WHICEB 2018 Proceedings

Wuhan International Conference on e-Business

Summer 6-30-2018

# A Hybrid Model to Analyze Air Pollution Spread Scales in Xi' an and Surrounding Cities

Qing Zhu
*International Business School of Shaanxi Normal University, Xi'an, 710000, China;Institute of Cross-Process Perception and Control, Shaanxi Normal University, Xi'an, 710000, China*

Ting Wang
*International Business School of Shaanxi Normal University, Xi'an, 710000, China*

Yiqiong Wu
*International Business School of Shaanxi Normal University, Xi'an, 710000, China*

Jian Chai
*Management and Economic School of Xidian University, Xi'an, 710000, China*

# A Hybrid Model to Analyze Air Pollution Spread Scales in Xi' an and Surrounding Cities

*QingZhu[12*], TingWang[1], Yiqiong Wu[1], JianChai[3]*
[1]International Business School of Shaanxi Normal University, Xi’an, 710000, China
[2]Institute of Cross-Process Perception and Control, Shaanxi Normal University, Xi’an, 710000, China
[3]Management and Economic School of Xidian University, Xi’an, 710000, China

**Abstract:** Air quality analysis and prediction are very important in environmental research as airborne pollution has become a significant health threat, especially in Chinese urban agglomerations. Most previous analysis systems have been based on direct factors, such as pollutant concentrations, wind speeds and direction, relative humidity, and temperature; however, the air quality in a city is also affected by the air quality conditions in surrounding areas. This paper proposes a novel strategy for the analysis and forecast of air quality levels, for which Artificial Neural Networks (ANNs) are employed to elucidate the complex relationships between air quality and meteorological predictor variables. The experimental results in the study demonstrated that the normalized EEMD-ANN model outperformed other models in terms of the Precise, MAE and MAPE. The proposed model, therefore, demonstrated its potential as an administrative tool for issuing air pollution forecasts and for designing suitable abatement strategies.

Keywords: air quality, EEMD, ANN, data mining

## 1. INTRODUCTION

Since the beginning of the economic reforms associated with the opening-up of China in the late 1970s, China's economy has grown at a remarkable speed; however, in concert with this growth, there has been a rapid rise in air pollution and poor air quality [1] due to the increased use of fossil energy. Air pollution and particularly health damaging air pollution has raised public awareness about the effects of poor air quality in both developing and developed countries. Air quality, which in China is the second worst of 180 countries according to the Environmental Performance Index (EPI) [2], can have severe negative effects on human health, the environment, agricultural quality, and ultimately, the economy. As there has been an increased focus on health in recent decades [3], the citizenry are becoming increasing concerned about the effects of air pollution on the quality of their daily lives [4]. In fact, particulate matter (PM) or particle pollution has been linked to asthma, cardiovascular and lung diseases, and even premature death, and airborne bacteria and fungi can act as triggers for various respiratory tract diseases and allergies [5]. The World Bank reported that in 2013, air pollution caused 5.5 million deaths worldwide, 1/10 of the total deaths that year (World Bank, 2016) [6] and an earlier study reported that in 2012, around 3.7 million people died as a result of exposure to outdoor air pollution (WHO, 2014) [7] deaths due to outdoor air pollution have been rising and are set to become the main environmental cause of premature deaths by 2050 [8]. Several environmental problems also have been found to have links with PM such as corrosion, soil pollution, and vegetation damage [9]. Global warming due to the greenhouse effect also results in long term consequences [10]. Pleijel found that the present ambient ozone $O_3$ concentrations have negatively affected wheat grain yields, and to a lesser extent, protein yields (protein mass per unit area) [11], [12]. This relationship between environmental damage and economic development has been widely studied. Voorhees et al. claimed that if the air quality in Shanghai reached the national secondary standard, the monetary value of averted deaths could be as much as 1.7 to 12 billion CNY [13]. Saidi and Hammami also found that there was a negative influence of $CO_2$ emissions on economic growth [14]; therefore, air pollution is affecting many aspects of the economy and people's lives.

It has become increasingly necessary to monitor and forecast air pollutants. In 2012, the air quality index (AQI) was proposed as the air quality indicator. Therefore, effective methods are needed to predict AQI and provide early warning systems. The AQI has been difficult to forecast because of its non-stationary chaotic original data, and because forecasting daily/hourly levels of air pollutants is difficult due to the physical and chemical processes involved in the formation, transportation, and elimination of PM [15]. Many studies have focused on air quality and pollutant prediction, but few have examined the geographical distance effect on the nearby target cities. Atakan et.al found that geographical distance influenced forecast accuracy [10]; therefore, as the air quality in a city is affected by pollutant concentrations, wind speed and direction, relative humidity, temperature and the AQI from surrounding areas, it is inefficient to take measures only in the polluted city as the geographical factors need to be added to the pollution scale prediction. If prediction accuracy can be improved, this could assist in the development of government policies for monitoring and improving air quality.

The purpose of this paper, therefore, is to propose a methodology that can measure the air pollution spread in a target city. The remainder of this paper is organized as follows. Section 2 briefly presents the related materials and methods, and in Section 3, the study region, data sources, and experimental design are described. Section 4 analyzes the data and discusses the results, and conclusions and future research directions are given in Section 5.

## 2. METHODOLOGY

In this section, Empirical mode decomposition (EEMD), Ensemble empirical mode decomposition (EEMD), Complementary ensemble empirical mode decomposition (CEEMD) and artificial neural networks (ANN) and related methods are briefly introduced.

### 2.1 Empirical mode decomposition (EMD)

As air quality has been difficult to forecast because of its non-stationary, chaotic data, many studies have proposed various techniques for forecasting air pollution indicators. Jiang et al. was able to predict $PM_{2.5}$ concentrations using a hybrid method that combined high-dimensional association rules (HDAR), a modified association framework (MAF) with temporal-spatial links, an LVQ network, and an adaptive fuzzy neural network (AFNN) [9]. Conventional methods such as the autoregressive integrated moving average (ARIMA), multiple linear models(MLR), and support vector regression(SVR) are not able to fully capture the information from initial signals [16]; however, Zhu et al. predicted the AQI using a model that combined EMD and SVR [16].

Empirical mode decomposition (EMD), which was proposed by Huang et al., is a self-adaptive signal decomposition method for complex non-linear, non-stationary time series data that can decompose a complicated signal into components of different frequencies and amplitudes called Intrinsic Mode Functions (IMFs) that contain important information from the original series [17]. IMFs have two characteristics that distinguish them from other signals:

*i)* The number of extreme values and zero-crossings must differ at most by one.

*ii)* At any point, the mean value between the upper and lower envelope is zero.

The EMD formulation is as follows.

$$x(t) = \sum_{i=1}^{N} M_i(t) + R_N \tag{1}$$

where $N$ is the number of IMFs, $x(t)$ represents the original signal, $M_i(t)$ is the $i_{th}$ IMF component, and $R_N$ is the final residual.

### 2.2 Ensemble empirical mode decomposition (EEMD)

However, EMD has been found to cause mode mixing, which can distort each IMF generated during the decomposition process. Therefore, new ensemble empirical model decomposition with added white noise has been found to resolve this problem. Ensemble empirical mode decomposition (EEMD) is an ensemble version of EMD that can decompose the copies from the original signal several times and adds different white Gaussian noise to avoid mode mixing. The formulation for EEMD is as follows.

$$x(t) = \frac{1}{I}\left(\sum_{i=1}^{I}\sum_{j=i}^{N} M_j^i(t) + R_N^i\right) \tag{2}$$

where $I$ is the number of the copies of the original signals, $M_j^i(t)$ is the $j_{th}$ component of the IMFs after $i_{th}$ Gaussian white noise $\varepsilon^i(t)$ is added, and $R_N^i$ is the $i_{th}$ residual.

### 2.3 Complementary ensemble empirical mode decomposition (CEEMD)

However, as white noise cannot be completely neutralized, some residual white noise remains in the IMF components. Complementary ensemble empirical mode decomposition (CEEMD) is a member of the empirical mode decomposition (EMD) family and can be utilized to remove the noise from original data by adding reverse signals during decomposition [2]. The basic CEEMD structure is the same as for EEMD except that a collection of independent Gaussian white noise and a complementary pair are created to cancel each other out.

$$\varepsilon^i(t) \in \{\varepsilon_+^{\frac{i}{2}}(t), \varepsilon_-^{\frac{i}{2}}(t)\} \tag{3}$$

Subject to, $\varepsilon_+^{\frac{i}{2}}(t) + \varepsilon_-^{\frac{i}{2}}(t) = 0, i \in \{1, \dots, I\}$

This paper employs all three methods to decompose the AQI series.

### 2.4 Artificial neural networks (ANN)

Artificial neural networks (ANNs) are a mathematical model that simulate the structure and function of the neural network and have been widely used in environmental processes [18], with many studies having predicted PM concentrations using ANNs as they are able to identify the correlated patterns between the input and the objective values through learning, memory, self-adaption, and intelligent processes [19]. ANNs have an input layer for receiving and initiating data, a hidden layer for information encryption and data transfer to the connected nodes in the next layer, and an output layer for the end output. The mathematical model is developed through a training step, after which a prediction step computes the output for a given set of input values using the model created in the training. Chellali et al. proved that ANNs were capable of faster execution speeds and were able to easily work with high-dimension data and capture the complex relationships between the pollutant and with good predictive results [15]. Figure 1 represents the process, the core formulation for which is as follows.

$$h_{w,b}(x) = f(W^T x) = f\left(\sum_{i=1}^{n} W_i x_i + b\right) \tag{4}$$

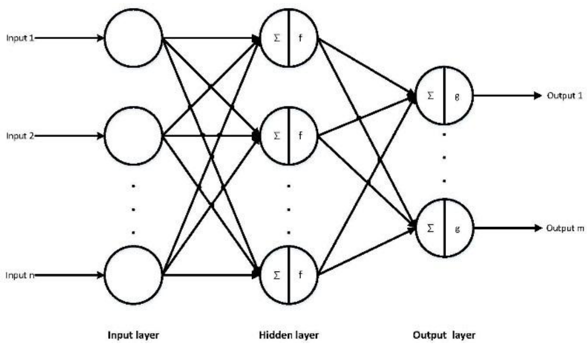



**Figure 1. ANN**

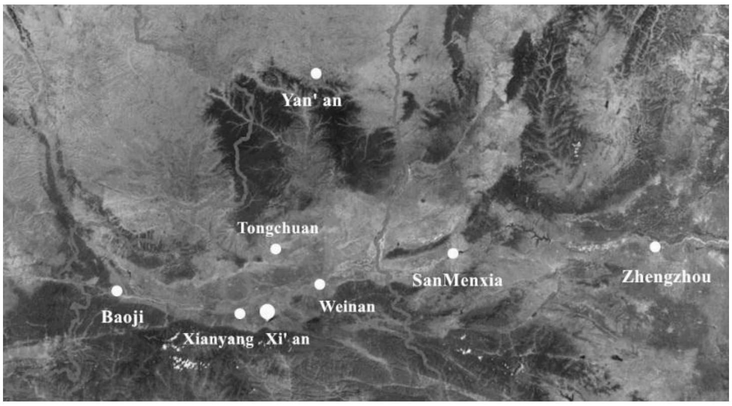
where $f$, $x_i$ denote the active functions in the total $n$ nodes in the hidden layers and input variables, $W_i$ is the weight between the two neurons, and $b$ reveals the bias.

Artificial neural networks (ANN) have obvious advantages for air pollutant concentration predictions especially in the non-linear systems traditional models are unable to handle [15]. Chellali et al. and Hurst et al. predicted $PM_{10}$ using an ANN model and proved its success [15], [20].

## 3. EXPERIMENTAL DESIGN

### 3.1 Study domains

The general aim of this research is to forecast air quality levels and study the influence on the target city of the air quality from peripheral cities. To do this, eight typical cities in the north of the Qinling mountains; Xi' an, Xianyang, Weinan, Tongcuan, Baoji, SanMenxia, Yan' an and Zhengzhou; were selected as the study sites; all of which have similar geographical locations. The geographical distributions are shown in Figure 2.




| Cities | Longitude | Latitude |
|---|---|---|
| Xi' an | 108. 95° E | 34. 27° N |
| Xianyang | 108. 7° E | 34. 33° N |
| Tongchuan | 109. 11° E | 35. 09° N |
| Weinan | 109. 5° E | 34. 52° N |
| Baoji | 107. 15° E | 34. 38° N |
| SanMenxia | 111. 12° E | 34. 47° N |
| Yan' an | 109. 47° E | 36. 6° N |
| Zhengzhou | 113. 4° E | 34. 46° N |

**Figure 2. Geographical distribution of the study domains**

Xi' an is one of China' s great ancient cities; however, it has been experiencing serious air quality problems because of increased industrial development and vehicle emissions. The most affective factor is Xi' an' s location under the Qinling Mountains, which prevents winds from the north. The haze from the northwestern areas accumulates above the Guanzhong Plain and is unable to disperse because of the Qinling Mountains. During winter, pollutants from the north are blown into the Xi' an region by the winter monsoon. Xianyang, Tongchuan, and Weinan are respectively located to the west, north and east of Xi' an, with Baoji and SanMenxia further to the east and Yan' an further to the north. As these cities have rapidly growing economies and rising populations because of the industrialization and urbanization, all suffer from airborne pollution. As these cities have similar geographical conditions to Xi' an, the pollutants are unable to easily disperse, worsening the regional haze. Further, increasing demand for electricity has overloaded the regional air loads. Zhengzhou, central city in Henan province is also experiencing rapid industrialization and urbanization and therefore also has a strong demand for electricity, steel, building materials and transport, all of which add to the pollutant emissions, which also exceed regional air loads. In The weather conditions in this region are characterized by high humidity and calm conditions, which prevent pollutant dispersal and exacerbate the regional haze.

These areas were chosen as the study sites for three reasons:

*i)* The air pollution is relatively serious but representative of air quality problems in China. In the air quality rankings for China's 74 major cities, Xi'an ranks in the bottom ten [2]; therefore, it is important to be able to scientifically predict air pollutant concentrations.

*ii)* As these cities are important, rapidly-developing urban agglomerations in China, they have better monitoring facilities and adequate available data.

*iii)* The climatic, meteorological, and geographical conditions in these cities are similar, making it possible to focus on the surrounding cities' influences on the target cities.

### 3.2 Data set

In each region, the historical daily AQIs were collected from the Ministry of Environmental Protection of the People’s Republic of China’s platform (http://datacenter.mep.gov.cn/). In this study, as the AQI was recently proposed, the earliest data records commence on 1 January, 2014; therefore, historical data from 1 January 2014 to 9 November 2017 in the eight cities was collected. In total there were 1403 data points collected as datasets, as shown in Figure 3. Eighty percent of the datasets in each area were used as the training and validation datasets, while the remaining 20% were the test data. Null data was represented as the mean of two data points from the two nearest days. To improve the accuracy of the results, the three AQI levels were grouped based on six original classes, in which the first level was defined as when the AQI was in a 0 to 100 interval, the second was when the AQI was in a 100 to 200 interval, and the third was when AQI was over 200 the numbers 1, 2 and 3 were used to represent these three levels.

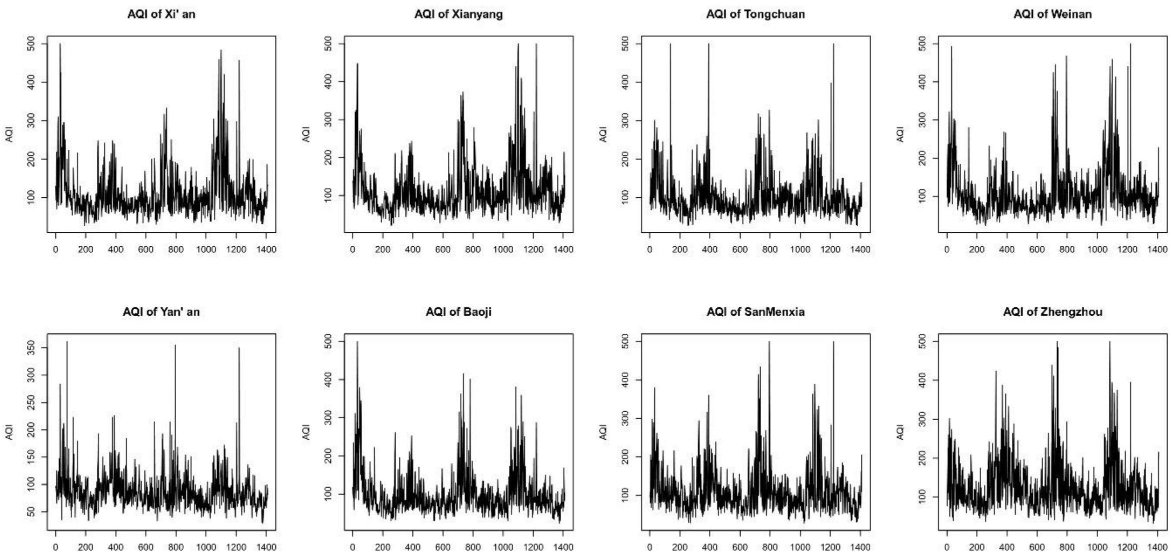



**Figure 3. Datasets**

In this experiment, Xi'an was deemed to be the center of the urban agglomerations. The study domains or datasets were divided into three groups based on their distances from the center city. The first was the nearest; X-T-W; which represented Xianyang, Tongchuan and Weinan, with the longest radius being 71.88 kilometers. The second was B-S, which represented Baoji and SanMenxia, which had the longest radius of 214.916 kilometers, and the third was Y-Z, which included Yan' an and Zhengzhou, with a 434.011 kilometer-radius.

### 3.3 Experimental setup

As this paper is focused on the air quality influences of the surrounding cities on the target city, the air quality levels each city were predicted based on the air quality conditions in nearby cities. As this paper is examining the influence in the center city of the surrounding cities, inner factors were not included in the variables. In this section, the air quality conditions in the surrounding cities are used to forecast the air quality conditions in the target city.

There were 5 stages in the experimental process:

*i)* Decomposition. The EMD, EEMD, and CEEMD were utilized to decompose the AQIs in each city, after which the obtained IMFs were divided into training datasets and test datasets and bound by the date 30 January 2017 for the training data.

*ii)* Normalization. Each IMF was decomposed using the EMD, EEMD, and CEEMD and then normalized within a 0 to 1 interval to serve as the input variables for the ANN training. The original IMFs were also put in the ANN net for comparison.

*iii)* ANN training. During the training process, the residuals were eliminated before inputting the normalized IMFs into the ANNs; however, later experiments added these residuals to the ANN for comparison. The three target city levels, 1, 2 and 3, served as the output. In the ANN model, the number of neurons in hidden layer was determined by the number of inputs and there were two hidden layers. The number of neurons in the first hidden layer was the same as the inputs, and in the second hidden layer, one third of the neurons came from the X-T-W condition and a half came from the B-S and Y-Z conditions. To ensure model accuracy and prevent over-fitting, after several trials, the ANN was trained over 1000 repetitions using the Std-Backpropagation weight-updating function to achieve dynamic learning.

*iv)* ANN testing. After training, the AQIs from the surrounding cities from 31 January 2017 to 9 December 2017 were used to test the accuracy of the proposed models trained in stage 3.

*v)* Experimental results. After decomposition, training, and testing, the results were obtained. All results were analyzed by performance comparison and several error indexes $R^2$, Precise, MAE and MAPE; were applied to measure the performances of the different models. The results are discussed in the next section.

## 4. DATA ANALYSIS AND DISCUSSION

The experimental results are presented in this section. The following three measures were used to analyze the accuracy of the air quality forecasting models; the mean absolute error (MAE), the mean absolute percentage error (MAPE) and the squared correlation coefficient ($R^2$),which were computed as shown in (5) to (7), as well as Precise, which is the ratio of the number of relevant documents retrieved to the total number of documents retrieved, to measure the accuracy of the retrieval system.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \widehat{y_i}| \tag{5}$$

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|y_i-\widehat{y_i}|}{|y_i|} \tag{6}$$

$$R^2 = \frac{[\sum_{i=1}^{n}(y_i-\bar{y})(y_i-\bar{y})]^2}{\sum_{i=1}^{n}(y_i-\bar{y})^2\sum_{i=1}^{n}(\widehat{y_i}-\bar{\hat{y}})^2} \tag{7}$$

where $y_i$ *(i=1,...,N)* is the observed value, N is the size of the test dataset, and $\widehat{y_i}$ is the $i_{th}$ predicted value.

In this section, the overall effect of the proposed strategy is discussed. The air quality in Xi'an, the target city, was forecast from 31 January 2017 to 9 December 2017 using the different models by predicting the air quality levels the surrounding cities. The specific results for $R^2$, Precise, MAE and MAPE were calculated for each model, and are shown in Table 3, which clearly shows the superiority of the proposed EEMD-ANN model.

Experiments were also conducted on the nearest X-T-W area using the different models and by adding the decomposed residuals from the original AQI data, the results of which are shown in the first to the fifth rows in Table 1. As can be seen, if the IMFs were not normalized, the $R^2$ for the EEMD-ANN and EMD-ANN had no value in the R 3.4.2 version, indicating that the models did not work. The comparisons between the groups including and not including the residuals showed that the residuals contributed to accuracy improvements in the EMD-ANN model but decreased the accuracy of the EEMD-ANN model, the forecasting from which was affected by white noise; however, in the CEEMD-ANN model, after the residuals were added, the $R^2$ was higher but the errors were also larger.

**Table 1. Model Performance. * represents the error index obtained by the models adding the residuals**

| | | EMD-ANN | | EEMD-ANN | | CEEMD-ANN | |
|---|---|---|---|---|---|---|---|
| | | Normalized | Raw | Normalized | Raw | Normalized | Raw |
| X-T-W* | R square* | 0.626 | 0.029 | 0.625 | - | 0.469 | - |
| | MAE* | 0.160 | 0.631 | 0.170 | 0.645 | 0.287 | 0.675 |
| | MAPE* | 0.090 | 0.271 | 0.159 | 0.276 | 0.275 | 0.276 |
| | Train Accuracy* | 0.876 | 0.625 | 0.862 | 0.620 | 0.867 | 0.620 |
| | Precise* | 0.840 | 0.501 | 0.830 | 0.494 | 0.713 | 0.494 |
| X-T-W | R square | 0.541 | 0.259 | **0.626** | 0.424 | 0.467 | 0.424 |
| | MAE | 0.209 | 0.383 | 0.149 | 0.238 | 0.259 | 0.259 |
| | MAPE | 0.109 | 0.301 | 0.095 | 0.168 | 0.131 | 0.156 |
| | Train Accuracy | 0.873 | 0.908 | 0.870 | 0.909 | 0.844 | 0.890 |
| | Precise | 0.791 | 0.617 | **0.851** | 0.762 | 0.741 | 0.741 |
| B-S | R square | 0.481 | 0.242 | 0.536 | 0.366 | 0.359 | 0.303 |
| | MAE | 0.284 | 0.369 | 0.184 | 0.284 | 0.284 | 0.316 |
| | MAPE | 0.232 | 0.280 | 0.134 | 0.232 | 0.147 | 0.200 |
| | Train Accuracy | 0.802 | 0.876 | 0.807 | 0.841 | 0.800 | 0.868 |
| | Precise | 0.720 | 0.667 | 0.816 | 0.723 | 0.720 | 0.684 |
| Y-Z | R square | 0.230 | 0.133 | 0.247 | 0.125 | 0.255 | 0.278 |
| | MAE | 0.351 | 0.411 | 0.348 | 0.433 | 0.340 | 0.326 |
| | MAPE | 0.223 | 0.235 | 0.195 | 0.246 | 0.196 | 0.200 |
| | Train Accuracy | 0.738 | 0.809 | 0.768 | 0.821 | 0.744 | 0.789 |
| | Precise | 0.652 | 0.610 | 0.652 | 0.582 | 0.670 | 0.677 |

After eliminating the residuals from the EMD-ANN model for the normalized X-T-W, the highest $R^2$ and Precise at 0.541 and 0.791 and the lowest MAE and MAPE at 0.209 and 0.109 were achieved. In the EEMD-ANN model, the MAE and MAPE for the normalized X-T-W were the lowest of all models at 0.149 and 0.095, and the $R^2$ and Precise were the highest at 0.626 and 0.851, which indicated that the normalized EEMD-ANN model was able to predict relatively accurate air quality levels in Xi' an and could be used by meteorological departments to predict haze and pollution levels. From the two models, X-T-W condition, were more accurate; and for the CEEMD-ANN model, the B-S condition was found to be the same as the previous model. For the X-T-W condition, the MAE and Precise for the normalized models were equal to the models not normalized and the $R^2$ was higher at 0.467 and the MAPE was lower at 0.131. In the Y-Z condition, the normalized model' s Precise were lower but the MAE was higher.

From Table 1, it can be seen that for all models, the normalized models better prevented over-fitting compared to differences between the training set accuracy and the test set accuracy and also had a lower error index and a higher $R^2$ in most cases, which indicated good model fitness and demonstrated the benefits of using normalization to process the data. Table 1 indicates that the air pollution in Xi' an influences the surrounding cities and that the further the distance, the smaller the influence.

While many studies have achieved accuracies over 90 percent, this paper only achieved accuracies over 85 percent, which may have been because this model did not consider inner variables such as wind speed or humidity. However, the predictions were found to agree with the actual measured data, and the performance of the developed normalized EEMD-ANN model was superior to the performances of the others.

From the error index analysis above, the following conclusions can be made. First, the EEMD-ANN model performed better than the EMD-ANN and CEEMD-ANN models when the data was normalized. Second, the differences in the model error values was possibly related to the changing distances; therefore, it is important to consider geographical factors when using decomposition methods. Third, it is important to select the proper model when seeking to predict different objects for different occasions.

It could be inferred that EEMD is more capable of decomposing signals according to the different physical characteristics, which means that each decomposed series denotes a sort of component with important information of the original signal. That makes the inputs of the ANN model are more representative so as to improve the accuracy compared with the ANN model used alone.

From the above analysis, it can be concluded that the air pollution regional effect increases as regional economic integration progresses. Coal generated electricity, motor vehicles, industrial processes, and dust are the main causes of regional air pollution, the severity of which depends on location, as weak wind, calm conditions and low humidity can prevent flow and cause high regional air pollution.

## 5. CONCLUSIONS AND FUTURE WORK

If properly designed and implemented, air quality forecasting model are important for air quality management and control. This work developed an air pollution spread scales measuring strategy.

Over the past few decades, China has seen a dramatic increase in airborne pollution, which has become a substantial threat to the environment and resident health. In the Chinese urban agglomerations that have intense industrialization, the air pollution problem is even worse. As diseases, environmental problems, and economic growth have been shown to be associated with air pollution problems, air pollution analysis and prediction have become an important focus in current environmental research.

This paper developed an effective method to measure air pollution spread scales that could be used by the relevant authorities to develop appropriate policies. Because air pollution in one place is affected by the air pollution in surrounding cities, it is necessary to understand the extent of the spread scales to be able to develop cooperative city government mitigation approaches. In this paper, an accurate model was developed; the normalized EEMD-ANN model; to forecast air quality in Xi' an from air pollution information from the surrounding cities.

From a monitoring and modeling point of view, the strategy proposed in this paper has three important applications

*i)* First the normalized EEMD-ANN model outperformed all others models considered in this paper.

*ii)* Second, the proposed models were able to determine the air pollution spread scales (Figure 4) and obtain a valid relationship between the surrounding cities.

*iii)* Finally, this model proved that it could be used by city planners to develop cooperative cross-city policies to manage regional air pollution and reduce the air pollution problems associated with regional economic integration.

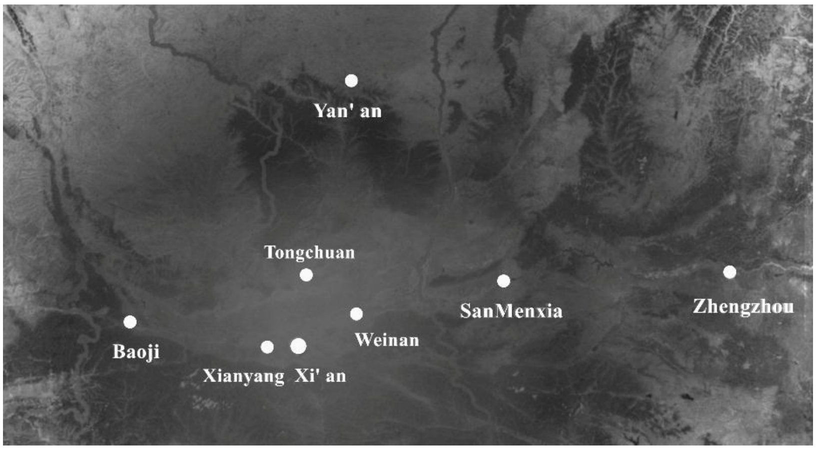



**Figure 4. Spread Scales**

At the same time, the proposed hybrid model called the normalized EEMD-ANN model can also be used in other conditions as the AQI is applied popularly in almost all cities of China except some small towns. Therefore, although different regions possess various datasets, the model is proper to analyze and predict the air quality considering the strong learning ability of this method.

As pollution content and the long-term effects of regional air pollution need to be considered when managing air quality, in the future, the interactions between the pollutants from the different areas and the long-term effects need to be assessed so as to develop more focused regional air pollution abatement plans.

## ACKNOWLEDGEMENT

This research was supported by the Science and Technology Department of Shaanxi Province 2011KRM108, the Chinese Ministry of Science and Technology 2012GXS2D027, the Shaanxi Land Construction Group SXDC 201208.

## REFERENCES

[1] Lu Z, Chen H, Hao Y, Wang J, Song X, Meng M T. (2017). The dynamic relationship between environmental pollution, economic development and public health: evidence from china. Journal of Cleaner Production, 166: 134-147.

[2] Yang Z, Wang J. (2017). A new air quality monitoring and early warning system: air quality assessment and air pollutant concentration prediction. Environmental Research, 158: 105–117.

[3] Deng Z, Liu S. (2017). Understanding consumer health information-seeking behavior from the perspective of the risk perception attitude framework and social support in mobile social media websites. International Journal of Medical Informatics, 105: 98–109

[4] X Zhang X, Liu S, Deng Z, Chen X. (2017). Knowledge sharing motivations in online health communities: a comparative study of health professionals and normal users. Computers in Human Behavior, 75: 797-810.

[5] Jahne M A, Rogers S W, Holsen T M, Grimberg S J, Ramler I P. (2015). Emission and dispersion of bioaerosols from dairy manure application sites: human health risk assessment. Environmental Science & Technology, 49(16): 9842–9849

[6] Ministry of Environmental Protection of the People's Republic of China, (2007). Action Plan on Environment and Health (2007-2015). www.china.org.cn/english/environment/238275.htm (Accessed 25 July 17)

[7] WHO, 2014. 4. Burden of disease from the joint effects of household and ambient air pollution for 2012. Report WHO (3 p)

[8] Cariolet J M, Colombert M, Vuillet M, Diab Y. (2018). Assessing the resilience of urban areas to traffic-related air pollution: Application in Greater Paris. Science of the Total Environment, 615: 588-596.

[9] Jiang P, Dong Q, Li P. (2017). A novel hybrid strategy for pm2.5 concentration analysis and prediction. Journal of Environmental Management, 196: 443-457.

[10] Kurt A, Oktay A B. (2010). Forecasting air pollutant indicator levels with geographic models 3days in advance using neural networks. Expert Systems with Applications, 37(12): 7986-7992.

[11] Broberg M C, Feng Z, Xin Y, Pleijel H. (2015). Ozone effects on wheat grain quality - a summary. Environmental Pollution, 197(1): 203-213.

[12] Grünhage L, Pleijel H, Mills G, Bender J, Danielsson H, Lehmann Y, et al. (2012). Updated stomatal flux and flux-effect models for wheat for quantifying effects of ozone on grain yield, grain mass and protein yield. Environmental Pollution, 165(6): 147-157

[13] Voorhees A S, Wang J, Wang C, Zhao B, Wang S, Kan H. (2014). Public health benefits of reducing air pollution in shanghai: a proof-of-concept methodology with application to benmap. Science of the Total Environment, 485-486(1): 396–405

[14] Saidi K, Hammami S. (2015). The impact of energy consumption and CO2, emissions on economic growth: fresh evidence from dynamic simultaneous-equations models. Sustainable Cities & Society, 14(1): 178-186.

[15] Chellali M R, Abderrahim H, Hamou A, Nebatti A, Janovec J. (2016). Artificial neural network models for prediction of daily fine particulate matter concentrations in algiers. Environmental Science & Pollution Research, 23(14): 14008-14017..

[16] Zhu S, Lian X, Liu H, Hu J, Wang Y, Che J. (2017). Daily air quality index forecasting with hybrid models: a case in china. Environmental Pollution. 231: 1232-1244

[17] Furlaneto D C, Oliveira L S, Menotti D, Cavalcanti G D C. (2017). Bias effect on predicting market trends with emd. Expert Systems with Applications, 82(C): 19-26.

[18] Park S, Kim M, Kim M, Namgung H G, Kim K T, Cho K H, et al. (2017). Predicting pm10 concentration in seoul metropolitan subway stations using artificial neural network (ANN). Journal of Hazardous Materials, 341: 75–82.

[19] Shi B, Wang P, Jiang J, Liu R. (2017). Applying high-frequency surrogate measurements and a wavelet-ANN model to provide early warnings of rapid surface water quality anomalies. Science of the Total Environment, 610-611: 1390–1399.

[20] Hrust L, Klaić Z B, Križan J, Antonić O, Hercog P. (2009). Neural network forecasting of air pollutants hourly concentrations using optimised temporal averages of meteorological variables and pollutant concentrations. Atmospheric Environment, 43(35): 5588-5596.