Computers, Materials & Continua CMC, vol.64, no.2, pp.1039-1049, 2020

Air Quality Prediction Based on Kohonen Clustering and ReliefF Feature Selection

Bolun Chen[1, 2], Guochang Zhu[1, *], Min Ji[1], Yongtao Yu[1], Jianyang Zhao[1] and Wei Liu[3]

Abstract: Air quality prediction is an important part of environmental governance. The accuracy of the air quality prediction also affects the planning of people's outdoor activities. How to mine effective information from historical data of air pollution and reduce unimportant factors to predict the law of pollution change is of great significance for pollution prevention, pollution control and pollution early warning. In this paper, we take into account that there are different trends in air pollutants and that different climatic factors have different effects on air pollutants. Firstly, the data of air pollutants in different cities are collected by a sliding window technology, and the data of different cities in the sliding window are clustered by Kohonen method to find the same tends in air pollutants. On this basis, combined with the weather data, we use the ReliefF method to extract the characteristics of climate factors that helpful for prediction. Finally, different types of air pollutants and corresponding extracted the characteristics of climate factors are used to train different sub models. The experimental results of different algorithms with different air pollutants show that this method not only improves the accuracy of air quality prediction, but also improves the operation efficiency.

Keywords: Air quality prediction, Kohonen clustering, ReliefF feature selection.

1 Introduction

With the continuous development of economy, environmental problems are increasingly prominent, and people's requirements for air quality are gradually improved. A previous study found that a comfortable living environment will enhance the happiness of residents and that people are even willing to use a particular portion of their income to reduce air pollution [Cuñado and De Gracia (2013)]. Severe air pollution can also affect human health [Pope 3rd, Bates and Raizenne (1995)]. Consequently, the prediction of air quality is of great significance.

---

[1] College of Computer Engineering, Huaiyin Institute of Technology, Huaian, 233003, China.

[2] Department of Physics, University of Fribourg, Fribourg, CH-1700, Switzerland.

[3] College of Information Engineering, Yangzhou University, Yangzhou, 225009, China.

[*] Corresponding Author: Guochang Zhu. Email: zhuguochang1996@163.com.

Received: 12 March 2020; Accepted: 16 April 2020.

The air quality index (AQI), an essential indicator for evaluating air quality is calculated by measuring the concentrations of carbon monoxide (CO), nitrogen dioxide (NO), ozone (O_3), PM10, PM2.5, and sulfur dioxide (SO_2) in the air. The higher the AQI, the more serious the air pollution, and vice versa. Therefore, predicting the AQI is equivalent to predicting the concentration of each of these air pollutants.

2 Related works

Initially, air quality was predicted based on researchers' knowledge of weather patterns and the local environment. Hutchison et al. [Hutchison, Smith and Faruqui (2005)] found that aerosol optical depth data from the Moderate Resolution Imaging Spectroradiometer (MODIS) are correlated with PM2.5 observations. In addition, Varghese et al. [Varghese, Langmann, Ceburnis et al. (2011)] noted that the horizontal resolution affects the prediction capabilities of regional climate/air quality models. Correspondingly, as the number of air quality monitoring points has increased, it has become possible to use a large amount of collected data to predict air quality using statistical analysis methods. For example, Cobourn [Cobourn (2010)] predicted air quality using atmospheric chemical and physical process methods to simulate the discharge, transport, diffusion, and conversion of atmospheric pollutants. Carbajal-Hernández et al. [Carbajal-Hernández, Sánchez-Fernández, Carrasco-Ochoa et al. (2012)] used a statistical index (sigma) to examine the influences of air quality parameters.

Moreover, machine learning is widely used in air quality prediction. Singh et al. [Singh, Gupta, Kumar et al. (2012)] used linear and nonlinear modeling to predict air quality. Wang et al. [Wang, Men and Lu (2008)] employed support vector machines to predict air quality. Rajput et al. [Rajput and Sharma (2017)] utilized multiple linear regression (MLR) to establish a model for predicting air pollutants in India. Nevertheless, because air quality trends are rarely linearly correlated with any influencing factors, most studies have used nonlinear models to predict the AQI. In particular, deep learning can produce an active nonlinear fit and has been applied by various researchers to predict air quality. Xia et al. [Xia, Hung and Hu (2018)] used fuzzy neural networks to predict air quality. Qi et al. [Qi, Wang, Song et al. (2018)] employed embedded feature selection and semi-supervised learning in different levels of deep learning networks and incorporated the interpolation, prediction and feature analysis of air quality into a single model.

Air quality data in the form of time series used in models with temporal properties to predict air quality effects accurately. Cheng et al. [Cheng, Shen, Zhu et al. (2018)] used a recurrent neural network (RNN) to extract information from an input sequence and measured the impact at other sites by using a fully connected layer to obtain the prediction results. Liu et al. [Liu, Yan, Li et al. (2019)] proposed an n-step recursive prediction method based on seq2seq to predict air quality. Pardo et al. [Pardo and Malpica (2017)] used a simple long short-term memory (LSTM) network to predict the air quality over the next 12 hours and 24 hours. However, air quality varies considerably depending on both the region and the climate, and thus, air quality prediction models generally exhibit poor generalizability. Accordingly, Wang et al. [Wang and Song (2018)] used Granger causality to distinguish between different weather patterns and differentiated among submodels for predicting the AQI.

3 Data and methods

The model proposed in this paper is schematically illustrated in Fig. 1. Here, air data and weather data are used as the inputs. First, the category is determined according to the air data, and then, weather factors are selected. The weather data and the air data obtained through feature selection are input into the corresponding submodel, and prediction results are produced.

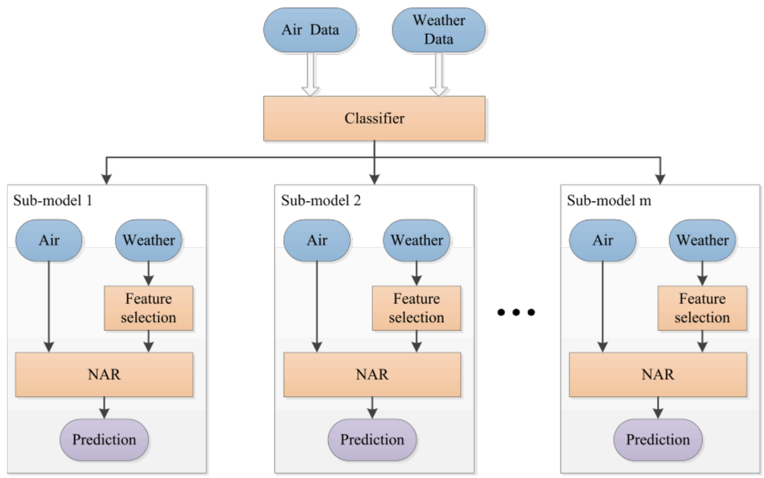

Figure 1: Algorithm model

3.1 Data

The data consist of two subsets: air data and weather data from 2019.1.16-2019.7.15 for 20 cities in Jiangsu province, China. The air data were collected from the data center of the Ministry of Ecology and Environment of China (http://datacenter.mee.gov.cn) on a daily scale. This database includes CO, NO, O_3, PM10, PM2.5, SO_2, and AQI data. air_j^i represents the air quality on the *j*-th day in the *i*-th city, where $air_j^i = \left[CO_j^i, NO_2{}_j^i, O_3{}_j^i, PM10_j^i, PM2.5_j^i, SO_2{}_j^i, AQI_j^i\right]$. The weather data were obtained from the China Weather Data Network (http://data.cma.cn). These data are collected on an hourly scale and include air pressure, wind speed, temperature, precipitation and other data. The average daily values were calculated to obtain the weather w_j^i on the *j*-th day in the *i*-th city. Weather factors are important because they influence the concentrations of air pollutants. Taking the forecasting of CO as an example, the data on the *j*-th day in the *i*-th city are denoted by c_j^i, where $c_j^i = \left(w_j^i, CO_j^i\right)$. For the time series, a sliding window of size *t* is used to sequentially sweep the entire data set, where each window represents *t* days of contiguous data in the *i*-th city, as follows: x_j^i, $x_j^i = (w_j^i, CO_j^i, w_{j+1}^i, CO_{j+1}^i, \dots, w_{j+t-1}^i, CO_{j+t-1}^i)$. Additionally, the CO data for the *i*-th city over *t* consecutive days starting from the *j*-th day are represented by a_j^i, where $a_j^i = (CO_j^i, CO_{j+1}^i, \dots, CO_{j+t-1}^i)$.

3.2 Kohonen clustering

In research, clustering is an important means of analyzing data [Li, Cui and Liu (2018)].

Clustering also has a certain influence on feature selection [Kanimozhi and Manjula (2018)]. The Kohonen neural network [Kohonen and Honkela (2007)] is a self-organizing competitive neural network that was proposed by Teuvo Kohonen. In this network, which is based on unsupervised learning, the weights between the winning neurons and their surrounding fields and the input layer can be modified through an iterative process. Neurons of the same type have the same weight coefficients, and the weights of different classes are different for clustering purposes. More specifically, the Kohonen neural network is a feedforward neural network with only two layers: an input layer and an output layer. The structure of the network is shown in Fig. 2.

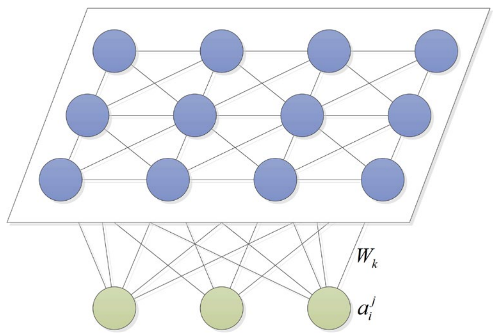

Figure 2: Kohonen network structure

The input to the Kohonen neural network is a_j^i, the weights between the input layer and the output layer are represented by W_k, the learning rate is $\delta(s) = 0.2 \times (1 - s/T)$, and the field width is $n(s) = 0.5 \times (1 - s/T)$, where the number of the current iteration is *s*, the maximum number of iterations is *T*, and $s \in [0, T]$. The distance between the input and output layers is d, where $d = ||a_j^i - W_k||$. The objective is to find the winning neuron, which is the neuron with the closest distance between the input layer and the output layer. Then, the domain width function value is calculated according to the following formula to adjust the corresponding weight:

$$\mathrm{W}_k(s+1) = \mathrm{W}_k(s) + n(s)\delta(\mathrm{s})(a_j^i - \mathrm{W}_k(s)) \tag{1}$$

Then, select the new input and repeat the process until the maximum number of iterations is reached.

3.3 ReliefF feature selection

The ReliefF [Robnik-Šikonja and Kononenko (2003)] algorithm is an improvement to the Relief algorithm that is suitable for feature weight calculations involving multiple samples. The basic concept of the ReliefF algorithm is that the characteristics of the target should be considered. The distances between a part of a sample and the rest of the sample should be as small as possible, whereas the distances from different samples should be as large as possible. Weather data have many features, some of which are less helpful than others in prediction scenarios [Hua, Chen, Yuan et al. (2019)]. If all features are selected for prediction, the operational efficiency will be reduced. Therefore, selecting the features that are most helpful for prediction without reducing the prediction accuracy is of great significance for improving the operational efficiency [Chen, Li, Zhang et al. (2018)].

First, a sample is randomly selected. Then, the q neighbours of the same type that are closest to that sample are chosen, denoted by H_k, where $H_k=\{h_1, h_2, \ldots, h_r, \ldots, h_q\}$. In addition, the q nearest neighbors M_k of different types are found, where $M_k=\{m_1, m_2, \ldots, m_r, \ldots, m_q\}$. Then, the weight of each feature f is calculated as follows:

$$W(f) = -\sum_{r=1}^{q} diff\left(f, x_j^i, h_r\right)^2 + \sum_{r=1}^{q} P_k \times diff\left(f, x_j^i, m_r\right)^2 \tag{2}$$

$$diff(f, x, y) = \begin{cases} 1 & if\ f\ is\ discrete,\ x(f) = y(f) \\ 0 & if\ f\ is\ discrete,\ x(f) \neq y(f) \\ \frac{|x(f) - y(f)|}{\max(f) - \min(f)} & if\ f\ is\ continuous \end{cases} \tag{3}$$

Here, P_k is the proportion of all samples that belong to the same category as x_j^i. Based on this calculation, the weights ranking of all features are obtained.

3.4 NAR neural network

A nonlinear autoregressive (NAR) neural network [Chow and Leung (1996)] uses its regression variables, that is, a linear combination of random variables at a given time, in the early stage to describe a nonlinear regression network of random variables at a later time. This method yields a common time series form that can be expressed as follows:

$$y(t) = b_0 + b_1 y(t-1) + b_2 y(t-2) + \cdots + b_n y(t-n) \tag{4}$$

where $y(t)$ is the output at time t and b is a coefficient. The network structure can be represented as shown in Fig. 3, where x is the input and H is the output of the hidden layer neurons.

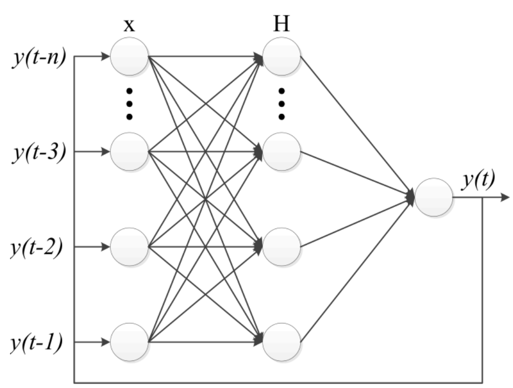

Figure 3: NAR network structure

3.5 Steps

The steps of the learning method are shown in Algorithm 1. Taking the prediction of the CO concentration as an example, the main steps of this method are as follows.

- Initialize the air data and weather data. Organize the windows of data from t consecutive days for the i-th city into x^i, where $x^i = (x_1^i, x_2^i, \ldots, x_j^i, \ldots, x_n^i)^T$. Correspondingly, sort the CO concentrations on the (j+t)-th day in the i-th city into

y^i, where $y^i = (CO^i_{1+t}, CO^i_{2+t}, \ldots, CO^i_{j+t}, \ldots, CO^i_{n+t})^T$, and sort the windows of CO concentrations on t consecutive days for the i-th city into a^i, where $a^i = (a^i_1, a^i_2, \ldots, a^i_j, \ldots, a^i_n)^T$. Then, the data windows over t consecutive days for all cities are given by X, where $X = (x^1, x^2, \ldots, x^i, \ldots, x^{20})^T$; and the CO concentrations on the (j+t)-th day in all cities are given by Y, where $Y = (y^1, y^2, \ldots, y^i, \ldots, y^{20})^T$; while the windows of CO concentrations over t consecutive days in all cities are given by A, where $A = (a^1, a^2, \ldots, a^i, \ldots, a^{20})^T$.

- First, cluster the data (lines 1-3). Apply the Kohonen network clustering to A to group the data with similar CO concentration trends into the same category; and then find the category corresponding to each window of CO data over t consecutive days starting from the j-th day in the i-th city. Accordingly, assign the category identified for city i and day j to the window of t consecutive days of data starting from the j-th day in the i-th city x^i_j and to the CO concentration on the (j+t)-th day in the i-th city CO^i_{j+t}. In this way, all the data from the same category are grouped into one class to obtain the sets of values X_k and Y_k for data of the k-th type (lines 4-7), where $X = (X_1, X_2, \ldots, X_k, \ldots, X_m)$ and $Y = (Y_1, Y_2, \ldots, Y_k, \ldots, Y_m)$.
- Next, select the features (lines 8-13). For each type of data, use the ReliefF feature selection method to obtain the feature selection ranking for that class, R_k. Delete the features of X_k in R_k with ranks higher than r that are not air pollutant concentrations to reduce the dimensionality of the input; thus, the data for the k-th class after feature selection are obtained, denoted by F_k, where $F = (F_1, F_2, \ldots, F_k, \ldots, F_m)$.
- Finally, train the learning models (lines 14-16). Use F_k as the input and Y_k as the output to train a submodel based on an NAR neural network for each type of data to obtain the final ensemble model.

Algorithm 1 Learning model.

Input: The sequence of air pollution concentration and weather features for consecutive t days, $\boldsymbol{X}$; the air pollution concentration for consecutive t days, $\boldsymbol{A}$; The sequence of predictive air quality, $\boldsymbol{Y}$;
Output: Ensemble model on the current data;
1: **for** each a^i_j **do**
2: Category(i,j)=Kohonen(a^i_j);
3: **end for**
4: **if** k= Category(i,j) **do**
5: $\boldsymbol{X_k}$=x^i_j;
6: $\boldsymbol{Y_k}$=y^i_j;
7: **end if**
8: **for** each $\boldsymbol{X_k}$ **do**
9: $\boldsymbol{R_k}$=ReliefF($\boldsymbol{X_k}$);
10: **if** $\boldsymbol{R_k(f)}$ is not air pollution concentration
11: $\boldsymbol{F_k}$=delete($\boldsymbol{X_k}$(f>r));
12: **end if**
13: **end for**
14: **for** each $\boldsymbol{F_k}$ **do**
15: Training each learning model $\boldsymbol{M_k}$ with $\boldsymbol{F_k}$ and $\boldsymbol{Y_k}$;
16: **end for**

4 Experiments

4.1 Performance metrics

An experiment was conducted to test the prediction results for the concentrations of four pollutants (CO, NO_2, O_3 and PM2.5) using four individual models and four hybrid models. The four individual models were NAR, back propagation (BP), extreme learning machine (ELM) and wavelet neural network (WaveNN) models. The four mixed models were Kohonen+ReliefF+NAR, Kohonen+ReliefF+BP, Kohonen+ReliefF+ELM, and Kohonen+ReliefF+WaveNN.

A BP neural network [Jin, Li, Wei et al. (2000)] is a type of multilayer feedforward neural network that is widely used in many research fields. The process of BP neural network analysis is divided into two stages. First, the signal is forward propagated from the input layer to the hidden layer to the output layer; then, the error is backpropagated from the output layer to the hidden layer to the input layer. During this sequential propagation through the layers, the weights and offsets of the hidden layer are adjusted in consideration of the output layer, and the weights and offsets of the input layer are adjusted in consideration of the hidden layer.

An ELM [Huang, Zhu, Siew et al. (2006)] is a feedforward artificial neural network with a single hidden layer. Unlike in a general neural network, the weights from the input layer to the hidden layer are randomly assigned following a particular distribution. Once the weights of the input layer have been determined, the weights from the hidden layer to the output layer are obtained in accordance with the least squares method to obtain the final trained model consisting of the entire network.

A WaveNN [Zhang and Benveniste (1992)] is similar to a BP neural network; the difference is that a nonlinear wavelet function is used in place of the normal nonlinear neuron excitation function. A WaveNN is a neural network model based on the wavelet transform. This combination of the wavelet transform with a neural network inherits the advantages of both methods.

In this experiment, four evaluation metrics were used, namely, the mean square error (RMSE), the root mean square error (RMSE), the mean absolute error (MAE), and the time required (expressed in seconds) [Chen, Hua, Yuan et al. (2018)]. The first three metrics are defined in the following equations:

$$\mathrm{MSE} = \frac{\sum(y-y')^2}{N} \tag{5}$$

$$\mathrm{RMSE} = \sqrt{\frac{\sum(y-y')^2}{N}} \tag{6}$$

$$\mathrm{MAE} = \frac{\sum|y-y'|}{N} \tag{7}$$

where y and y' denote the real and predicted values, respectively.

4.2 Experimental results

The experiment considered short-term forecasting, medium-term forecasting, and long-term forecasting. The data from the first three days of the forecasting period were used to predict the pollutant concentration on the fourth day, the data from the first six days were used to predict the pollutant concentration on the seventh day, and the data from the first twelve days were, used to predict the pollutant concentration on the thirteenth day. As shown in Fig. 4, for the short-term, medium-term, and long-term forecasts, the Kohonen+ReliefF+NAR model achieves the lowest error and the highest accuracy by mining the air quality characteristics, clustering data with the same pollutant concentration trends, and using different submodels for different types of data.

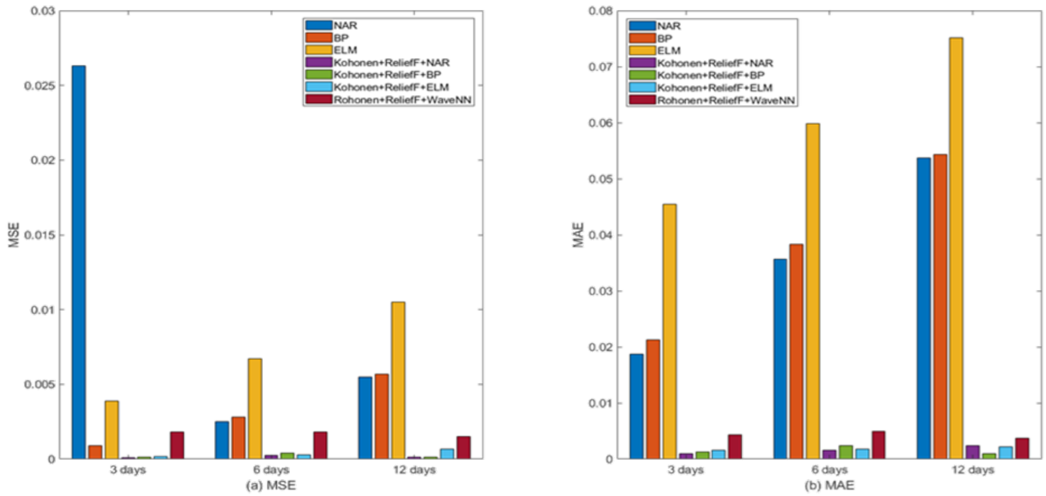

Figure 4: Comparison of the prediction accuracy of the different algorithms

By applying ReliefF feature selection to reduce the feature dimensionality, the operational efficiency can be improved. Due to the high complexity of the WaveNN model, its running time is much longer than those of the other models; therefore, it is not included in the comparison provided in Fig. 5. As shown in this figure, after implementing feature selection, the efficiency is significantly improved for all experimental cases, and the accuracy is also improved. Among the investigated models, the NAR model shows the most significant improvement in efficiency after feature selection, with a greatly reduced running time. The BP and ELM models are not as efficient, as seen from their low operational efficiencies. The running time of the ELM model is much longer than those of the other models due to its detailed calculations; therefore, the corresponding results are difficult to visualize in Fig. 5.

For the models trained individually on each category of data, as shown in Fig. 6, the prediction accuracies for each category exhibit the same trend, but overall, the NAR model performs best. For the categories with large numbers of samples, the accuracies of each model are rather similar, in contrast, for the categories with small numbers of samples, the accuracy of different models varies greatly.

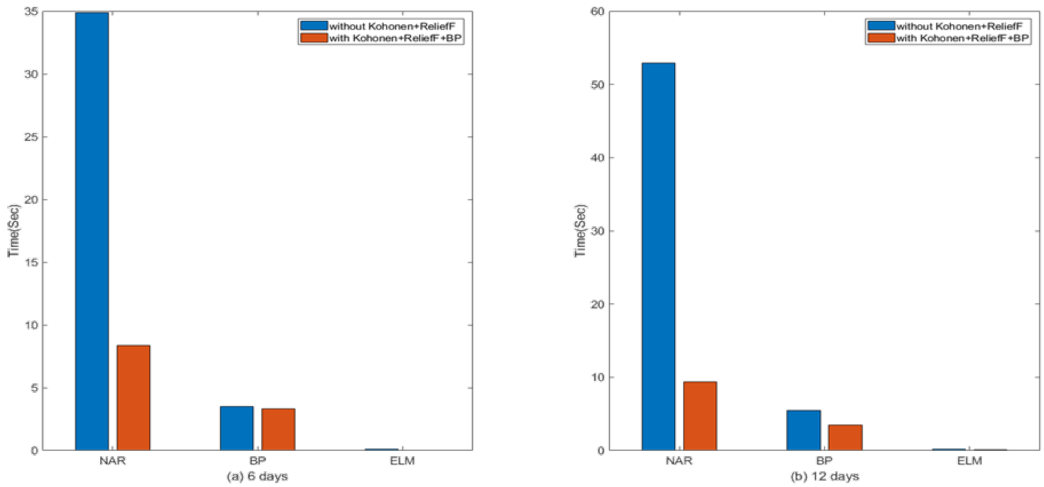

Figure 5: Comparison of run times, demonstrating the advantage of using Kohonen+ReliefF

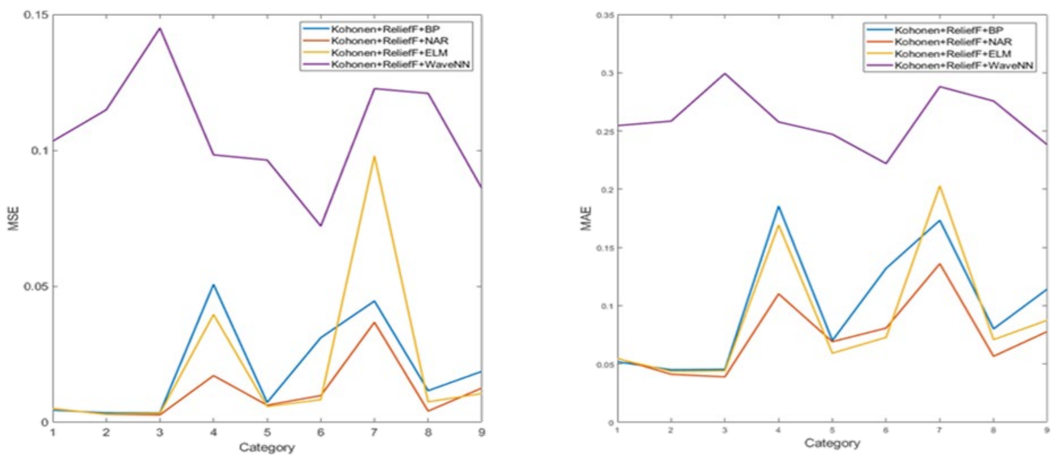

Figure 6: Comparison of the prediction accuracy of the different algorithms for different categories of data

5 Conclusion

Air quality predictions, which broadly influence human lives, are as important as weather forecasts. The internationally accepted air quality evaluation standard is the AQI, which is a unified measure of the concentrations of six pollutants in the air. However, the concentrations of air contaminants are affected by weather factors, and considering an excessive number of weather factors can reduce the operational efficiency of prediction models. Compared with traditional air quality prediction strategies, the proposed method more comprehensively considers AQI trends and characteristics, is more general in its application, and can better predict the AQI. Additionally, the ReliefF feature selection method is used to eliminate weather features that contribute little to the prediction, thereby greatly reducing the dimensionality of the input and improving the operational efficiency of forecasting.

Funding Statement: This research was supported in part by the National Natural Science Foundation of China under grant Nos. 61602202 and 61603146, the Natural Science Foundation of Jiangsu Province under contracts BK20160428 and BK20160427, the Six talent peaks project in Jiangsu Province under contract XYDXX-034 and the project in Jiangsu Association for science and technology.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

Carbajal-Hernández, J. J.; Sánchez-Fernández, L. P.; Carrasco-Ochoa, J. A.; Martínez-Trinidad, J. F. (2012): Assessment and prediction of air quality using fuzzy logic and autoregressive models. *Atmospheric Environment*, vol. 60, pp. 37-50.

Chen, B. L.; Li, F. F.; Zhang, Y. J.; Ma, J. L. (2018): Information filtering in evolving online networks. *Physics Letters A*, vol. 382, no. 5, pp. 265-271.

Chen, B. L.; Hua, Y.; Yuan, Y.; Jin, Y. (2018): Link prediction on directed networks based on AUC optimization. *IEEE Access*, vol. 6, pp. 28122-28136.

Cheng, W.; Shen, Y.; Zhu, Y.; Huang, L. (2018): A neural attention model for urban air quality inference: learning the weights of monitoring stations. *Thirty-Second Association for the Advancement of Artificial Intelligence Conference on Artificial Intelligence*, pp. 2151-2158.

Chow, T. W. S.; Leung, C. T. (1996): Neural network based short-term load forecasting using weather compensation. *IEEE Transactions on Power Systems*, vol. 11, no. 4, pp. 1736-1742.

Cobourn, W. G. (2010): An enhanced PM2.5 air quality forecast model based on nonlinear regression and back-trajectory concentrations. *Atmospheric Environment*, vol. 44, no. 25, pp. 3015-3023.

Cuñado, J.; De Gracia, F. P. (2013): Environment and happiness: new evidence for Spain. *Social Indicators Research*, vol. 112, no. 3, pp. 549-567.

Hua, Y.; Chen, B. L.; Yuan, Y.; Zhu, G. C.; Ma, J. L. (2019): An influence maximization algorithm based on the mixed importance of nodes. *Computers, Materials & Continua*, vol. 59, no. 2, pp. 517-531.

Huang, G. B.; Zhu, Q. Y.; Siew, C. K. (2006): Extreme learning machine: theory and applications. *Neurocomputing*, vol. 70, no. 1-3, pp. 489-501.

Hutchison, K. D.; Smith, S.; Faruqui, S. J. (2005): Correlating modis aerosol optical thickness data with ground-based pm2.5 observations across Texas for use in a real-time air quality prediction system. *Atmospheric Environment*, vol. 39, no. 37, pp. 7190-7203.

Jin, W.; Li, Z. J.; Wei, L. S; Zhen, H. (2000): The improvements of BP neural network learning algorithm. *WCC 2000-ICSP 2000. 5th International Conference on Signal Processing Proceedings. 16th World Computer Congress, IEEE*, vol. 3, pp. 1647-1649.

Kanimozhi, U.; Manjula, D. (2018): An intelligent incremental filtering feature selection and clustering algorithm for effective classification. *Intelligent Automation and Soft Computing*, vol. 24, no. 4, pp. 701-709.

Kohonen, T.; Honkela, T. (2007): Kohonen network. *Scholarpedia*, vol. 2, no. 1, pp. 1568.

Li, S. S.; Cui, T. J.; Liu, J. (2018): Research on the clustering analysis and similarity in factor space. *Computer Systems Science and Engineering*, vol. 33, no. 5, pp. 397-404.

Liu, B.; Yan, S.; Li, J.; Qu, G.; Li, Y. et al. (2019): A sequence-to-sequence air quality predictor based on the n-step recurrent prediction. *IEEE Access*, vol. 7, pp. 43331-43345.

Pardo, E.; Malpica, N. (2017): Air quality forecasting in Madrid using long short-term memory networks. *International Work-Conference on the Interplay Between Natural and Artificial Computation*, pp. 232-239.

Pope 3rd, C. A.; Bates, D. V.; Raizenne, M. E. (1995): Health effects of particulate air pollution: time for reassessment? *Environmental Health Perspectives*, vol. 103, no. 5, pp. 472-480.

Qi, Z.; Wang, T.; Song, G.; Hu, W.; Li, X. et al. (2018): Deep air learning: interpolation, prediction, and feature analysis of fine-grained air quality. *IEEE Transactions on Knowledge and Data Engineering*, vol. 30, no. 12, pp. 2285-2297.

Rajput, T. S.; Sharma, N. (2017): Multivariate regression analysis of air quality index for Hyderabad city: forecasting model with hourly frequency. *International Journal of Applied Research*, vol. 3, no. 8, pp. 443-447.

Robnik-Šikonja, M.; Kononenko, I. (2003): Theoretical and empirical analysis of reliefF and rreliefF. *Machine Learning*, vol. 53, no. 1-2, pp. 23-69.

Singh, K. P.; Gupta, S.; Kumar, A.; Shukla, S. P. (2012): Linear and nonlinear modeling approaches for urban air quality prediction. *Science of the Total Environment*, vol. 426, pp. 244-255.

Varghese, S.; Langmann, B.; Ceburnis, D.; O'Dowd, C. D. (2011): Effect of horizontal resolution on meteorology and air-quality prediction with a regional scale model. *Atmospheric Research*, vol. 101, no. 3, pp. 574-594.

Wang, J.; Song, G. (2018): A deep spatial-temporal ensemble model for air quality prediction. *Neurocomputing*, vol. 314, pp. 198-206.

Wang, W.; Men, C.; Lu, W. (2008): Online prediction model based on support vector machine. *Neurocomputing*, vol. 71, no. 4-6, pp. 550-558.

Xia, Y.; Hung, M. H.; Hu, R. (2018): Performance prediction of air-conditioning systems based on fuzzy neural network. *Journal of Computers*, vol. 29, no. 2, pp. 7-20.

Zhang, Q.; Benveniste, A. (1992): Wavelet networks. *IEEE Transactions on Neural Networks*, vol. 3, no. 6, pp. 889-898.