The pollutant concentration prediction model of NNP-BPNN based on the INI algorithm, AW method and neighbor-PCA

Hong Zhao¹ · Yi Wang² · Jiahui Song² · Ge Gao²

Received: 7 January 2018 / Accepted: 28 April 2018 / Published online: 1 June 2018
© The Author(s) 2018

Abstract

At present, the numerical prediction models fail to predict effectively due to the lack of basic data of pollutant concentration in a short term in China. Therefore, it is necessary to study the statistical prediction methods based on historical data. The traditional Back Propagation Neural Network (BPNN) has been used to predict the pollutant concentration. The missing data also has an impact on modeling, and how to use historical data effectively of multiple monitoring stations in a city should be concerned. In this study, the Improved Newton Interpolation (INI) algorithm has been adopted to solve the problem of missing data, and assigning weight (AW) method has been proposed to enrich data of per station. The Neighbor-Principal Component Analysis (Neighbor-PCA) algorithm has been employed to reduce the dimension of data in order to avoid overfitting caused by high dimension and linear correlation of multiple factors. The strategy of early stopping and gradient descent algorithm have been utilized to avoid the slow convergence speed and overfitting by the traditional BPNN. The methods (INI, AW, Neighbor-PCA) have been integrated as a prediction model named NNP-BPNN. Forecasting experiments of PM_{2.5} have shown that the NNP-BPNN model can improve the accuracy and generalization ability of the traditional BPNN model. Specifically, the average root mean square error (RMSE) has been reduced by 24% and the average correlation relevancy has been increased by 9.4%. It took 20 s to implement BPNN model, it took 170 s to implement NN-BPNN model and it took 47 s to implement NNP-BPNN model. The time used by NNP-BPNN model is reduced by 72% than that of NN-BPNN model.

Keywords Newton interpolation · Improved Newton interpolation algorithm · Assigning weight method · Neighbor-PCA · BPNN · Pollutant concentration prediction

1 Introduction

Currently, environmental pollution poses a serious threat to human health (Zheng et al. 2016). The air quality prediction model mainly adopts the classic regression statistical model and the Numerical Air Prediction model (NAP) (Demuzere et al. 2009). The NAP uses mathematical models of the atmosphere and oceans to predict the air quality based on current atmosphere conditions and the pollutant sources (Zhang et al. 2014). Using cloud computing to implement high performance of computing of NAP would be the key

research area in the future (Li et al. 2017; Liu et al. 2017; Li et al. 2017; Cui et al. 2016). However, the NAP can not be adopted among many cities in China due to the lack of the basic data of pollutant concentration in a short time (Huchao et al. 2015; Niska et al. 2005). Hence, statistical prediction methods based on historical data should be studied in order to obtain a high prediction performance.

Shi et al. (2012) have proved that Artificial Neural Network (ANN) can provide better results than the traditional multiple linear regression models about pollutant concentration prediction based on historical monitoring data. The combined model of ARMA and BPNN based on historical monitoring data studied by Zhu and Lu (2016) has a smaller prediction error than that of the traditional BPNN. The time series data is decomposed into wavelet coefficients, and the prediction experiments based on three types of neural network models (Multi-layer Perceptron, Elman and Support Vector Machine) based on historical data have shown that the improved models have provided

✉ Yi Wang
wynk@mail.nankai.edu.cn
Hong Zhao
zhaoh@nankai.edu.cn

¹ Nankai University, Weijin Road, Tianjin, China

² Nankai University, Tongyan Road, Tianjin, China

low RMSE and mean absolute error in comparison to the original models, which is developed by Feng et al. (2013). Grivas and Chaloulakou (2006) has employed a novel ANN which selected factors of historical monitoring data via genetic algorithm, proving the model has produced smaller error than that of the regression model in the prediction of PM_{10} concentration. Shi et al. (2012) has used Feed-forward BPNN with the hyperbolic tangent sigmoid activation function and the Levenberg–Marquardt optimization method to predict PM_{10} concentration in each station. However, there is no explanation that the interaction between central station and neighbor stations in their studies.

The monitoring data that is often lost from multiple stations in a city. Nejadkoorki and Baroutian (2011) have removed the missing data, resulting in the waste of data. The Newton Interpolation algorithm has good interpolation effect when the data loss ratio is not too high (Breu et al. 2016). Thus, in this study, the missing data is effectively interpolated by INI algorithm. In order to enlarge the data, we have taken neighbor stations data into consideration. Because different distance makes different effects on central station, weight has been assigned to each stations data. The oversize of the dimension of input data would affect the generalization ability of the BPNN, therefore, the PCA algorithm is utilized to extract the eigenvector from neighbor stations effectively in the study (Skrobot et al. 2016). The extracted eigenvector from neighbor stations and the eigenvectors of the central station are regarded as input to the BPNN, which can not only keep information as much as possible, but also ensure that the dimension of the input to the BPNN is suitable, increasing the generalization ability of the BPNN. Finally, early stop strategy (Yang and Zhang 2017) has been performed during training. That is, when the error reached the value set by program, the training should be terminated in order to avoid overfitting. The learning rate gradient descent strategy (Huang and Lin 2009) has been employed to avoid the slow convergence speed.

NEXT, INI algorithm would be proved in detail for dealing with missing data. Multiple stations data and AW method have enlarge data for modeling. However, this affects the speed and efficiency of BPNN. Thus, Neighbor-PCA algorithm has been employed to reduce the dimension of data in order to avoid overfitting. BP model and NN_BPNN model are used for contrast. The effectiveness of NN_BPNN model has showed that the integration of INI algorithm and AW method proved to be persuasive and the effectiveness of NNP_BPNN model has showed that the combination of INI algorithm, AW method and Neighbor-PCA algorithm proved to be persuasive.

Fig. 1 The neuron map of the BPNN

Fig. 2 The structure of BPNN

2 Related algorithm

In this section, the basic principles of BPNN, Newton interpolation algorithm and PCA algorithm are respectively introduced.

2.1 BPNN principles

BPNN is composed by a series of simple unit connected with each other densely. In data mining, neural network has been employed by Chang and Yang (2017). Each unit has a certain amount of input and output. The structure of neuron is shown in Fig. 1 and the structure of BPNN is shown in Fig. 2.

The stimulation delivered by neuron is called x_i . The connection weight is called w_i . This accumulation is called a_j . Sigmoid(x) is put as the activation function. The error is utilized to modify the connection weights for feedback to complete the learning process, which is the feedback mechanism of the BPNN (Guan et al. 2016). The inner of neuron also interferes with the output result, so an extra bias named as b is brought. So we get the formula (1).

$$a_j = f\left(\sum_{i=1}^n (x_i \times w_i) + b\right) \quad (1)$$

But in practice, the traditional BPNN has some shortcomings, such as the long training time, the slow convergence speed, local minimum and the poor stability. It is hard to adjust the initial weight and learning rate parameters (Xiao et al. 2012). If the dimension of the input is high enough, there is a greater challenge for training.

2.2 Newton interpolation algorithm

Interpolation function has many different types. Li et al. has introduced multiple methods of interpolation (Li et al. 2017). Using basis function to get the Lagrange interpolation polynomial is common in the theoretical analysis. The basis function would be changed with the change of nodes, which results in the change of formula. Newton interpolation algorithm can overcome this shortcoming (Varsamis and Karampetakis 2012). Newton interpolation algorithm function is determined by independent variables and dependent variables. The first-order function, second-order function, kth-order function are shown in formulas (2), (3) and (4) respectively. Newton interpolation formula defined in formula (5) is deduced by formulas (2), (3) and (4).

$$f[x_i, x_j] = \frac{f(x_j) - f(x_i)}{x_j - x_i} \quad (2)$$

$$f[x_i, x_j, x_k] = \frac{f[x_j, x_k] - f[x_i, x_j]}{x_k - x_i} \quad (3)$$

$$f[x_0, x_1, \dots, x_k] = \frac{f[x_0, \dots, x_{k-2}, x_k] - f[x_0, x_1, \dots, x_{k-1}]}{x_k - x_{k-1}} \quad (4)$$

$$f(x) = f(x_0) + f[x_0, x_1](x - x_0) + \dots + f[x_0, x_n](x - x_0) \dots (x - x_{n-1}) + f[x, x_0, \dots, x_n](x - x_0) \dots (x - x_n) = N_n(x) + R_n(x) \quad (5)$$

Whereas, the instability of the interpolation results is affected by the high power of formula (5) during the process of interpolation (Hilbach 1979).

2.3 PCA algorithm

Mapping high dimensional data to low dimensional data through PCA algorithm (Nejadkoorki and Baroutian 2011). It seems convenient to introduce the following steps of PCA algorithm (Table 1).

3 NNP-BPNN model

In order to improve the data utilization and prediction accuracy, the NNP-BPNN model is proposed for pollutant concentration prediction. In this paper, the INI algorithm has been utilized to handle missing data. The AW method is calculated based on geographical location. The Neighbor-PCA algorithm is used to deal with the data of the neighbor stations.

Table 1 The steps of PCA algorithm

Input	data $\phi = \{X_1 X_2 \dots X_p\}$ low dimensional: k
Steps	1. Centralize all the data: $X_i - \frac{1}{p} \sum_{p=1}^p X_i$ 2. Calculate the covariance matrix for the dataset: XX^T 3. Eigen value decomposition for the matrix XX^T 4. Take the eigenvectors of the largest k eigenvalues: $v_1, v_2, \dots, v_k$
Output	Projection matrix: $V = (v_1, v_2, \dots, v_k)$

3.1 INI algorithm

There will be some outliers or even missed value in the data because of all kinds of reasons. If we use the Newton interpolation algorithm directly, large error would be brought. The INI algorithm is seen as formula (6).

$$f(x) = \begin{cases} f_{\text{newton}}(x) & \frac{f_{\text{newton}}(x) - \overline{f(x)}}{f(x)} \leq \varepsilon(I) \\ \overline{f(x)} = \sum_{i=1}^k x_i p_i & \frac{f_{\text{newton}}(x) - \overline{f(x)}}{f(x)} > \varepsilon(II) \end{cases} \quad (6)$$

$$\overline{f(x)} = \frac{1}{k} \sum_{i=1}^k f(x_i) \quad (7)$$

$f_{\text{newton}}(x)$ is the interpolation results of formula (5). $\overline{f(x)}$ is the expectation of k samples. x_i ($i = 1, 2, 3, \dots, k-1$) is independent variables from the nearest k hours. The data of kth hour is the dependent variable. The p_i is the probability and it is a constant of $\frac{1}{k}$.

If the formula I is true, then the samples of $(k-1)$ is used as the independent variables and the kth sample is used as the dependent variable, getting the $f_{\text{newton}}(x)$. The expectation value of $(k-1)$ samples is as the input to $f_{\text{newton}}(x)$. If the formula II is true, then $f(x)$ is calculated by formula (7). The data before interpolation is named as D_{original} . The data after the interpolation is named of D.

3.2 Multiple stations data and the dimension reduction

In this part, we mainly introduce the data composition from multiple stations, Neighbor-PCA algorithm and the training process of BPNN model.

3.2.1 Multiple stations data and AW method

Due to different distance, neighbor stations have a different impact on the central station. The air pollutant concentration

of all monitoring stations is recalculated by AW method defined as formula (8).

$$k_{ij} = 1 - \frac{d_{ij}}{\sum_{j=1}^n d_{ij}} \quad (8)$$

The k_{ij} is the weight between the i th station and the j th station. The d_{ij} is the distance between the i th station and the j th station. It is calculated based on the latitude and longitude and defined in formula (9).

$$d_{ij} = 637814 \times \text{ACOS} \left(1 - \left(\sin \left(\frac{(90 - C) \times PI}{180} \right) \times \cos \left(\frac{B \times PI}{180} \right) - \sin \left(\frac{(90 - F) \times PI}{180} \right) \times \cos \left(\frac{E \times PI}{180} \right) \right)^2 \right) \quad (9)$$

The C is the latitude of the i th station, and The B is the longitude of the i th station, The F is the latitude of the j th station, and The E is the longitude of the j th station.

The closer distance is, the greater the impact on the central station. The k_{ij} from formula (8) is the minus function. The neighbor stations that are closer from central station can get greater weight from k_{ij} , enhancing the ability of BPNN model.

The D_i is the data of i th station and its composition is seen as formula (10). The D is the collection of data of k stations and is seen as formula (11). ϕ_i is the data of $D_i \times k_{ij}$ where n ranges from 0 to n and is seen as formula (12). $\phi_{neighbor}$ is the data of $D_i \times k_{ij}$ where n ranges from 1 to n .

$$D_i = (F_{pollutant}, F_{neighbor}) \quad (10)$$

$$D = [D_1, \dots, D_i, \dots, D_n] \quad (11)$$

$$\phi_i = [D_0 \times 1, D_1 \times k_{i1}, \dots, D_i \times k_{ij}, \dots, D_n \times k_{in}] \quad (12)$$

Compared with the traditional methods, we have considered the neighbor stations based on geographic information, which improves the accuracy of BPNN model.

3.2.2 Dimension reduction process

If the dimension is too high or the sample size is too small, the BPNN would be unable to learn the general rule (Meng and Meng 2010). In this paper, the PCA algorithm has been used to handle data $\phi_{neighbor}$.

The extracted eigenvector from data $\phi_{neighbor}$ using PCA algorithm is named as V defined as formula (13). ϕ defined in formula (14) has been used as the input to BPNN. A three-layer structure BPNN mentioned above has been applied in this paper. The number of nodes of the hidden layer are determined by the formula (15). The m is the number of nodes of the input layer and the q is the number of nodes of the output layer. The a is the constant from 1 to 10 (Wang

et al. 2017). The relevancy, error and RMSE between the prediction and actual value are respectively defined as formulas (16), (17) and (18).

$$V = (v_1, v_2, \dots, v_m) \quad (13)$$

$$\phi = [D_0, V] \quad (14)$$

$$N = \sqrt{m + q} + a \quad (15)$$

$$R = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}} \quad (16)$$

$$\text{error} = (y - \hat{y}) \quad (17)$$

$$\text{RMSE} = \sqrt{\frac{\sum_1^n \text{error}^2}{n}} \quad (18)$$

4 Modeling methods

In order to verify the effectiveness of the NNP-BP model, three models have been established, including the BPNN model, the NN-BPNN model with INI algorithm and AW method and the NNP-BPNN model with INI algorithm, AW method and Neighbor-PCA algorithm.

Modeling is performed in Matlab7.0, and data D is as input to the models. We choose traingda function as gradient descent function and set 100 as interval. Learning rate is 0.1. We set 1000 as the largest number of training and set 0.0001 as target error. The input nodes is 15, the number of output nodes is 1, and a is 8. So, the hidden layer node is 12 according to the formula (15). One half of data has been applied to train and another half of data has been treated to evaluate the model.

The learning rate (lr) valued between 0 and 1, which determines the step size for updating in each iteration. If lr is too

Table 2 The steps of BPNN model

Input	Original data $D_{original}$
Steps	1. Calculate the number of hidden layer nodes of BPNN according to formula (15) and set parameters 2. Training the BPNN model
Output	The sum of error and RMSE

Table 3 The steps of NN-BPNN model

Input	Original data $D_{original}$
Steps	<ol style="list-style-type: none"> 1. The missing value is calculated to obtain the data D according to formulas (6) and (7). 2. Calculate the distance d_{ij} of each station according to the formula (9) 3. Calculate the weight k_{ij} according to d_{ij} and formula (10) 4. We can get the data ϕ according to the k_{ij} and formula (11) 5. The data ϕ is used as the input for BPNN model 6. Calculate the number of hidden layer nodes of BPNN according to formula (15) and set other parameters 7. Training the NN-BPNN model
Output	The sum of error and RMSE

Table 4 The steps of NNP-BPNN model

Input	Original data $D_{original}$
Steps	<ol style="list-style-type: none"> 1. Use PCA algorithm to deal with data $\phi_{neighbor}$ to obtain data V 2. Get the input to BPNN according to the data V and formulas (13) and (14) 3. Calculate the number of hidden layer nodes of BPNN according to formula (15) and set other parameters 4. Training the NN-BPNN model
Output	The sum of error and RMSE

big, it is easy to oscillate, and if lr is too small, it converges too slowly. The gradient descent method can not only solve the problem, to a certain extent, but also make results closer to the global minimum (Table 2).

4.1 Modeling of BPNN model

4.2 Modeling of NN-BPNN model

Assuming data of the i th hour is missed, we take the data named x from the $(i-4)$ th to the $(i-2)$ th and the $(i-1)$ th data is named as y , which is used to calculate Newton interpolation polynomial. The value of $f_{newton}(\bar{x})$ can be calculated. The ranged from 0 to 1. The missing data can be obtained by formulas (6) and (7). Different weights have been assigned for each station to get the data ϕ (Table 3).

4.3 Modeling of NNP-BPNN model

Throughout this section, The PCA algorithm is performed to deal with data $\phi_{neighbor}$ to obtain data V and train NNP-BPNN model as follows (Table 4).

5 Experiments and analysis

Here, we mainly talk about data in experiment and discussions about results.

5.1 Data description

16 stations data in one city, including pollutant concentration and meteorological parameters of 254 days from July 30, 2015 to April 10, 2016, have been employed in this paper. The basic information is consist of stations name, time, longitude, latitude. Pollutants are consist of CO , SO_2 , O_3 , NO , NO_2 , NOX , $PM_{2.5}$, PM_{10} , VOC . The meteorological parameters are consist of relative humidity, wind direction, wind speed, temperature, pressure, visibility, total rainfall. Among them, the longitude and latitude are used to calculate the distance. A continuous 24 h data is lost among 16 stations.

5.2 Results and discussions

Three models have been established, including BPNN model, NN-BPNN model, NNP-BPNN model, which the same data and parameters are used. Here, we discuss the effectiveness of NN-BPNN model with INI algorithm and the AW method. Then we discuss the validity of NNP-BPNN model with INI algorithm, AW method and Neighbor-PCA algorithm. The RMES has been regarded as the evaluation criterion shown in formula (18).

5.2.1 The effectiveness of NN-BPNN model

Figure 3 is the comparison of RMES between BPNN model and NN-BPNN model. In the neural network, the local

Fig. 3 The RMES between BPNN model and NN-BPNN model

Fig. 4 The RMES of three models

minimum value of the error function is not the global minimum. Ten times of experiments have been performed to prove that the NN-BPNN model has a smaller value of sum of RMES than that of BPNN model, as shown in Fig. 5. The RMSE of three models is shown in Fig. 4. Meanwhile, experimental results on the concentration of $PM_{2.5}$ show that the sum of RMES of NNP-BPNN model is relatively minimal, as shown in Fig. 5. The relevancy of the three models is shown in Table 5.

In ten times of experiments, the area of RMSE produced by BPNN model is the largest among that of three models, the average RMSE of NN-BPNN model is reduced by 18% than that of BPNN model, the average RMSE of NNP-BPNN model is reduced by 24% than that of BPNN model and the average RMSE of NNP-BPNN model is reduced by 7% than that of NN-BPNN model. It took 20 s to implement BPNN model, it took 170 s to implement NN-BPNN model

Fig. 5 The RMES of three models

Table 5 The value of relevancy

ID	BPNN	NN-BPNN	NNP-BPNN
1	0.75	0.77	0.79
2	0.61	0.70	0.81
3	0.73	0.76	0.82
4	0.70	0.76	0.78
5	0.71	0.61	0.76
6	0.77	0.76	0.81
7	0.78	0.78	0.80
8	0.81	0.71	0.80
9	0.73	0.77	0.79
10	0.68	0.78	0.80

and it took 47 s to implement NNP-BPNN model. The time used by NNP-BPNN model is reduced by 72% than that of NN-BPNN model. Empirically, the NNP-BPNN model of statistical forecasting method based on historical monitoring data is simple and practical.

6 Conclusions

In this paper, we study the problem that how to handle missing data and how to utilize historical data effectively for prediction with BPNN. There are some achievements in this paper. Firstly, the INI algorithm is adopted to deal with missing data, so as to avoid the waste of data. Secondly, the AW method not only enriches the experimental data, improving the utilization rate of data, but also reduces the RMSE. The average RMSE of NN-BPNN model is reduced by 18% compared with BPNN model. Thirdly, the NNP-BPNN model with INI algorithm, AW method and Neighbor-PCA algorithm has improved the prediction accuracy and relevancy. The average RMSE of NNP-BPNN model is reduced by 7% compared with NN-BPNN model. The average RMSE of NNP-BPNN model is reduced by 24% compared with BPNN model.

There are still some limitation in this research. More methods of dimension reduction should be studied in the future work. The research of NNP-BPNN model to find out the global minimum error. Multiple machine learning models should be explored in air pollutant concentration prediction area.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Breuß M, Kemm F, Vogel O (2016) A Numerical Study of Newton Interpolation with Extremely High Degrees
- Chang X, Yang Y (2017) Semisupervised feature analysis by mining correlations among multiple tasks. *IEEE Trans Neural Netw Learn Syst* 28(10):2294–2305
- Chang X, Yu YL, Yang Y et al (2017) Semantic pooling for complex event analysis in untrimmed videos. *IEEE Trans Pattern Anal Mach Intell* 39(8):1617–1632
- Cui B, Liu Z, Wang L (2016) Key-aggregate searchable encryption (KASE) for group data sharing via cloud storage. *IEEE Trans Comput* 65(8):2374–2385
- Demuzere M, Trigo RM, Arellano VGD et al (2009) The impact of weather and atmospheric circulation on O₃ and PM₁₀ levels at a mid-latitude station. *J Atmos Chem Phys* 9(2009):2695–2714
- Feng Q, Wu S, Du Y et al (2013) Improving neural network prediction accuracy for PM₁₀ individual air quality index pollution levels. *Environ Eng Sci* 30(12):725
- Grivas G, Chaloulakou A (2006) Artificial neural network models for prediction of PM hourly concentrations, in the Greater Area of Athens, Greece. *Atmos Environ* 40(7):1216–1229
- Guan Z, Tian Z, Xu Y, et al (2016) Rain fall predict and comparing research based on Arcgis and BP neural network. In: International conference on materials engineering, manufacturing technology and control
- Hlbach G (1979) The general recurrence relation for divided differences and the general Newton-interpolation-algorithm with applications to trigonometric interpolation. Springer, New York
- Huang N, Lin L (2009) An improved BP neural network model based on quasic-Newton algorithm. In: International conference on natural computation, IEEE, pp 352–356
- Huchao LI, Shao A, Dengxin HE et al (2015) Application of back-propagation neural network in predicting non-systematic error in numerical prediction model. *J Plateau Meteorol* 42(6):1198–1201
- Li Z, Nie F, Chang X et al (2017) Beyond trace ratio: weighted harmonic mean of trace ratios for multiclass discriminant analysis. *IEEE Trans Knowl Data Eng* PP(99):1
- Li J, Zhang Y, Chen X et al (2018) Secure attribute-based data sharing for resource-limited users in cloud computing. *J Comput Secur* 75:1–12
- Li T, Liu Z, Li J et al (2016) CDPS: A cryptographic data publishing system. *J Comput Syst Sci* 89
- Liu Y, Zhu Q, Yao D et al (2015) Forecasting urban air quality via a back-propagation neural network and a selection sample rule. *Atmosphere* 6(7):891–907
- Liu Z, Li T, Li P et al (2017) Verifiable searchable encryption with aggregate keys for data sharing system. *J Future Gener Comput Syst*
- Meng X, Meng X (2010) Nonlinear system simulation based on the BP neural network. In: Third international conference on intelligent networks and intelligent systems. IEEE Computer Society, pp 334–337
- Nejadkoorki F, Baroutian S (2011) Forecasting extreme PM₁₀ concentrations using artificial neural networks. *Int J Environ Res* 6(1):277–284
- Niska H, Rantamki M, Hiltunen T et al (2005) Evaluation of an integrated modelling system containing a multi-layer perceptron model and the numerical weather prediction model HIRLAM for the forecasting of urban airborne pollutant concentrations. *J Atmos Environ* 39(35):6524–6536
- Shi LZ, Deng QH, Lu C et al (2012) Prediction of PM₁₀ mass concentrations based on BP artificial neural network. *J Cent South Univ* 43(5):1969–1974
- Skrobot VL, Castro EVR, Pereira RCC et al (2016) Use of principal component analysis (PCA) and linear discriminant analysis (LDA) in gas chromatographic (GC) data in the investigation of gasoline adulteration. *Energy Fuels* 21(6):5–19
- Ul-Saufie AZ, Shukri A, Nor Y et al (2011) Comparison between multiple linear regression and feed forward back propagation neural network models for predicting PM₁₀ concentration level based on gaseous and meteorological parameters. *Int J Appl Sci Technol* 1:42–49
- Varsamis DN, Karampetakis NP (2012) On a special case of the two-variable Newton interpolation polynomial. In: International conference on communications, computing and control applications. IEEE, pp 1–6
- Wang J, Shi P, Jiang P et al (2017) Application of BP neural network algorithm in traditional hydrological model for flood forecasting. *Water* 9(1):48
- Xiao F, Wu M, Huang H et al (2012) Novel node localization algorithm based on nonlinear weighting least square for wireless sensor networks. *Int J Distrib Sens Netw*:1238–1241 (2012)
- Yang M, Zhang X (2017) A novel travel adviser based on improved back-propagation neural network. In: International conference on intelligent systems, modelling and simulation, IEEE, pp 283–288
- Zhang L, Wang S, Yu Z et al (2014) Development of an instant correction and display system of numerical weather prediction products in China. *J Chin Geogr Sci* 24(6):682–693
- Zheng S, Pozzer A, Cao CX et al (2016) Long-term (2001–2012) fine particulate matter (PM_{2.5}) and the impact on human health in Beijing, China. *J Atmos Chem Phys* 14(21):5715–5725
- Zhu H, Lu X (2016) The prediction of PM_{2.5} value based on ARMA and improved BP neural network model. In: International conference on intelligent NETWORKING and collaborative systems, pp 515–517 (2016)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.