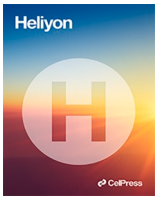
Heliyon 10 (2024) e35580

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Real-time prediction of the chemical oxygen demand component parameters in activated sludge model using backpropagation neural network

Ping Wang [a], Yanqiong Chen [a], Chen Zhang [a], Yuzhen Shi [a,*], Bin Wang [b], Chaochao Lai [a], Huan He [a], Bin Huang [a]

[a] Faculty of Environmental Science and Engineering, Kunming University of Science and Technology, Kunming, 650500, China
[b] School of Environment and Resource, Southwest University of Science and Technology, Mianyang, Sichuan, 621010, China

ARTICLE INFO

Keywords:
Activated sludge models
Backpropagation neural networks
Chemical oxygen demand components
Dynamic model

ABSTRACT

Activated sludge models are increasingly being adopted to guide the operation of wastewater treatment plants. Chemical oxygen demand (COD) is an indispensable input for such models. To ensure that the activated sludge mathematical model can adapt to various water quality conditions and minimize prediction errors, it is essential to predict the parameters of the COD components in real-time based on the actual influent COD concentrations. However, conventional methods of determining the components' contributions are too intricate and time-consuming to be really useful. In this study, the chemical oxygen demand in the actual waste water treatment plant was disassembled and analyzed. The research involved determining the proportions of each COD component, assessing the reliability of the measurement parameters, and examining potential factors affecting measurement accuracy, including weather conditions, pipeline conditions, and residents' habits. Then, a backpropagation neural network was developed which can deliver real-time predictions for five important contributors to COD in real time. In addition, using the receiver operating characteristics curve and prediction accuracy to evaluate the performance of the prediction model. For all five components, which S_S, X_S, S_I, X_A, and X_H, the prediction accuracy of model was more than 80 %. The maximum deviation values of these parameters fall within the range of the actual detected values, suggesting that the model's predictions align well with real-world observations, and demonstrated prediction performance adequate for practical application in wastewater treatment. This article can provide research basis for the engineering application of activated sludge model and help for the intelligent upgrading of waste water treatment plants.

1. Introduction

Wastewater treatment plants (WWTPs) protect aqueous environments and help to ensure adequate sanitation of normal human life in modern cities. In response to progressively stringent national and local government regulations on wastewater treatment, energy consumption thresholds, and resource recovery mandates [1–3]. To ensure the stable operation, WWTPs encounter challenges

* Corresponding author.
E-mail address: yzs870307@163.com (Y. Shi).

https://doi.org/10.1016/j.heliyon.2024.e35580
Received 7 May 2024; Received in revised form 20 July 2024; Accepted 31 July 2024
Available online 8 August 2024

stemming from significant influent fluctuations and the inherent complexity of nonlinear treatment processes [4]. To effectively monitor the operational parameters of WWTPS, operators are required to track changes in various variables throughout the treatment process using instrumentation. This monitoring process often results in significant delays due to the inherent sequential nature of operations, including detection, reading, feedback, and subsequent adjustment. Furthermore, the operation and maintenance of WWTPs have become increasingly complex because of the absence of standardized operating guidelines that can be uniformly applied to all influent conditions. As a result, staff members are often compelled to rely on empirical or semi-empirical control methods, which can lead to inefficient resource utilization and suboptimal treatment outcomes [5]. To address these challenges, activated sludge models (ASMs) are gaining popularity as a tool in WWTP operations. Their accuracy and adaptability have become widely accepted. A computer-controlled ASM can simulate diverse conditions rapidly, accommodating the complexity and variability inherent in wastewater quality and volume [6]. The ASM offers intuitive demonstrations tailored to the needs of WWTP staff, helping them choose reliable responses to variations [7]. It alleviates the difficulties related to experience-based judgment, manual calculations, and manual control for WWTPs managers, leading to a less complex management process.

The are now various ASM models accepted by the International Water Association [8,9]. Accurate prediction using them relies on good input data such as good chemical oxygen demand (COD) component parameters. Researchers have proposed various techniques for collecting the required data, including respirometry, biochemical oxygen demand analysis, physical-chemical methods and the batch activated sludge method. However, the complexity and time-consuming nature of component measurements, combined with the significant resources required for collecting dynamic influent data, significantly hinder the ability to update COD component parameters in real time. Consequently, these challenges restrict the practical application of ASMs [10]. Influences like catchment area size, wastewater system type, the population served, industrial emissions, soil type, rainfall patterns, and temperature all contribute to variations in the influent COD detected [11,12]. These factors lead to a discrepancy between the actual COD components and those assumed during the model's initial construction. In the currently prevalent ASMs, initial COD component values are fixed and do not adjust to variations in influent water quality. This limitation undermines the accuracy of the models' predictions for wastewater treatment processes. Predicting the COD generated by each component of WWTP influent has therefore become essential for getting accurate predictions from activated sludge mathematical models.

Today multiple linear regression and autoregressive integrated moving averages [13,14] are used to predict wastewater quality, but their inherent linearity limits their ability to handle the nonlinearity and complexity of COD components. Artificial neural networks have made rapid progress in predicting wastewater quality prediction [15–18]. They can handle complex functional relationships and require fewer hypothetical variable to achieve good accuracy. Their ability to deal with nonlinear relationships and uncertainty gives them a significant advantage in wastewater quality prediction [19–21] thus made them a hot topic in wastewater quality research [22]. Reinforcement learning and random forest are commonly used machine learning techniques for wastewater quality prediction. Reinforcement learning, in contrast, distinguishes itself through its method of evaluating actual output parameters and formulating a reward strategy. Random forest is chiefly applied to data classification. Variability, randomness, and classification difficulties in COD components are inherent challenges that need to be addressed in the selection of prediction algorithms. Despite the prevalent use of both reinforcement learning and random forest in wastewater quality prediction, each of these methods has limitations in the separation of COD components.

More recently, backpropagation artificial neural networks (BPNNs) have emerged as promising tools for wastewater quality prediction due to their ability to capture complex functional relationships with fewer a hypothetical variable. They have found extensive application in predicting wastewater quality in reservoirs [23], wastewater treatment plants [24], aquaculture farms [25], rivers and elsewhere. This can be attributed to their speed of convergence, robust nonlinear mapping, and adaptable network structure [26–29]. Currently, common BPNNs can only broadly predict single indicators like total COD [30]. Research is still needed to enable detailed prediction of COD contributors such as dissolved COD and particulate COD. with their complex, nonlinear characteristics. But the powerful self-teaching ability of BPNNs promises that accurate prediction of influent characteristics can be achieved, providing support for precise application of activated sludge models.

This study addressed some of the main challenges in decomposing COD components in ASM model parameter detection. Real wastewater treatment plant influent was tested in the laboratory, separating its COD components, and developing a COD component prediction model using a BPNN. The BPNN model was then employed in simulating and real-time prediction of the COD component parameters crucial for building useful activated sludge mathematical models. The study's experimental data and theoretical advances should contribute to the engineering application of the ASM.

2. Materials and methods

2.1. Wastewater samples

The study used wastewater sampled from a WWTP located in the southern urban area of Kunming China. The plant is a domestic wastewater treatment facility serving a population of 579,000. Its designed treatment capacity is 300,000m^3/d. The WWTP serves a large population across a wide area, and experiences significant fluctuations in water quality. Therefore, selecting this WWTP to construct the COD component prediction model is representative. Wastewater samples were collected three times a day—morning, noon, and evening—over a period of 25 days. To address model errors arising from variations in data due to residents' water consumption habits, influent fluctuations, weather conditions, and rainfall in the service area. The samples were stored at 4 °C until tested. The samples' quality-determining components were analyzed and influent fluctuations were recorded.

2.2. Analytical methods

The influent's total COD was measured (COD_T) along with its biochemical oxygen demand (BOD), soluble inert organic matter concentration (S_I), soluble biodegradable substrate concentration (S_S), slowly biodegradable organic matter concentration (X_S), active heterotrophic biomass concentration (X_H), active autotrophic biomass concentration (X_A), particulate inert organic matter concentration (X_I), and the yield coefficient for heterotrophic biomass (Y_H). The sample handling and detection methods are elaborated in the supporting information. COD_T and BOD were quantified using China's national standard methods (State Environmental Protection Administration of China, Ministry of environmental protection of China, 2002). The determination of S_I and S_S applied physico-chemical method [31] and flocculation method [32], respectively. Detailed sample handling methods were shown in (supporting information, SI). S_I could be detected by testing, and S_S was then estimated using Eq. (1):

$$S_s = COD_f - S_I \tag{1}$$

The X_S levels were determined using the ultimate biochemical method. The measured limiting biochemical oxygen demand (BOD_u) is known to account for approximately 88 % of readily biodegradable COD [33]. Equation (2) then allows estimating Xs as

$$X_S + S_S = \text{readily biodegradable COD} = \frac{BOD_u}{0.88} = \frac{BOD_5}{0.616} \tag{2}$$

In this context the five days' biochemical oxygen demand (BOD_5) in the urban wastewater was 0.7 times the BOD_u, allowing X_S to be calculated using Eq. (2). X_H and X_A were determined using the initial the oxygen uptake rate of the raw water. Since the initial heterotrophic bacteria in the raw water do not grow rapidly, the initial dissolved oxygen (DO) consumption rate r (at t_0) was obtained by the oxygen uptake rate line, and X_H could then be calculated using Eq. (3) [34]. For X_A detection the sample was added to a 2L reactor without allylthiourea (ATU), and briefly subjected to high-intensity aeration. The DO changes were recorded to obtain r at time t. The r at time t observed without ATU was subtracted from the r at t_0 obtained with ATU to obtain the DO consumption rate $\Delta r(t_0)$. The concentration of autotrophic bacteria X_A could then be obtained using Eq. (4) [35].

$$X_H = r(t_0)\left[\left(\frac{1 - Y_H}{Y_H}\right)\mu_H\right]^{-1} \tag{3}$$

$$X_A = \Delta r(t_0)\left[\left(\frac{4.57 - Y_A}{Y_A}\right)\mu_A\right]^{-1} \tag{4}$$

where μ_H is the maximum specific growth rate of the heterotrophic bacteria, taken as 5.0/d [36]; Y_A, the yield coefficient of the autotrophic bacteria, was taken as 0.24 [36]; with μ_A, the maximum specific growth rate of the autotrophic bacteria, taken as 0.8/d [37].

X_I could be calculated based on the material balance shown in Eq. (5) [35].

$$X_I = COD_T - X_S - S_S - X_H - X_A - S_I \tag{5}$$

Y_H was determined using aerobic respirometry with the experimental device shown in Fig. S-2 of the supplementary information. The supporting material also details the detection methods for each variable. Y_H was then calculated according to Eq. (6) [38].

$$Y_H = 1 - \frac{\Delta DO}{SD}\frac{V + V_W}{V_W} \tag{6}$$

where ΔDO was the change in dissolved oxygen (mg/L), SD was the COD concentration detected by acetic acid as a biodegradable substrate (mg/L), V was the volume of the container (L) and V_W was the volume of wastewater (L).

2.3. Constructing and training the BPNN

Creating a BPNN involves two primary phases. In the first phase, signals propagate forward, starting from the input layer through the hidden layer and ultimately reaching the output layer. In the second phase, errors propagate backward, moving from the output layer to the hidden layer and finally to the input layer. The errors guide adjusting the weights and biases linking the hidden layer to the output layer and the input layer to the hidden layer. Information transfer between layers uses activation functions. In this case, the *tansig* activation function (Eq. (7)) was employed between the input layer and the hidden layer to normalize the data, activate neurons, and facilitate the creation of connections between the input and output data. The *purelin* activation function (Eq. (8)), was used between the hidden layer and the output layer to map the hidden layer's output to a predefined result within a specified range. This process involves converting dimensionless quantities into a specific dimensional representation, which allows for the calculation of errors in the model. The backpropagation of errors is described by Eqs. (9)–(11).

$$O_j = f\left(\sum_{i=1}^{I} W_{ij}X_I - q_j\right) \tag{7}$$

$$Y_K = f\left(\sum_{j=1}^{J} T_{jk} O_j - q_k\right) \tag{8}$$

$$E_K = \frac{1}{2}\sum_{j=1}^{J}\left(y_j^{k'} - y_j^{k}\right)^2 \tag{9}$$

$$\triangle w_{hj} = -\eta \frac{\partial E_k}{\partial w_{hj}} \tag{10}$$

$$\triangle \theta_j = -\eta \frac{\partial E_K}{\partial \theta_j} \tag{11}$$

i, j, and k here represented the number of neurons in the input, hidden, and output layers, respectively. W_{hj} and q_j denote the weighting and threshold of the jth node in the hidden layer involved with the input layer, while T_{jk} and q_k represent the weighting and threshold of "neuron" k in the output layer communicating with the hidden layer. E_K is the error: the difference between the actual output value y_j^k and the expected output $y_j^{k'}$. η is a learning rate.

The BPNN in this study had such a three-layer structure, with one hidden layer. The input and output layers both had one "neuron" to give a one-to-one correspondence. Five BPNNs were created for model training relating COD with S_I, S_S, X_S, X_H and X_A, respectively. Fig. 1 illustrates their structure.

The number of nodes in the hidden layer is a crucial determinant of prediction performance. Too few hidden layer nodes can weaken the nonlinear processing capability of a BPNN. Too many, however, complicate network training and promote overfitting. This study used a trial-and-error method to determine the optimum number of hidden layer nodes, adjusting the constant a in Eq. (12)

$$l = \sqrt{m+n} + a \tag{12}$$

where l is the number of hidden layer nodes, m is the number of nodes in the input layer and n is the number in the output layer. a is the adjustment constant, and ranged from 1 to 10. Through experimentation with varying numbers of neuron nodes, this constant helps in determining the optimal number of hidden layer neurons. The process involves evaluating operational efficiency and accuracy across different configurations to identify the most suitable neural network structure.

Modeling with a BPNN involves first data collection and analysis, and problem formulation. The network's structure and learning process must then be decided. That is followed by the training, guided by evaluation of the model's predictions. The modeling process is depicted in Fig. 2. Additional details are available in the supporting information.

Training samples are one of the key factors affecting the nerve. When the number of samples required by the network during training is related to the complexity of the nonlinear mapping relationship, the accuracy of the network mapping will generally be enhanced with the increase of the number of samples, but there will be a critical point. When the number of training samples accumulates to a certain amount, the accuracy of the neural network will not change. Therefore, the selection of training samples should be representative, and overlap should be avoided, and the scope of working characteristics of the research object should be covered and satisfied as far as possible. Due to the complexity and time consuming of COD component analysis method, the amount of data obtained is limited. The data selected is the average value of daily samples. Based on the above standards, the sample data in this paper are shown in Table S-1 (Supporting information, SI).

Network training needs to set learning rate lr, training step size and training accuracy. Training accuracy generally decreases as the training step length increases, though this relationship is not entirely restrictive. A lower learning rate results in a smaller step size, which means that weight adjustments during the learning process are more cautious, leading to slower convergence. Conversely, a higher learning rate allows for more aggressive weight adjustments, which can accelerate convergence but may also cause oscillations around the optimal solution or even lead to divergence if the learning rate is too high. Under the condition of system stability, the learning rate is generally set between 0.01 and 0.9. In this training, *purelin* linear transfer function is selected for the network output

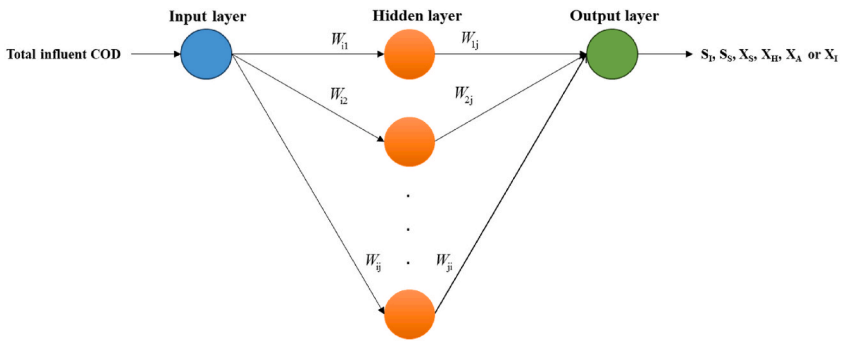


Fig. 1. BPNN structure.

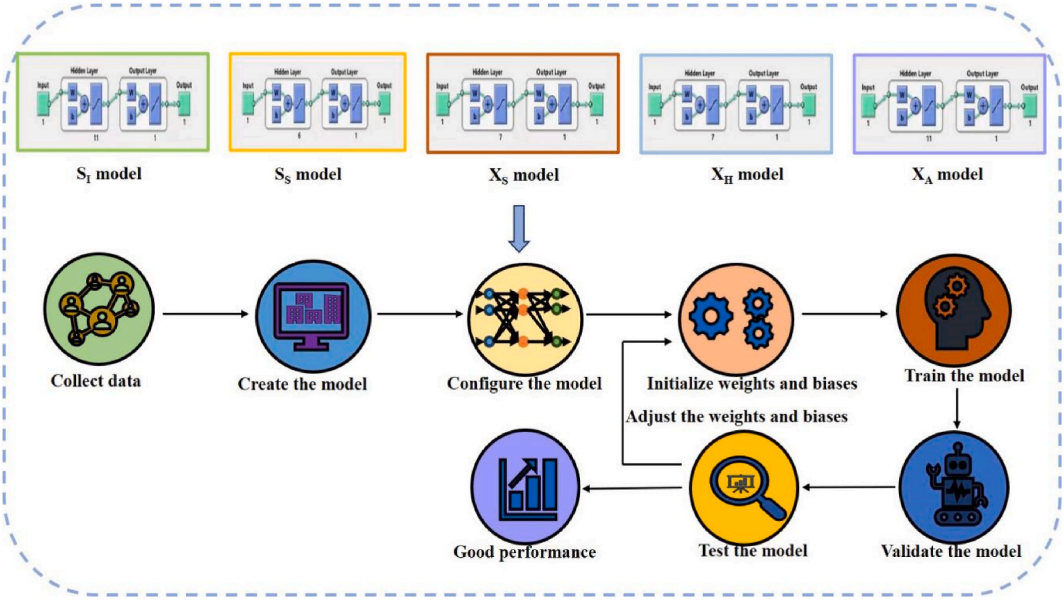


Fig. 2. Flowchart of building and training a BPNN.

layer and *sigmoid* activation function is selected as transfer function for the hidden layer.

The parameters of the neural network are selected as follows: the maximum learning times of S_I, S_S, X_S, X_H is 1000, and the maximum learning times of X_A is 6000. The five prediction models had a learning rate of 0.01 and a training target of 0.001. *traingdm* was used for training function, *learngdm* was used for learning function, and *mse* was used for performance function. The *newff ()* function is used to create the network.

Several indices were used to assess the accuracy of the BPNN's predictions: root mean square error (RMSE) [39], mean square error (MSE) [40], mean absolute error (MAE) [41], and mean absolute percentage error (MAPE) [42]. Their calculation is specified in equations (13)–(16). In each case, t represents the predicted output, o is the actual output, j is jth sample and n represents the total number of samples. Each index ranges from 0 upward, with smaller values indicating better prediction performance.

$$\mathrm{RMSE} = \sqrt{\left(\frac{1}{n}\right) \times \sum_{j=1}^{n} \left(t_j - o_j\right)^2} \tag{13}$$

$$\mathrm{MAPE} = \left(\frac{o - t}{o}\right) \times 100 \tag{14}$$

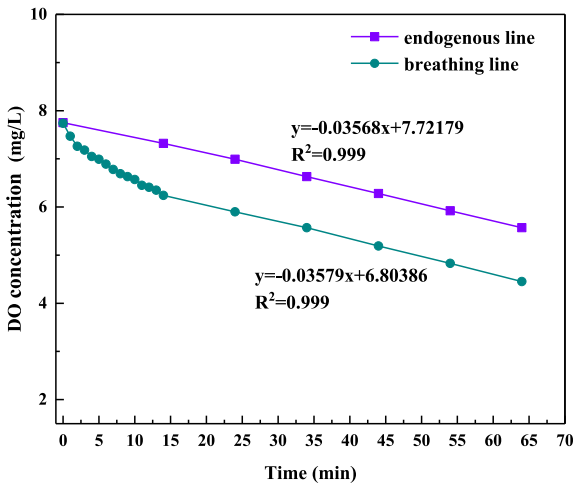


Fig. 3. Oxygen uptake rate graph.

$$\mathrm{MSE} = \left(\frac{1}{n}\right) \times \sum_{j=1}^{n} \left((t_j - o_j)\right)^2 \tag{15}$$

$$\mathrm{MAE} = \left(\frac{1}{n}\right) \times \sum_{j=1}^{n} \left|(t_j - o_j)\right| \tag{16}$$

3. Results and discussion

3.1. Decomposition of influent COD components in the WWTP

The yield coefficient of heterotrophic bacteria (Y_H) is a key variable in calculating components, and its accuracy determines the accuracy of the model [43]. Y_H values vary in different regions and must be measured experimentally. Fig. 3 shows that the yield

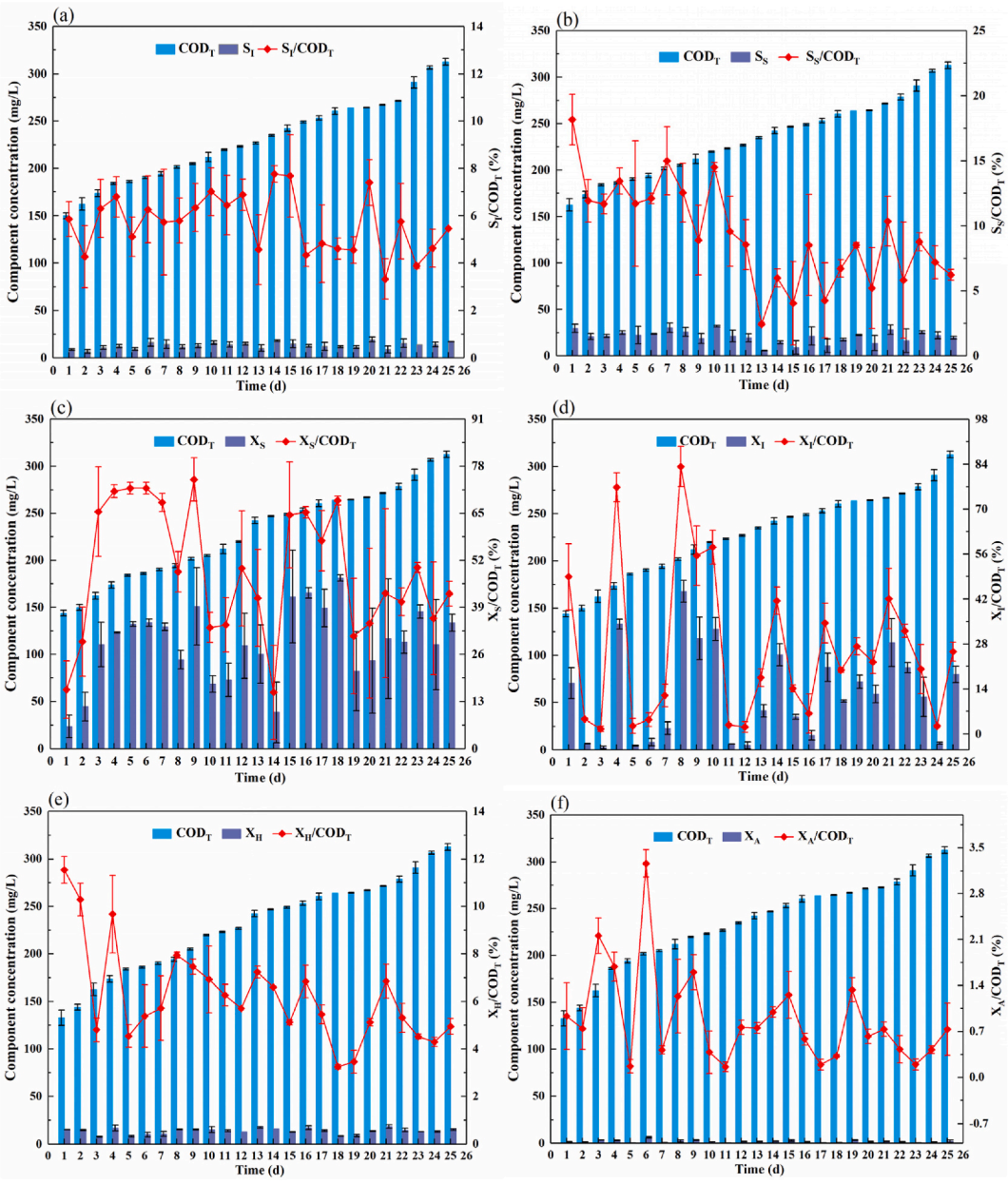


Fig. 4. The proportion of the total COD attributable to each component.

coefficient of heterotrophic bacteria in this study's WWTP. When endogenous line was basically consistent with the breathing line slope, the microorganisms consumed the rapidly biodegradable organic matter, and the DO content at this time could be substituted into Eq. (6) to calculate Y_H, which was 0.66 gVSS/(gCOD). This falls within the range of 0.5–0.67 gVSS/(gCOD) reported in previous studies [44]. The initial slope of the respiration line is slightly greater than that of the endogenous line. This is due to the fact that sodium acetate in the solution is not fully biodegradable, resulting in a slow biological degradation process [45]. Consequently, this leaded to a smaller change in △DO and a lower overall result. Nevertheless, the high correlation coefficient of 0.999 indicates that the detection accuracy was reliable.

Fig. 4(a)–(f) present the proportions of the total COD attributable to S_I, S_S, X_H and X_A. They show that the measured values of the four influent components varied little. The range of S_I was 6.85–19.57 mg/L, with an average of 13.30 mg/L. The range of S_I/COD_T was 3.87–7.76 %, with an average of 5.74 %. Some researchers have come to the same conclusion [33]. The range of S_S was 5.75–31.90 mg/L, with an average of 21.16 mg/L. S_S/COD_T ranged from 2.45 to 18.16 %, with an average of 9.51 %, lower than the range reported by relevant study [46]. This might be attributed to the long-distance pipeline transport of wastewater processed in the WWTP studied [47]. During the transportation process the vigorous activity of heterotrophic bacteria would consume some of the S_S. A certain amount of X_H was also found—7.77 to 18.60 mg/L with an average of 13.53 mg/L. X_H/COD_T ranged from 3.24 to 11.54 %, with an

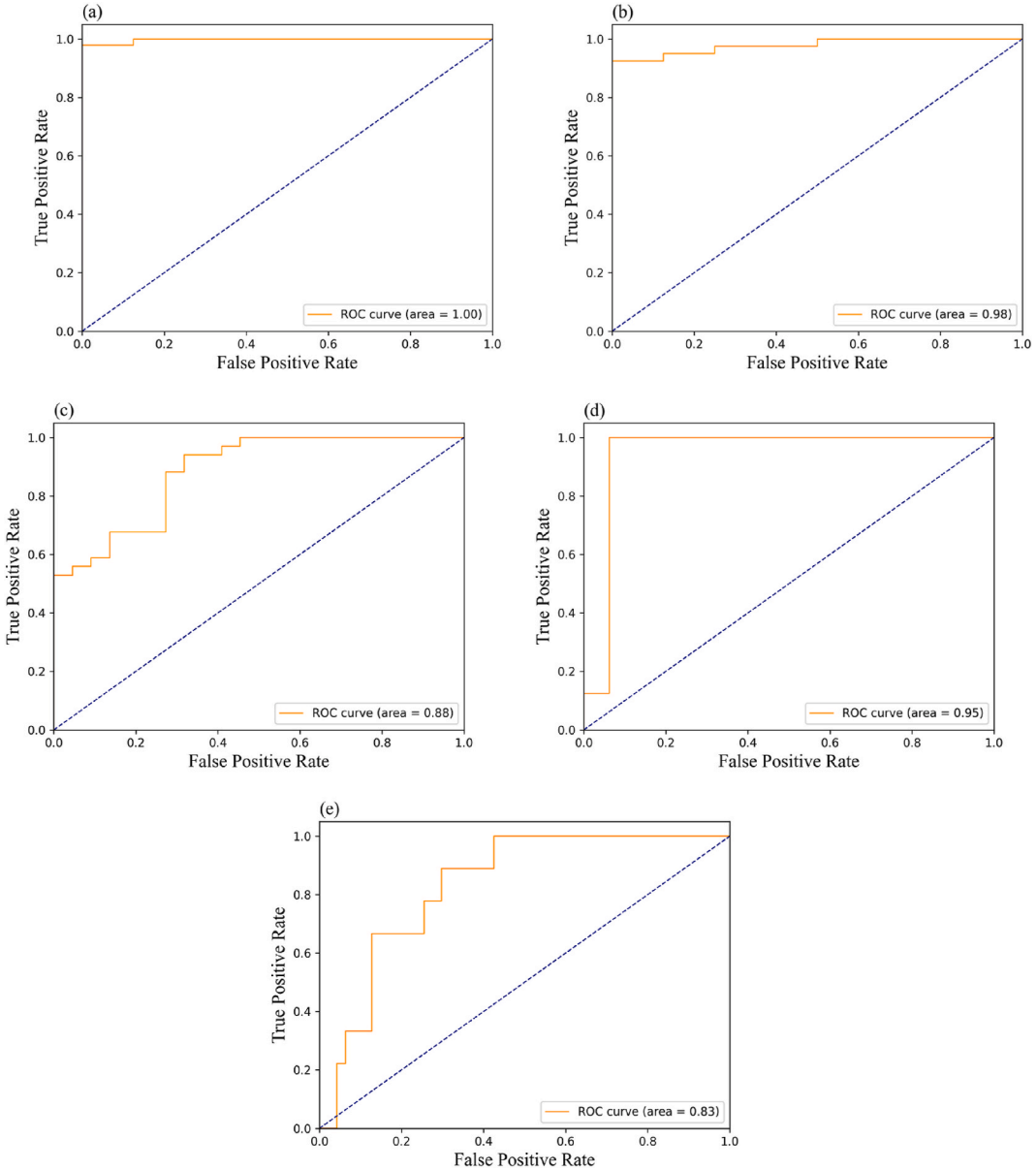


Fig. 5. ROC curves for the BPNN prediction model.

average of 6.06 %. That is within the reported range by previous research [33]. The range of X_A was 0.33–6.59 mg/L, with an average of 1.95 mg/L. The variation in X_A/COD_T ranged from 0.16 to 3.26 %, with an average of 0.89 %. The negligible proportion of this component led to its exclusion from further analysis. The fluctuations in the concentrations of the four remaining model components were low. The experimental results confirm that S_I, S_S, and X_H were reliable indicators, which offered both initial values and a training set for constructing the COD component prediction model. This validation improves the model's adaptability to S_I, S_S, and X_H, thereby aiding in the design of wastewater treatment processes and the creation of more refined forecasting models.

In Fig. 4 (c) and 4 (d) X_S and X_I show substantial fluctuations as COD_T contributors. COD_T ranged from 132.90 to 312.50 mg/L, with an average of 233.42 mg/L. The range of X_S was 29.30–184.30 mg/L, with an average of 112.05 mg/L and X_S/COD_T ranged from 8.37 to 81.55 %, averaging 49.29 %. That is a substantial range, but it agrees with the range reported by previous study. The range of X_I was 1.13–171.04 mg/L, so also very large. The range of X_I/COD_T was 0.34–64.80 %, with an average of 27.37 %. Compared with some reported similar studies, the data range of this study was larger. The difference probably arises from the fact that the wastewater received by this study's WWTP includes rainwater runoff from urban roads carrying a lot of inert particulate matter. More inert particulate matter reduces the wastewater's biochemical oxygen demand and its biodegradable COD [48]. The large variation in these

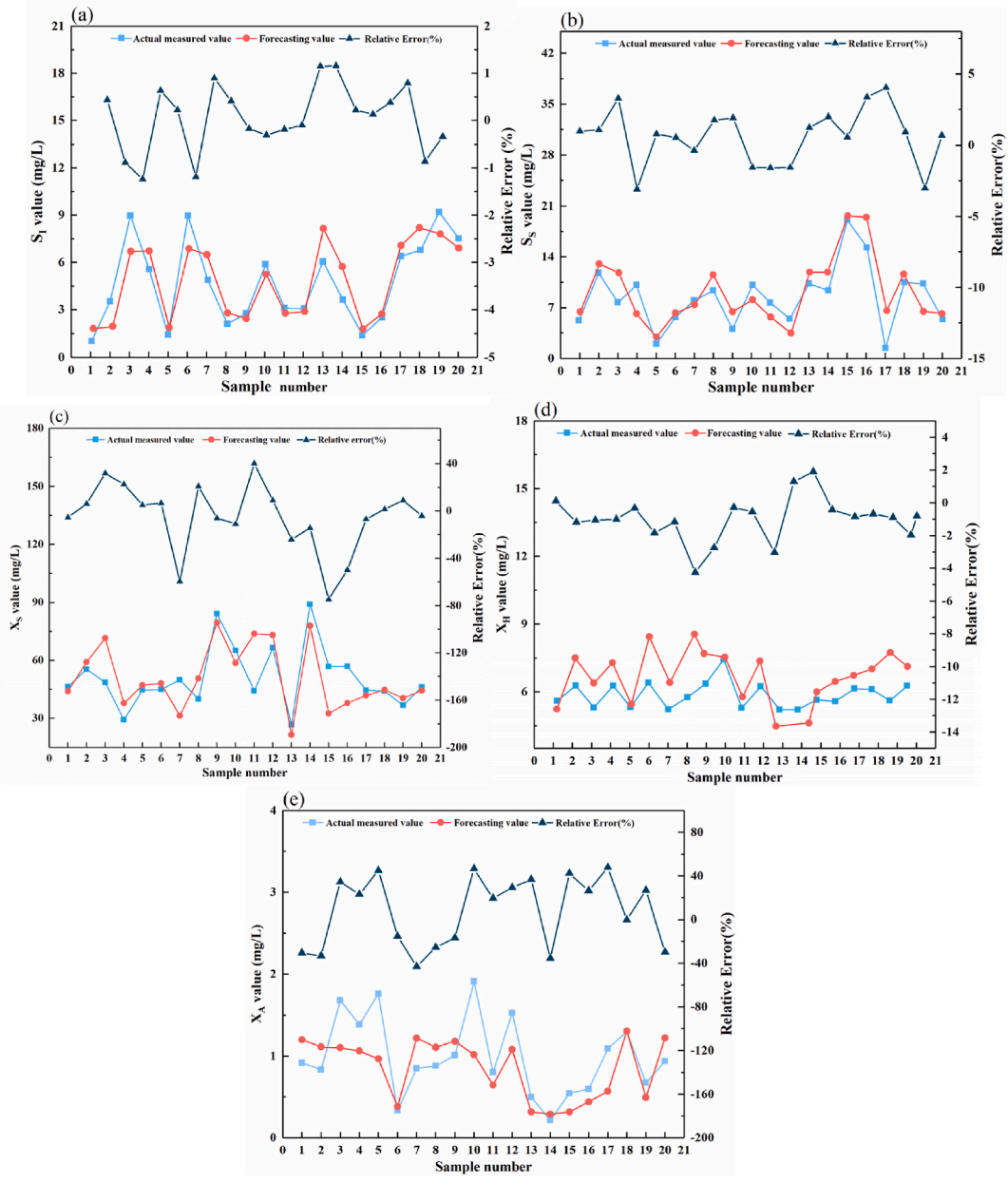


Fig. 6. Accuracy of the neural network's prediction models.

two components influenced by external factors makes it difficult to predict any patterns of variation. The BPNN dealt with the small fluctuations in S_I, S_S, X_H, and X_A easily, but COD_T, X_S, and X_I were much more difficult. Nevertheless, the BPNN's has powerful autonomous learning capabilities effectively addressed the nonlinear relationships involved and achieved useful predictive performance for the two components.

3.2. The BPNN's prediction performance

COD_T was used as the input variable for the BPNN, while S_I, S_S, X_S, X_H, and X_A were the outputs. Due to the variability and instability of COD components, normalization was applied, resulting in discretization of the detected value distribution for each component (Figure S-4, supplementary information). Consequently, the prediction models for each component needed to be independent of each other, and each model required adjustment based on the characteristics of the components. During the network training, testing, and validation process, the optimal number of hidden layer nodes for the five models was determined. It was found to be 11 for S_I, 6 for S_S, 11 for X_S, 7 for X_H, and 7 for X_A. The number of these nodes was set for both operational efficiency and prediction accuracy.

The receiver operating characteristics (ROC) curve (Fig. 5) was the metric used to assess the efficacy of the prediction model. This methodology involves comparing the prediction outcomes of the evaluation model with actual data to derive the correct rate and error rate of the training model's predictions on the test set. In this study, a prediction was deemed correct if the error between the predicted result and the actual data fell within a 10 % margin; otherwise, it was considered incorrect. And Fig. 5(a)–(e) shows the ROC curve of S_I, S_S, X_S, X_H, and X_A. The area under the curve (AUC) serves as an indicator of the model's accuracy, with a larger AUC signifying that the model's predictions were closely aligned with the actual data, thus denoting a superior prediction model [49]. In Fig. 5(a) and (b) and (d) the AUC exceeds 95 %, indicating a high-quality prediction model. The AUC in Fig. 5 (c) and (e) does not reach the 90 % threshold. That is probably due to the variability inherent in the analyses and to intrinsic differences among the components, such as the dataset quality is compromised due to the volatility in the measurement results of X_S, and within the data set of X_A, there is an issue with overly concentrated data distribution, giving rise to overfitting in the prediction model. That complicates the development of a sufficiently accurate prediction model.

In Fig. 6, the red curve represents the predicted values, the blue curve rep-resents the actual values, and the green curve represents the prediction error. It shows that S_I, S_S, X_S, X_H and X_A were all predicted with accuracy good enough to meet the needs of practical wastewater treatment. The predictive accuracy of S_I, S_S, X_S, X_H, and X_A, were 85.84 %, 84.25 %, 81.78 %, 83.54 %, 83.38 %, respectively. Only Xs showed an average pre-diction error exceeding 1.1 %, and average prediction error for all other component parameters such as S_I, S_S, X_H, X_A were 0.58 %, 1.05 %, 1.09 %, 0.07 %, respectively. The discrepancy was likely due to various influencing factors affecting X_S, such as weather conditions, pipeline conditions, and residents' living habits [50]. Additionally, the fact that the actual detected values of X_A fluctuate less than those in the training set suggested that the prediction model may misclassify stable input values as "wrong," which contributed to prediction errors. Despite this, a prediction accuracy of 83.38 % was acceptable for real-world applications. Overall, this level of accuracy demonstrated the viability of BPNN model for component prediction and their potential for future improvements.

4. Conclusions

A backpropagation neural network was created and trained using information on the influent wastewater quality components at a wastewater treatment plant. Six contributors to influent COD were modeled: S_S, X_S, S_I, X_A, and X_H. And their prediction accuracy was 84.25 %, 81.78 %, 85.84 %, 83.38 %, and 83.54 %, respectively. Monitoring data revealed that the levels of the particulate components X_S and X_I varied significantly, affected by factors such as weather conditions, pipeline conditions and residents' habits. The levels of the other four components studied varied less. The maximum deviation values of these parameters fall within the range of the actual detected values, suggesting that the model's predictions align well with real-world observations, and demonstrated prediction performance adequate for practical application in wastewater treatment. This research has demonstrated a new method for predicting the contributors to COD in wastewater, enabling more accurate predictions of the quality of the wastewater leaving a wastewater treatment plant.

Data availability statement

Data will be made available on request.

CRediT authorship contribution statement

Ping Wang: Writing – review & editing, Methodology, Investigation, Formal analysis. **Yanqiong Chen:** Software, Investigation, Data curation. **Chen Zhang:** Writing – original draft, Visualization, Methodology. **Yuzhen Shi:** Data curation, Conceptualization. **Bin Wang:** Project administration. **Chaochao Lai:** Validation. **Huan He:** Supervision. **Bin Huang:** Resources, Funding acquisition.

Declaration of competing interest

The authors have no competing interests to declare that are relevant to the content of this article.

Acknowledgements

This study was funded by the Foundation for Distinguished Young Talents of Yunnan Province (Grant No. 202101AV070006).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e35580.

References

[1] J. Filipe, R.J. Bessa, M. Reis, R. Alves, P. Povoa, Data-driven predictive energy optimization in a wastewater pumping station, Appl. Energy 252 (2019) 11343, https://doi.org/10.1016/j.apenergy.2019.113423.
[2] S. Reifsnyder, M. Garrido Baserba, F. Cecconi, L. Wong, P. Ackman, N. Melitas, et al., Relationship between manual air valve positioning, water quality and energy usage in activated sludge processes, Water Res. 173 (2020) 115537, https://doi.org/10.1016/j.watres.2020.115537.
[3] D. Torregrossa, G. Schutz, A. Cornelissen, F. Hernandez-Sancho, J. Hansen, Energy saving in WWTP: daily benchmarking under uncertainty and data availability limitations, Environ. Res. 148 (2016) 330–337, https://doi.org/10.1016/j.envres.2016.04.010.
[4] X. Li, S. Lei, G. Wu, Q. Yu, K. Xu, H. Ren, et al., Prediction of pharmaceuticals removal in activated sludge system under different operational parameters using an extended ASM-PhACs model, Sci. Total Environ. 871 (2023) 162065, https://doi.org/10.1016/j.scitotenv.2023.162065.
[5] J. Yang, B. Chen, Energy efficiency evaluation of wastewater treatment plants (WWTPs) based on data envelopment analysis, Appl. Energy 289 (2021) 116680, https://doi.org/10.1016/j.apenergy.2021.116680.
[6] A. Nuhoglu, B. Keskinler, E. Yildiz, Mathematical modelling of the activated sludge process—the Erzincan case, Process Biochem. 40 (2005) 2467–2473, https://doi.org/10.1016/j.procbio.2004.09.011.
[7] K.V. Gernaey, M.C.M. Van Loosdrecht, M. Henze, M. Lind, S.B. Jorgensen, Activated sludge wastewater treatment plant modelling and simulation: state of the art, Environ. Model. Software 19 (2004) 763–783, https://doi.org/10.1016/j.envsoft.2003.03.005.
[8] M. Henze, W. Gujer, T. Mino, M. Van Loosedrecht, Activated sludge models ASM1, ASM2, ASM2d and ASM3, Water Intell. Online 5 (2000), https://doi.org/10.2166/9781780402369, 9781780402369-9781780402369.
[9] M. Henze, W. Gujer, T. Mino, T. Matsuo, M.C. Wentzel, G. Marais, et al., Activated sludge model No. 2d, ASM2d, Water Sci. Technol. 39 (1999) 165–182, https://doi.org/10.1016/S0273-1223(98)00829-4.
[10] C. Martin, P.A. Vanrolleghem, Analysing, completing, and generating influent data for WWTP modelling: a critical review, Environ. Model. Software 60 (2014) 188–201, https://doi.org/10.1016/j.envsoft.2014.05.008.
[11] P. Kanthale, R. Pandey, D. Thakur, S.K. Gujar, P.R. Gogate, S. Thakre, et al., Application of combined hydrodynamic cavitation and Fenton reagent for COD reduction of cellulosic fiber industry effluents, J. Water Proc. Eng. 56 (2023) 104500, https://doi.org/10.1016/j.jwpe.2023.104500.
[12] J.M. Bidu, K.N. Njau, M. Rwiza, B. Van Der Bruggen, Textile wastewater treatment in anaerobic reactor: influence of domestic wastewater as co-substrate in color and COD removal, S. Afr. J. Chem. Eng. 43 (2023) 112–121, https://doi.org/10.1016/j.sajce.2022.10.007.
[13] T. Rajaee, A. Boroumand, Forecasting of chlorophyll-a concentrations in South San Francisco Bay using five different models, Appl. Ocean Res. 53 (2015) 208–217, https://doi.org/10.1016/j.apor.2015.09.001.
[14] G. Zhang, Time series forecasting using a hybrid ARIMA and neural network model, Neurocomputing 50 (2003) 159–175, https://doi.org/10.1016/s0925-2312(01)00702-0.
[15] F. Chang, P. Chen, L. Chang, Y. Tsai, Estimating spatio-temporal dynamics of stream total phosphate concentration by soft computing techniques, Sci. Total Environ. 562 (2016) 228–236, https://doi.org/10.1016/j.scitotenv.2016.03.219.
[16] S. Sadeghi Tabas, N. Humaira, S. Samadi, N.C. Hubig FlowDyn, A daily streamflow prediction pipeline for dynamical deep neural network applications, Environ. Model. Software 170 (2023), https://doi.org/10.1016/j.envsoft.2023.105854.
[17] S. Aslan, F. Zennaro, E. Furlan, A. Critto, Recurrent neural networks for water quality assessment in complex coastal lagoon environments: a case study on the Venice Lagoon, Environ. Model. Software 154 (2022), https://doi.org/10.1016/j.envsoft.2022.105403.
[18] D.B. May, M. Sivakumar, Prediction of urban stormwater quality using artificial neural networks, Environ. Model. Software 24 (2009) 296–302, https://doi.org/10.1016/j.envsoft.2008.07.004.
[19] J. Anmala, O.W. Meier, A.J. Meier, S. Grubbs, GIS and Artificial neural network–based water quality model for a stream network in the upper green river basin, Kentucky, USA, J. Environ. Eng. 141 (2015) 04014082, https://doi.org/10.1061/(ASCE)EE.1943-7870.0000801.
[20] M. El Kateb, C. Trellu, A. Darwich, M. Rivallin, M. Bechelany, S. Nagarajan, et al., Electrochemical advanced oxidation processes using novel electrode materials for mineralization and biodegradability enhancement of nanofiltration concentrate of landfill leachates, Water Reasearch 162 (2019) 446–455, https://doi.org/10.1016/j.watres.2019.07.005.
[21] S. Chen, G. Fang, X. Huang, Y. Zhang, Water quality prediction model of a water diversion project based on the Improved Artificial Bee Colony–backpropagation neural network, Water 10 (2018) 806, https://doi.org/10.3390/w10060806.
[22] L. Zhang, Z. Zou, W. Shan, Development of a method for comprehensive water quality forecasting and its application in Miyun reservoir of Beijing, China, J. Environ. Sci. 56 (2017) 240–246, https://doi.org/10.1016/j.jes.2016.07.017.
[23] Y. Zhao, J. Nan, F. Cui, L. Guo, Water quality forecast through application of BP neural network at Yuqiao reservoir, J. Zhejiang Univ. - Sci. 8 (2007) 1482–1487, https://doi.org/10.1631/jzus.2007.A1482.
[24] H. Lin, Y. Kang, D. Wang, Z. Lu, W. Tian, S. Wang, Forecast of water quality along the luanhe river line based on BP neural network, IOP Conf. Ser. Earth Environ. Sci. 267 (2019), https://doi.org/10.1088/1755-1315/267/3/032075.
[25] H. Xu, B. Lv, J. Chen, L. Kou, H. Liu, M. Liu, Research on a prediction model of water quality parameters in a marine ranch based on LSTM-BP, Water 15 (2023) 2760, https://doi.org/10.3390/w15152760.
[26] R. Huang, C. Ma, J. Ma, X. Huangfu, Q. He, Machine learning in natural and engineered water systems, Water Res. 205 (2021) 117666, https://doi.org/10.1016/j.watres.2021.117666.
[27] L. Zhao, T. Dai, Z. Qiao, P. Sun, J. Hao, Y. Yang, Application of artificial intelligence to wastewater treatment: a bibliometric analysis and systematic review of technology, economy, management, and wastewater reuse, Process Saf. Environ. Protect. 133 (2020) 169–182, https://doi.org/10.1016/j.psep.2019.11.014.
[28] P. Antwi, D. Zhang, L. Xiao, F.T. Kabutey, F.K. Quashie, W. Luo, et al., Modeling the performance of single-stage nitrogen removal using anammox and partial nitritation (SNAP) process with backpropagation neural network and response surface methodology, Sci. Total Environ. 690 (2019) 108–120, https://doi.org/10.1016/j.scitotenv.2019.06.530.
[29] W. Zhang, N.B. Tooker, A. Mueller, V. Enabling wastewater treatment process automation: leveraging innovations in real-time sensing, data analysis, and online controls, Environmental Science-Water Research & Technology 6 (2020) 2973–2992, https://doi.org/10.1039/d0ew00394h.
[30] D. Hanbay, I. Turkoglu, Y. Demir, Prediction of chemical oxygen demand (COD) based on wavelet decomposition and neural networks, Clean: Soil, Air, Water 35 (2007) 250–254, https://doi.org/10.1002/clen.200700039.

[31] P.A. Vanrolleghem, H. Spanjers, B. Petersen, P. Ginestet, I. Takacs, Estimating (combinations of) activated sludge model no. 1 parameters and components by respirometry, Water Sci. Technol. 39 (1999) 195–214, https://doi.org/10.2166/wst.1999.0042.
[32] A.S. Ahmed, A. Khalil, Y. Ito, M.C.M. Van Loosdrecht, D. Santoro, D. Rosso, et al., Dynamic impact of cellulose and readily biodegradable substrate on oxygen transfer efficiency in sequencing batch reactors, Water Res. 190 (2021) 116724, https://doi.org/10.1016/j.watres.2020.116724.
[33] J. Makinia, S.A. Wells, A general model of the activated sludge reactor with dispersive flow—I. model development and parameter estimation, Water Res. 34 (2000) 3987–3996, https://doi.org/10.1016/S0043-1354(00)00150-0.
[34] M. Mainardis, M. Buttazzoni, M. Cottes, A. Moretti, D. Goi, Respirometry tests in wastewater treatment: why and how? A critical review, Sci. Total Environ. (2021) 793, https://doi.org/10.1016/j.scitotenv.2021.148607.
[35] M. Mainardis, M. Buttazzoni, M. Cottes, A. Moretti, D. Goi, Respirometry tests in wastewater treatment: why and how? A critical review, Sci. Total Environ. 793 (2021) 148607, https://doi.org/10.1016/j.scitotenv.2021.148607.
[36] S.M. Tiar, M. Bessedik, C. Abdelbaki, N.B. Elsayed, A. Badraoui, A. Slimani, et al., Steady-State and dynamic simulation for wastewater treatment plant management: case study of Maghnia City, North-West Algeria, Water 16 (2024), https://doi.org/10.3390/w16020269.
[37] Y. Racault, A.E. Stricker, A. Marquot, ASM1 dynamic calibration and long-term validation for an intermittently aerated WWTP, Water Sci. Technol. 53 (2006) 247–256, https://doi.org/10.2166/wst.2006.427.
[38] M.J. Moya-Llamas, M.G. Pacazocchi, A. Trapote, Respirometric tests in a combined UASB-MBR system treating wastewater containing emerging contaminants at different OLRs and temperatures: biokinetic analysis, J. Environ. Manag. 345 (2023) 118643, https://doi.org/10.1016/j.jenvman.2023.118643.
[39] L. Fang, B. He, A deep learning framework using multi-feature fusion recurrent neural networks for energy consumption forecasting, Appl. Energy 348 (2023) 121563, https://doi.org/10.1016/j.apenergy.2023.121563.
[40] S.U. Khan, N. Khan, F.U.M. Ullah, M.J. Kim, M.Y. Lee, S.W. Baik, Towards intelligent building energy management: AI-based framework for power consumption and generation forecasting, Energy Build. 279 (2023) 112705, https://doi.org/10.1016/j.enbuild.2022.112705.
[41] N. Jin, F. Yang, Y. Mo, Y. Zeng, X. Zhou, K. Yan, et al., Highly accurate energy consumption forecasting model based on parallel LSTM neural networks, Adv. Eng. Inf. 51 (2022), https://doi.org/10.1016/j.aei.2021.101442.
[42] H. Dong, J. Zhu, S. Li, W. Wu, H. Zhu, J. Fan, Short-term residential household reactive power forecasting considering active power demand via deep Transformer sequence-to-sequence networks, Appl. Energy 329 (2023) 120281, https://doi.org/10.1016/j.apenergy.2022.120281.
[43] W.L. Lv, S.Z. Gou, Y.W. Chen, Y.F. Sun, X.D. Gao, Z.L. Tan, et al., Determination and analysis of kinetic parameters of heterotrophic bacteria in mathematical simulation, Appl. Ecol. Environ. Res. 16 (2018) 1507–1518, https://doi.org/10.15666/aeer/1602_15071518.
[44] X. Lu, T.D.S. Pereira, H.E. Al-Hazmi, J. Majtacz, Q. Zhou, L. Xie, et al., Model-based evaluation of N2O production pathways in the anammox-enriched granular sludge cultivated in a sequencing batch reactor, Environ. Sci. Technol. 52 (2018) 2800–2809, https://doi.org/10.1021/acs.est.7b05611.
[45] S. Kliem, M. Kreutzbruck, C. Bonten, Review on the biological degradation of polymers in various environments, Materials 13 (2020) 4586, https://doi.org/10.3390/ma13204586.
[46] X. Wu, Y. Yang, G. Wu, J. Mao, T. Zhou, Simulation and optimization of a coking wastewater biological treatment process by activated sludge models (ASM), J. Environ. Manag. 165 (2016) 235–242, https://doi.org/10.1016/j.jenvman.2015.09.041.
[47] P. Liu, J. Dai, K. Huang, Z. Yang, Z. Zhang, X. Guo, Sources of micro(nano)plastics and interaction with co-existing pollutants in wastewater treatment plants, Crit. Rev. Environ. Sci. Technol. 53 (2023) 865–885, https://doi.org/10.1080/10643389.2022.2095844.
[48] C. Fall, E. Millán-Lagunas, K.M. Bâ, I. Gallego-Alarcón, D. García-Pulido, C. Díaz-Delgado, et al., COD fractionation and biological treatability of mixed industrial wastewaters, J. Environ. Manag. 113 (2012) 71–77, https://doi.org/10.1016/j.jenvman.2012.08.032.
[49] K. Attwood, S. Hou, A. Hutson, Application of the skew exponential power distribution to ROC curves, J. Appl. Stat. 50 (2023) 1709–1724, https://doi.org/10.1080/02664763.2022.2037528.
[50] A. Damayanti, Z. Ujang, M.R. Salim, G. Olsson, A.Z. Sulaiman, Respirometric analysis of activated sludge models from palm oil mill effluent, Bioresour. Technol. 101 (2010) 144–149, https://doi.org/10.1016/j.biortech.2009.08.034.