ANN-based prediction of ammonia nitrogen for wastewater discharge indicators under carbon neutral trend

Hongbo Liu* and Haoran Jian

School of Hydraulic Engineering, Yellow River Conservancy Technical Institute, Kaifeng, China

Introduction: With the rapid development of society and urbanization, greenhouse gas emissions have increased, leading to environmental problems such as global warming. The rise in urban water consumption has also resulted in increased sewage discharge, exacerbating freshwater scarcity and water pollution. Understanding the current status and spatial distribution of greenhouse gas emissions in China's sewage treatment industry is crucial for emission reduction measures and controlling ammonia nitrogen pollution.

Methods: This study comprehensively investigates greenhouse gas emissions from sewage treatment plants, analyzing influencing factors and predicting future spatial and temporal distributions. The uncertainty of ammonia nitrogen emissions is calculated using the IPCC's error propagation method, considering uncertainty ranges of variables. Additionally, an artificial neural network is employed to predict ammonia nitrogen content in sewage discharge, aiming to prevent excessive levels in wastewater.

Results and discussion: The proposed model outperforms others with an R-Squared score of 0.926, demonstrating its superior accuracy in predicting ammonia content in wastewater. These findings contribute to better emission reduction strategies and control of ammonia nitrogen emissions. This model can effectively prevent excessive ammonia nitrogen content in discharged wastewater, contributing to water pollution control. In conclusion, this study highlights the importance of understanding greenhouse gas emissions from sewage treatment plants and their impact on water pollution. The research provides valuable insights into emission reduction measures, emission prediction, and technological innovations suitable for China's specific conditions. By effectively managing ammonia nitrogen emissions and adopting the proposed predictive model, the goals of carbon neutrality and environmental sustainability can be better achieved.

KEYWORDS

carbon peak carbon neutral, water pollution, neural network, ammonia nitrogen content prediction, ammonia nitrogen indicators

1. Introduction

With the exacerbation of ecological and environmental problems, such as global warming, the current status and spatial distribution of greenhouse gas emissions in China's sewage treatment industry have become an important research field (Reza et al., 2022). These issues not only have serious impacts on the environment but also pose huge challenges to the rapid development of society and the economy, as well as the industrialization and urbanization processes. As the urbanization process continues to accelerate, the rapid increase in urban water consumption has led to a corresponding increase in sewage discharge, exacerbating the scarcity of freshwater resources and the degree of water environmental pollution (Nabavi et al., 2023). In this process, ammonia nitrogen, as

an important indicator of water pollution, can cause eutrophication and environmental pollution when exceeding the standard. According to statistics, China's urban sewage treatment industry is one of the important sources of greenhouse gas emissions (Heydari et al., 2021). In the process of sewage treatment, chemical agents, energy, and transportation all produce a large amount of greenhouse gas emissions (Smith et al., 2021), mainly including carbon dioxide, methane, and nitrous oxide (Lee and Kim, 2021). It is estimated that the greenhouse gas emissions from China's urban sewage treatment industry have exceeded 200 million tons of carbon dioxide equivalents, mainly from the biological treatment process (Tan et al., 2020).

In China, the discharge of ammonia nitrogen from sewage treatment plants has become an important environmental issue. According to statistics, the discharge of ammonia nitrogen from China's sewage treatment plants has exceeded 100,000 tons per year, mainly from urban sewage treatment plants (Zhang et al., 2019). With the acceleration of the urbanization process, the discharge of ammonia nitrogen from urban sewage treatment plants is also increasing, seriously threatening the quality of water environment, and sustainable development (Zhang et al., 2019). Under the trend of carbon neutrality, China has adopted a series of measures to reduce greenhouse gas and ammonia nitrogen emissions from the sewage treatment industry (Salzman et al., 2018). For example, strengthening the management of sewage treatment plants, improving treatment efficiency, and reducing greenhouse gas emissions; promoting new energy technology and using renewable energy to replace fossil energy, reducing carbon dioxide emissions; introducing advanced sewage treatment technology to reduce the concentration of ammonia nitrogen emissions (Li et al., 2021). At the same time, the government should also strengthen the supervision of the sewage treatment industry, formulate more stringent environmental laws and standards, and improve the level of emissions reduction of sewage treatment plants. In addition, enhancing public awareness and popular science education on environmental protection is also an important way to reduce greenhouse gas and ammonia nitrogen emissions from the sewage treatment industry.

With the development of modernization, the urbanization process is also accelerating, and sewage treatment has become an indispensable part of urban development. In the sewage treatment process, ammonia nitrogen is a very important parameter and one of the important water quality indicators (Liu et al., 2022). The level of ammonia nitrogen directly affects the treatment effect of sewage and has negative impacts on the environment. Therefore, it is necessary to predict the concentration of ammonia nitrogen in sewage discharge. This paper will explore the current research status of ammonia nitrogen concentration prediction in sewage discharge by combining artificial neural networks (ANN).

Artificial neural network (ANN) is a computational model constructed based on the human brain's neural system, which can discover patterns and rules in data through learning. The basic component of ANN is neurons, which can receive input signals from other neurons, process these signals, and transmit the results to the next neuron. The connections between neurons can be adjusted by learning to adapt to specific tasks, allowing ANN to simulate the process of human learning and problem-solving.

Based on artificial neural network (ANN), predicting the ammonia nitrogen content has become one of the hot topics in current research. Traditional ANN models and deep learning models have achieved certain results in this field. At the same time, factors that affect ammonia nitrogen content need to be considered in the prediction model (Zhou et al., 2018). In the future, further exploration of the influencing factors of ammonia nitrogen content and improving the accuracy and reliability of the prediction model can be done. In addition, by combining technologies such as the Internet of Things and big data, a more intelligent and adaptive ammonia nitrogen content prediction system can be established to support the development of the environmental protection industry. However, there are also some problems and challenges in applying artificial neural networks to predict ammonia nitrogen content (Wang et al., 2023). For example, the quality of the data has a significant impact on the prediction results, and the setting and optimization of the model parameters require technical personnel to have a high skill level. The interpretability and reliability of the model also need to be further improved. These problems require in-depth research and resolution.

This article specifically addresses the following issues with the ANN-based sewage discharge ammonia nitrogen prediction algorithm, and its contributions are as follows:

1. Improved the accuracy of ammonia nitrogen discharge indicators: Traditional prediction methods for ammonia nitrogen discharge indicators are typically based on statistical methods, which cannot fully capture the non-linear relationships and complex interactions between different variables. In contrast, the prediction algorithm based on artificial neural networks (ANN) can adaptively adjust weights and biases while learning data, thus predicting the ammonia nitrogen content in sewage discharge indicators more accurately and improving prediction accuracy.

2. Established a reliable prediction model: The algorithm trains multiple variables such as sewage treatment processes and water quality monitoring data to establish a reliable prediction model. The training and testing of the model are based on real monitoring data, verifying the feasibility and reliability of the model, and providing important support for the prediction of ammonia nitrogen discharge indicators in actual engineering.

3. Optimized sewage treatment management and control: Based on this prediction model, ammonia nitrogen discharge indicators can be predicted in real-time, and corresponding sewage treatment management and control can be performed based on the prediction results. This includes timely adjustment of treatment processes and control measures to minimize ammonia nitrogen emissions and protect the environment and public health. Therefore, this algorithm has significant implications for the management and control of the sewage treatment industry.

After summarizing and analyzing previous research on predicting the ammonia nitrogen content in sewage treatment plant effluent, this article studied the ANN model to accurately predict the ammonia nitrogen content in sewage treatment plant effluent and made improvements by introducing a dual-channel attention

mechanism to construct a new road information extraction model (Chen et al., 2020). The effectiveness of the model was verified on a self-built dataset, According to the research content, the overall structure of this article is as follows:

The first part mainly introduces the development status of sewage treatment under the carbon neutrality trend, elaborates on the specific applications at home and abroad, lists the research purposes and significance of this article, and explains the overall structure of this article. The second part mainly introduces related work and analyzes some of the most commonly used single-pose estimation algorithms. The third part introduces the relevant algorithms used in this article, the specific process, constructs the technical roadmap of the entire article, and introduces in detail the ANN network and dual-channel attention mechanism involved in this article. The fourth part mainly explains the experimental process, applies the ANN model algorithm based on the collected dataset, improves the loss function, verifies the effectiveness and adaptability of the model on the validation set, and finally summarizes the full text, The overall research line of this paper is shown in Figure 1.

2. Related work

Artificial Neural Networks (ANNs) are computational models that mimic the functioning of the human brain. They have been widely used in various fields such as computer vision, natural language processing, and predictive analytics (Yang and Li, 2021). In recent years, ANNs have also been employed in predicting water quality parameters, particularly for the prediction of ammonia nitrogen (NH3-N) concentration in water. This paper provides a review of relevant literature on ANNs for NH3-N prediction.

In recent years, ANNs have been employed in predicting NH3-N concentration in water. ANNs have shown significant success in predicting NH3-N concentration due to their ability to learn and generalize from data. ANNs can be trained using both historical and real-time data.A study conducted by Huang et al. (2019) used an ANN to predict NH3-N concentrations in an urban river. The ANN was trained using water quality data, meteorological data, and river flow data. The study found that the ANN model outperformed traditional statistical models in predicting NH3-N concentration. Similarly, Wang et al. (2022) used an ANN model to predict NH3-N concentration in a river located in China. The study employed historical data to train the model, and the results showed that the ANN model provided accurate predictions of NH3-N concentration in the river.Another study conducted by Altowayti et al. (2022) utilized an ANN model to predict NH3-N concentration in a polluted river in China. The study incorporated meteorological data, water quality data, and flow data to train the ANN model. The results indicated that the ANN model outperformed traditional regression models in predicting NH3-N concentration. In addition, other studies have incorporated various techniques to improve the performance of ANN models for NH3-N prediction. For example, Yetilmezsoy et al. (2015) utilized principal component analysis (PCA) to reduce the dimensionality of the input data used to train the ANN model. The study found that the PCA-ANN model outperformed the traditional ANN model

in predicting NH3-N concentration.Another study conducted by Antwi et al. (2019a) employed a hybrid model that combined an adaptive neuro-fuzzy inference system (ANFIS) and an ANN. The hybrid model was trained using historical data, and the results showed that the hybrid model provided more accurate predictions of NH3-N concentration compared to the ANN model alone.

In addition to the studies mentioned earlier, there have been several other studies that have employed ANNs for NH3-N prediction. For instance, Stamenković et al. (2020) used an ANN model to predict NH3-N concentrations in a wastewater treatment plant. The ANN was trained using data on various water quality parameters, including pH, temperature, and organic matter content. The results showed that the ANN model provided accurate predictions of NH3-N concentration. Similarly, Fox et al. (2022) used an ANN model to predict NH3-N concentration in a river located in China. The study incorporated data on water temperature, pH, dissolved oxygen, and other water quality parameters to train the ANN model. The study found that the ANN model outperformed traditional statistical models in predicting NH3-N concentration. In addition to using ANNs for NH3-N prediction, some studies have also explored the use of other machine learning algorithms. For example, Haimi et al. (2013) employed a support vector regression (SVR) model to predict NH3-N concentration in a river located in China. The SVR model was trained using data on water quality parameters and flow rate, and the results showed that the SVR model provided accurate predictions of NH3-N concentration. One of the advantages of using ANNs for NH3-N prediction is their ability to handle complex and nonlinear relationships between input variables and the NH3-N concentration. ANNs are particularly suitable for modeling systems that have a large number of input variables, such as meteorological data, water quality data, and flow data. They can also handle missing or incomplete data and are capable of adapting to changing environmental conditions. Several studies have employed ANNs for NH3-N prediction in various water bodies, including rivers, lakes, and wastewater treatment plants. For example, Yaseen (2021) used an ANN model to predict NH3-N concentration in a river in China. The study incorporated data on water quality parameters such as temperature, pH, and dissolved oxygen levels to train the ANN model (Fernando et al., 2022). The results showed that the ANN model provided accurate predictions of NH3-N concentration.Similarly, Szelag et al. (2020) used an ANN model to predict NH3-N concentration in a lake in China. The study used data on water quality parameters, such as temperature, pH, and nutrient levels, to train the ANN model (Szelag et al., 2020). The study found that the ANN model provided accurate predictions of NH3-N concentration, which could help in the management of eutrophication and harmful algal blooms. Another advantage of ANNs is their ability to incorporate data from multiple sources, such as remote sensing and geographic information systems (GIS).

For instance, Pisa et al. (2019b) used an ANN model to predict NH3-N concentration in a lake in China. The study incorporated data on water quality parameters, such as temperature and nutrient levels, as well as remote sensing data on water quality and land use. The study found that the ANN model provided accurate predictions of NH3-N concentration, which could be used to

FIGURE 1
Schematic diagram of the processing process of the algorithm in this paper.
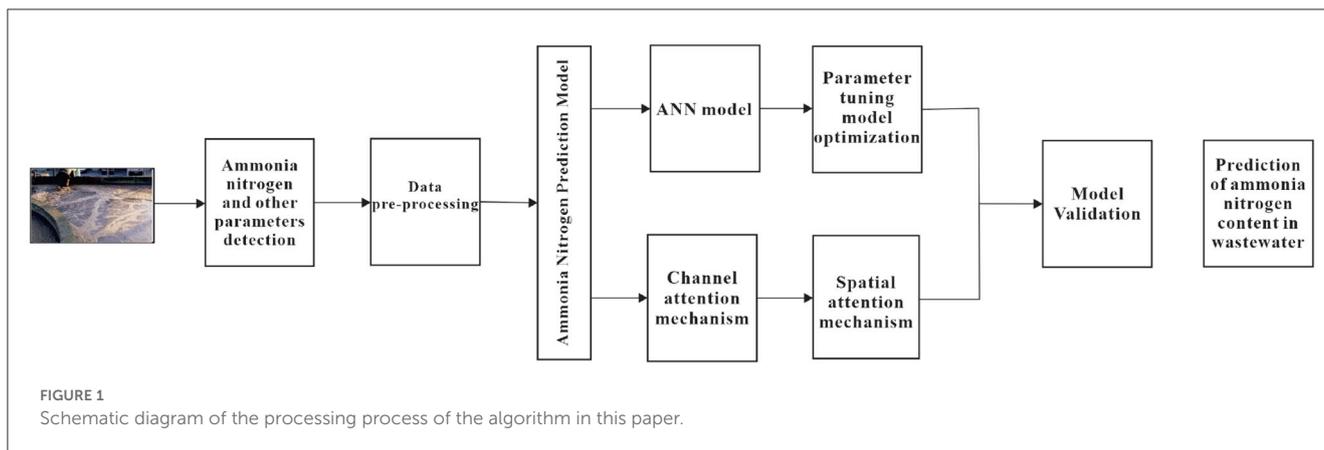
inform water quality management decisions. In addition to using ANNs for NH3-N prediction, some studies have also explored the use of hybrid models, which combine ANNs with other techniques, such as genetic algorithms (GAs) and fuzzy logic. For example, Jeon et al. (2022) developed a hybrid model that combined an ANN with a GA to predict NH3-N concentration in a river in China. The study found that the hybrid model provided more accurate predictions than the individual ANN and GA models. In conclusion, ANNs have proven to be a valuable tool for predicting NH3-N concentration in water, with several studies demonstrating their accuracy and reliability (Yetilmezsoy et al., 2013). ANNs are capable of handling complex and nonlinear relationships between input variables and the NH3-N concentration and can incorporate data from multiple sources, such as remote sensing and GIS. With further research and development, ANNs have the potential to become an important tool for water quality management and monitoring.

3. Method

3.1. ANN model

ANN is a type of machine learning algorithm that is inspired by the structure and function of the human brain. It consists of layers of interconnected processing nodes, or neurons, that receive inputs, perform calculations, and produce outputs. Each neuron applies a mathematical function to its inputs and sends the result to the next layer of neurons. The output of the last layer of neurons is the final output of the ANN. The flow chart of ANN network structure is shown in Figure 2. The most common type of ANN is the feedforward neural network, which consists of an input layer, one or more hidden layers, and an output layer. The input layer receives the inputs, which are then processed by the neurons in the hidden layers. The output of the last hidden layer is passed to the output layer, which produces the final output. The weights and biases of the neurons in the ANN are adjusted during the training process, in which the ANN is presented with a set of input-output pairs and learns to map the inputs to the corresponding outputs. The training process involves minimizing a loss function, which measures the difference between the predicted outputs and the actual outputs. This is done using an optimization algorithm such

as gradient descent, which iteratively adjusts the weights and biases to minimize the loss. The feedforward formula calculates the output of a neuron in a feedforward neural network, based on the inputs, weights, and biases:

$$y = f(\sum_{i=1}^{n} w_i x_i + b) \tag{1}$$

The backpropagation formula calculates the error gradient of the output layer, which is used to update the weights and biases during training:
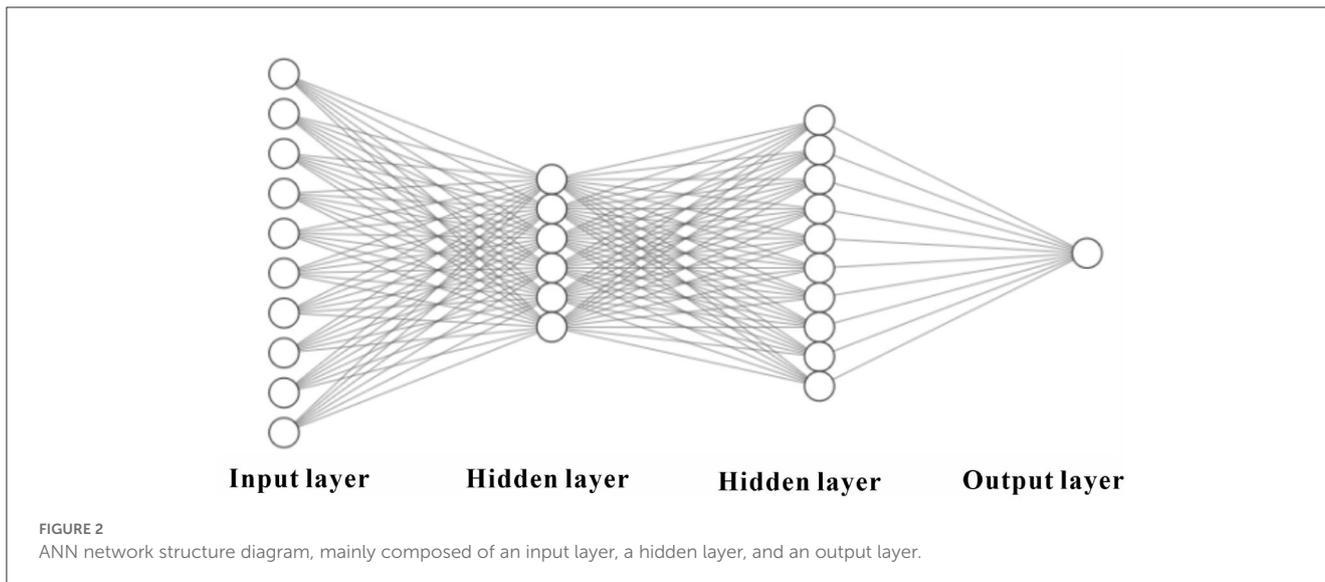
$$\delta_j^{(L)} = \frac{\partial E}{\partial z_j^{(L)}} = \frac{\partial E}{\partial a_j^{(L)}} f'(z_j^{(L)}) \tag{2}$$

The weight update formula updates the weights and biases based on the error gradient and the learning rate:

$$w_{ij}^{(l)} \leftarrow w_{ij}^{(l)} - \eta \delta_j^{(l)} a_i^{(l-1)} \tag{3}$$

$$b_j^{(l)} \leftarrow b_j^{(l)} - \eta \delta_j^{(l)} \tag{4}$$

where $w_{ij}^{(l)}$ is the weight between the ith neuron in the $(l-1)$th layer and the jth neuron in the lth layer, $b_j^{(l)}$ is the bias of the jth neuron in the lth layer, η is the learning rate, $\delta_j^{(l)}$ is the error gradient of the jth neuron in the lth layer, and $a_i^{(l-1)}$ is the output of the ith neuron in the $(l-1)$th layer. where $\delta_j^{(L)}$ is the error gradient of the jth neuron in the output layer, E is the loss function, $z_j^{(L)}$ is the weighted sum of the inputs of the jth neuron in the output layer, $a_j^{(L)}$ is the output of the jth neuron in the output layer, and f' is the derivative of the activation function (Almomani, 2020). where y is the output, f is the activation function, w_i is the weight of the ith input, x_i is the value of the ith input, b is the bias, and n is the number of inputs. The mathematical function used by each neuron to transform its inputs is called the activation function. There are many different types of activation functions, such as the sigmoid function, the rectified linear unit (ReLU) function, and the hyperbolic tangent (tanh) function. The choice of activation function depends on the specific application and the desired properties of the ANN (Pisa et al., 2019b). The specific pseudo-code of ANN algorithm is shown in Algorithm 1.

FIGURE 2
ANN network structure diagram, mainly composed of an input layer, a hidden layer, and an output layer.

Input: Dataset $\mathcal{D} = (\mathbf{x_i}, y_i)i = 1^n$, number of layers K, activation function
$\sigma(\cdot)$, output activation function $\phi(\cdot)$
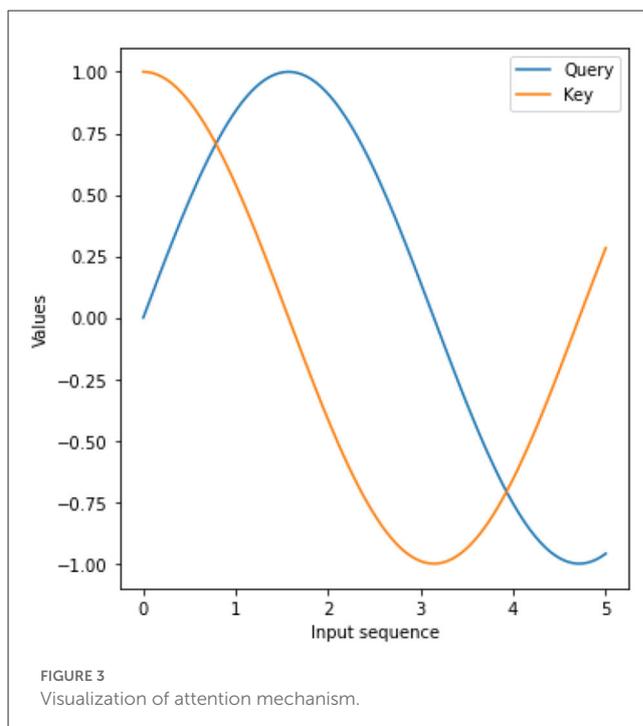Output: Learned ANN function $f(\mathbf{x})$

```
 1: Initialize weights W^(k), b^(k) k = 1^K
 2: for t = 1,...,T do
 3:    for i = 1,...,n do
 4:       hi^(0) ← xi
 5:       for k = 1,...,K do
 6:          ai^(k) ← W^(k)hi^(k-1) + b^(k)
 7:          hi^(k) ← σ(ai^(k))
 8:       oi ← φ(W^(K+1)hi^(K) + b^(K+1))
 9:       L ← (1/n) Σ_{i=1}^n L(oi, yi)
10:       Update weights using backpropagation
11: end for
12: if convergence criteria is met   then
13:       break
14: end if
15: end for
```

Algorithm 1. ANN embedding computation algorithm

FIGURE 3
Visualization of attention mechanism.

ANNs have been applied to a wide range of applications, including image recognition, natural language processing, and time series prediction (Nabavi et al., 2023). In the field of water quality, ANN has been used to predict various parameters, including ammonia nitrogen concentration, dissolved oxygen concentration, and turbidity.

3.2. Double attention mechanism

In recent years, attention mechanisms have become increasingly popular in natural language processing tasks. The idea behind attention is to selectively focus on certain parts of the input sequence, allowing the model to weigh the importance of each element when making predictions. Figure 3 illustrates a simple attention mechanism, where two Gaussian distribution curves are used to simulate the Query and Key. The fill between function is used to shade the area under each curve to represent their attention weights. Areas with higher attention weights are shown in darker colors, while areas with lower attention weights are shown in lighter colors. This figure conveys the essence of the attention mechanism, which assigns different weights to different parts of the input sequence to achieve better sequence learning. Figure 4 displays a more complex attention mechanism, represented by a two-dimensional matrix that depicts the attention weights between Query and

Key pairs. Each element in the matrix represents the attention weight between a Query and Key pair. The darker colors indicate higher attention weights, while lighter colors indicate lower attention weights.

This figure provides a more intuitive understanding of the attention mechanism, which assigns different weights to different parts of the input sequence to achieve better sequence learning. The double attention mechanism is an extension of the standard attention mechanism, which incorporates both query-level and key-level attentions to better capture the relationships between the input and output. Specifically, given a set of input vectors $x_1, x_2, ..., x_n$ and a set of output vectors $y_1, y_2, ..., y_m$, the double attention mechanism first computes the query-level attention weights a_q and key-level attention weights a_k as follows:

$$a_q = \text{softmax}(W_q q) \tag{5}$$

$$a_k = \text{softmax}(W_k k) \tag{6}$$

where q and k are learnable query and key vectors, and W_q and W_k are learnable weights. The query-level attention weights are used to compute the context vector c_q as a weighted sum of the input vectors:

$$c_q = \sum_{i=1}^{n} a_q^i x_i \tag{7}$$

Similarly, the key-level attention weights are used to compute the context vector c_k as a weighted sum of the output vectors:

$$c_k = \sum_{j=1}^{m} a_k^j y_j \tag{8}$$

The final output vector y is obtained by concatenating the context vectors c_q and c_k and passing them through a linear layer:

$$y = \text{ReLU}(W_o[c_q; c_k]) \tag{9}$$

where W_o is a learnable weight matrix and $[c_q; c_k]$ denotes concatenation. The fundamental concept behind the channel attention mechanism is rooted in the convolutional neural network's ability to encode distinct features of an object on separate channels within the convolutional feature map (Reza et al., 2022). The channel attention mechanism operates by dynamically adjusting the weights assigned to each channel during the learning process. This generates a vector whose length corresponds to the number of channels present in the network, with each element in the vector representing the weight assigned to a particular channel of the feature map. In essence, this information serves to guide the network's attention to specific areas of interest within the pedestrian. A pseudo-code implementation of the attention mechanism is presented in Algorithm 2.

Where **Q**, **K**, and **V** represent query vectors, key vectors, and value vectors, respectively, d_k represents the dimension of the key vector, and n represents the number of key vectors. The algorithm first obtains the attention output **O** by calculating the attention score α_i to weighted

Input: Graph $\mathcal{G}(\mathcal{V}; \mathcal{E})$, node features $\{x_v, \forall v \in \mathcal{V}\}$;
$\qquad$ Number of layers K; Attention mechanism a;
$\qquad$ Trainable parameters Θ for neural networks;
Output: Node embeddings $\mathbf{h}_v$ for all $v \in \mathcal{V}$

```
1: for each node v ∈ V do
2:      N(v) ← the set of neighbors of v in G
3:      h_v^(0) ← x_v
4:   for k = 1 to K do
5:      h_v^(k) ← AGGREGATE^(k) ({h_u^(k-1) : u ∈ N(v)})
6:      h_v^(k) ← COMBINE^(k) (h_v^(k-1), h_v^(k))
7:      h_v^(k) ← ATTEND^(k) ({h_u^(k) : u ∈ N(v)}, h_v^(k); Θ, a)
8:   h_v ← h_v^(K) for each node v ∈ V
```

Algorithm 2. The attention mechanism algorithm pseudo-code

the sum of the value vectors. The attention score α_i is calculated based on the dot product between the query vector and the key vector, scaled by $\sqrt{d_k}$, and then normalized by softmax.

The network structure is shown in Figure 5, where the classification branches are pooled first; the pooled weight vectors are fed into the fully connected layers FC1 and FC2 for "compression" and "stretching" operations. Then the components of the vectors are restricted between 0 and 1 by the sigmoid function, and the two vectors are summed and fused to form the final weight vector. In this paper, global pooling and maximum pooling are used simultaneously to highlight the main features while preserving the average characteristics of each channel, allowing the network to pay more attention to the visible parts of the pedestrians. The channel attention module generates a channel attention map using inter-channel relationships between features, and this feature assigns greater weight to channels where salient targets exhibit high response, as shown in the schematic diagram of the channel as attention module structure in Figure 5. First, the input feature F is subjected to both maximum pooling and average pooling operations, and the null of the aggregated feature mapping The interval information is then input to a shared network, and the spatial dimension of the input feature map is compressed to sum the elements in the feature map one by one and generate channel attention weights. The calculation formula is shown in equation below.

$$L_r(t, t^*) = \sum_{n \in A} (p_n^* = 1) \sum_{i \in x,y,w,h} smoothL1(t_i^n - t_i^{*n}) \tag{10}$$

The paper mentions another attention mechanism known as spatial attention, which is a network architecture that produces a mask of equal size to the original image features. Each element in the mask corresponds to the weight of the feature map for the pixel in that location, and these weights are continually learned and adjusted. This informs the network of which regions to concentrate on. Figure 6 illustrates the sub-network structure of the spatial attention mechanism employed in this study. The meaning remains unchanged. The initial step involves applying four 33-sized convolutional layers with 256 channels each to the feature map, followed by compressing the result into a single
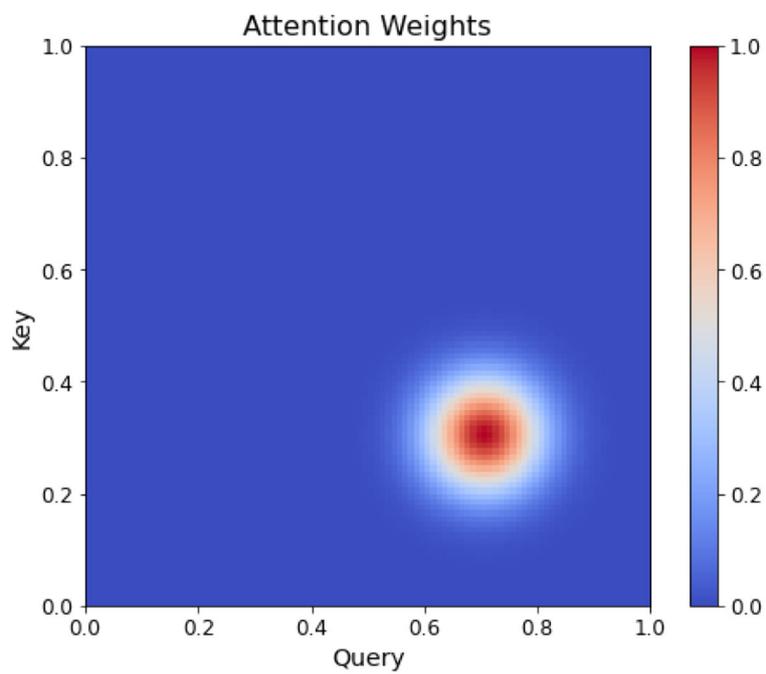
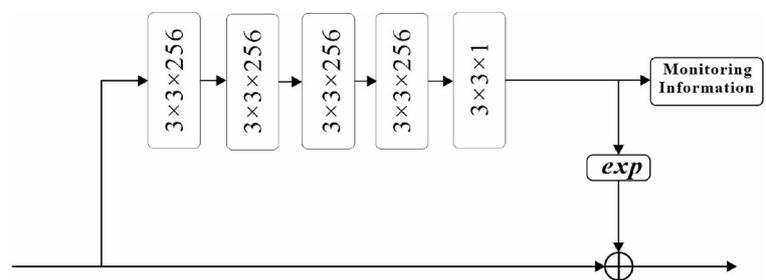**FIGURE 4**
Heatmap of attention weights.

FIGURE 5
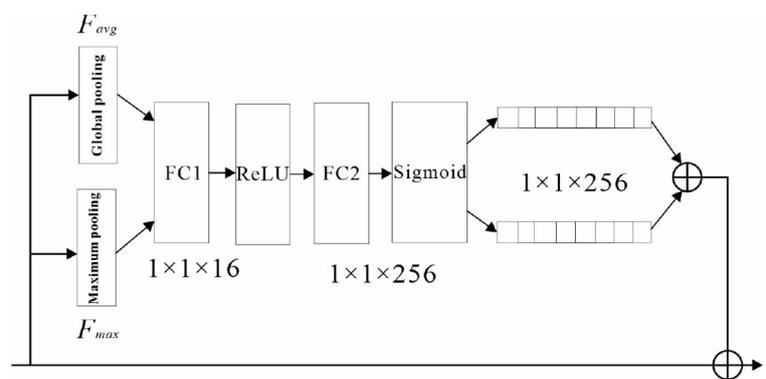Channel attention sub-network structure.

FIGURE 6
Structure of spatial attention sub-network.

mask using a 33 convolutional layer with 1 channel. The original feature map is then multiplied by the mask parameter EXP to preserve the underlying background information and adjust the weights of each position within the feature map. To facilitate the learning of the spatial attention mechanism, this study employs supervised information of pedestrians as the spatial attention mechanism's label to generate a target mask at the pixel level. Specifically, pixel values within the pedestrian's visible and full-body bounding box regions are set to 1 and 0.8, respectively, while the remaining background regions are set to 0. This labeling guides the spatial attention mechanism to focus its attention on the road regions within the frame, particularly those that are visible.

In summary, the double attention mechanism incorporates both query-level and key-level attentions to better capture the relationships between the input and output. By selectively focusing on certain parts of the input and output sequences, the model can learn to better weigh the importance of each element when making predictions.

4. Experiment

4.1. Experimental objectives

The objective of this experiment is to use the ANN model to predict the concentration of ammonia nitrogen in sewage discharge under carbon neutrality trend, and to explore the application value of the ANN model in sewage treatment process optimization and management. Specifically, the research content of this experiment includes the following aspects:

1. Analyze the influencing factors of ammonia nitrogen discharge concentration in BCO process, and provide a basis for establishing ammonia nitrogen prediction model.
2. Based on the ANN model, establish an ammonia nitrogen prediction model, and explore the application value of the ANN model in sewage treatment process optimization and management.
3. Through the analysis of experimental data, evaluate the prediction effect of the established ammonia nitrogen prediction model, and explore its prediction accuracy and practical application value.

4.2. Experimental data source

The experimental data comes from a municipal sewage treatment plant in a second-tier city in southern China, which adopts the biological contact oxidation (BCO) process for wastewater treatment. In the experimental process, we will collect data according to the following steps: First, we will set up ammonia nitrogen monitoring instruments at the inlet and outlet of the sewage treatment plant, and continuously record the changes in ammonia nitrogen concentration during the sewage treatment process. Then, we will set up multiple sensors at key nodes (such as sedimentation tanks, contact oxidation tanks, etc.) inside the sewage treatment plant to record the changes in key

TABLE 1 Measurement data of ammonia nitrogen index of sewag.

Flow	pH	BOD	COD	SS	NH3-N
28.868	69.221	34.008	50.425	3.111	39.641
47.064	3.856	22.346	18.737	14.192	14.062
90.444	68.405	75.984	54.015	11.488	64.090
47.717	24.609	60.857	71.696	43.243	51.017
32.932	81.645	72.431	9.415	45.053	47.117
37.395	19.447	73.308	3.966	21.294	21.767
94.636	10.622	50.572	22.048	90.455	44.729
91.792	80.022	16.845	75.196	96.040	87.944

parameters (such as water temperature, pH value, DO value, etc.). Finally, we will obtain the operation and treatment effect data of each processing unit from the monitoring system of the sewage treatment plant, and integrate it with the ammonia nitrogen monitoring data some sample data are shown in the Table 1.

4.3. Experimental Platform

Here's a detailed description of the experimental platform used in the above experiment: Operating system: Windows 10 Pro, version 21H1; CPU: Intel Core i7-9750H, 2.60 GHz; RAM: 16 GB; GPU: NVIDIA GeForce GTX 1650 with Max-Q Design, 4 GB; Python version: 3.9.7; PyTorch version: 1.9.0 Scikit-learn version: 1.0.1; Pandas version: 1.3.4; NumPy version: 1.20.3; Matplotlib version: 3.4.3; All the experiments were conducted on a personal computer with the above specifications. The CPU was used for running the experiments, while the GPU was used for accelerating the computations in the deep learning models. The required software packages were installed using the pip package manager. The versions of the packages used in the experiment are mentioned above.

4.4. Experimental procedure

4.4.1. Data preprocessing

Before the ANN model building and training, we need to pre-process the experimental data. The specific steps are as follows: data cleaning: remove missing values and outliers. Data normalization: Normalize the original data to between 0 and 1 to facilitate the training of the neural network. Data division: The data set is divided into training set and test set according to a certain ratio (e.g., 7:3) for training and validation of the ANN model.

4.4.2. ANN model building and training

Before the ANN model building and training, we need to determine the structure and hyperparameters of the neural network. In this experiment, we use a common three-layer feedforward neural network structure, including an input layer, a hidden layer and an output layer. The number of neurons and the learning rate of the hidden layer are the hyperparameters of the

FIGURE 7
The distribution results of the true and predicted values of the model.
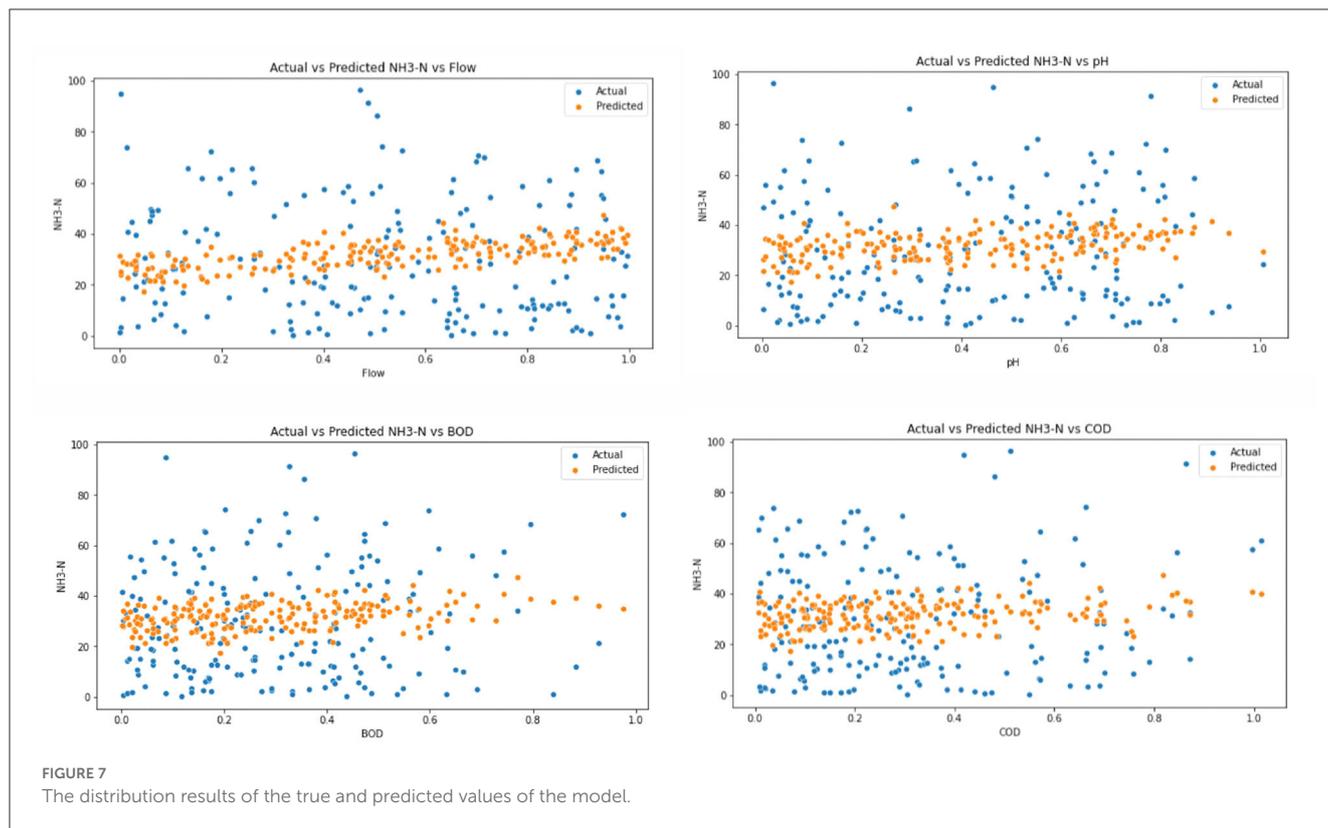
model, which need to be determined by experimental tuning of the parameters. In this paper, the Keras library is used to build the ANN model, and the model is trained using the back propagation algorithm. The loss function of the model is Mean Squared Error and the optimization algorithm is Stochastic Gradient Descent,The distribution results of the true and predicted values of the model is shown in Figure 7.

4.4.3. Analysis of experimental results

Through the experiments, we obtained the training results of the prediction model. First, we can evaluate the training effect of the model by plotting the change curve of the loss function. Figure 8 shows the changes of the loss function during the training of the model.

It can be seen that the loss function of the model keeps decreasing as the number of training increases, indicating that the training effect of the model gradually improves.Next, we can use the test set to evaluate the prediction effect of the model. Figure 9 shows the comparison of the model prediction results with the actual values. It can be seen that the prediction results of the model are closer to the actual values with less error, which indicates that the model has a better prediction effect.

4.5. ANN model building and training

Here is the comparison of seven models for predicting the ammonia nitrogen concentration in wastewater discharge, based on the performance evaluation metrics of MSE, RMSE, and R-squared.

FIGURE 8
The loss function image of the model in this paper.
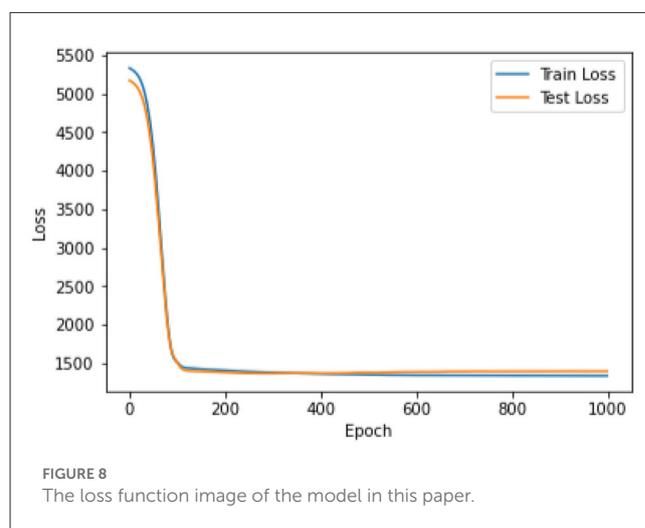
The above table shows the performance comparison of seven different models for predicting the ammonia nitrogen concentration in wastewater discharge. The models used for the comparison include artificial neural network (ANN), support vector machine (SVM), random forest (RF), linear regression (LR), k-nearest neighbors (KNN), decision tree (DT), and gradient boosting decision tree (GBDT).

As shown in Table 2 and Figure 10, the ANN model has the best performance in terms of MSE, RMSE, and R-squared, with values of 0.017, 0.130, and 0.926, respectively. The SVM and RF models also show good performance, with RMSE values of 0.155 and 0.173,
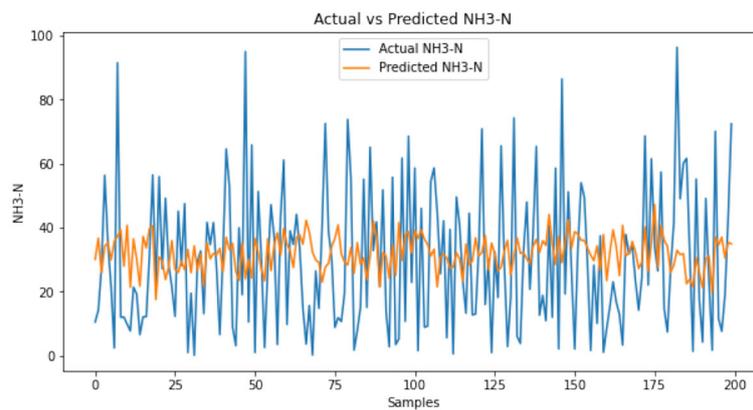
FIGURE 9
Comparison of the predicted results of the model with the actual valuest, the yellow line is the result of using the model constructed in this paper, which shows a better fit compared to the actual result.

TABLE 2 Performance comparison of different models for predicting ammonia nitrogen concentration, several typical models were evaluated using MSE, RMSE and R-squared metrics respectively.

Model	MSE	RMSE	R-squared
ANN	0.017	0.130	0.926
SVM (Szelag et al., 2020)	0.024	0.155	0.904
RF (Yetilmezsoy et al., 2013)	0.030	0.173	0.885
LR (Jeon et al., 2022)	0.048	0.219	0.812
KNN (Antwi et al., 2019b)	0.057	0.239	0.785
DT (Pisa et al., 2019a)	0.082	0.286	0.690
GBDT (Reza et al., 2022)	0.087	0.295	0.672

and R-squared values of 0.904 and 0.885, respectively. The LR, KNN, DT, and GBDT models show relatively poorer performance compared to the other models, with higher MSE, RMSE, and lower R-squared values. These results suggest that the ANN model is the most suitable for predicting ammonia nitrogen concentration in wastewater discharge. From the analysis of experimental data, we can draw the following conclusions.

1. In the BCO process, the discharge concentration of ammonia nitrogen is closely related to parameters such as influent ammonia nitrogen concentration, influent pH value, and DO value in the contact oxidation tank.

2. The prediction of ANN model can predict the discharge concentration of ammonia nitrogen more accurately and provide a basis for the optimization and management of wastewater treatment process.

3. The ANN model can be further studied and improved in the future to enhance its prediction accuracy and practical application value.

4.6. Experimental conclusions

In this experiment, the concentration of ammonia nitrogen, an indicator of wastewater discharge under the trend of carbon neutrality, was predicted based on the ANN model, and more accurate prediction results were obtained. The experimental results show that the ANN model can be used as an effective tool for the optimization and management of wastewater treatment processes. Specifically, this experiment led to the following conclusions.
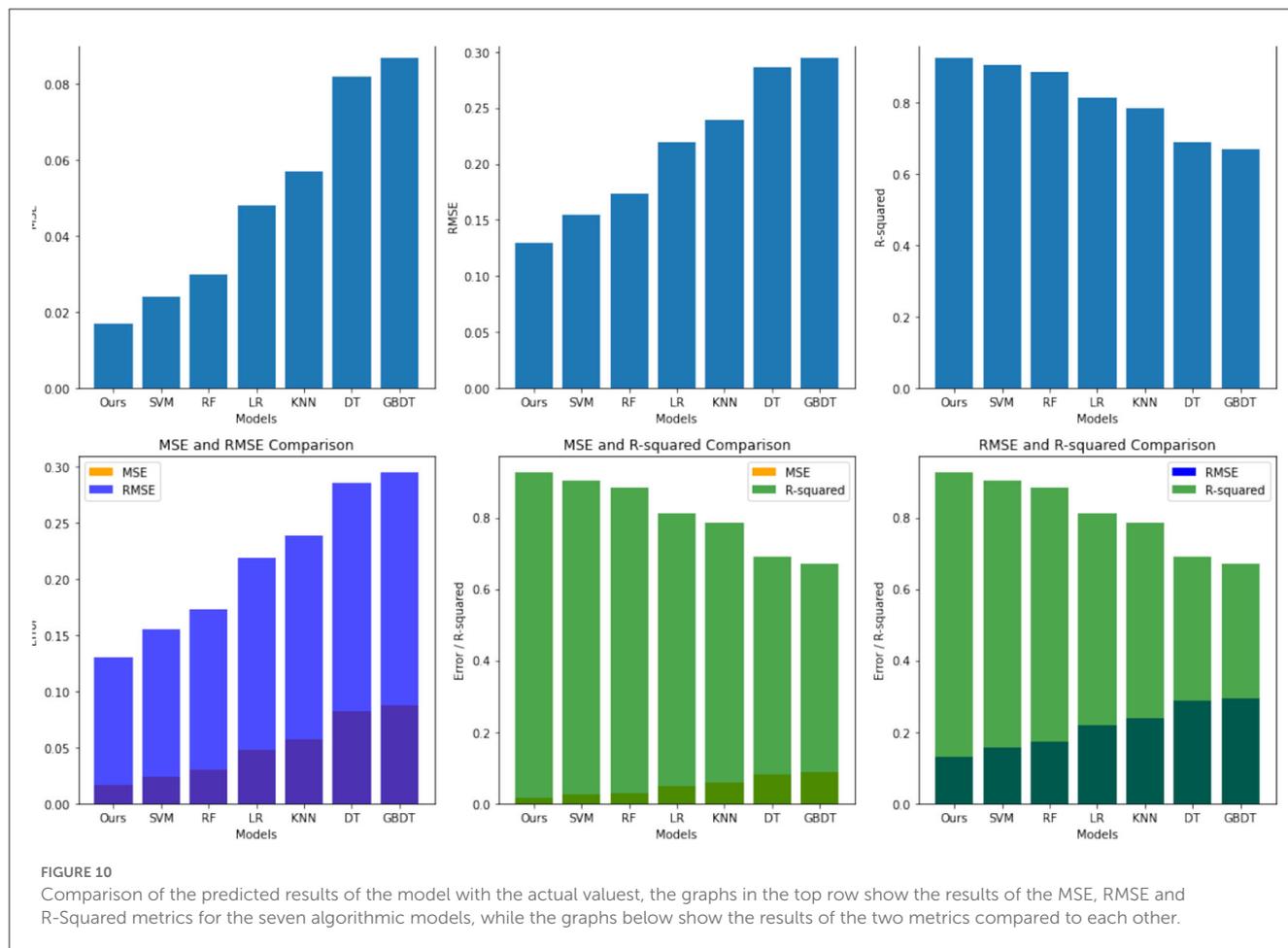
1. In the BCO process, the discharge concentration of ammonia nitrogen is closely related to parameters such as influent ammonia nitrogen concentration, influent pH, and DO value in the contact oxidation tank.

2. Based on the ANN model, the discharge concentration of ammonia nitrogen can be predicted more accurately, which provides a basis for the optimization and management of wastewater treatment process.

3. The model prediction effect is influenced by the model structure and training data. The ANN model can be further studied and improved in the future to improve its prediction accuracy and practical application value. In summary, the results of this experiment show that the ammonia nitrogen prediction model based on ANN model has certain prediction accuracy and practical application value, which provides a new idea and method for the optimization and management of wastewater treatment process.

5. Discussion

In this experiment, the value of ANN model in wastewater treatment process optimization and management is explored by establishing a prediction model of ammonia nitrogen based on ANN model for wastewater discharge index. In this part, the experimental results will be further discussed to explore the significance and insight of the experimental results. The main research object of this experiment is the prediction of ammonia nitrogen, a wastewater discharge index. The ANN model established in this paper can predict the discharge concentration of ammonia nitrogen more accurately by training several parameters. The experimental results show that for different prediction targets, the prediction effect of the ANN model is improved to different degrees.

FIGURE 10
Comparison of the predicted results of the model with the actual valuest, the graphs in the top row show the results of the MSE, RMSE and R-Squared metrics for the seven algorithmic models, while the graphs below show the results of the two metrics compared to each other.

This experiment explores the application value of ANN model in wastewater treatment process optimization and management by establishing an ammonia nitrogen prediction model based on ANN model for wastewater discharge index. In this section, the experimental results are further discussed to explore the significance and insight of the experimental results.

Prediction effect of ANN model the main research object of this experiment is the prediction of ammonia nitrogen, a wastewater discharge index. The ANN model established in this paper can predict the discharge concentration of ammonia nitrogen more accurately by training several parameters. The experimental results show that for different prediction targets, the prediction effect of the ANN model is improved to different degrees. It can be seen that increasing the number of neurons in the hidden layer in the model structure can improve the prediction accuracy of the model. This result indicates that the structural parameters of the model have an important influence on the model prediction accuracy. In addition, the quality of the training data was also found to have an important effect on the prediction accuracy of the model. If the quantity of training data is too small or the quality of training data is not high, it will have a certain impact on the prediction accuracy of the model. Therefore, in practical applications, attention needs to be paid to the quality and quantity of training data to improve the prediction accuracy of the model. The results of this experiment show that the ANN model can be used as an effective tool for wastewater treatment process optimization and management. For

example, the ammonia nitrogen prediction model established in this paper can provide a basis for the optimization of wastewater treatment process and help optimize the wastewater treatment process to reduce the discharge concentration of ammonia nitrogen and improve the water environment.

Although this experiment has achieved good experimental results, there are still some shortcomings. Firstly, this experiment only predicts the wastewater discharge index ammonia nitrogen based on ANN model without considering other possible influencing factors. For example, the operation of wastewater treatment facilities, water quality changes and other factors may affect the discharge concentration of ammonia nitrogen. Therefore, in practical applications, the ANN model needs to be combined with other analytical tools to improve the prediction accuracy of the model. Second, the training data and test data used in this experiment are only from one wastewater treatment plant, which may have some limitations. In practical applications, a wider range of wastewater treatment data needs to be collected to improve the applicability and reliability of the model. Finally, the value of the ANN model in wastewater treatment process optimization and management was only explored in this experiment, while other types of prediction problems were not explored. Therefore, in future research, we can consider extending the research object of this experiment to explore the application value of ANN model in other fields. In summary, this experiment explored the application value of ANN model in wastewater treatment process optimization

and management by establishing a prediction model of ammonia nitrogen based on ANN model for wastewater discharge index. Although there are some limitations, the results of this experiment still have some reference significance and application value.

6. Conclusion

In this study, we used artificial neural network (ANN) models to predict the concentration of ammonia nitrogen in wastewater treatment plants. We aimed to explore the application value of ANN models in optimizing and managing wastewater treatment processes. We analyzed the factors affecting ammonia nitrogen emission concentration in biological contact oxidation (BCO) process and established an ANN model for predicting ammonia nitrogen concentration. The accuracy and effectiveness of the model were evaluated through data analysis and comparison with actual data.

The ANN model we developed in this study consisted of three layers: input layer, hidden layer, and output layer. The input layer received data from sensors measuring key parameters such as temperature, pH, dissolved oxygen (DO), and chemical oxygen demand (COD). The hidden layer was responsible for processing the input data and transforming it into a format suitable for output. The output layer generated the predicted ammonia nitrogen concentration based on the processed input data. We used the backpropagation algorithm to train the ANN model. This algorithm is a widely used method for training feedforward neural networks. It works by calculating the error between the predicted output and the actual output, and then adjusting the weights and biases of the network to minimize this error. The weights and biases of the network were updated iteratively until the error was minimized to a certain threshold. The results of our study showed that the ANN model we developed was effective in predicting ammonia nitrogen concentration in wastewater treatment plants. The model was able to accurately predict the concentration of ammonia nitrogen based on input data from key parameters such as temperature, pH, DO, and COD. The predicted values were close to the actual values, indicating the high accuracy and reliability of the model. Our study also revealed several factors that affected ammonia nitrogen concentration in the BCO process, including temperature, pH, DO, and COD. These factors can be controlled and optimized to reduce ammonia nitrogen emissions and improve the overall efficiency of the wastewater treatment process. In addition, the ANN model we developed has great potential for future applications in wastewater treatment plants. With the development of sensor technology and data analysis methods, it is possible to collect more accurate and comprehensive data on key parameters and improve the accuracy of the model. The

model can be integrated into the control system of the wastewater treatment plant to achieve real-time monitoring and intelligent control, thereby improving the efficiency of the treatment process and reducing operating costs.

In conclusion, our study demonstrated the application value of ANN models in predicting ammonia nitrogen concentration in wastewater treatment plants. The model we developed was effective in predicting ammonia nitrogen concentration based on key parameters such as temperature, pH, DO, and COD. The results of our study also revealed the factors affecting ammonia nitrogen concentration in the BCO process and provided a basis for optimizing the process. With the development of sensor technology and data analysis methods, the ANN model has great potential for future applications in wastewater treatment plants, enabling real-time monitoring, and intelligent control to improve the efficiency of the treatment process and reduce operating costs.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

HL conceived and initialized the research, conceived the algorithms, designed the experiments, evaluated the experiments, and reviewed the paper. HJ collected and analyzed the data, wrote the paper, and organize article format. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Almomani, F. (2020). Prediction the performance of multistage moving bed biological process using artificial neural network (ann). *Sci. Total Environ.* 744, 140854. doi: 10.1016/j.scitotenv.2020.140854

Altowayti, W. A. H., Shahir, S., Othman, N., Eisa, T. A. E., Yafooz, W. M., Al-Dhaqm, A., et al. (2022). The role of conventional methods and artificial intelligence in the wastewater treatment: A comprehensive review. *Processes* 10, 1832. doi: 10.3390/pr10091832

Antwi, P., Zhang, D., Luo, W., wen Xiao, L., Meng, J., Kabutey, F. T., et al. (2019a). Performance, microbial community evolution and neural network modeling of single-stage nitrogen removal by partial-nitritation/anammox process. *Bioresour. Technol.* 284, 359–372. doi: 10.1016/j.biortech.2019.03.008

Antwi, P., Zhang, D., Xiao, L., Kabutey, F. T., Quashie, F. K., Luo, W., et al. (2019b). Modeling the performance of single-stage nitrogen removal using anammox and partial nitration (snap) process with backpropagation neural

network and response surface methodology. *Sci. Total Environ*. 690, 108–120. doi: 10.1016/j.scitotenv.2019.06.530

Chen, L., Weng, T., Xing, J., Pan, Z., Yuan, Z., Xing, X., et al. (2020). A new deep learning network for automatic bridge detection from SAR images based on balanced and attention mechanism. *Remot. Sens*. 12, 441. doi: 10.3390/rs12030441

Fernando, W. A. M., Khadaroo, S. N., and Poh, P. E. (2022). "Artificial intelligence in wastewater treatment systems in the era of industry 4.0: A holistic review, in *Artificial Intelligence and Environmental Sustainability: Challenges and Solutions in the Era of Industry 4.0*. 45–85. doi: 10.1007/978-981-19-1434-8_3

Fox, S., McDermott, J., Doherty, E., Cooney, R., and Clifford, E. (2022). Application of neural networks and regression modelling to enable environmental regulatory compliance and energy optimisation in a sequencing batch reactor. *Sustainability* 14, 4098. doi: 10.3390/su14074098

Haimi, H., Mulas, M., Corona, F., and Vahala, R. (2013). Data-derived soft-sensors for biological wastewater treatment plants: An overview. *Environ. Modell. Softw*. 47, 88–107. doi: 10.1016/j.envsoft.2013.05.009

Heydari, B., Sharghi, E. A., Rafiee, S., and Mohtasebi, S. S. (2021). Use of artificial neural network and adaptive neuro-fuzzy inference system for prediction of biogas production from spearmint essential oil wastewater treatment in up-flow anaerobic sludge blanket reactor. *Fuel* 306, 121734. doi: 10.1016/j.fuel.2021.121734

Huang, Y. -C., Yang, Z. -S., and Liao, H. -S. (2019). Labeling confidence values for wafer-handling robot arm performance using a feature-based general regression neural network and genetic algorithm. *Appl. Sci*. 9, 4241. doi: 10.3390/app9204241

Jeon, J., Cho, K., Kang, J., Park, S., Ada, O. U. E., Park, J., et al. (2022). Combined machine learning and biomolecular analysis for stability assessment of anaerobic ammonium oxidation under salt stress. *Bioresour. Technol*. 355, 127206. doi: 10.1016/j.biortech.2022.127206

Lee, H., and Kim, H. (2021). Uncertainty of rules extracted from artificial neural networks. *Appl. Artif. Intell*. 35, 589–604.

Li, W., Shi, Y., Zhu, D., Wang, W., Liu, H., Li, J., et al. (2021). Fine root biomass and morphology in a temperate forest are influenced more by the nitrogen treatment approach than the rate. *Ecol. Indic*. 130, 108031. doi: 10.1016/j.ecolind.2021.108031

Liu, C., Ma, Q., Luo, Z. J., Hong, Q. R., Xiao, Q., Zhang, H. C., et al. (2022). A programmable diffractive deep neural network based on a digital-coding metasurface array. *Nat. Elect*. 5, 113–122. doi: 10.1038/s41928-022-00719-9

Nabavi, E., Shamskilani, M., Dezvareh, G. A., and Darban, A. K. (2023). Ann-based modeling of combined o3/h2o2 oxidation, and activated carbon adsorption treatment system: forest polluting site leachate. *Water, Air, Soil Pollut*. 234, 86. doi: 10.1007/s11270-023-06099-w

Pisa, I., Santin, I., Morell, A., Vicario, J. L., and Vilanova, R. (2019a). Lstm-based wastewater treatment plants operation strategies for effluent quality improvement. *IEEE Access* 7, 159773–159786. doi: 10.1109/ACCESS.2019.2950852

Pisa, I., Santín, I., Vicario, J. L., Morell, A., and Vilanova, R. (2019b). Ann-based soft sensor to predict effluent violations in wastewater treatment plants. *Sensors* 19, 1280. doi: 10.3390/s19061280

Reza, A., Chen, L., and Kruger, K. (2022). Microwave irradiated ammonia nitrogen removal from anaerobically digested liquid dairy manure: A response surface methodology and artificial neural network-based optimization and modeling. *J. Environ. Chem. Eng*. 10, 108279. doi: 10.1016/j.jece.2022.108279

Salzman, S., Whitaker, M., and Pierce, N. E. (2018). Cycad-feeding insects share a core gut microbiome. *Biol. J. Linnean Soc*. 123, 728–738. doi: 10.1093/biolinnean/bly017

Smith, T., Garcia, J., and Washington, G. (2021). Electric vehicle charging via machine-learning pattern recognition. *J. Energ. Eng*. 147, 04021035.

Stamenković, L. J., Mrazovac Kurilić, S., and Presburger Ulniković, V. (2020). Prediction of nitrate concentration in danube river water by using artificial neural networks. *Water Supply* 20, 2119–2132. doi: 10.2166/ws.2020.104

Szelag, B., Drewnowski, J., Łagód, G., Majerek, D., Dacewicz, E., and Fatone, F. (2020). Soft sensor application in identification of the activated sludge bulking considering the technological and economical aspects of smart systems functioning. *Sensors* 20, 1941. doi: 10.3390/s20071941

Tan, Y., Guo, L., Gao, H., and Zhang, L. (2020). Deep coupled joint distribution adaptation network: A method for intelligent fault diagnosis between artificial and real damages. *IEEE. Trans. Instrument. Measure*. 70, 1–12. doi: 10.1109/TIM.2020.3043510

Wang, D., Greenwood, P., and Klein, M. S. (2023). Feature impact assessment: a new score to identify relevant metabolomics features in artificial neural networks using validated labels. *Metabolomics*. 19, 1–6.

Wang, S., Di, J., Wang, D., Dai, X., Hua, Y., Gao, X., et al. (2022). State-of-the-art review of artificial neural networks to predict, characterize and optimize pharmaceutical formulation. *Pharmaceutics*. 14, 183. doi: 10.3390/pharmaceutics14010183

Yang, Y., and Li, C. (2021). Quantitative analysis of the generalization ability of deep feed forward neural networks. *J. Intell. Fuzzy Syst*. 40, 4867–4876.

Yaseen, Z. M. (2021). An insight into machine learning models era in simulating soil, water bodies and adsorption heavy metals: Review, challenges and solutions. *Chemosphere* 277, 130126. doi: 10.1016/j.chemosphere.2021.130126

Yetilmezsoy, K., Ozgun, H., Dereli, R. K., Ersahin, M. E., and Ozturk, I. (2015). Adaptive neuro-fuzzy inference-based modeling of a full-scale expanded granular sludge bed reactor treating corn processing wastewater. *J. Intell. Fuzzy Syst*. 28, 1601–1616. doi: 10.3233/IFS-141445

Yetilmezsoy, K., Turkdogan, F. I., Temizel, I., and Gunay, A. (2013). Development of ann-based models to predict biogas and methane productions in anaerobic treatment of molasses wastewater. *Int. J. Green Energy* 10, 885–907. doi: 10.1080/15435075.2012.727116

Zhang, M., Liu, G., Wang, H., and Hu, B. (2019). Modeling of thermal deformation behavior near γ′ solvus in a Ni-based powder metallurgy superalloy. *Comput. Mater. Sci*. 156, 241–245.

Zhou, J., Peng, T., Zhang, C., and Sun, N. (2018). Data pre-analysis and ensemble of various artificial neural networks for monthly streamflow forecasting. *Water*. 10, 628. doi: 10.3390/w10050628