Combined forecasting model of urban water consumption based on adaptive filtering and BP neural network

Fuchen Ban

Faculty of Municipal and Environmental Engineering,
Shenyang Construction University,
9 Hunnan Avenue, Hunnan District,
Shenyang City, Liaoning Province, China
Email: banfc@163.com

Dan Wu* and Yueming Hei

Faculty of Municipal and Environmental Engineering,
Shenyang Construction University,
9 Hunnan Avenue, Hunnan District,
Shenyang City, Liaoning Province, China
Email: 2457865659@qq.com
Email: 502588525@qq.com
*Corresponding author

Abstract: In order to solve the problem of improving the precision of urban short-term water consumption forecasting, the idea of combination forecasting is put forward. According to the water use data of a city, the time series prediction method and the explanatory prediction method are used to forecast the water use in the short-term. In order to combine the advantages of the two forecasting methods, this paper proposes a combination forecasting method based on weight coefficient optimisation theory. Compared with the single prediction model, the combined forecasting model has higher accuracy and stability.

Keywords: water demand prediction; adaptive filtering method; BP neural network method; combined forecasting model.

Reference to this paper should be made as follows: Ban, F., Wu, D. and Hei, Y. (2018) 'Combined forecasting model of urban water consumption based on adaptive filtering and BP neural network', *Int. J. Social and Humanistic Computing*, Vol. 3, No. 1, pp.34–45.

Biographical notes: Fuchen Ban's research direction is the optimisation of water and drainage system and graduated at the Shenyang Construction University, Faculty of Municipal and Environmental Engineering.

Dan Wu's research direction is water supply and drainage system optimisation and graduated from the Shenyang Construction University, Faculty of Municipal and Environmental Engineering.

Yueming Hei's research direction is water supply and drainage system optimisation and graduated from the Shenyang Construction University, Faculty of Municipal and Environmental Engineering.

1 Introduction

The prediction of urban water consumption is based on the water data information of the past period of time. By establishing a certain quantitative relationship, the water consumption of the future hours, days, months and even years is predicted (Yanchun et al., 2015). According to the prediction time of water consumption, it can be divided into two categories: long-term prediction and short-term prediction. The short-term forecast is based on the water data of a few days or months of the city, and the factors affecting the weather and water use habits to predict the water consumption of the city in the next few days or months. Short-term prediction is the basis of urban operation and scheduling (Jian, 2007). Therefore, the accuracy of water consumption prediction is very important to the normal operation and maintenance of water supply network.

At present, there are many forecasting methods of urban water consumption, which can be generally summarised as qualitative prediction and quantitative prediction. The adaptive filtering method and BP neural network method belong to the two methods of quantitative prediction of water consumption, namely time series method and regression analysis prediction method. Time series method is a method of statistical statistics and curve fitting for the historical data of a variety of factors through mathematical statistics. Because the system is regarded as a black box for prediction, the accuracy of the prediction results will be greatly affected if there are unpredictable unexpected events in the future. The regression analysis prediction method is a method that can reflect a certain relationship between the input and output. In the prediction, the effect of the influence factors will be considered, but the selection of the factors and the quantised value of the influencing factors are both subjective and uncertain, and the accuracy of the prediction results will also be affected. Therefore, proper combination of the two methods will improve their prediction accuracy by combining their advantages and complemented defects. In this paper, the adaptive filtering method and BP neural network algorithm are used to predict the short-term water consumption of the city in a single term. On this basis, a combined prediction model of the two methods is established.

2 The principle of establishing and evaluating the prediction model of water consumption

2.1 Basic principles of adaptive filtering

Adaptive filtering is one of the prediction methods of time series, and it forecasts the future water data by fitting the rule of historical water consumption data. Therefore, adaptive filtering is suitable for forecasting short-term water consumption data with a certain period rule, such as monthly water consumption, daily water consumption and time water consumption. The basic principle of adaptive filtering is to set a set of weights for the predicted time series. Through the continuous training of historical water data, the weight number is constantly adjusted. After several iterations, a group of weights which have little change are finally obtained. On this basis, the water consumption is solved.

The basic formula of the adaptive filtering method is as follows (Mingyang et al., 2012):

$$\hat{y}_{t+1} = \omega_1 y_t + \omega_2 y_{t-1} + \cdots + \omega_N y_{t-N+1} = \sum_{i=1}^{N} \omega_i y_{t-i+1}$$

where

$\hat{y}_{t+1}$ the predictive value of phase $t + 1$

ω_i the number of observational weights in phase $t - i + 1$

y_{t-i+1} the observational values of phase $t - i + 1$

N the number of weights, $i = 1, 2, \ldots N$; $t = N, N + 1, \ldots N$, n is the number of the sequence data.

The weight adjustment formula is as follows:

$$\omega_i' = \omega_i + 2ke_{t+1} y_{t-i+1}$$

where

e_{t+1} phase $t + 1$ prediction error $e_{t+1} = y_{t-i+1} - \hat{y}_{t+1}$

ω'_i the adjusted number of i weights

ω_i the number of i weights before the adjustment

k the size of learning constant determines the speed of weight adjustment.

2.2 Basic principle of BP neural network algorithm

Artificial neural network (ANN) is an algorithm model that imitates the information pattern of the biological neural structure. It is similar to the biological neural network. It is also composed of the input layer, the output layer, the hidden layer and the weight of the connections between the layers. BP neural network is a widely used algorithm in ANN. It is an error feedforward network model based on ANN (Wei et al., 2009).

The basic formula for calculating BP neural network is as follows:

$$net_{pj} = \sum_{i=1}^{N} W_{ij} O_{pi}$$

$$O_{pi} = f\left(net_{pj}\right)$$

$$f(x) = \frac{1}{1 + e^{-x}}$$

where

net_{pj} the sum of the input of the unit J

W_{ij} the weight of neurons and between them

O_{pi} the input of the unit

f function

N processing unit number

p the number of training samples

j the number of neurons – neural network.

The BP neural network will have an error every time when the output value is calculated. The network will adjust the weight according to the error, the formula for the error E_p and the formula for the weight W_{ij} are as follows:

$$E_p = \frac{1}{2}\sum_j \left(d_{pj} - O_{pj}\right)^2$$

$$W_{ij}(t+1) = W_{ij}(t) + \eta\delta_{pj}O_{pj}$$

where

δ_{pj} training error

t the number of learning times

η learning factors.

2.3 Basic principle of combined prediction method

The most common method used in the combinatorial prediction model is the linear grouping method, which is the sum of the individual prediction results multiplied by the corresponding weights.

The specific formula is as follows (Deyang, 2014):

$$y = \sum_{i=1}^{n} w_i y_i$$

where

y the result of combination prediction

w_i the weight coefficient of the first i forecasting results

y_i the prediction results of single i

n the number of single prediction model.

The most important thing in the combination forecasting model is to determine the weight coefficients of each single prediction result (Jing and Jun, 2016).

At present, the commonly used methods are arithmetic average method, reciprocal method and mathematical optimisation method (Pu et al., 2010).

The optimisation model of combination weight coefficient is established as follows:

$$F = \min \sum_{i=1}^{n} \left[\frac{w_i\left(y_i - q\right)}{q}\right]^2$$

where

q the true water consumption at the i moment.

The constraints are as follows:

$$\sum_{i=1}^{n} w_i = 1$$

$$0 < w_i < 1$$

Before using the combined model to predict the water consumption of cities, we first need to predict single models, and then combine the single prediction results linearly. The combination weight coefficient period of the single model is 24 hours, so the combination weight coefficient of the forecast result is 24 for one period, and the optimal combination weight coefficient of a period and 24 groups is finally obtained. On this basis, a single model is combined.

2.4 Comparison method of prediction results

In order to compare the three models for city water consumption forecast accuracy, this paper uses three kinds of commonly used evaluation index model as a standard of comparison, the model evaluation index are the average absolute error and relative error and model validity (Hui et al., 2011; Bin et al., 2002). Assuming that the model is used to forecast the hourly water consumption of N, ft is the first t hours of water consumption forecast, qt is the actual water consumption of the first t hours and the three model evaluation index calculation formula is as follows:

1 The average absolute error

$$MAE = \frac{1}{n}\sum_{t=1}^{n}\left|f_t - q_t\right|$$

2 The average relative error

$$MAPE = \frac{1}{n}\sum_{t=1}^{n}\frac{\left|f_t - q_t\right|}{q_t}$$

3 The accuracy of variance

The prediction accuracy of the model at t hours is calculated as follows:

$$A_t = 1 - \frac{\left|f_t - q_t\right|}{q_t}$$

The accuracy of variance is calculated as follows:

$$\sigma(A_t) = \left(\frac{1}{N}\sum_{t=1}^{N} A_t^2 - \left(\frac{1}{N}\sum_{t=1}^{N} A_t\right)^2\right)^{1/2}$$

4 The validity of model

The validity of the model is calculated as follows:

$$S = E(A_t)\left(1 - \sigma(A_t)\right)$$

The $E(A_t)$, $\sigma(A_t)$ are the mean and mean variance of the predicted accuracy of the model at t hours respectively.

3 Example verification of model prediction

In order to verify the feasibility of the model, we can take five days, 120 hours data of water consumption, to verify the accuracy of the model and stability. The water consumption data of the model are shown in Table 1.

According to the change of water consumption in Table 1 for five days, the changing rule of water consumption data is basically the same in 24 hours. The trend of water consumption began at 4 o'clock and peaked at 9 o'clock. At 9 o'clock, the trend of water consumption began to decline at 10 o'clock. The trend of water consumption began to pick up again at 14 o'clock at 15 o'clock, and reached its peak at 19 o'clock at 18 again. The trend of water consumption began to decline gradually, which is basically in line with the people's habit of using water.

3.1 An examples of time water consumption forecasting by adaptive filtration method

The first step to predict water consumption by adaptive filtering is to select a certain number of time series and determine the weight number and the value. This example will be predicted 24 hours of water, so the weight number is 24 and the initial weight value is 1/24. We can calculate the water consumption in the 25th hour according to 24 weights and the first 24 water data. Calculate the 25th weight according to the formula of weight adjustment and the predicted and real values, then using the last 24 weights and the corresponding 24 water data to calculate the 26th water consumption, and so on. Finally, the new weight is obtained by using 96 hours of water consumption, and the number of iterations is set until the 24 weights are basically stabilised, using the latest weight to obtain a total of 24 hours of water consumption on the 5th day. Fortran prepared using a good program for prediction, The results are shown in Figure 1.

From Figure 1, we can see that the prediction results of the adaptive filtering model are basically consistent with the trend of the real water data, but the prediction accuracy is poor. According to the data, the maximum error of the 24 hour prediction results is 20.65% at 17 o'clock in the 13th day, and the minimum error of the prediction results is 0.94% at 10 o'clock in the 13th day. Therefore, the accuracy of time series prediction model is not very high, and it is not controlled within 15%, which does not meet the accuracy requirement of model prediction.

Table 1 the data of water consumption

Time	*The first day*	*The second day*	*The third day*	*The forth day*	*The fifth day*
1	23,163	25,332	24,584	25,685	24,022
2	24,000	22,315	22,158	23,568	22,581
3	25,324	23,419	23,155	22,984	20,584
4	24,627	26,331	23,184	22,569	20,896
5	37,501	35,327	30,562	31,541	34,598
6	43,231	44,528	40,548	42,621	38,589
7	48,441	49,312	46,859	49,521	40,985
8	51,441	52,461	50,225	48,598	50,654
9	50,219	51,008	52,785	53,624	55,256
10	47,112	48,216	50,337	51,652	52,572
11	48,335	46,315	49,524	50,245	48,569
12	43,632	44,012	46,258	45,123	47,589
13	41,335	42,316	46,589	43,856	40,879
14	37,224	37,892	40,358	38,585	43,456
15	33,021	36,515	38,569	35,892	38,568
16	34,443	36,859	39,225	39,262	38,456
17	41,012	42,156	41,158	41,562	38,074
18	45,217	45,263	43,394	47,235	45,525
19	44,315	43,546	46,892	45,535	46,879
20	39,126	41,995	43,254	42,323	48,689
21	40,521	39,871	40,623	42,121	42,358
22	35,132	34,516	38,562	40,132	39,624
23	32,710	31,698	36,215	36,169	33,652
24	27,352	26,859	30,251	30,158	27,526

3.2 An example of time water consumption forecasting by BP neural network method

The prediction method of urban water consumption by BP neural network belongs to the interpretive prediction method. The BP neural network algorithm is used to decompose the influencing factors of the time consumption and make comprehensive prediction when the time consumption is forecasted. There are two main influencing factors of water consumption, climate factors and water using habits (Hongbo et al., 2002). The quantisation value of temperature is measured according to the proportion of water used in the corresponding temperature range in historical data. The quantitative value of cloudy and clear is based on the common sense that people use more water on sunny days and use less water on rainy days. The weekly quantisation value is based on the ratio of the average weekly consumption to Monday consumption. The quantitative results are shown in Table 2.

Figure 1 Comparison of predicted results and real values of adaptive filtering method (see online version for colours)

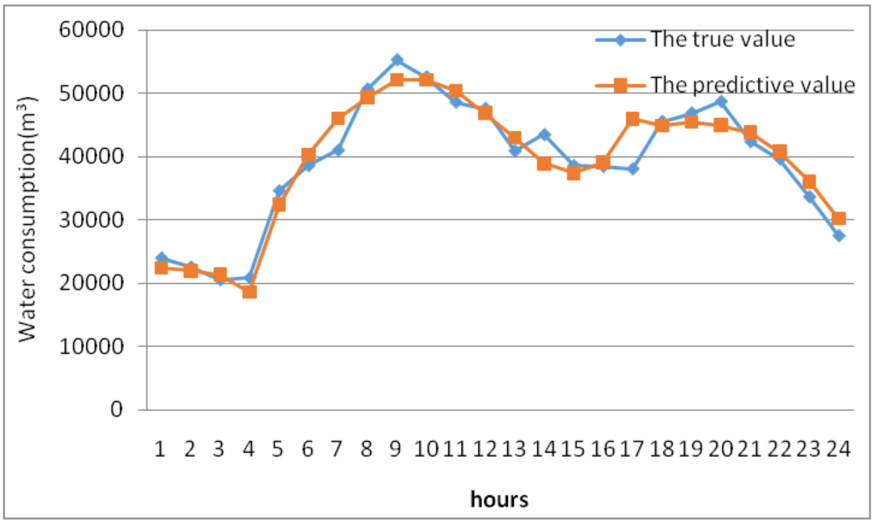
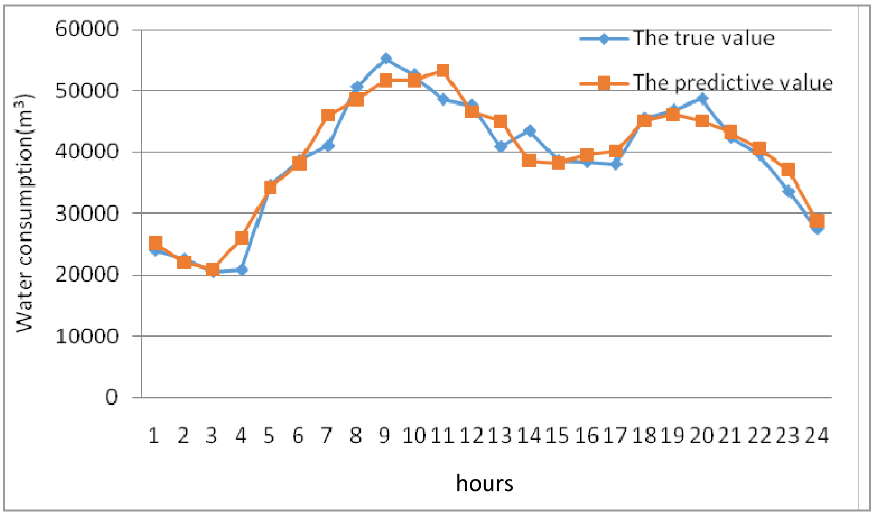


Figure 2 Comparison of predicted results and real values of BP neural network (see online version for colours)



According to the above formula, we can write the function by MATLAB, and then we can solve the model. The BP network structure has three input units, two hidden layers and one output units. The three input units are holiday conditions, sunny conditions and temperature, the output unit is time consuming water, the first hidden layer has seven nodes, and the second hidden layer has five nodes. The network learning rate is 0.01, the momentum factor is 0.08, the network training times are 40,000 times, the training error and the target error are 0.001 and 0.00005 respectively, and the synthetic threshold is 0.01. The hidden layer action function is Sigmoid, the output layer action function is purelin, the network training function is traingdm, and the network learning function is learngdm. The first three days were used as training samples, and the 4th day were used as test samples to predict the 5th day. The forecast results are shown in Figure 2.

Table 2 quantified value comparison table for water consumption factors

Temperature range	*Quantified values*	*Sunny conditions*	*Quantified values*	*Week*	*Quantified values*
0°C–5°C	0.15~0.25	Fine	2.25	Monday	1.00
5°C–10°C	0.25~0.35	Fine to cloudy	2	Tuesday	1.017
10°C–15°C	0.35~0.45	Cloudy	1.75	Wednesday	1.01
15°C–20°C	0.45~0.55	Cloudy to overcast	1.5	Thursday	1.014
20°C–25°C	0.55~0.65	Overcast	1.25	Friday	1.032
25°C–30°C	0.65~0.75	Light rain	1	Saturday	1.036
30°C–35°C	0.75~0.85	Medium rain	0.75	Sunday	1.040
35°C–40°C	0.85~0.95	Heavy rain	0.5		

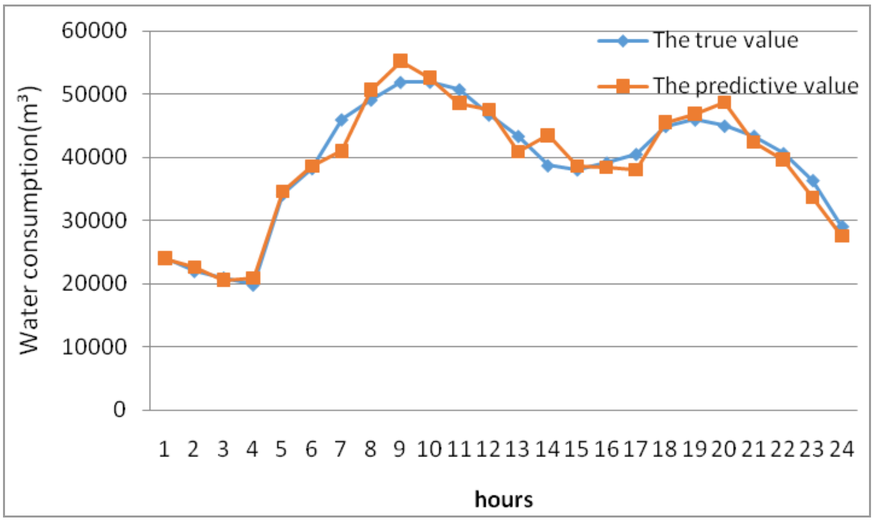
From Figure 2, we can see that the prediction result of BP neural network model is basically consistent with the trend of real water data, but the prediction accuracy is poor. According to the data, the maximum error of the 24 hour prediction results is 24.49% at 4 o'clock in the 14th day, and the minimum error of the prediction results is 1.04% at 15 o'clock in the 14th day. Therefore, the accuracy of BP neural network model is not very high, and it is not controlled within 15%, which does not meet the accuracy requirement of model prediction.

3.3 An example of time water consumption forecasting by combination prediction method

In the process of solving the objective function, we can solve the weight coefficient by the method of solving the multivariate function's most value by Lagrange, and then we can write the MATLAB function to calculate the weight coefficient. In this paper, the Lagrange method is used to solve the weight coefficient. Finally, 24 sets of optimal weight coefficients are obtained. Take seven points and eight points as an example, the weight coefficients are respectively 0.0684 and 0.9316 with 0.5004 and 0.4996. The predicted results of the specific combination are shown in Figure 3.

From Figure 3, we can see that the prediction results of the combination forecasting model are basically the same with the trend of the real water data, and the two curves are well fitted. According to the data, the maximum error of the 24 hour prediction results is 12.26% at 8 o'clock in the 14th day, and the minimum error of the prediction results is 0.10% at 1 o'clock in the 14th day. The prediction accuracy of the combined forecasting model is much higher than that of the single model, and all of them are within 15%, which is in line with the prediction accuracy requirement of the model.

Figure 3 Comparison of predicted results and real values of combined forecasting method (see online version for colours)



4 Comparison of prediction results

According to the calculation results of the model and the formula of the model evaluation index, four evaluation indexes of three models are calculated, and the model is evaluated synthetically.

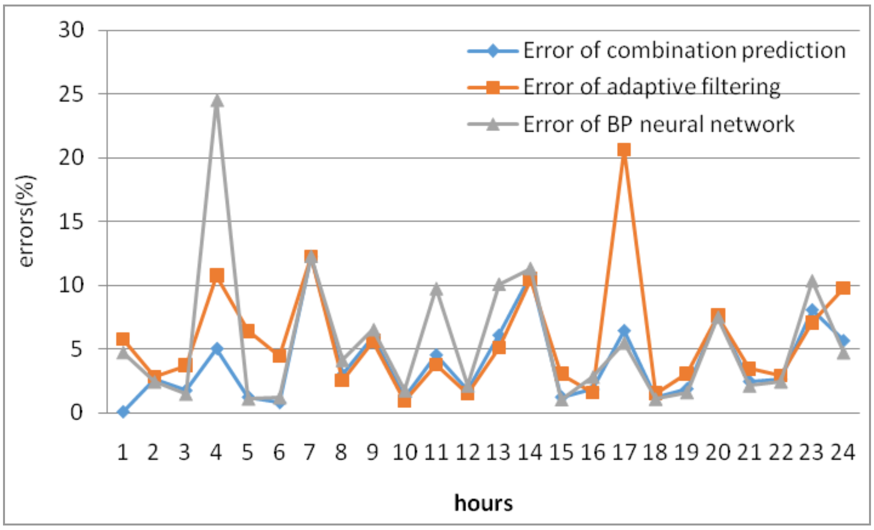
From Table 3, it can be seen that in the average relative error index, the average relative error of the combined prediction model is less than two single prediction models, and the average error of the prediction is less than 5%, which is in line with the precision requirement of the prediction model, so the accuracy of the combined prediction model is higher. At the same time, in the model precision index, the effectiveness of the combined prediction model is close to 1, which is higher than the validity of the two single models. It is proved that the combined prediction model has better prediction accuracy.

Table 3 comparison table for evaluation of three prediction models

The model comparison index	*Adaptive filtering model*	*BP neural network model*	*Combination prediction model*
Average absolute error	2,147.21	2,060.05	1,611.43
Average relative error (%)	5.47	5.26	4.11
The accuracy of variance	0.2650	0.2670	0.0314
The validity of model	0.7403	0.7201	0.9297

It can be seen from Figure 4 that error fluctuation of combination forecasting model is smooth, and the error level is smaller than the single model. At the same time, the variance of the accuracy of the combined forecasting model is smaller than the two single models. Therefore, the combined forecasting model has better prediction accuracy and stability than the single model.

Figure 4 The chart of three prediction model error (see online version for colours)



5 Conclusions

In this paper, the adaptive filtering method and BP neural network algorithm are used to establish the single prediction model of water consumption. On the basis of this, by solving the optimal weight minimisation prediction model error, we established the combined demand forecasting model, which is based on adaptive filtering method and BP neural network method. Through the comparison of the prediction error results of the three models, compared to the single forecasting model, combination forecasting model prediction results with higher accuracy and better stability, which provide a more scientific method for prediction of water for the city.

References

Bin, L., Shirong, X., Guang, B. and Wu, L. (2002) 'Grey neural network combined model for predicting urban water consumption', *China Water Supply and Drainage*, Vol. 18, No. 2, pp.66–68.

Deyang, L. (2014) *Research on Time Series Combination Forecasting Model based on SARIMA and BP Neural Network*, Lanzhou University, China.

Hongbo, L., Hongwei, Z. and Tian (2002) 'Prediction of water consumption by artificial neural network', *Chinese Water Supply and Drainage*, Vol. 18, No. 12, pp.39–41.

Hui, S., Yulan, W. and Fu, W. (2011) 'BP neural network and fuzzy time series forecasting model and its application', *Computer Applications*, Vol. 31, No. 2, pp.90–91.

Jian, L. (2007) *Study on Optimal Scheduling Model of Urban Water Supply System based on BP Neural Network and Genetic Algorithm*, Hohai University, China.

Jing, Z. and Jun, C. (2016) *Time Series ARIMA and BP Neural Network Prediction Model based on the Statistics and Decision*, Vol. 2016, No. 4, pp.29–32.

Mingyang, A., Wei, H., Shuang, L., Sia, F.W. and Junling, W. (2012) 'Comparison of three methods for forecasting water consumption of urban water supply', *Water Supply and Drainage*, Vol. 2012, No. 2, pp.229–231.

Pu, W., Rongyan, C., Xiaonan, S. et al. (2010) 'Application of weighted combination model in urban water consumption forecasting', *China Water Supply and Drainage*, Vol. 18, No. 3, pp.428–434.

Wei, C., Jian, L. and Zhicheng, W. (2009) 'Prediction model of urban water consumption based on BP neural network', *Journal of Harbin Institute of Technology*, Vol. 2009, No. 6, pp.197–200.

Yanchun, Z., Shuping, L., Ziwei, Z., Jilong, S. and Bilan, W. (2015) 'Short term water consumption forecasting based on BP neural network toolbox', *Water Supply and Drainage*, Vol. 2015, No. S1, pp.375–377.