


A neural network approach for short-term water demand forecasting based on a sparse autoencoder

Haidong Huang, Zhenliang Lin, Shitong Liu and Zhixiong Zhang *

College of Civil Engineering and Architecture, Beibu Gulf University, Qinzhou, Guangxi 535011, China

*Corresponding author. E-mail: bbg_u_1234@163.com

 ZZ, 0000-0002-9327-2608

ABSTRACT

This paper presents a backpropagation neural network (BPNN) approach based on the sparse autoencoder (SAE) for short-term water demand forecasting. In this method, the SAE is used as a feature learning method to extract useful information from hourly water demand data in an unsupervised manner. After that, the extracted information is employed to optimize the initial weights and thresholds of the BPNN. In addition, to enhance the effectiveness of the proposed method, data reconstruction is implemented to create suitable samples for the BPNN, and the early stopping method is employed to overcome the BPNN overfitting problem. Data collected from a real-world water distribution system are used to verify the effectiveness of the proposed method, and a comparison with the BPNN and other BPNN-based methods which integrate the BPNN with particle swarm optimization (PSO) and the mind evolutionary algorithm (MEA), respectively, is conducted. The results show that the proposed method can achieve fairly accurate and stable forecasts with a 2.31% mean absolute percentage error (MAPE) and 320 m³/h root mean squared error (RMSE). Compared with the BPNN, PSO-BPNN and MEA-BPNN models, the proposed method gains MAPE improvements of 5.80, 3.33 and 3.89%, respectively. In terms of the RMSE, promising improvements (i.e., 5.27, 2.73 and 3.33%, respectively) can be obtained.

Key words: backpropagation neural network, Guigang region, sparse autoencoder, water demand forecasting, water distribution system

HIGHLIGHTS

- To enhance the performance of the BPNN, the SAE is introduced to extract useful features in an unsupervised manner.
- An effective framework which integrates the BPNN with the SAE and early stopping technique is proposed for water demand forecasting.
- The proposed method is verified by comparing with the BPNN and similar methods which integrate the BPNN with PSO and the MEA, respectively.

1. INTRODUCTION

Water demand forecasting is the basis of smart scheduling for water distribution systems. Since prediction accuracy can directly affect the reliability and practicability of management decisions, reliable and accurate forecasts are of significance for effective water management. According to different forecast horizons and forecast frequencies, water demand forecasting can be divided into long-term, medium-term and short-term forecasting (Pacchin *et al.* 2019). Short-term demand prediction generally forecasts water demand over limited time horizons (e.g., 1 month or 1 day) with a time step ranging from daily to sub-hourly (e.g., 15 or 5 min) (Bárdossy *et al.* 2009; Tabesh & Dini 2009). In this paper, we focus on hourly water demand forecasting.

Over the past decades, a wide variety of methods have been proposed for water demand forecasting based on different principles (Donkor *et al.* 2014). Artificial neural networks (ANNs) have always been a research hotspot in this field due to their ability to handle nonlinear data (Ghalekhondabi *et al.* 2017). Among these ANNs, the most commonly used type is the backpropagation neural network (BPNN), where a backpropagation algorithm is used for training (Bougadis *et al.* 2005). Some previous studies have shown that it is possible to yield fairly accurate forecasts using BPNNs to predict short-term water demand (Adamowski & Karapataki 2010; Herrera *et al.* 2010). Although BPNNs perform well in some cases, they easily

This is an Open Access article distributed under the terms of the Creative Commons Attribution Licence (CC BY 4.0), which permits copying, adaptation and redistribution, provided the original work is properly cited (<http://creativecommons.org/licenses/by/4.0/>).

fall into the local optimal solutions due to the randomness of the initial weights and thresholds, which results in poor generalization, especially for complex prediction problems. To address this problem, some studies have adopted optimization algorithms (e.g., the genetic algorithm (GA), particle swarm optimization (PSO) or the mind evolutionary algorithm (MEA)) to optimize the initial weights and thresholds of the BPNN. The results show that the prediction performance of water demand forecasting can improve to different degrees compared with BPNNs without optimization (Pulido-Calvo & Gutierrez-Estrada 2009; Huang *et al.* 2022). In other words, this approach is a feasible way to improve the BPNN performance by optimizing its initial weights and thresholds. However, previous studies have usually emphasized supervised learning to optimize the BPNN's initial weights and thresholds, while unsupervised learning is ignored.

In recent years, many deep learning-based methods have been proposed to address complex prediction problems (Alipanahi *et al.* 2015; Lv *et al.* 2015; Hu *et al.* 2019; Cao *et al.* 2021). As a kind of machine learning method, deep learning methods can learn more useful features when the machine learning models are constructed with many hidden layers and massive training data, thereby improving prediction accuracy. Given this advantage, deep learning methods have gained great interest in short-term water demand forecasting (Salloom *et al.* 2021, 2022; Chen *et al.* 2022; Sharma 2022). The above studies suggest that deep learning methods may be a promising alternative for improving the prediction performance of short-term water demand forecasting. However, these deep learning-based methods usually require professional knowledge during their construction. Unfortunately, the required expert knowledge is not always available, which means that these methods are not general enough to tackle prediction tasks. In addition, deep learning-based methods are usually time-consuming in terms of tuning the related parameters and training the networks due to their much more complex structures.

In this context, a simple method that couples a BPNN with a sparse autoencoder (SAE), named the SAE-BPNN model, is proposed for short-term water demand forecasting in this paper. Note that there is no general method/model that can obtain forecasts with high accuracy and good stability in all cases. Therefore, this study does not pursue the best method/model for prediction problems but seeks to predict short-term water demand in a simple and effective way without losing prediction performance. Thus, the main purposes of this study include the following aspects: (1) investigating the potential of the proposed method in short-term water demand forecasting and (2) examining whether the proposed method can achieve some improvements in prediction performance compared with some similar methods.

The main contributions of this paper are as follows: (1) to enhance the performance of the BPNN, the SAE is introduced to extract useful features in an unsupervised feature manner; (2) an effective framework which integrates the BPNN with the SAE and early stopping technique is proposed for water demand forecasting and (3) the proposed method is verified by comparing with the BPNN and similar methods which integrate the BPNN with PSO and the MEA, respectively.

2. METHODOLOGY

In the proposed method, to enhance the prediction performance, hourly water demand data are first reconstructed to create suitable samples for the BPNN. The samples obtained from reconstruction can effectively describe the variation patterns of hourly water demand data, which may be conducive to improving the prediction performance. Next, the SAE is used to extract features from samples in an unsupervised manner. After that, the useful features obtained from the SAE are employed to initialize the weights and thresholds to overcome the aforementioned deficiency of BPNNs. Finally, the BPNN with optimized parameters is trained for prediction. During the training process, the early stopping method is employed to avoid the BPNN's overfitting problem. Figure 1 shows the flow chart for water demand prediction using the proposed method.

2.1. Data reconstruction

Input variables are important factors that have a considerable impact on the accuracy of short-term forecasting (Arjmand *et al.* 2020), and the input variables that can effectively describe the data features are beneficial in terms of improving the prediction performance. Therefore, it is necessary to reconstruct the hourly water demand data to achieve input variables that can reflect the relationships among data. In general, hourly water demand data have the characteristics of periodicity and short-term correlation. To represent these characteristics, data reconstruction is implemented by using the framework proposed by Huang *et al.* (2022); the framework is provided in Table 1. After that, the BPNN's inputs and outputs can be obtained.

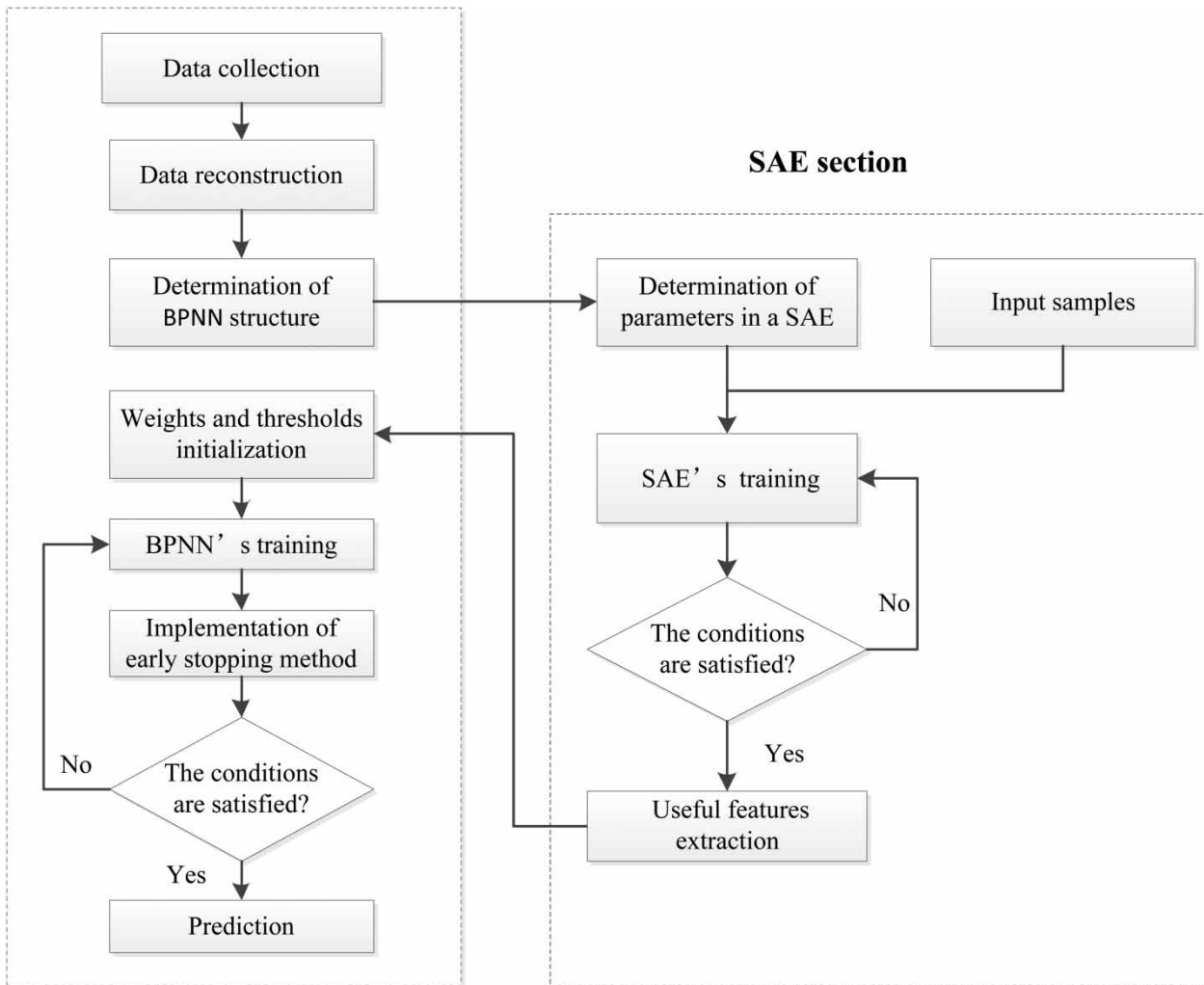


Figure 1 | Flow chart for water demand prediction using the proposed method.

Table 1 | Sample structure obtained from data reconstruction

Component of a sample	Data	Description
Input variables	$T(i - j) - T(i - 1)$ $D_1(i) - D_n(i)$ $W_1(i) - W_m(i)$	$T(i)$ denotes hourly water demand at time i ; $D_n(i)$ denotes the hourly water demand at time i for the previous day n ; $W_m(i)$ is the hourly water demand at time i on the same day for the previous week m .
Output variables	$T(i)$	

2.2. BPNN model

The BPNN is a kind of multilayer feed-forward neural network with single propagation. Due to its ability to handle nonlinear complex problems and simple structures, the BPNN has been widely used in many fields (Bougadis *et al.* 2005; Guo *et al.* 2021). Generally, it consists of one input layer, one or more hidden layers and one output layer, and in this paper, the BPNN-based methods adopt a network structure with one hidden layer. Figure 2 depicts the basic structure of the BP network model with one hidden layer.

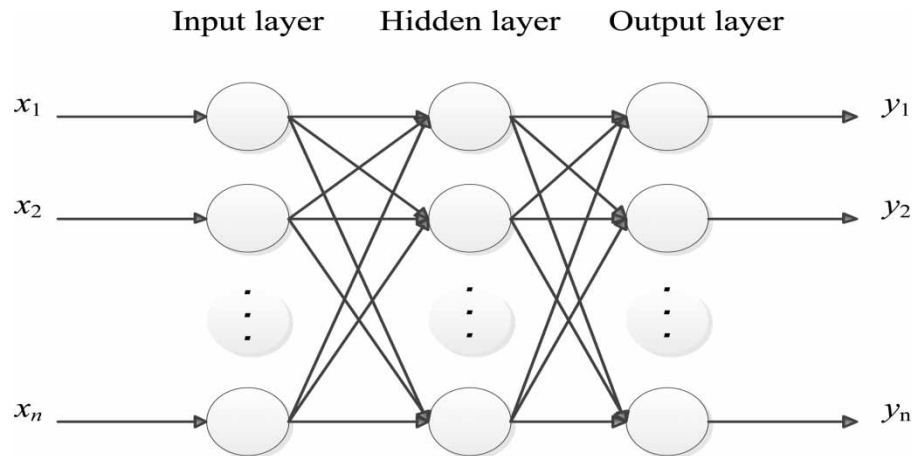


Figure 2 | BPNN structure with one hidden layer, where x_1, x_2, \dots, x_n are the inputs of the BPNN, and y_1, y_2, \dots, y_n are the outputs of the BPNN.

2.3. Sparse AE

An AE is essentially a neural network that can extract feature representations of data in an unsupervised learning manner (Shin *et al.* 2013). The basic AE structure is similar to a three-layer neural network, which consists of an input layer, a hidden layer and an output layer, as shown in Figure 3.

From Figure 3, it can be seen that the input and output layers have the same size. In practice, the goal of AE is to reconstruct an approximation of the input at the output layer. For this purpose, an AE is implemented through two stages: encoding and decoding. The encoding stage is executed in the encoder, where the input vector X is transformed into

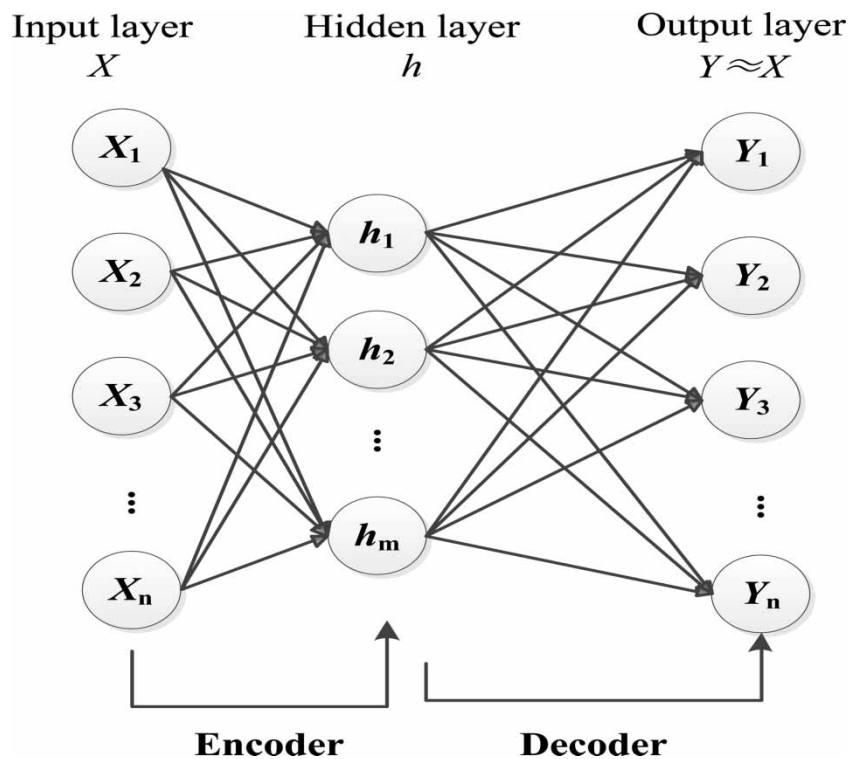


Figure 3 | Standard AE structure.

the intermediate vector \mathbf{h} through a given transfer function. The decoding stage is performed in the decoder, where \mathbf{h} is reconstructed into the output vector \mathbf{Y} . The process of encoding and decoding can be described by Formulas (1) and (2), respectively.

$$\mathbf{h} = f_1(\mathbf{W}_1\mathbf{X} + \mathbf{b}_1) \quad (1)$$

$$\mathbf{Y} = f_2(\mathbf{W}_2\mathbf{h} + \mathbf{b}_2) \quad (2)$$

where \mathbf{W}_1 is a weight matrix for the encoder, \mathbf{W}_2 is a weight matrix for the decoder, \mathbf{b}_1 is a bias vector for the encoder, \mathbf{b}_2 is a bias vector for the decoder and f_1 and f_2 represent the transfer functions for the encoder and the decoder, respectively. Generally, the sigmoid function is chosen as the transfer function in an AE. It can be expressed as follows:

$$f_1(x) = f_2(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

when training an AE, the key to replicating the input at the output is to minimize the reconstruction error between the input vector \mathbf{X} and its reconstruction \mathbf{Y} at the output layer. The reconstruction error can be measured by a cost function, which is defined as follows:

$$J_E(\mathbf{W}, \mathbf{b}) = \frac{1}{n} \sum_{i=1}^n \frac{1}{2} \|\mathbf{X} - \mathbf{Y}\|^2 \quad (4)$$

where n is the dimension of the input vector \mathbf{X} . When the reconstruction error obtained from Formula (4) reaches a specified value or the training termination condition is satisfied, \mathbf{Y} can be obtained.

Similar to a multilayer perceptron, AE can also train the network in a backpropagation fashion. If the reconstruction error is small enough, it can be considered that most of the information in the training sample data is preserved. However, the model will be ineffective if the output is just a simple copy of the input. To learn meaningful features, a sparsity constraint is usually introduced into a traditional AE, and therefore, an SAE can be obtained. Due to this sparsity constraint, an SAE can control the number of ‘active’ neurons in the hidden layer. While training an SAE, the weight matrix \mathbf{W} and the bias vector \mathbf{b} are continuously tuned to reduce the reconstruction error, and finally, comprehensive and useful features can be obtained. Due to the addition of the sparsity constraint, the SAE cost function can be expressed as follows (Sun *et al.* 2016):

$$J_S(\mathbf{W}, \mathbf{b}) = J_E(\mathbf{W}, \mathbf{b}) + \beta \sum_{i=1}^m \left[\rho \log \frac{\rho}{\hat{\rho}_i} + (1 - \rho) \log \frac{1 - \rho}{1 - \hat{\rho}_i} \right] \quad (5)$$

where β is the weight of the sparsity penalty; m is the number of neurons in the hidden layer; ρ is a sparsity parameter that is a constant close to zero and $\hat{\rho}_i$ is the average activation of the i th neuron in the hidden layer.

To date, the SAE has become a popular unsupervised feature learning method because it has a powerful ability to effectively find succinct and high-level representations in complex data (Xu *et al.* 2016; Wang *et al.* 2020).

2.4. Implementation of the proposed method

The implementation procedure for the proposed method includes two stages: unsupervised pretraining and supervised fine-tuning. From this point of view, the proposed method can be considered a semi-supervised learning method, and the proposed method’s specific procedure is described as follows:

(1) Unsupervised pretraining

The main goal of this stage is to provide the initial weights and thresholds for the BPNN. The SAE is first trained using sufficient unlabeled training samples. Suppose a training sample of the BPNN is denoted as $\{x_1, x_2, \dots, x_n, y\}$, where $\{x_1, x_2, \dots, x_n\}$ is the input variable, y is the output variable. Then, $\{x_1, x_2, \dots, x_n\}$ can be used as the unlabeled training sample of the SAE. Once the SAE training process is completed, the useful features can be extracted from the unlabeled samples. After that, the extracted features are used as the initial parameters for the BPNN. Since the initial weights and thresholds

influence the BPNN's prediction performance, it is crucial to train the SAE well to successfully extract meaningful information. In addition, to make use of the information obtained from the SAE, the size of the SAE's hidden layer must be the same as that of the BPNN's hidden layer.

(2) Supervised fine-tuning

This stage aims to find the final parameters for the BPNN. Although the information obtained from the SAE can be used to optimize the BPNN's initial parameters, there is no guarantee that the BPNN will yield good predictions using these parameters. As a consequence, it is still necessary to search for appropriate parameters. For this purpose, the BPNN needs to be trained using sufficient labeled training samples. As part of the proposed method, early stopping, which is an effective method used to avoid the BPNN overfitting problem (Gurbuz *et al.* 2003; Cheng *et al.* 2016), is adopted during the BPNN's training. When the training termination condition is satisfied, the BPNN will have suitable parameters and can be used for water demand forecasting.

2.5. Forecasting performance evaluation

Appropriate criteria are crucial for rationally and effectively evaluating the performance of a forecasting method. It is therefore critically important to select the appropriate types and number of criteria for evaluation. Although there are many criteria used to evaluate whether a method has a good forecasting ability (Donkor *et al.* 2014), no generally accepted criteria exist. This paper used the following measures commonly adopted by other studies (Alizadeh *et al.* 2017; Zhang *et al.* 2022): mean absolute percentage error (MAPE), absolute percent error (APE), root mean squared error (RMSE), coefficient of determination (R^2) and computational load which consists of training and prediction time. The MAPE and APE are used to measure the forecasting accuracy, while the RMSE is adopted to evaluate the prediction stability by measuring the variance of errors between the observed values and the predicted values. As for R^2 , it is used to measure the fitting degree among observed values and forecasts. In general, the lower the MAPE, APE and RMSE are, the better the performances of forecasting methods become. The mathematical expressions for these indicators can be denoted as follows (Adamowski & Karapataki 2010; Donkor *et al.* 2014; Valipour 2017):

$$\text{APE} = \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (6)$$

$$\text{MAPE} = \frac{1}{n} \times \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \times 100\% \quad (7)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \times \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (8)$$

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (9)$$

where y_i and \hat{y}_i represent the observed values and the predicted values, respectively, n is the number of observed values, and \bar{y} is the mean value of the observed values.

3. CASE STUDY

3.1. Data description

The data used in this study were obtained from a real-world water distribution system located in Guigang region, southern China (Huang *et al.* 2022), and the dataset contains hourly water demand data over 175 days. Therefore, a total of 4,200 observed values are included in the dataset. Based on the framework shown in Table 1, the hourly water demand data were reconstructed to generate proper samples for the BPNN. To obtain more suitable samples, correlation analysis was conducted to determine appropriate values for parameters j , n and m , which are shown in Table 1. The values of j , n and m were finally set to 4, 4 and 2, respectively. After that, a sample set containing 3,864 samples was created for the BPNN. To evaluate

the performance of the forecasting methods, the sample set was divided into a training set (including 3,092 samples, approximately 80% of the total samples) and a testing set (772 samples, approximately 20% of the total samples).

3.2. Forecasting methods used for comparisons

In this paper, two scenarios were created to evaluate the proposed method. The purpose of Scenario 1 is to verify the effectiveness of the SAE and the superiority of the proposed method over other methods. In Scenario 1, the proposed method was compared with a traditional BPNN and two BPNN-based models, where the initial weights and thresholds of the BPNN were optimized by PSO and the MEA, respectively. Both PSO and the MEA are intelligent optimization algorithms that are usually used to search for optimal values when optimizing BPNNs (Wang *et al.* 2015; Chang *et al.* 2020). For convenience, the BPNN models that are based on PSO and the MEA are denoted as the PSO-BPNN model and MEA-BPNN model, respectively. In Scenario 2, to investigate the effectiveness of the samples obtained from data reconstruction, the proposed method was compared with the SAE-BPNN model without data reconstruction. The parameter settings for forecasting methods are discussed in detail in the following sections.

3.3. Scenario 1 parameter settings

A total of four methods are considered in Scenario 1. Note that the forecasting methods in this study were performed using the MATLAB software. Furthermore, all of the methods in Scenario 1 adopted the same samples for prediction. The process used to determine the related parameters is presented as follows:

(1) BPNN model

To utilize the BPNN model, its structure should be specified in advance. The basic parameters of the BPNN structure include the number of input nodes, hidden nodes and output nodes. The numbers of input nodes and output nodes are equal to the numbers of inputs and outputs, respectively. According to the 'Data description' subsection, the number of input nodes can be set to 10, while the number of output nodes can be set to 1, and for the number of hidden nodes, the optimal value can be obtained by using a grid search method. A summary of the parameter settings for the BPNN is listed in Table 2.

(2) Other methods

To provide a fair comparison, the other BPNN-based methods adopted the same parameters described in the BPNN section. For the PSO-BPNN model, the PSO parameters include two acceleration factors, iteration times, velocity ranges, position ranges, and population sizes. The first four parameters were determined based on published values (Wang *et al.* 2015) and several trials, which are described in this paper. Moreover, a grid search method was adopted to obtain the optimal population size value according to the minimum MAPE on the test data. The PSO parameters are shown in Table 3.

In the MEA-BPNN model, the MEA parameters include the number of superior subgroups, number of temporary subgroups, number of iterations and population size. The first three parameters were specified based on published values (Huang *et al.* 2022) and a trial-and-error procedure. The optimal population size can also be obtained by using the same method as that adopted for PSO. The MEA parameters are listed in Table 4.

Table 2 | BPNN parameter settings

Number of input nodes	Number of hidden nodes	Number of output nodes	Search range for optimal number of hidden nodes	Other parameters
10	5	1	[5, 20]	Default

Table 3 | PSO parameter settings

Acceleration factors c_1, c_2	Number of iterations	Velocity range	Position range	Population size	Search range for optimal population size
1.49445	10	[-1, 1]	[-5, 5]	70	[20, 100]

Table 4 | MAE parameter settings

Population size	Number of iterations	Number of superior subgroups and temporary subgroups	Search range for optimal population size
200	20	(5, 5)	[20, 200]

With regard to the proposed method, the SAE should have the same number of hidden nodes as the BPNN to ensure that the proposed method operates properly. In the MATLAB environment, three hyperparameters, termed the coefficient of the L2 weight regularization, the coefficient of sparsity regularization and the sparsity proportion, are usually used to control the SAE sparsity. To achieve better performance, a grid search method was conducted to obtain the optimal values for the aforementioned parameters based on the minimal MAPE on the test data. The SAE parameters are displayed in Table 5.

3.4. Scenario 2 parameter settings

In Scenario 2, the proposed method and the SAE–BPNN model without data reconstruction are considered, and the parameters of the proposed method are the same as those in Scenario 1. Due to the difference in samples, the parameters of the SAE–BPNN model without data reconstruction need to be specified separately. The samples were constructed by using a rolling method as shown in Figure 4.

In Figure 4, $x(i)$ is the i th data point in the water demand time series.

To provide a fair comparison, the inputs in a sample for the SAE–BPNN model without data reconstruction consist of 10 sequential data. In other words, the sample size of the SAE–BPNN model without data reconstruction is the same as that of the proposed method. According to the framework shown in Figure 4, a sample dataset including a total of 4,190 samples can be obtained. Based on similar procedures to Scenario 1, the parameters of the SAE–BPNN model without data reconstruction can be determined. Detailed information about the parameter settings is illustrated in Table 6.

Table 5 | SAE parameter settings

Parameter	Value	Parameter	Value
Coefficient of L2 weight regularization	0.1	Search range for optimal coefficient of L2 weight regularization	[0.05, 1]
Coefficient of sparsity regularization	1	Search range for optimal coefficient of sparsity regularization	[1, 5]
Sparsity proportion	0.2	Search range for sparsity proportion	[0.05, 1]
Transfer function for the decoder	purelin	Other parameters	Default

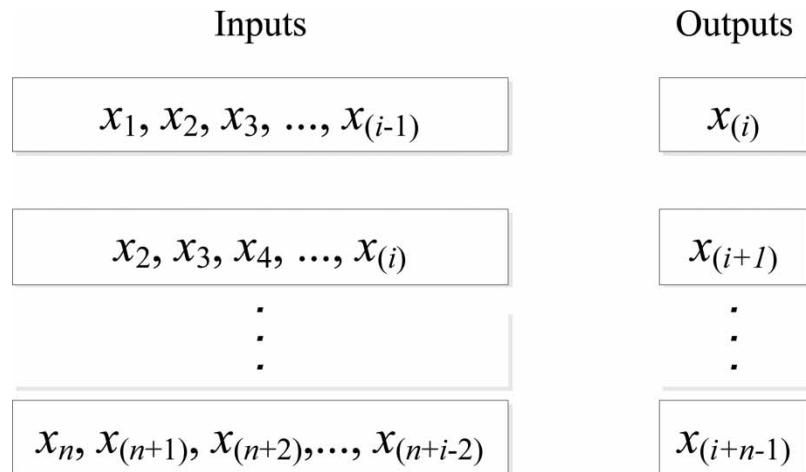
**Figure 4** | Sample construction of the SAE–BPNN model without data reconstruction.

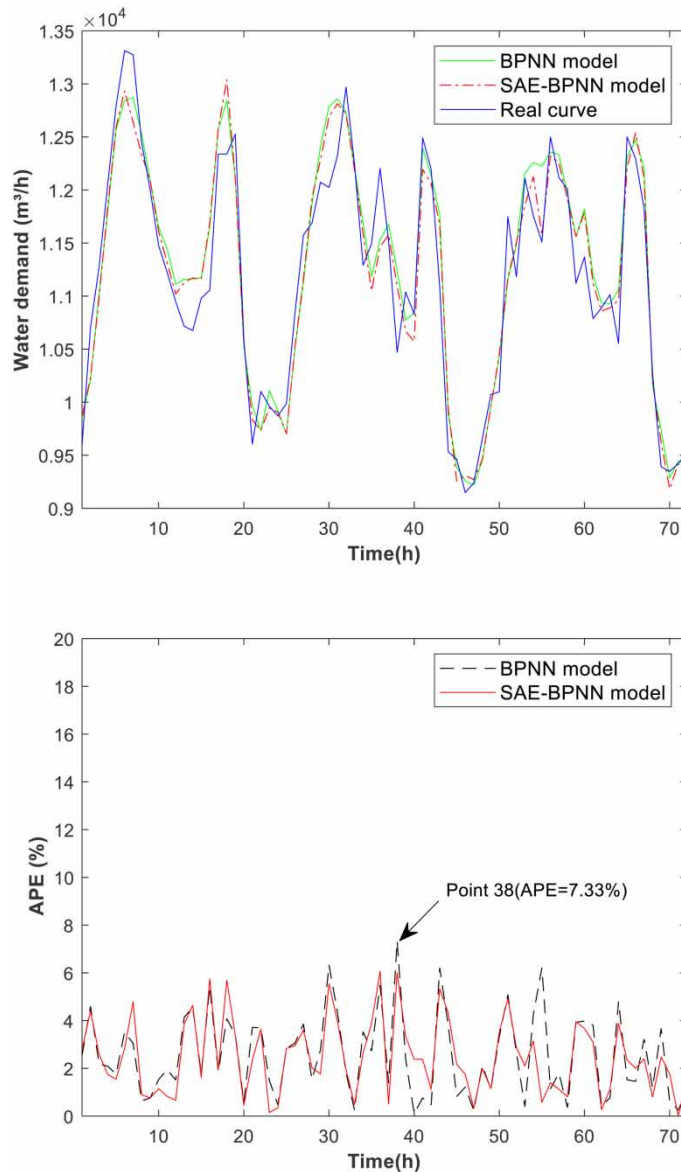
Table 6 | Parameters of the SAE-BPNN model without data reconstruction

SAE section	BPNN section
Coefficient of L2 weight regularization: 0.55	Number of input nodes: 10
Coefficient of sparsity regularization: 5	Number of hidden nodes: 18
Sparsity proportion: 0.6	Number of output nodes: 1
Transfer function for the decoder: logsig	Other parameters: default
Other parameters: default	

3.5. Results and discussion

In this study, a laptop was used to implement the MATLAB programs for all the prediction methods. The main parameters of the laptop are as follows: Core™ i7-7500 μ 2.70 GHz, RAM 8.00.

Figures 5–7 illustrate a comparison between the BPNN-based models in terms of the prediction accuracy in Scenario 1. From Figure 5, it can be seen that the proposed method has a lower APE value than the BPNN in most cases, which indicates

**Figure 5** | Prediction results and APE values for the BPNN model and the proposed method.

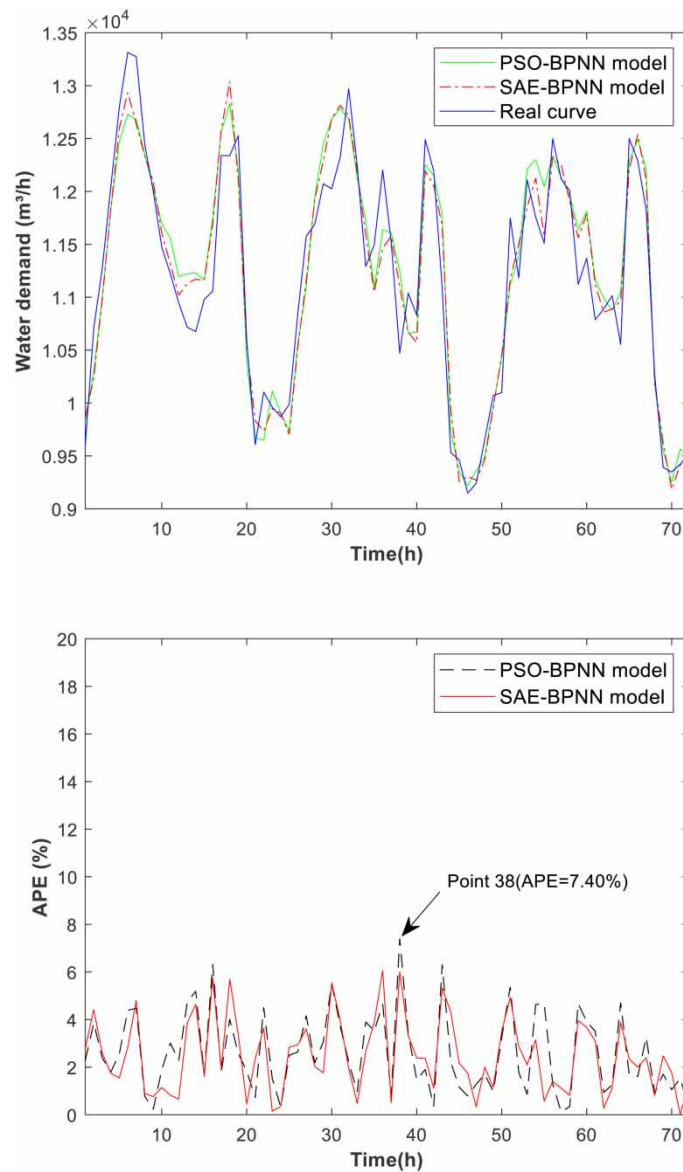


Figure 6 | Prediction results and APE values for the PSO-BPNN model and the proposed method.

that more accurate forecasts can be obtained by using the proposed method. Furthermore, the BPNN produced some points with high APE values (e.g., the APE value of 7.33% at point 38), while no such extreme values were found when using the proposed method. This result reveals that the SAE may mitigate the error at some extreme points. Figures 6 and 7 illustrate similar information about the prediction accuracy. Overall, the proposed method can obtain more accurate forecasts than the PSO-BPNN and MEA-BPNN models. Interestingly, the APE value at point 38 is still high when using the PSO-BPNN and MEA-BPNN models; this result suggests that the proposed method is superior to the PSO-BPNN and MEA-BPNN models.

Table 7 shows the comprehensive comparison results of the forecasting performance. The improvements gained from the comparison between the proposed method and three other BPNN-based methods are provided in Table 8.

From Table 7, it is clear that the proposed method has the lowest values in terms of both the MAPE and RMSE. The proposed method achieves a 2.31% MAPE and 320 m³/h RMSE. Compared with the BPNN, the proposed method gained MAPE improvements of 5.80%. In terms of the RMSE, promising improvements (i.e., approximately 5.27%) can also be obtained. These observations demonstrate that the proposed method is effective in improving the performance of the BPNN. One reasonable explanation is that optimizing the BPNN through the SAE can effectively overcome the defects caused by the

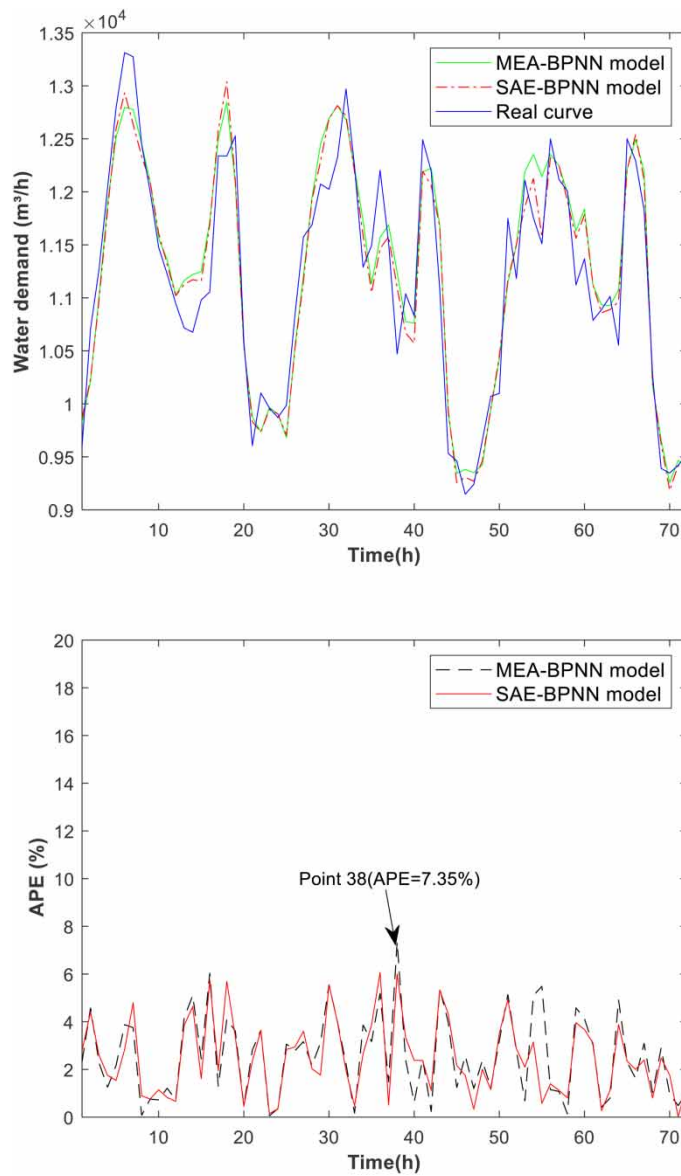


Figure 7 | Prediction results and APE values for the MEA-BPNN model and the proposed method.

Table 7 | Comprehensive comparison of the forecasting performance

Method	MAPE (%)	RMSE (m ³ /h)	R ²	Computational load (s)	APE		
					Number (APE > 5%)	Maximum value (%)	Standard deviation
BPNN	2.45	338	0.93	2.17	79	10.27	1.81
PSO-BPNN	2.39	329	0.91	73.13	73	10.13	1.76
MEA-BPNN	2.40	331	0.92	3.98	71	9.43	1.77
SAE-BPNN	2.31	320	0.91	3.57	61	8.95	1.71

randomness of initial weights and thresholds. Similar improvements can also be found in terms of the APE. Compared with the BPNN, the number of APE (>5%) declined significantly from 79 to 61 when using the proposed method. Furthermore, the maximum value of APE decreased by 1.32% from 10.27 to 8.95%, and the standard deviation of APE gained

Table 8 | Improvements gains in different comparisons

Indicator	BPNN	PSO-BPNN	MEA-BPNN
MAPE improvements (%)	5.80	3.33	3.89
RMSE improvements (%)	5.27	2.73	3.33

improvements of 5.52%. One main reason for achieving these improvements is that the SAE can improve the signal-to-noise ratio to a certain extent. In other words, the SAE can be used for data denoising, which is conducive to removing the impact of noise on modeling. As shown in Table 7, the RMSE value obtained from the proposed method is fairly low, there is still room to further reduce uncertainties. In future work, it is suggested that weather variables should be taken into consideration during the procedure of input variables selection. However, it is expected to be a challenge for water utilities to collect and keep record of weather-related data, which would deteriorate the practicality of the proposed method to some extent.

In Table 7, it can be seen that the proposed method performed better than the PSO-BPNN and MEA-BPNN models in the majority of indicators. Compared with the PSO-BPNN and MEA-BPNN models, the proposed method gained MAPE improvements of 3.33 and 3.89%, respectively. In terms of the RMSE, improvements of 2.73 and 3.33%, respectively, can be achieved. These results suggest that the proposed method outperforms the PSO-BPNN and MEA-BPNN models in terms of both forecasting accuracy and stability. Similar information can also be found in other indicators shown in Table 7. Compared with the BPNN, although the PSO-BPNN and MEA-BPNN models also gained some improvements in APE, these improvements are inferior to that obtained from the proposed method. The underlying reasons for these findings are mainly as follows: First, unlike the SAE, PSO and the MEA have no ability to reduce data noise. This is the key reason why the proposed method is more effective to reduce the number of extreme APE values. Next, the SAE can effectively extract useful input features and use them as the basis for optimizing the BPNN's thresholds and weights. Finally, PSO and the MEA only rely on their global search ability to obtain optimal thresholds and weights without input features extraction.

Unlike the aforementioned indicators, the BPNN has an advantage over the other three methods, which is mainly because the BPNN's structure is simpler than that of the three other methods. While the computational load of the proposed method is slightly higher than that of the BPNN, it has no impact on the practicality of the proposed method because, in our experiments, it took only 3.57 s to predict 772 data points. Furthermore, compared with the PSO-BPNN and MEA-BPNN models, the proposed method took less time to achieve the desired forecasts. Considering that these three methods have the same parameters as the BPNN, it can be inferred that the proposed method is more efficient in terms of searching for the optimal parameters than the PSO-BPNN and MEA-BPNN models.

Similar to the computational load, the BPNN has the best performance in R^2 . This means that the BPNN has a better goodness of fit. However, as shown in Table 7, all R^2 values for the other BPNN-based methods are greater than 0.90, which indicates that the forecasts obtained from these methods also fit observed values well. That is to say, it is acceptable and feasible to use these methods for prediction.

Figure 8 shows the comparison results between the proposed method and the SAE-BPNN model without data reconstruction.

From Figure 8, it is clear that the proposed method performs much better than the SAE-BPNN model without data reconstruction in terms of prediction accuracy. On the whole, the curve of the APE values for the proposed method is relatively stable, while the APE values for the SAE-BPNN model without data reconstruction fluctuate dramatically. This observation at least implies that (1) sample construction plays an important role in the proposed method and (2) as part of the proposed method, the method used to construct samples in this study is effective. More evidence for the above inferences can be found in the statistics in Table 7. In terms of the results in Table 7, all the methods achieves a fairly low MAPE and RMSE. What's more, the BPNN seemed to perform quite well (e.g., the MAPE value is only 2.45%, and the RMSE value is 338 m³/h), and the proposed method still achieved more than 5.0% improvements in terms of both MAPE and RMSE. These findings can be considered as more convincing evidence that the SAE is effective in improving the performance of the BPNN.

4. CONCLUSIONS

In this paper, a new method that integrates the SAE into the BPNN is proposed for short-term water demand forecasting. In this method, the SAE module conducts feature extraction in an unsupervised learning manner, whereas the BPNN

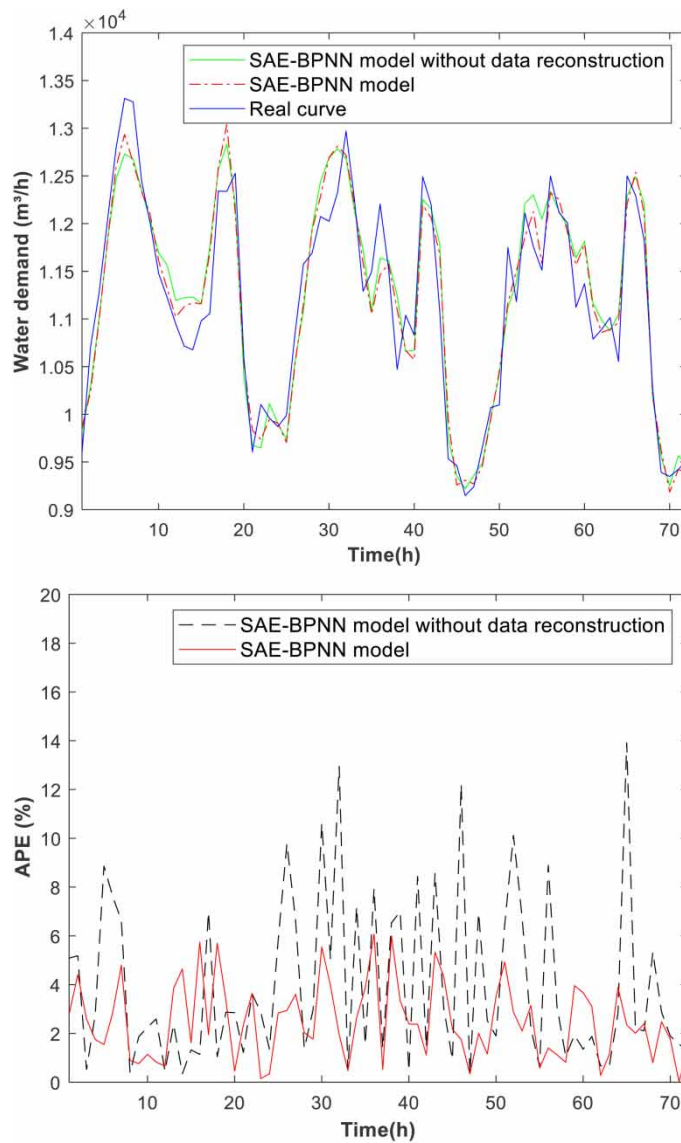


Figure 8 | Comparison between the SAE-BPNN model with and without data reconstruction.

module is used to forecast the water demand. To enhance the proposed method's forecasting performance, data reconstruction is adopted to generate suitable samples for the BPNN module. Hourly water demand data obtained from a real-world water distribution system were used to verify the effectiveness of the proposed method, and comparisons with other similar methods were also considered. The results show that the proposed method has an advantage over the BPNN, PSO-BPNN and MEA-BPNN models in both prediction accuracy and stability. In addition, the findings also prove that the proposed method demonstrates promise as a powerful tool for short-term water demand forecasting in a simple but effective way.

Extracting useful features by unsupervised learning is an advantage of the proposed method over other methods in this study. However, the proposed method also has some limitations. First, adequate unlabeled samples are required to achieve useful features through unsupervised learning. The forecasting performance of the proposed method may be subject to unlabeled samples. Next, as a kind of BPNN-based method, the proposed method also has requirements for massive training samples. As a consequence, the application of the proposed method may be limited in some scenarios where sufficient samples are not available.

ACKNOWLEDGEMENTS

This work is supported in part by the Middle-Aged and Young Teachers' Basic Ability Promotion Project of Guangxi (Grant No. 2021KY0438) and the Natural Science Foundation of Guangxi Province (Grant No. 2022GXNSFAA035582).

DATA AVAILABILITY STATEMENT

All relevant data are available from an online repository or repositories: <https://kdocs.cn/l/cbiLlwiRFAi1>.

CONFLICT OF INTEREST

The authors declare there is no conflict.

REFERENCES

- Adamowski, J. & Karapataki, C. 2010 Comparison of multivariate regression and artificial neural networks for peak urban water-demand forecasting: evaluation of different ANN learning algorithms. *Journal of Hydrologic Engineering* **15** (10), 729–743.
- Alipanahi, B., Delong, A., Weirauch, M. T. & Frey, B. J. 2015 Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nature Biotechnology* **33**, 831–838.
- Alizadeh, M. J., Shahheydari, H., Kavianpour, M. R., Shamloo, H. & Barati, R. 2017 Prediction of longitudinal dispersion coefficient in natural rivers using a cluster-based Bayesian network. *Environmental Earth Sciences* **76**, 86.
- Arjmand, A., Samizadeh, R. & Saryazdi, M. D. 2020 Meta-learning in multivariate load demand forecasting with exogenous meta-features. *Energy Efficiency* **13** (3), 871–887.
- Bárdossy, G., Halász, G. & Winter, J. 2009 Prognosis of urban water consumption using hybrid fuzzy algorithms. *Journal of Water Supply: Research and Technology-AQUA* **58** (3), 203–211.
- Bougadis, J., Adamowski, K. & Diduch, R. 2005 Short-term municipal water demand forecasting. *Hydrological Processes* **19** (1), 137–148.
- Cao, Y., Zhou, X. & Yan, K. 2021 Deep learning neural network model for tunnel ground surface settlement prediction based on sensor data. *Mathematical Problems in Engineering* **2021** (1), 1–14.
- Chang, Y., Yue, J., Guo, R., Liu, W. & Li, L. 2020 Penetration quality prediction of asymmetrical fillet root welding based on optimized BP neural network. *Journal of Manufacturing Processes* **50**, 247–254.
- Chen, L., Yan, H., Yan, J., Wang, J., Tao, T., Xin, K., Li, S., Pu, Z. & Qiu, J. 2022 Short-term water demand forecast based on automatic feature extraction by one-dimensional convolution. *Journal of Hydrology* **606**, 127440.
- Cheng, J., Wang, X., Si, T., Zhou, F., Wang, Z., Zhou, J. & Cen, K. 2016 Maximum burning rate and fixed carbon burnout efficiency of power coal blends predicted with back-propagation neural network models. *Fuel* **172** (MAY15), 170–177.
- Donkor, E., Mazzuchi, T., Soyer, R. & Roberson, J. 2014 Urban water demand forecasting: review of methods and models. *Journal of Water Resources Planning and Management* **140** (2), 146–159.
- Ghalekhondabi, I., Ardjmand, E., Young, W. & Weckman, G. R. 2017 Water demand forecasting: review of soft computing methods. *Environmental Monitoring and Assessment* **189** (7), 313.
- Guo, R., Hu, W., Song, Q., Ji, S., Qi, W. & Yu, H. 2021 Improving the tensile shear load of Al-Mg-Si alloy FSLW joint by BPNN-GA. *Transactions of the Indian Institute of Metals* **74** (6), 1521–1528.
- Gurbuz, H., Kivrak, E., Soyupak, S. & Yerli, S. V. 2003 Predicting dominant phytoplankton quantities in a reservoir by using neural networks. *Hydrobiologia* **504** (1–3), 133–141.
- Herrera, M., Torgo, L., Izquierdo, J. & Pérez-García, R. 2010 Predictive models for forecasting hourly urban water demand. *Journal of Hydrology* **387** (1–2), 141–150.
- Hu, R., Fang, F., Pain, C. C. & Navon, I. M. 2019 Rapid spatio-temporal flood prediction and uncertainty quantification using a deep learning method. *Journal of Hydrology* **575**, 911–920.
- Huang, H., Zhang, Z., Lin, Z. & Liu, S. 2022 Hourly water demand forecasting using a hybrid model based on mind evolutionary algorithm. *Water Supply* **22** (1), 917–927.
- Lv, Y., Duan, Y., Kang, W. & Li, Z. X. 2015 Traffic flow prediction with big data: a deep learning approach. *IEEE Transactions on Intelligent Transportation Systems* **16** (2), 865–873.
- Pacchin, E., Gagliardi, F., Alvisi, S. & Franchini, M. 2019 A comparison of short-term water demand forecasting models. *Water Resources Management* **33** (4), 1481–1497.
- Pulido-Calvo, I. & Gutierrez-Estrada, J. C. 2009 Improved irrigation water demand forecasting using a soft-computing hybrid model. *Biosystems Engineering* **102** (2), 202–218.
- Salloom, T., Kaynak, O. & He, W. 2021 A novel deep neural network architecture for real-time water demand forecasting. *Journal of Hydrology* **599**, 126353.
- Salloom, T., Kaynak, O., Yu, X. & He, W. 2022 Proportional integral derivative booster for neural networks-based time-series prediction: case of water demand prediction. *Engineering Applications of Artificial Intelligence* **108**, 104570.

- Sharma, S. K. 2022 A novel approach on water resource management with multi-criteria optimization and intelligent water demand forecasting in Saudi Arabia. *Environmental Research* **208**, 112578.
- Shin, H. C., Orton, M. R., Collins, D. J., Doran, S. J. & Leach, M. O. 2013 Stacked auto-encoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **35** (8), 1930–1943.
- Sun, W., Shao, S., Zhao, R., Yan, R., Zhang, X. & Chen, X. 2016 A sparse auto-encoder-based deep neural network approach for induction motor faults classification. *Measurement* **89** (ISFA), 171–178.
- Tabesh, M. & Dini, M. 2009 Fuzzy and neuro-fuzzy models for short-term water demand forecasting in Tehran. *Iranian Journal of Science and Technology Transaction B: Engineering* **33** (1), 61–77.
- Valipour, M. 2017 Calibration of mass transfer-based models to predict reference crop evapotranspiration. *Applied Water Science* **7**, 625–635.
- Wang, Q., Li, Y., Diao, M., Gao, W. & Qi, Z. 2015 Performance enhancement of INS/CNS integration navigation system based on particle swarm optimization back propagation neural network. *Ocean Engineering* **108**, 33–45.
- Wang, X., Guo, G., Liu, S., Wu, Y. & Smith, K. 2020 Burst detection in district metering areas using deep learning method. *Journal of Water Resources Planning and Management* **146** (6), 04020031.
- Xu, J., Xiang, L., Liu, Q., Gilmore, H., Wu, J., Tang, J. & Madabhushi, A. 2016 Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images. *IEEE Transactions on Medical Imaging* **35** (1), 119–130.
- Zhang, G., Hu, Y., Yang, D., Ma, L., Zhang, M. & Liu, X. 2022 Short-term bathwater demand forecasting for shared shower rooms in smart campuses using machine learning methods. *Water* **14** (8), 1291.

First received 16 June 2022; accepted in revised form 15 November 2022. Available online 22 November 2022