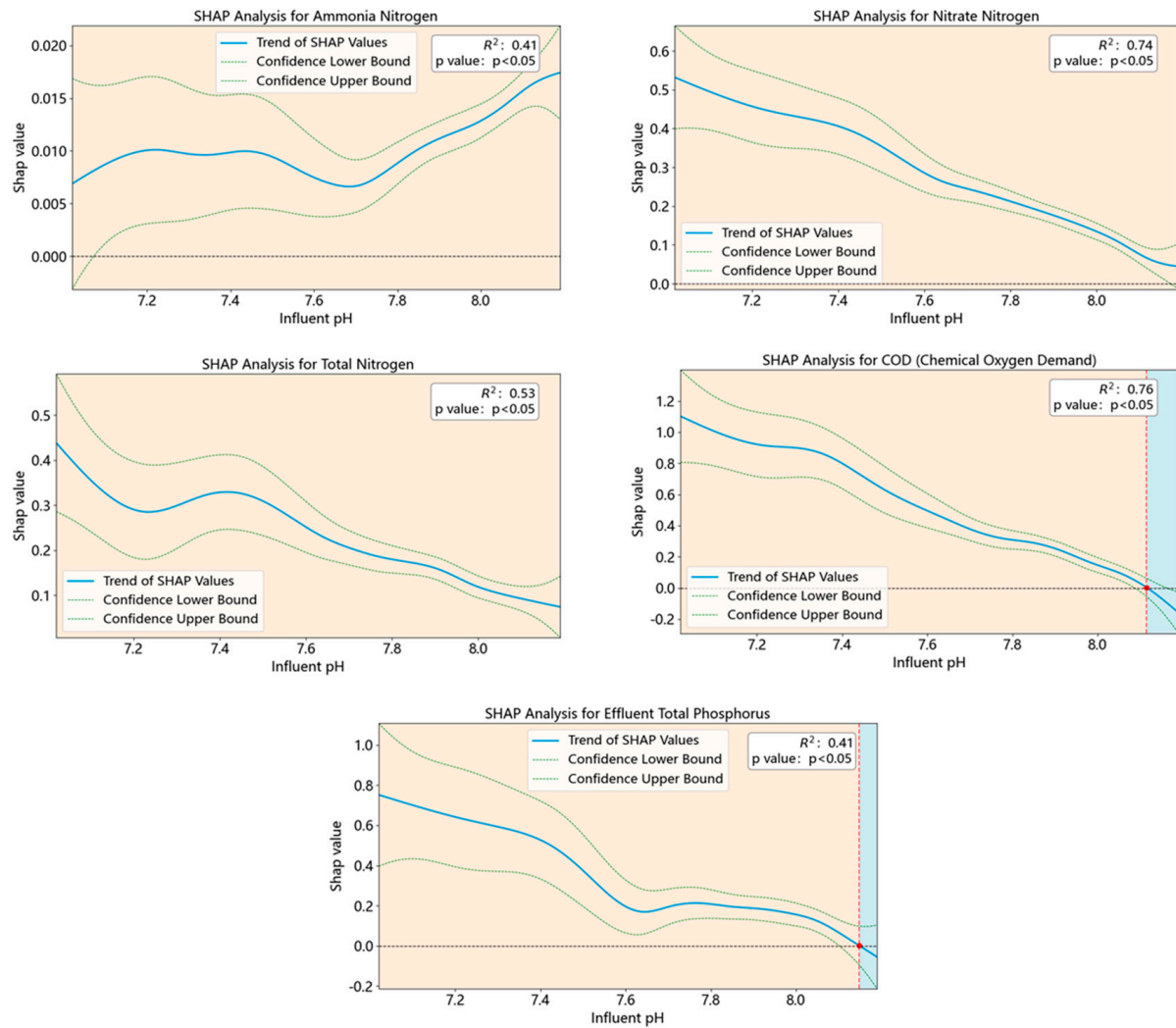


Figure 4. SHAP fitting result of pH value. (These dotted lines are set up to mark turning points).

As shown in Figure 5, the SHAP value increased continuously with the concentration ratio, while the SHAP value gradually stabilized after reaching a concentration ratio of 5. This variation might reflect that the cooling system is reaching a dynamic equilibrium state. As the concentration ratio increased, dissolved solids and nutrients in the circulating water were repeatedly concentrated due to the evaporation of circulating water, leading to a gradual increase in the concentration of these pollutants in the water. This increase in pollutant concentrations could influence the chemical balance and biodegradation processes in the circulating tower. For example, the increase in dissolved solids may inhibit microbial activity in the water [32], slowing the degradation rate and then decreasing the removal efficiency of ammonia nitrogen and total nitrogen. Additionally, high concentration ratios can alter the physicochemical conditions for certain reactions in the water, such as redox potential and pH [33], which further affect the transformation and degradation of these nutrients and organic matter. When the concentration ratio reached a certain value, the cooling circulation tower system entered a dynamic equilibrium state, at which the SHAP value tended to stabilize, indicating that the sensitivity of pollutant concentration changes to the concentration ratio is reduced, and the reaction rates and concentrations in the system moved towards equilibrium.

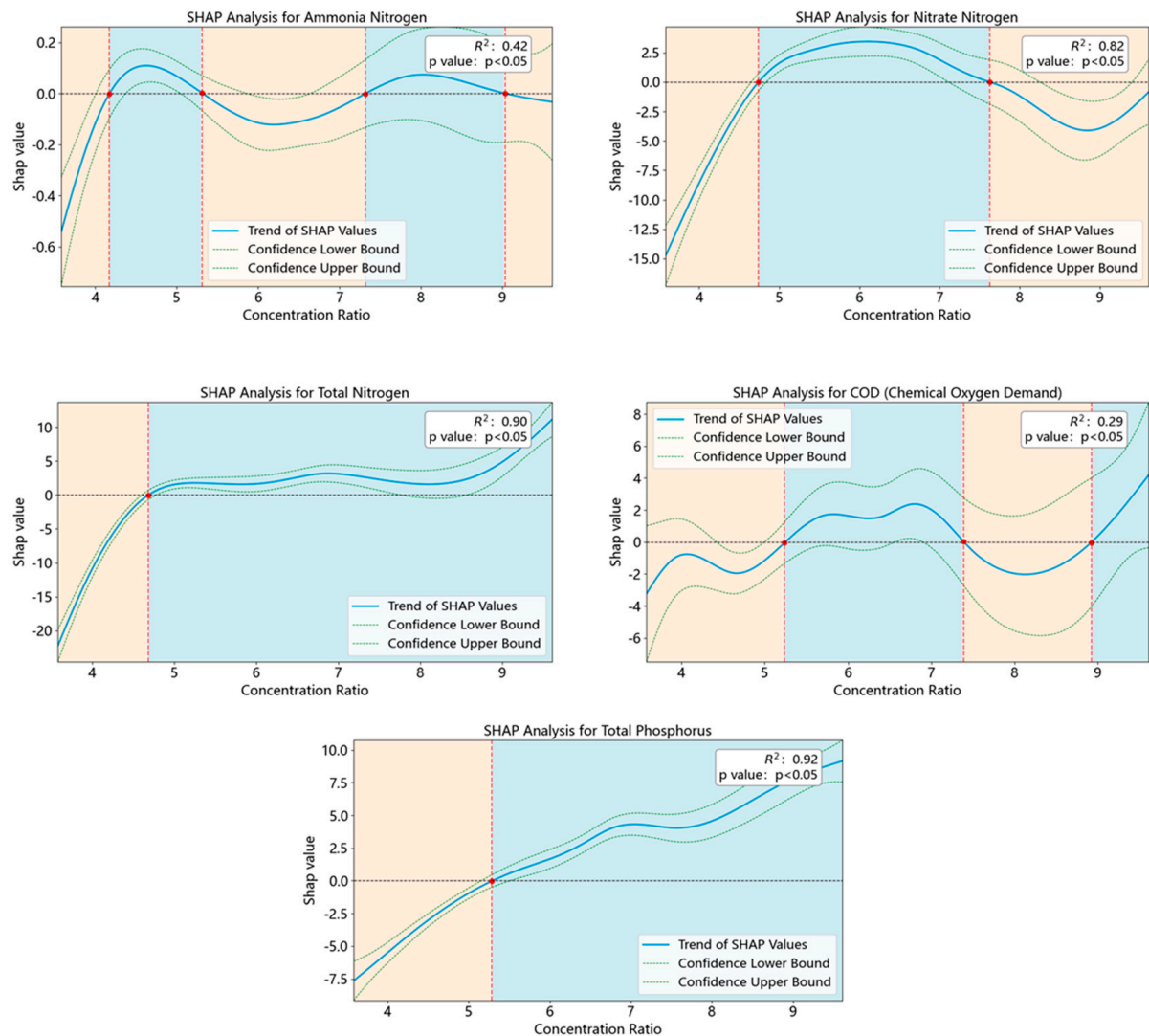


Figure 5. SHAP fitting result of concentration ratio. (The different colors represent areas of positive (blue) and negative (yellow) shifts).

However, while this study employs the SHAP method to interpret model predictions, the recent development of FLIT-SHAP [34] (Feature-Localized-Intercept-Transformed SHAP) has been regarded as a meaningful advancement in model explainability, which provides absolute feature contributions with interpretable units aligned with the predicted variable and may enhance the physical interpretability of local explanations. Although it was not applied in the current work, FLIT-SHAP could be a valuable extension in future studies aiming for more direct quantitative insights into model outputs.

4. Conclusions

- (1) The TCN model demonstrated superior predictive performance for ammonia nitrogen, nitrate nitrogen, total nitrogen, COD, and total phosphorus in the effluent of the power plant's circulating cooling tower, achieving a high fitting accuracy with low RMSE and MAE values.
- (2) Compared to traditional models such as XGBoost and SVR, the TCN achieved a higher R^2 and maintained prediction times under 1 s, supporting its suitability for real-time water quality monitoring despite slightly longer training times.
- (3) Causal inference analysis revealed that effluent water quality is most strongly influenced by the corresponding parameters in the makeup water, followed by the

concentration ratio, highlighting the dominant role of influent quality in determining discharge characteristics.

- (4) SHAP analysis indicated that higher inflow pH (>7) reduces effluent concentrations of ammonia nitrogen, nitrate nitrogen, total nitrogen, COD, and total phosphorus, likely due to enhanced volatilization, microbial inhibition, and improved degradation. While increasing the concentration ratio raises impurity levels and inhibits microbial activity, its influence plateaus around a ratio of 5, suggesting a dynamic equilibrium state.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/pr13061917/s1>: Table S1: The used water quality indicators of the influent and effluent of the circulating cooling tower system.

Author Contributions: Conceptualization, writing—original draft preparation, Y.W.; methodology, software, X.T.; validation, formal analysis, H.H.; investigation, methodology, resources, P.T.; resources, data curation, writing—review and editing, R.G.; visualization, supervision, project administration, funding acquisition, X.J.; validation, formal analysis, S.L. (Shaojie Li), S.L. (Shan Luo) and W.L.; writing—review and editing, Z.C. All authors have read and agreed to the published version of the manuscript.

Funding: The authors declare that this study received funding from China Energy Engineering Group Guangdong Electric Power Design Institute Co., Ltd. (Research and application of key technologies for achieving standard discharge of deep treatment of circulating cooling water drainage, grant number EX11131W). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this article or the decision to submit it for publication.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Acknowledgments: We thank Songwei Zhou for the technical support and supply of materials used for experiments.

Conflicts of Interest: Authors Yongjie Wan, Xing Tian, Hanhua He, Ruiying Gao, Xiaohui Ji, Shaojie Li, Shan Luo, Wei Li were employed by the company China Energy Engineering Group Guangdong Electric Power Design Institute Co., Ltd. Author Peng Tong was employed by the company Huaneng Dongguan Gas Turbine Thermoelectricity Co., Ltd. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

XGBoost	eXtreme Gradient Boosting
SVM	Support Vector Machine
CNNs	Convolutional Neural Networks
LSTM	Long Short-Term Memory
TCN	Temporal Convolutional Network
SHAP	Shapley Additive Explanations
COD	Chemical Oxygen Demand
RMSE	Root Mean Square Error
MAE	Mean Absolute Error
TT	Inference Time
EconML	Causal Inference Model

References

1. Mekonnen, M.M.; Gerbens-Leenes, P.W.; Hoekstra, A.Y. The consumptive water footprint of electricity and heat: A global assessment. *Environ. Sci.-Water Res. Technol.* **2015**, *1*, 285–297. [\[CrossRef\]](#)
2. Macknick, J.; Meldrum, J.; Nettles-Anderson, S.; Heath, G.; Miara, A. Life cycle water use for photovoltaic electricity generation: A review and harmonization of literature estimates. In Proceedings of the 2014 IEEE 40th Photovoltaic Specialist Conference (PVSC), Denver, CO, USA, 8–21 June 2014; pp. 1458–21460.
3. Pan, S.-Y.; Snyder, S.W.; Packman, A.I.; Lin, Y.J.; Chiang, P.-C. Cooling water use in thermoelectric power generation and its associated challenges for addressing water-energy nexus. *Water-Energy Nexus* **2018**, *1*, 26–41. [\[CrossRef\]](#)
4. Chafa, A.T.; Chirinda, G.P.; Matope, S. Design of a real-time water quality monitoring and control system using Internet of Things (IoT). *Cogent Eng.* **2022**, *9*, 2143054. [\[CrossRef\]](#)
5. Nong, X.Z.; He, Y.; Chen, L.H.; Wei, J.H. Machine learning-based evolution of water quality prediction model: An integrated robust framework for comparative application on periodic return and jitter data. *Environ. Pollut.* **2025**, *369*, 125834. [\[CrossRef\]](#)
6. Libera, D.A.; Sankarasubramanian, A. Multivariate bias corrections of mechanistic water quality model predictions. *J. Hydrol.* **2018**, *564*, 529–541. [\[CrossRef\]](#)
7. Zhu, J.J.; Yang, M.Q.; Ren, Z.J. Machine Learning in Environmental Research: Common Pitfalls and Best Practices. *Environ. Sci. Technol.* **2023**, *57*, 17671–17689. [\[CrossRef\]](#)
8. Safeer, S.; Pandey, R.P.; Rehman, B.; Safdar, T.; Ahmad, I.; Hasan, S.W.; Ullah, A. A review of artificial intelligence in water purification and wastewater treatment: Recent advancements. *J. Water Process Eng.* **2022**, *49*, 102974. [\[CrossRef\]](#)
9. Alvi, M.; Batstone, D.; Mbamba, C.K.; Keymer, P.; French, T.; Ward, A.; Dwyer, J.; Cardell-Oliver, R. Deep learning in wastewater treatment: A critical review. *Water Res.* **2023**, *245*, 120518. [\[CrossRef\]](#)
10. El-Rawy, M.; Abd-Ellah, M.K.; Fathi, H.; Ahmed, A.K.A. Forecasting effluent and performance of wastewater treatment plant using different machine learning techniques. *J. Water Process Eng.* **2021**, *44*, 102380. [\[CrossRef\]](#)
11. Xiang, Y.R.; Jiang, L.Z. Water Quality Prediction Using LS-SVM with Particle Swarm Optimization. In Proceedings of the 2009 Second International Workshop on Knowledge Discovery and Data Mining, Moscow, Russia, 23–25 January 2009; pp. 900–904.
12. Huang, R.X.; Ma, C.X.; Ma, J.; Huangfu, X.L.; He, Q. Machine learning in natural and engineered water systems. *Water Res.* **2021**, *205*, 117666. [\[CrossRef\]](#)
13. Huan, J.; Fan, Y.X.; Xu, X.G.; Zhou, L.W.; Zhang, H.; Zhang, C.; Hu, Q.C.; Cai, W.X.; Ju, H.R.; Gu, S.L. Deep learning model based on coupled SWAT and interpretable methods for water quality prediction under the influence of non-point source pollution. *Comput. Electron. Agric.* **2025**, *231*, 109985. [\[CrossRef\]](#)
14. Ahmed, A.A.; Sayed, S.; Abdoulhalik, A.; Moutari, S.; Oyedele, L. Applications of machine learning to water resources management: A review of present status and future opportunities. *J. Clean. Prod.* **2024**, *441*, 140715. [\[CrossRef\]](#)
15. APHA. *Standard Methods for the Examination of Water and Wastewater*, 21st ed.; American Public Health Association: Washington, DC, USA, 2012.
16. Yang, H.; Zhao, X.; Wang, L. Review of Data Normalization Methods. *Comput. Eng. Appl.* **2023**, *59*, 13–22. [\[CrossRef\]](#)
17. Zuo, H.Y.; Gou, X.T.; Wang, X.; Zhang, M.Y. A Combined Model for Water Quality Prediction Based on VMD-TCN-ARIMA Optimized by WSWOA. *Water* **2023**, *15*, 4227. [\[CrossRef\]](#)
18. Hodson, T.O. Root-mean-square error (RMSE) or mean absolute error (MAE): When to use them or not. *Geosci. Model Dev.* **2022**, *15*, 5481–5487. [\[CrossRef\]](#)
19. Chernozhukov, V.; Chetverikov, D.; Demirer, M.; Duflo, E.; Hansen, C.; Newey, W.; Robins, J. Double/debiased machine learning for treatment and structural parameters. *Econom. J.* **2018**, *21*, C1–C68. [\[CrossRef\]](#)
20. Yao, L.Y.; Chu, Z.X.; Li, S.; Li, Y.L.; Gao, J.; Zhang, A.D. A Survey on Causal Inference. *ACM Trans. Knowl. Discov. Data* **2021**, *15*, 74. [\[CrossRef\]](#)
21. Bordt, S.; von Luxburg, U. From Shapley Values to Generalized Additive Models and back. In Proceedings of the 26th International Conference on Artificial Intelligence and Statistics, PMLR, Valencia, Spain, 25–27 April 2023.
22. Lundberg, S.M.; Lee, S.I. A Unified Approach to Interpreting Model Predictions. In Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017), Long Beach, CA, USA, 4–9 December 2017.
23. Qian, S.T.; Qiao, X.Y.; Zhang, W.M.; Yu, Z.J.; Dong, S.A.; Feng, J.G. Machine learning-based prediction for settling velocity of microplastics with various shapes. *Water Res.* **2024**, *249*, 121001. [\[CrossRef\]](#)
24. Zhang, X.; Li, D.S. Multi-input multi-output temporal convolutional network for predicting the long-term water quality of ocean ranches. *Environ. Sci. Pollut. Res.* **2023**, *30*, 7914–7929. [\[CrossRef\]](#)
25. Dong, S.; Wang, P.; Abbas, K. A survey on deep learning and its applications. *Comput. Sci. Rev.* **2021**, *40*, 100379. [\[CrossRef\]](#)
26. Chen, J.S.; Ran, X.K. Deep Learning With Edge Computing: A Review. *Proc. IEEE* **2019**, *107*, 1655–1674. [\[CrossRef\]](#)
27. Quick, A.M.; Reeder, W.J.; Farrell, T.B.; Tonina, D.; Feris, K.P.; Benner, S.G. Nitrous oxide from streams and rivers: A review of primary biogeochemical pathways and environmental variables. *Earth-Sci. Rev.* **2019**, *191*, 224–262. [\[CrossRef\]](#)

28. Capson-Tojo, G.; Moscoviz, R.; Astals, S.; Robles, A.; Steyer, J.P. Unraveling the literature chaos around free ammonia inhibition in anaerobic digestion. *Renew. Sustain. Energy Rev.* **2020**, *117*, 109487. [[CrossRef](#)]
29. Zhang, F.Z.; Du, Z.Y.; Wang, J.H.; Du, Y.J.; Peng, Y.Z. Acidophilic partial nitrification (pH < 6) facilitates ultra-efficient short-flow nitrogen transformation: Experimental validation and genomic insights. *Water Res.* **2024**, *260*, 121921.
30. Low, S.J. Determination of Activation Energy of Saponification Reaction through pH Analysis. *Res. Rev. J. Chem.* **2017**, *6*, 1–5.
31. Amaro, H.M.; Salgado, E.M.; Nunes, O.C.; Pires, J.C.M.; Esteves, A.F. Microalgae systems- environmental agents for wastewater treatment and further potential biomass valorisation. *J. Environ. Manag.* **2023**, *337*, 117678. [[CrossRef](#)]
32. Hanson, A.J.; Luek, J.L.; Tummings, S.S.; McLaughlin, M.C.; Blotevogel, J.; Mouser, P.J. High total dissolved solids in shale gas wastewater inhibit biodegradation of alkyl and nonylphenol ethoxylate surfactants. *Sci. Total Environ.* **2019**, *668*, 1094–1103. [[CrossRef](#)]
33. Liu, J.; Zhi, Q.; Hu, X. Analysis and measures of pH rising for circulating water in indirect air cooling unit. In *International Conference on Mechatronics and Intelligent Control (ICMIC 2024)*; SPIE: St Bellingham, WA, USA, 2024; Volume 13447, pp. 625–628.
34. Esu, O.C.; Pyo, J.; Cho, K. Machine learning-derived dose-response relationships considering interactions in mixtures: Applications to the oxidative potential of particulate matter. *J. Hazard. Mater.* **2024**, *475*, 134864. [[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.