Urban Climate 41 (2022) 101078

Contents lists available at ScienceDirect

Urban Climate

journal homepage: www.elsevier.com/locate/uclim

Smog prediction based on the deep belief - BP neural network model (DBN-BP)

Jiawei Tian [a], Yan Liu [a], Wenfeng Zheng [a,*], Lirong Yin [b,*]

[a] School of Automation, University of Electronic Science and Technology of China, Chengdu, 610054, PR China
[b] Department of Geography and Anthropology, Louisiana State University, Baton Rouge, LA 70803, USA

ARTICLE INFO

Keywords:
PM2.5
PM10
Deep belief -back propagation neural network
In-depth prediction
Haze
Air pollution

ABSTRACT

Smog pollution is becoming a significant problem for people worldwide, becoming an essential threat to the global environment. Many studies on haze already exist, which still need to continue in-depth research to better deal with haze problems. Due to its unique geographical environment, Sichuan has become one of the areas with severe smog pollution. Therefore, the research and prediction of smog pollution in Sichuan has become an urgent need. This paper proposes a deep learning technology based on a Deep Belief-Back Propagation neural network. It makes in-depth prediction research by using the air pollution data of PM2.5, PM10, O_3, CO NO_2, and SO_2 in Sichuan smog to provide a decision-making basis for predicting and preventing smog polluted weather. According to the prediction results of the model, the concentrations of PM2.5 and PM10 in Chengdu were predicted. The analysis shows that the larger the number of hidden layers in the belief network, the higher the prediction accuracy. Under the same network, the prediction accuracy of PM2.5 is significantly higher than that of PM10. Compared with the traditional Back Propagation neural network, the prediction effect of the Deep Belief-Back Propagation neural network is better.

1. Introduction

People worldwide have suffered or are currently suffering from smog pollution, including China, the United States, Britain, Germany, Japan, India, and France (Narayanan, 2002; Sun et al., 2006; Zheng et al., 2015). Smog pollution has become an essential threat to the global environment (Friedlander, 1977; Ma et al., 2012; Sun et al., 2006; Zhang et al., 2013; Wang et al., 2014; Yin et al., 2019; Zheng et al., 2021a, 2021b, 2021c, 2021d; Zheng et al., 2021a, 2021b; Ma et al., 2021). In the mid-20th century, hundreds of people died in the photochemical smog event in Los Angeles (Tiao et al., 1975) due to air quality. The smog incident in London in 1952 (Hunt et al., 2003) was also caused by smog. In the following two months, 12,000 British people died of respiratory diseases. There were different degrees of coronary heart disease, bronchitis, tuberculosis, and even cancer. In the early days of industrialization, Japan also suffered from air pollution caused by smog (Imura, 2013). Hundreds of people were suffering from respiratory diseases, and their lives were seriously threatened. As a developing country, India also suffers from the haze. The research of Ambade, B. et al. (Ambade et al., 2021a, 2021b; Ambade et al., 2020; Kumar et al., 2020) Shows that PM2.5 is closely related to carcinogens such as polycyclic aromatic hydrocarbons, which is a significant threat to human health. China has also suffered from massive and severe smog pollution (Ma et al., 2012; Li et al., 2016; Chen et al., 2020). In some cities, the air pollution level reached six levels of heavy pollution, and the air pollution

* Corresponding authors.
E-mail addresses: winfirms@uestc.edu.cn (W. Zheng), lyin5@lsu.edu (L. Yin).

https://doi.org/10.1016/j.uclim.2021.101078
Received 27 August 2021; Received in revised form 14 November 2021; Accepted 27 December 2021
Available online 18 January 2022
2212-0955/

index in some areas exceeded 500 (Zheng et al., 2017; Li et al., 2015; Tang et al., 2020a, 2020b; Zheng et al., 2015; Tang et al., 2021).

Existing studies have focused on haze already, which still need to continue in-depth research to better deal with haze problems (Chen et al., 2020; Li et al., 2017a, 2017b; Liu et al., 2018; Zheng et al., 2016a, 2016b; Zheng et al., 2015; Zheng et al., 2013). Yang X et al. (Yang et al., 2016a, 2016b) built a Long-Term Prediction Model of Beijing Haze Episodes Using Time Series Analysis in 2016 and received good predictions. Lee et al. (Lee et al., 2017) predicted haze events in southeast Asia using machine learning algorithms in 2017. Chen, Y (Chen, 2018; Chen et al., 2019) proposed a prediction algorithm of PM2.5 mass concentration based on adaptive BP neural network. Ni, X. Y. et al. (Ni et al., 2017) used the BP neural network to predict the short-term PM2.5 concentration in Beijing. In recent years, deep belief neural networks have also been used in the field of smog (Xie, 2017; Xing et al., 2019).

Due to the haze deposition caused by the basin's topography, Sichuan has become one of the areas with severe smog pollution (Zhang et al., 2019; Li et al., 2017a, 2017b; Zheng et al., 2016a). Therefore, the research and prediction of smog pollution in Sichuan has become an urgent need.

In order to realize the short-term prediction of two critical factors in smog (PM2.5/PM10), this paper designs and constructs a fog and haze prediction model based on a Deep Belief-BP neural network. By inputting the data set: including historical haze data (PM2.5, PM10, O_3, CO, NO_2, SO_2 concentration) and future haze data (PM2.5/PM10 concentration), we predicted the PM2.5 and PM10 concentrations in Chengdu urban area, compared and analyzed the actual values. By selecting different hidden layers to carry out simulation experiments, the back-propagation algorithm is used to optimize the model parameters. Finally, we compared the haze prediction model based on the Deep Belief-Back Propagation neural network with the existing mainstream Back Propagation neural network prediction method.

2. The study area and data

2.1. Data descriptions

Sichuan is characterized by inland basins. Chengdu is located in the center of the Sichuan Basin, with an average elevation of 500 m in the urban area of Chengdu. Due to the significant difference in the meteorological background in mountainous, hilly, and plain areas, the terrain of Chengdu Plain is relatively closed. Therefore, this study selects Chengdu as the research area to avoid the impact of different meteorological backgrounds.

This study collected the real-time concentration data of PM2.5, PM10, O_3, CO, NO_2, and SO_2 at six detection points (Sanwayao, Caotangsi, Junpingjie, Liangjiaxiang, Shahepu, and Shilidian) in Chengdu urban area from June 1, 2014, to June 30, 2017, by compiling Python crawler code.

The detection data of the detection point is updated once every hour and 24 times a day. Detection data include PM2.5, PM10, O_3,

Fig. 1. Distribution of detection points.

SO_2, CO, and NO_2. The unit of PM2.5, PM10, O_3, and SO_2 is μg/m^3, the unit of CO and NO_2 is mg/m^3. The average concentration at each time of the six monitoring points is calculated as the historical concentration data in this study as effective data of the haze prediction model.

The geographical location of the six detection points is shown in Fig. 1. The detection points are evenly distributed in Chengdu, which can fully reflect the air quality in Chengdu.

Smog's performances vary through seasons. Generally, the autumn and winter seasons are the high-frequency seasons of smog (Zheng et al., 2016b). Smog occurs relatively at a low frequency in the spring and summer (Tan et al., 2009). Therefore, considering the representativeness of different seasons, we select the haze data (PM2.5, PM10, O_3, CO, NO_2, SO_2 concentration) and meteorological data (atmospheric temperature, humidity, and wind data) from 0:00 on July 4, 2016, to 23:00 on July 10, 2016, and from 0:00 on December 30, 2016, to 23:00 on January 5, 2017. Meteorological data comes from China Weather Network.

2.2. Correlation analysis

Correlation coefficients were analyzed for weather factors and PM2.5 using MATLAB R2014a tools in conjunction with correlation analysis theory. The results are shown in Table 1 below:

According to the PM2.5 correlation coefficient table, the correlation between PM2.5 and other factors in winter is PM10≧NO_2≧CO≧SO_2≧O_3; in summer, the correlation between PM2.5 and other factors is ranked as PM10≧CO≧O_3≧SO_2≧NO_2. When determining the correlation, if the | correlation coefficient value | < 0.4, it is low correlation; if 0.4 ≤ | correlation coefficient value | < 0.7, is significant linear correlation; if 0.7 ≤ | correlation coefficient value | <1, is highly correlated, so in general, PM2.5 at the Chengdu detection point is low-correlation with meteorological factors, significant correlation with CO, SO_2, NO_2, O_3, and high correlation with PM10.

Because both PM2.5 and PM10 affect the critical factors of haze, the prediction of smog is converted into the prediction of PM2.5 and PM10, and according to the correlation analysis results, PM2.5 in the short-term haze weather time period. Correlation with meteorological factors is not significant. It is considered that the weather parameters are stable in the short term. Therefore, high correlation CO, SO2, NO2, O3, PM10, and PM2.5 are selected as input data when modeling the haze prediction model. Used to train the haze prediction model and achieve the goal of predicting the concentration value of PM2.5/PM10.

2.3. Mean completion data

The experimentally collected PM2.5, PM10, O_3, CO, NO_2, and SO_2 concentration effective data at each test point in Chengdu from June 1, 2014, to June 30, 2017, totaled 26,120 effective data. In order to fill the missing pollutant concentration data at some time, this study uses the mean interpolation method to fill the data. After the completion method completes the data, a total of 27,000-time data is obtained, which ensures sufficient data for constructing a deep neural network. As shown in Eq. 1:

$$X_t = \frac{1}{2}(X_{t-1} + X_{t+1}) \tag{1}$$

Among them, X_t represents the missing pollutant concentration data at the current time, and X_{t-1} represents the pollutant concentration data at the previous moment, X_t represents the pollutant concentration data at a time before and after.

2.4. Standardized processing

In the neural network, large-value data tends to increase the proportion of influence on the model, thus losing the characteristic properties of the data with low value (Hagan et al., 1997). Therefore, to avoid errors caused by different numerical ranges in this study, all historical weather data are first calculated according to the following formula. Normalized to between −1 and 1, as shown in Eq. 2:

$$X' = \frac{X - \overline{X}}{X_{\max} - X_{\min}} \tag{2}$$

Among them, X' indicates the weather data set after the standardization process is completed. X represents the original weather data set, $\overline{X}$ indicates the mean of the weather data set. X_{max} indicates the maximum value of the weather dataset, and X_{min} represents the minimum value of the weather dataset.

At the same time, the historical weather data after the pre-processing is divided into training data, verification data, and test data according to the proportion of 80%, 10%, and 10%.

Table 1
Correlation coefficient values.

Correlation coefficient	Highest temperature	Lowest temperature	humidity	Wind power	O_3	CO	NO_2	PM10	SO_2
Winter	0.29	−0.01	−0.25	−0.35	−0.13	0.49	0.54	0.79	0.48
Summer	0.38	−0.05	−0.22	−0.38	−0.56	0.67	0.38	0.95	0.39

3. Deep belief - back propagation neural network prediction model

By using relevant characteristics in the formation process of PM2.5 and PM10, learning the essential characteristics of the formation and the law of change can achieve the purpose of prediction. Since the causes of PM2.5 and PM10 are not very clear, and the factors leading to them are various, the factors with high correlation with PM2.5 and PM10 and historical concentrations are selected to predict the concentration of PM2.5 and PM10 in the short term. Furthermore, because the historical concentrations of CO, SO_2, NO_2, O_3, PM10, and PM2.5 are closely related to the predicted target PM2.5/PM10 but can be distinguished significantly, the prediction can be completed accurately. Therefore, the PM2.5/PM10 prediction method based on Deep Belief neural networks (DBNs) proposed in this paper is feasible.

3.1. BP neural network

Deep belief neural network prediction model adds Back Propagation (BP) neural network into the last layer to accept the output vector of the topmost restricted Boltzmann machine and as the input of the BP neural network (Zou and Conzen, 2005), thus realizing supervised training sample data. It operates in this way because the deep belief neural network is the superposition of multilayer RBM (Restricted Boltzmann Machine). RBM training algorithm can only map the feature vector of this layer optimally. In order to ensure the optimal feature vector of the whole deep belief network prediction model, the function of the BP neural network is reflected in the whole. The prediction model is fine-tuned, and the error values are passed back to all RBM layers (Huang et al., 2006).

3.2. Activation function

The selection of the activation function is related to the final optimization result of the whole prediction model. It affects the values of the neural nodes of each layer except for the input layer. This article selects the sigmoid function as the activation function of the neural network.

The Sigmoid function is shown in Eq. (3):

$$f_{(x)} = \frac{1}{1 + e^{-x}} \tag{3}$$

The derivative of the Sigmoid function is shown in Eq. (4):

$$f'_{(x)} = f_{(x)}\left(1 - f_{(x)}\right) \tag{4}$$

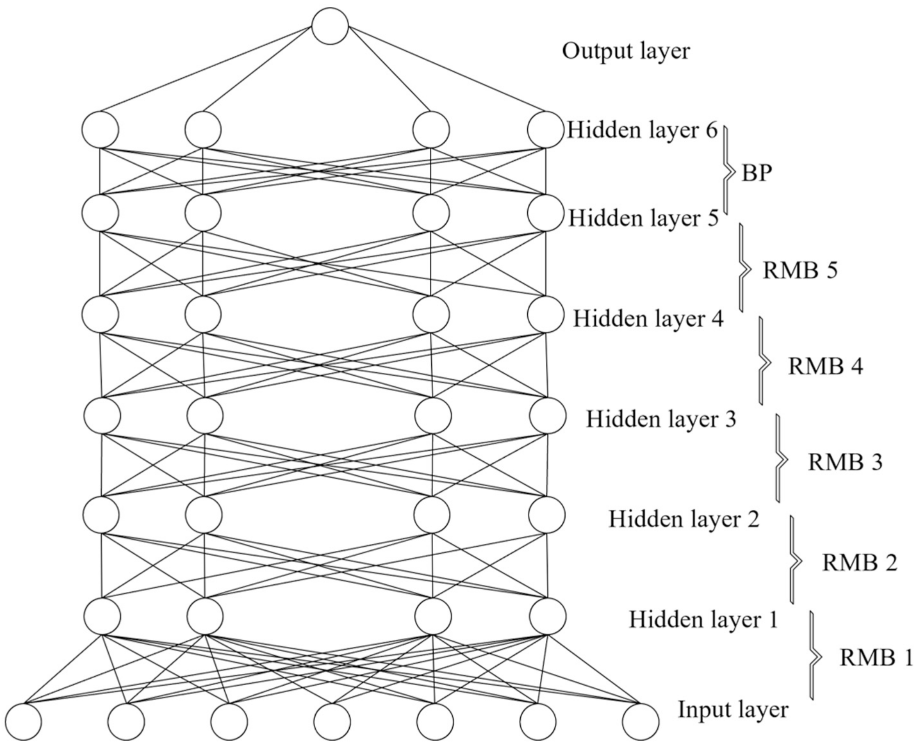

Fig. 2. The structure of the prediction model based on deep belief-BP neural network.

3.3. Deep belief-BP network prediction model structure

In this paper, the deep belief neural network prediction model is constructed by using deep learning theory to complete the construction of a multilayer deep neural network (Yang et al., 2016a). The superposition of the underlying multilayer RBM makes the characteristics of the sample data set more obvious and effective, and it can be used directly in BP neural network. Furthermore, the training provides excellent initial parameters, speeds up the model's training, and uses the supervised back-propagation algorithm to train the BP neural network to complete the parameter micro-adjustment of the haze prediction model. Therefore, multilayer RBM and multilayer BP neural networks can be complemented to each other.

The deep belief-BP neural network haze prediction model can be viewed as combining multiple unsupervised restricted Boltzmann machines and multiple supervised BP neural networks. In the research process of this article, the deep belief neural network is combined with BP neural network to predict the PM2.5/PM10 concentration in the future. The neural network input layer nodes of the deep belief neural network PM2.5/PM10 prediction model are constructed. The number of nodes in the input layer is 29, the number of nodes in the output layer is 1, and the number of neural network layers is a total of n layers (n variable), which is superimposed by a layer-limited Boltzmann machine and adds a layer BP neural network at the top layer, plus the output layer. The neural network structure diagram design is shown in Fig. 2 below.

The deep network structure composed of the stacked multilayer RBM and BP neural network in the above figure extracts the concentration characteristics of PM2.5 / PM10. The input raw data is the complete and effective time-series concentration obtained in advance. The output data is expressed as Target data, that is, the value of PM2.5/PM10 concentration predicted at the future time.

The deep network structure is composed of the multilayer RBM and BP neural network extracts the concentration characteristics of PM2.5/PM10. The input raw data is the complete effective time-series concentration obtained after pre-processing. The output data is expressed as the target data, that is, the value of the PM2.5/PM10 concentration predicted at the future time. The prediction model is used to predict the PM2.5/PM10 concentration data after 24 h.

4. Experiment

4.1. Training process

In this section, the deep belief network model training is simplified as a three-layer RBM and a layer of BP neural network. The model's training process is divided into two steps, namely the pre-training process and the fine-tuning process.

(1) Pre-training process: This process mainly uses the CD (Contrastive Divergence) algorithm to train RBM layer by layer quickly. The lowest visible layer vector of RBM is the input training sample vector. Then, the hidden layer output vector calculated by the weight matrix and bias vector can be input to the next RBM. The weight update relies on the correlation difference between the hidden layer and the visual layer vector for the visual layer. Repeat the above process, and the error will be passed down until the last layer of RBM. At this time, the unsupervised pre-training process of multilayer RBM is completed.

(2) Fine-tuning process: This process mainly uses the back-propagation algorithm to fine-tune the deep belief neural network prediction model, and the model converges to the global optimum. Because of the existence of the BP neural network, the output vector of the last top RBM is passed as input to the multilayer BP neural network. The output of the BP neural network is compared with the marker data, the prediction model error is calculated, and the back-propagation algorithm is used to correct

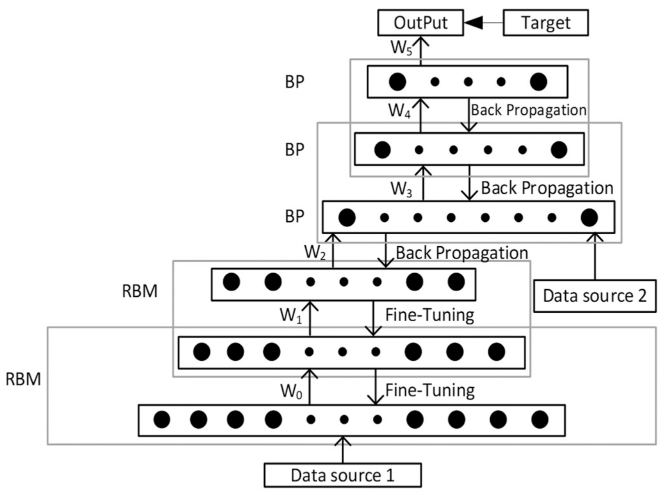

Fig. 3. The training process of the prediction model based on Deep belief-BP neural network.

the error. Under the transfer, while adjusting the network parameter values, complete the fine-tuning process of the deep belief neural network prediction model.

The training process of the prediction model is shown in Fig. 3:

4.2. Determination of a hidden layer of neural network

The network structure of the deep belief-BP neural network smog prediction model determines the ability to solve practical problems (Shao, 2009). It reflects the network's ability to express complex functions. The number of input and output layer neurons in the prediction model is related to the problem to be solved and is thus determined. For example, in the study of this article, when PM2.5 concentrations are predicted, the input is the PM2.5 concentrations in the past 24 h. The value and the PM10, O_3, CO, NO_2, and SO_2 concentration values at the current time are output as the PM2.5 concentration value at the time to be predicted. Therefore, the number of neurons in the input layer is 29, and the number of neurons in the output layer is 1, prediction. The same is true for the PM10 concentration, except that the PM10 in the input data is changed to PM2.5. In constructing the deep belief-BP neural network haze prediction model, the trouble is determining the number of hidden layers and the number of neurons in each hidden layer unit. However, there is no complete maturity theory in the current research process that can be applied to this.

4.2.1. Determination of the number of hidden layers

The prediction accuracy of the deep belief-BP neural network prediction model can theoretically increase the number of hidden layers, but increasing the number of hidden layers increases the complexity of the prediction model and lengthens the training time, and may lead to over-fitting. Therefore, in the research process of this paper, the maximum number of layers in the hidden layer is chosen to be 7. The number of neurons in each hidden layer is not particularly large to ensure the depth of the neural network while avoiding computational complexity.

4.2.2. Determination of the number of hidden layer neurons

The number of hidden layer nodes is also significant for the deep belief-BP neural network prediction model. Too few hidden layer nodes may lose some of the characteristics of the training data set, thus affecting the prediction effect. However, too many hidden layer nodes will prolong the network training time. Unfortunately, in the application of practical problems, there is no complete theory for temporarily determining the number of hidden layer nodes, and most of them are determined according to the size and complexity of the training data set. In the study of this article, the number of hidden layer nodes is determined according to the empirical formula: m represents the number of input layer nodes, and n represents the number of output layer nodes, which is a positive constant less than 10, and l represents the number of hidden layer neurons. Because the output of this paper is divided into two parts, the input of the multilayer RBM is 24, and the input of the multilayer BP network is the number of outputs of the multilayer RBM plus 5. Therefore, the number of hidden neurons is between 6 and 16. In this study, the prediction neural network with a different number of hidden layers is tested, and the prediction effect determines the most appropriate number.

4.3. Parameter setting

Different parameters will affect the final prediction value in the simulation experiment process based on the deep belief neural network smog prediction model. Therefore, to make the deep belief neural network more suitable for PM2.5/PM10 prediction, the following parameters are needed in the simulation experiment.

(1) Define the number of input layer nodes and output layer nodes, respectively. According to the characteristics of the deep belief neural network combined with the collected data set format, the number of input layer nodes is set to 29, the number of output layer nodes is set to 2; when the PM2.5 concentrations are predicted, the input data is PM10, O_3, CO, NO_2. The concentration of SO_2 at the n^{th} time and the PM2.5 concentrations at the first 24 h of the n time are output as the PM2.5 concentrations at the n + 6th time after the labeling; when the PM10 concentration is predicted, the input data is PM2.5, O_3, CO. The NO concentration at the n^{th} time of NO_2 and SO_2 and the PM10 concentration at the 24th time before n is output as the PM10 concentration at the n + 6th time after the mark.
(2) Set and initialize the parameters of the prediction model. This paper is based on the Python-Tensor Flow framework to realize the simulation experiment of the prediction model. The weight gradient learning rate is set to 0.01, the visible layer node offset is initialized to 0.05, and the hidden layer node is biased. The initialization is 0.1, the target error is 0.005, and the cutoff iteration is 500.

4.4. Evaluating indicator

This article mainly studies the number of hidden layers in the deep belief neural network prediction model and the influence of the number of neural units in each hidden layer on the prediction accuracy of PM2.5/PM10. The root means square error RMSE is used as the evaluation index. The formula is as Eq. (5):

$$RMSE = \sqrt{\frac{1}{m}\sum_{i}^{m}(T_i - P_i)^2} \quad (5)$$

In the above formula, i refers to the test sample point number, m refers to the total number of days to be predicted; T_i represents the actual concentration of the test sample point, which is the labeled data, the unit is μg/m^3; P_i represents the predicted concentration during the test sample point training, which is valued in μg/m^3.

In order to more accurately reflect the prediction effect of PM2.5/PM10 and to calculate the accuracy of the observation experiment, the experimental results in this article mainly select the prediction data of 360 time and the real data for comparison and select different hidden layers and hidden layers, respectively. Furthermore, the number of neurons was used to analyze the PM2.5/PM10 prediction model. Each experiment was performed five times, and the best one was selected as the experimental result for analysis. At the same time, to further explain the experimental results, the PM2.5 and PM10 concentration values are divided into 6 grades in the research process. Because the grades are more detailed, they can be applied to both PM2.5 and PM10. The specific results are shown in Table 2:

If the predicted result and the real result are at the same level, the predicted result is judged to be excellent; if the predicted result and the real result are two adjacent levels, the predicted result is judged to be acceptable; if the predicted result and the real result differ by two levels and above, the prediction result is determined to be unacceptable.

5. Results

The output of the deep belief neural network prediction model is respectively the predicted PM2.5 and PM10 concentrations, and the corresponding real PM2.5 and PM10 concentrations are shown in Fig. 4. In judging the accuracy of the prediction model, the mean square error is calculated based on this standard.

5.1. PM2.5 forecast results

In this paper, the computer processor model used to train the deep belief-BP neural network prediction model is Intel(R) Core (TM) i7–4790. The training of the full-text predictive model is performed on this computer. In this article, the input data (1) (PM2.5 concentration 24 h before time n) is set to 24, the input data (2) (Concentration of PM10, O_3, CO, NO_2, and SO_2 at the nth time) is 5, and the number of output layers is 1, which is the predicted PM2.5 or PM10 concentration value. The number of layers of the BP neural network is determined to be 4. Then the number of layers of the deep belief neural network, the number of hidden layer units, and the prediction result of PM2.5 are shown in Table 3. The experimental results compared with the BP network are shown in Fig. 5:

According to the above table in the PM2.5 prediction experiment, the deep belief-BP neural network has a hidden layer of 5 in the belief network. The number of neurons is 10, 9, 8,7, and 6. At this time, the Root Mean Square Error (RMSE) is at least 14.07. Therefore, the prediction effect is the best, as shown by the blue line above. When directly using the BP network for prediction, the number of input data is 29, and the output data is 1. The number of hidden layer neurons is 10, 9, 8, 7, and 6. The green dotted line shows the prediction results in Fig. 5.

From the experimental results in Figs. 4 and 5, we can know:

In the case of determining the number of hidden layers and the number of nodes of the BP network, and the number of hidden layer nodes of the trusted network, the prediction accuracy is related to the number of hidden layers of the trusted network. For example, when the number of hidden layers of the belief network is 5, the RMSE is at least 14.07. At this time, the prediction accuracy is the highest. However, as the number of hidden layers increases, the prediction accuracy is slightly floating, or even the RMSE increases. The reason may be that the number of hidden layers in the network increases, and the network structure is complex, which causes the local optimal solution in the network training process, resulting in over-fitting. Therefore, it can be considered that when the number of hidden layers of the belief network is 5, the deep belief-BP neural network prediction constructed in this paper is constructed. As a result, the model has the best predictive effect on PM2.5.

5.2. PM10 prediction results

The experimental results of the predicted PM10 concentration values for the Deep Belief-BP neural network are shown in Table 4:

According to Table 4 in the PM10 prediction experiment, the deep belief-BP neural network has 6 hidden layers in the belief network. The number of hidden layer neurons is 10, 9, 8, 7, 6, and 5. At this time, the RMSE is at least 23.28. The prediction effect is the best, as shown by the blue line in Fig. 6. When directly using the BP network for prediction, the number of input data is 29, and the output data is 1. The number of hidden layer neurons is 10, 9, 8, 7, 6, and 5. The green dotted line shows the prediction results in Fig. 6.

From the experimental results in Fig. 6, we can know:

Table 2
PM2.5 and PM10 Levels.

Level	1	2	3	4	5	6
Level range (μg/m^3)	0–35	36–75	76–115	116–150	151–250	>250

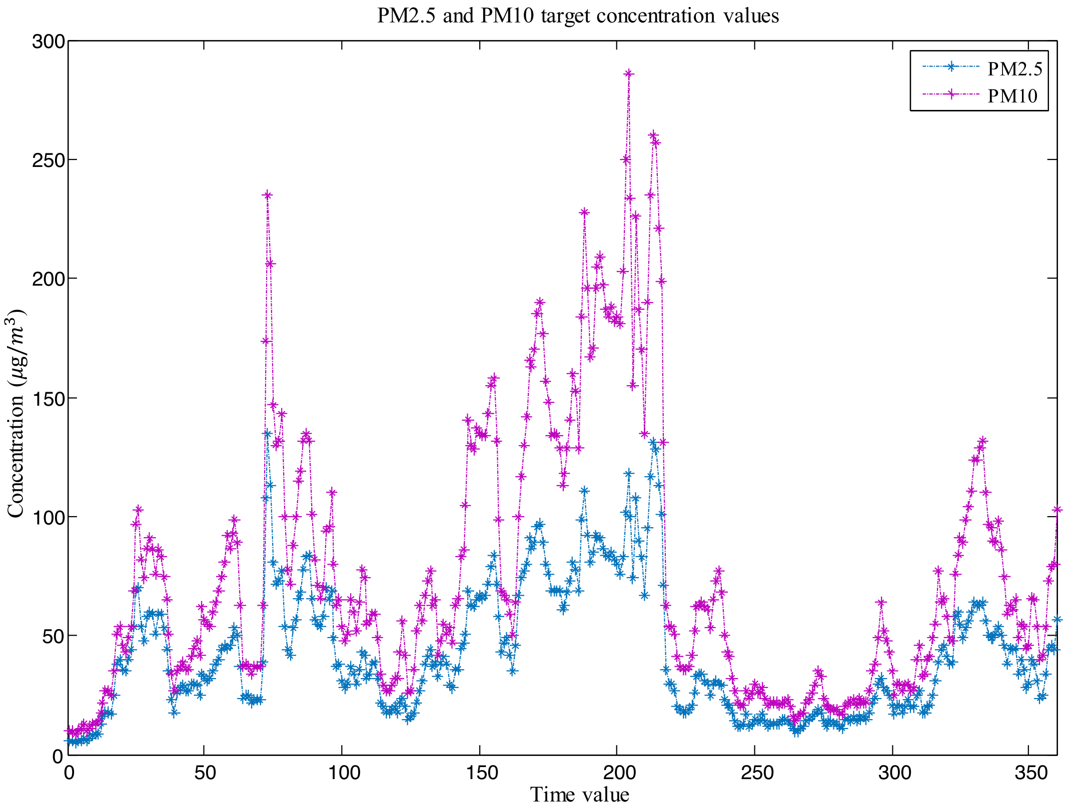

Fig. 4. Actual PM2.5 and PM10 concentration values.

Table 3
PM2.5 prediction experiment results of deep belief-BP neural network.

Number of hidden layers	Number of hidden layer units	PM2.5 RMSE (μg/m^3)	excellent	Acceptable	Unacceptable
1	10	18.20	63.89%	33.61%	2.50%
2	10 9	16.13	71.11%	27.50%	1.39%
3	10 9 8	15.49	68.61%	30.83%	0.56%
4	10 9 8 7	14.69	70.28%	28.89%	0.56%
5	10 9 8 7 6	14.07	72.50%	26.94%	0.56%
6	10 9 8 7 6 5	14.11	73.33%	26.11%	0.56%
7	10 9 8 7 6 5 4	14.08	73.61%	25.83%	0.56%
8	10 9 8 7 6 5 4 3	14.06	73.33%	26.11%	0.56%

(1) The prediction results of PM2.5 and PM10 are related to the number of hidden layers. The larger the number of hidden layers, the higher the prediction accuracy and the prediction effect of PM2.5 is higher than that of PM10, which is related to the data characteristics of PM2.5 and PM10. The concentration data of PM2.5 is lower than PM10, and the concentration range is smaller, so the RMSE is relatively small, and the precision is relatively high.
(2) For PM10, when the number of hidden layers of the belief network is 6, the RMSE is at least 23.28. At this time, the prediction accuracy is the largest. However, as the number of hidden layers continues to increase, the prediction accuracy does not continue to increase. However, the change It is not obvious that it can be considered that when the number of hidden layers is 6, the deep belief-BP neural network prediction model constructed in this paper has the best prediction effect on PM10.

6. Discussion

Based on the theory of deep learning, combining the Boltzmann machine and BP neural network, this study proposes a Deep Belief-BP neural network haze prediction model. In the Chengdu area, accurate prediction of PM2.5/PM10 concentration in a short period of time has been achieved. Then we will discuss the community contribution and future direction of this research.

In this article, our proposed model combines the two models of DBN and BP. In the community, both the deep belief network (Xie, 2017; Xing et al., 2019) model and the BP neural network (Chen, 2018; Chen et al., 2019; Ni et al., 2017) have appeared, and many researchers have optimized them. The forecasting model we built takes full advantage of the respective advantages of the two models.

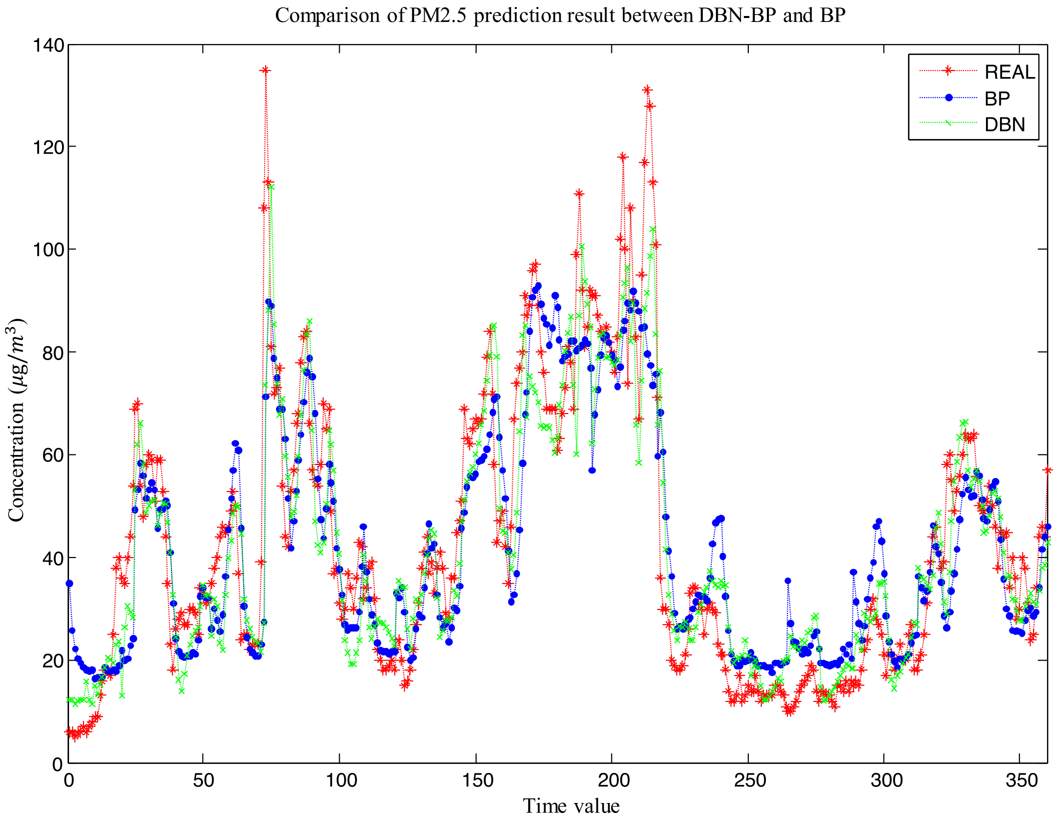

Fig. 5. Comparison of prediction results between DBN-BP and BP on PM2.5 (in 5 hidden layers).

Table 4
PM10 prediction experiment results of deep belief-BP neural network.

Number of hidden layers	Number of hidden layer units	PM2.5 RMSE(μg/m3)	excellent	Acceptable	Unacceptable
1	10	32.21	49.72%	45.28%	5.00%
2	10 9	29.25	55.83%	39.44%	4.44%
3	10 9 8	24.43	62.78%	33.89%	3.33%
4	10 9 8 7	23.49	64.44%	34.72%	0.83%
5	10 9 8 7 6	23.54	65.28%	33.33%	1.39%
6	10 9 8 7 6 5	23.28	68.33%	30.00%	1.67%
7	10 9 8 7 6 5 4	23.37	70.83%	27.78%	1.39%
8	10 9 8 7 6 5 4 3	23.32	70.56%	28.33%	1.11%

The learning process of the Boltzmann machine is unsupervised, and the learning ability is powerful. Therefore, DBN can quickly obtain the characteristics of the sample data set with high accuracy. However, the adjustment of its internal parameters is very complicated and requires manual optimization. To optimize the parameters, the function of the BP neural network is to fine-tune the entire prediction model and pass the error value back to all RBM layers. In the above cases, we realized the use of DBN to optimize the input parameters for BP; the use of BP to realize the supervised parameter adjustment of DBN. It solves the defects of the above two kinds of neural networks in independently predicting the concentration of haze. Compared with the optimization methods mentioned above, combining the two networks we proposed has better results.

For this study, we need to explore more aspects. When predicting the haze concentration, we abandon the less relevant meteorological data. In future research, we can try to add more data types to the model. Explore the combination of input data to make the predicted value more accurate. The haze forecast in this paper is only a short-term forecast, and no forecasting model is designed to be slidable. When creating the prediction model, we took into account the particularity of Chengdu's geographical environment. Although the model successfully predicts the smog in the Chengdu area from the results, its usability in other areas has not been tested yet.

7. Conclusion

Aiming at the problem of PM2.5 and PM10 concentration prediction, this study constructs a Deep Belief-BP neural network fog and

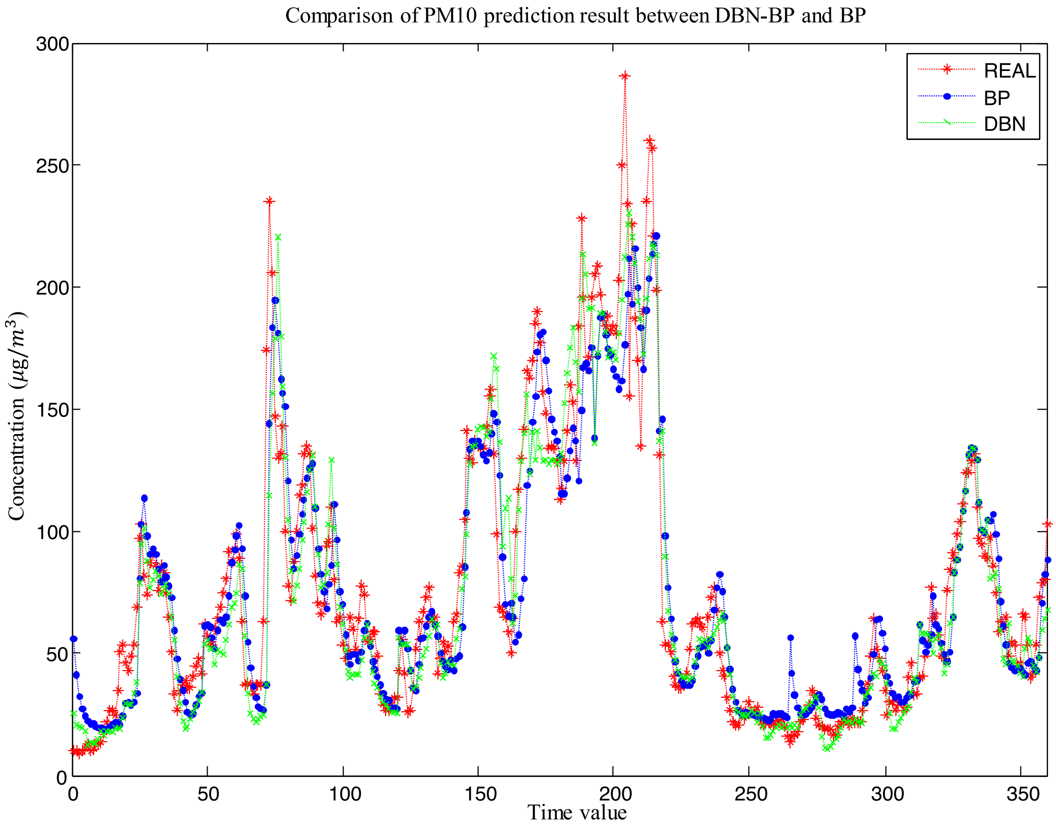

Fig. 6. Comparison of prediction results between DBN-BP and BP on PM10 (in 6 hidden layers).

haze prediction model. Using pollutant data from 6 meteorological detection stations in the Chengdu urban area as the model's input, the short-term PM2.5 and PM10 concentration values were predicted. And set up a comparative experiment with the traditional BP neural network prediction model. Due to the existence of multilayer RBM, the PM2.5/PM10 concentration characteristics of the previous 24 moments can be quickly extracted and passed to the multilayer BP neural network. It speeds up the training speed of the multilayer BP neural network and avoids the shortcomings of the BP neural network that it is easy to fall into the optimal local solution.

According to the experimental results, the DBN-PB neural network we proposed has more accurate results for haze prediction. For the setting of the number of layers of the deep neural network, we also set up a comparative experiment, analyzed, and selected the optimal value. When predicting PM2.5, DBN-BP has a better overall prediction performance. Especially when the pollution concentration is high, it is significantly improved compared with the traditional BP prediction model. It also has a good performance in predicting PM10. The autumn and winter seasons are the seasons with a high incidence of haze pollution in Chengdu, and the pollution level has been maintained above Level 3 for a long time. For this situation, our proposed DBN-BP prediction model can provide more accurate prediction results.

Funding statement

This work was jointly supported by the Sichuan Science and Technology Program (2021YFQ0003).

Ethics statement

The Ethic statement is not applicable. This study does not include any animal or human studies.

Declaration of Competing Interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Acknowledgments

The authors express their sincere appreciation and profound gratitude to research assistants Lihong Song and Weizheng Huang for their help and support in collecting and sorting the data.

References

Ambade, B., Sankar, T.K., Kumar, A., Sethi, S.S., 2020. Characterization of PAHs and n-alkanes in atmospheric aerosol of Jamshedpur city, India. J. Hazardous Toxic Radioactive Waste 24 (2), 04020003.
Ambade, B., Kumar, A., Sahu, L.K., 2021a. Characterization and health risk assessment of particulate bound polycyclic aromatic hydrocarbons (PAHs) in indoor and outdoor atmosphere of central East India. Environ. Sci. Pollut. Res. 1–12.
Ambade, B., Kurwadkar, S., Sankar, T.K., Kumar, A., 2021b. Emission reduction of black carbon and polycyclic aromatic hydrocarbons during COVID-19 pandemic lockdown. Air Qual. Atmos. Health 1–15.
Chen, Y., 2018. Prediction algorithm of PM2.5 mass concentration based on adaptive BP neural network. Computing 100 (8), 825–838.
Chen, Y., An, J., Yanhan., 2019. A novel prediction model of PM2.5 mass concentration based on back propagation neural network algorithm. J. Intell. Fuzzy Syst. 37, 3175–3183.
Chen, X., Yin, L., Fan, Y., Song, L., Ji, T., Liu, Y., Zheng, W., 2020. Temporal evolution characteristics of PM2. 5 concentration based on continuous wavelet transform. Sci. Total Environ. 699, 134244.
Friedlander, S.K., 1977. Smoke, dust and haze: fundamentals of aerosol behavior. wi.
Hagan, M.T., Demuth, H.B., Beale, M., 1997. Neural network design. PWS Publishing Co.
Huang, C.-L., Shih, H.-C., Chao, C.-Y., 2006. Semantic analysis of soccer video using dynamic Bayesian network. IEEE Trans. Multimedia 8 (4), 749–760.
Hunt, A., Abraham, J.L., Judson, B., Berry, C.L., 2003. Toxicologic and epidemiologic clues from the characterization of the 1952 London smog fine particulate matter in archival autopsy lung tissues. Environ. Health Perspect. 111 (9), 1209–1214.
Imura, H., 2013. Air pollution: Concern rising about health effects. In: Environmental Issues in China Today. Springer, pp. 87–97.
Kumar, A., Ambade, B., Sankar, T.K., Sethi, S.S., Kurwadkar, S., 2020. Source identification and health risk assessment of atmospheric PM2. 5-bound polycyclic aromatic hydrocarbons in Jamshedpur, India. Sustain. Cities Soc. 52, 101801.
Lee, H.-H., Chulakadabba, A., Tonks, A., Yang, Z., Wang, C., 2017. Predicting the occurrence of haze events in Southeast Asia using machine learning algorithms. AGUFM 2017 (PA31C-02).
Li, X., Zheng, W., Wang, D., Yin, L., Wang, Y., 2015. Predicting seismicity trend in southwest of China based on wavelet analysis. Int. J. Wavelets Multiresolution Inf. Process. 13 (02), 1550011.
Li, X., Lam, N., Qiang, Y., Li, K., Yin, L., Liu, S., Zheng, W., 2016. Measuring county resilience after the 2008 Wenchuan earthquake. Int. J. Disaster Risk Sci. 7 (4), 393–412. https://doi.org/10.1007/s13753-016-0109-2.
Li, X., Zheng, W., Lam, N., Wang, D., Yin, L., Yin, Z., 2017a. Impact of land use on urban water-logging disaster: a case study of Beijing and New York cities. Environ. Eng. Manag. J. 16 (5).
Li, X., Zheng, W., Yin, L., Yin, Z., Song, L., Tian, X., 2017b. Influence of social-economic activities on air pollutants in Beijing, China. Open Geosciences 9 (1), 314–321.
Liu, S., Wang, L., Liu, H., Su, H., Li, X., Zheng, W., 2018. Deriving bathymetry from optical images with a localized neural network algorithm. IEEE Trans. Geosci. Remote Sens. 56 (9), 5334–5342.
Ma, J., Xu, X., Zhao, C., Yan, P., 2012. A review of atmospheric chemistry research in China: photochemical smog, haze pollution, and gas-aerosol interactions. Adv. Atmos. Sci. 29 (5), 1006–1026.
Ma, Z., Zheng, W., Chen, X., Yin, L., 2021. Joint embedding VQA model based on dynamic word vector. PeerJ Comput. Sci. 7, e353.
Narayanan, S., 2002. Assessing the economic damage from Indonesian fires and the haze: a conceptual note. Singapore Econ. Rev. 47 (02), 229–241.
Ni, X.Y., Huang, H., Du, W.P., 2017. Relevance analysis and short-term prediction of PM2.5 concentrations in Beijing based on multi-source data. Atmos. Environ. 150, 146–161.
Shao, S., 2009. Application of BP neural network model in sports aerobics performance evaluation. In: Paper presented at the 2009 Pacific-Asia Conference on Knowledge Engineering and Software Engineering.
Sun, Y., Zhuang, G., Tang, A., Wang, Y., An, Z., 2006. Chemical characteristics of PM2. 5 and PM10 in haze− fog episodes in Beijing. Environ. Sci. Technol. 40 (10), 3148–3155.
Tan, J.-H., Duan, J.-C., Chen, D.-H., Wang, X.-H., Guo, S.-J., Bi, X.-H., et al., 2009. Chemical characteristics of haze during summer and winter in Guangzhou. Atmos. Res. 94 (2), 238–245.
Tang, Y., Liu, S., Deng, Y., Zhang, Y., Yin, L., Zheng, W., 2020a. Construction of force haptic reappearance system based on Geomagic touch haptic device. Comput. Methods Prog. Biomed. 190, 105344.
Tang, Y., Liu, S., Li, X., Fan, Y., Deng, Y., Liu, Y., Yin, L., 2020b. Earthquakes spatio–temporal distribution and fractal analysis in the Eurasian seismic belt. Rendiconti Lincei. Scienze Fisiche e Naturali 31 (1), 203–209.
Tang, Y., Liu, S., Deng, Y., Zhang, Y., Yin, L., Zheng, W., 2021. An improved method for soft tissue modeling. Biomed. Signal Process. Control 65, 102367.
Tiao, G., Box, G., Hamming, W., 1975. Analysis of Los Angeles photochemical smog data: a statistical overview. J. Air Pollut. Control Assoc. 25 (3), 260–268.
Wang, Y., Yao, L., Wang, L., Liu, Z., Ji, D., Tang, G., Xin, J., 2014. Mechanism for the formation of the January 2013 heavy haze pollution episode over central and eastern China. Sci. China Earth Sci. 57 (1), 14–25.
Xie, J., 2017. Deep Neural Network for PM2.5 Pollution Forecasting Based on Manifold Learning. 2017 International Conference on Sensing, Diagnostics, Prognostics, and Control (SDPC).
Xing, Y., Yue, J., Chen, C., Xiang, Y., Chen, Y., Shi, M., 2019. A deep belief network combined with modified Grey wolf optimization algorithm for PM2.5 concentration prediction. Appl. Sci. 9 (18).
Yang, X., Molchanov, P., Kautz, J., 2016a. Multilayer and multimodal fusion of deep neural networks for video classification. In: Paper Presented at the Proceedings of the 24th ACM International Conference on Multimedia.
Yang, X., Zhang, Z., Zhang, Z., Sun, L., Xu, C., Yu, L., 2016b. A long-term prediction model of Beijing haze episodes using time series analysis. Comput. Intel. Neurosci. 2016.
Yin, L., Li, X., Zheng, W., Yin, Z., Song, L., Ge, L., Zeng, Q., 2019. Fractal dimension analysis for seismicity spatial and temporal distribution in the circum-Pacific seismic belt. J. Earth Syst. Sci. 128 (1), 22.
Zhang, Y., Ma, G., Yu, F., Cao, D., 2013. Health damage assessment due to PM2. 5 exposure during haze pollution events in Beijing-Tianjin-Hebei region in January 2013. Zhonghua Yi Xue Za Zhi 93 (34), 2707–2710.
Zhang, L., Guo, X., Zhao, T., Gong, S., Xu, X., Li, Y., Zheng, Y., 2019. A modelling study of the terrain effects on haze pollution in the Sichuan Basin. Atmos. Environ. 196, 77–85.
Zheng, W., Li, X., Lam, N., Wang, X., Liu, S., Yu, X., Yao, J., 2013. Applications of integrated geophysical method in archaeological surveys of the ancient Shu ruins. J. Archaeol. Sci. 40 (1), 166–175.
Zheng, W., Li, X., Xie, J., Yin, L., Wang, Y., 2015. Impact of human activities on haze in Beijing based on grey relational analysis. Rendiconti Lincei 26 (2), 187–192.
Zheng, W., Li, X., Yin, L., Wang, Y., 2016a. The retrieved urban LST in Beijing based on TM, HJ-1B and MODIS. Arab. J. Sci. Eng. 41 (6), 2325–2332.
Zheng, W., Li, X., Yin, L., Wang, Y., 2016b. Spatiotemporal heterogeneity of urban air pollution in China based on spatial analysis. Rendiconti Lincei 27 (2), 351–356.
Zheng, W., Li, X., Yin, L., Yin, Z., Yang, B., Liu, S., Li, Y., 2017. Wavelet analysis of the temporal-spatial distribution in the Eurasia seismic belt. Int. J. Wavelets Multiresolution Inf. Process. 15 (03), 1750018.
Zheng, W., Liu, X., Ni, X., Yin, L., Yang, B., 2021a. Improving visual reasoning through semantic representation. IEEE Access 9, 91476–91486. https://doi.org/10.1109/ACCESS.2021.3074937.

Zheng, W., Yin, L., Chen, X., Ma, Z., Liu, S., Yang, B., 2021b. Knowledge base graph embedding module design for visual question answering model. Pattern Recogn. 108153 https://doi.org/10.1016/j.patcog.2021.108153.
Zheng, W., Liu, X., Yin, L., 2021c. Research on image classification method based on improved multi-scale relational network. PeerJ Comput. Sci. 7, e613 https://doi.org/10.7717/peerj-cs.613.
Zheng, W., Liu, X., Yin, L., 2021d. Sentence representation method based on multi-layer semantic network. Appl. Sci. 11 (3), 1316.
Zou, M., Conzen, S.D., 2005. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. Bioinformatics 21 (1), 71–79.