Application of a novel early warning system based on fuzzy time series in urban air quality forecasting in China

Jianzhou Wang[a], Hongmin Li[a*], Haiyan Lu[b]

[a] *School of Statistics, Dongbei University of Finance and Economics, Dalian, China*

[b] *School of Software, Faculty of Engineering and Information Technology, University of Technology, Sydney, Australia*

[*] *Corresponding author. Address: School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China*

Tel.:18742082511

E-mail address: hongminli0911@126.com

Abstract

With atmospheric environmental pollution becoming increasingly serious, developing an early warning system for air quality forecasting is vital to monitoring and controlling air quality. However, considering the large fluctuations in the concentration of pollutants, most previous studies have focused on enhancing accuracy, while few have addressed the stability and uncertainty analysis, which may lead to insufficient results. Therefore, a novel early warning system based on fuzzy time series was successfully developed that includes three modules: deterministic prediction module, uncertainty analysis module, and assessment module. In this system, a hybrid model combining the fuzzy time series forecasting technique and data reprocessing approaches was constructed to forecast the major air pollutants. Moreover, an uncertainty analysis was generated to further analyze and explore the uncertainties involved in future air quality forecasting. Finally, an assessment module proved the effectiveness of the developed model. The experimental results reveal that the proposed model outperforms the comparison models and baselines, and both the accuracy and the stability of the developed system are remarkable. Therefore, fuzzy logic is a better option in air quality forecasting and the developed system will be a useful tool for analyzing and monitoring air pollution.

Key words: Hybrid pollutants forecasting model; Fuzzy time series; Interval analysis; Data preprocessing; Forecasting accuracy

1. Introduction

With increasing urbanization, industrial development, vehicle use and industrial emissions, more fossil fuels are being burned, resulting in increasing emissions of sulfur dioxide (SO_2), carbon monoxide (CO), nitrogen dioxide (NO_2), ozone and particulate matter (PM), and the side effects of economic development are being exacerbated. Air pollution is a serious detriment to the health of humans and other animals, and it is increasingly destructive to vegetation and monuments [1]. Air pollution is a significant environmental issue in many parts of the world [2], and numerous Chinese cities have suffered from serious air pollution in recent years [3-4]; among them, the Beijing-

Tianjin-Hebei (Jing-Jin-Ji) region, which is an important part of China's economy with an annually average PM2.5 concentration of 106 μg/m^3, was one of the most polluted regions in China [5]. In recent years, increasing research on air quality in the Jing-Jin-Ji regions has been undertaken [6-8]; therefore, the problem of air pollution cannot be ignored.

The atmosphere is one of the most basic elements supporting human survival; a good atmosphere is necessary for human health [9]. Moreover, the pollutants that occur from emissions, and are universal around the world, include dust, CO, SO_2, NO_2, hydrocarbons, oxides and arsenic, lead, cadmium and other Heavy metals. Even more concerning is that the pollutants can bring about numerous diseases, including lung cancer, respiratory disease, cardiovascular disease and so on; furthermore, some studies have found evidence of a relationship between exposure to air pollutants and the occurrence of numerous diseases [10-12]. Therefore, the problem of air pollution has attracted a wide range of attention from people and the government. China has never relaxed air pollution controls and has released a number of air pollution control policies to improve air quality [13]. Therefore, accurate forecasting of primary pollutant concentrations not only has practical significance but also has important policy implications for the future air quality improvement.

Although the air pollution projections are grim, this does not mean this situation is not preventable. Fortunately, many researchers have proposed many approaches to analyze, estimate and forecast the pollutant concentration data to assist decision makers in monitoring air pollutant data, which can be classified into two groups: deterministic prediction and uncertainty prediction. Deterministic prediction focuses on point forecast in the future state while the goal of a uncertainty prediction is to provide that the future state of pollutant concentration will fall in an interval defined by a confidence level [14]. In monitoring air pollution data, deterministic prediction provided the definite pollutant concentration series in the future state, which is conducive to the relevant environmental protection agencies to do a good job in air warning and formulating an air pollution control plan in a timely manner [1, 15]. Nevertheless the uncertainty prediction mainly focuses on probabilistic interval prediction and thus contains more information compared to deterministic prediction. As for uncertainty prediction, many scholars usually apply proper models for conducting deterministic forecasts and integrate the algorithm for improving the distribution fitting so that different levels of intervals are estimated with the identified distributions and deterministic forecasts. Uncertainty prediction is thereby supposed a powerful tool to find out the degree and direction of the air pollution development [16-18]. It is quite clear that uncertainty prediction, which is based on deterministic prediction, is essential to forecasting pollutant concentrations. The better the forecasting performance uncertainty prediction, the higher the accuracy of deterministic prediction [17]. Therefore, a crucial step is to select an appropriate deterministic model. With respect to deterministic models, many researchers have applied time series methods to successfully forecast pollutant concentrations, and these approaches fundamentally include statistical models, chemical transport models (CTMs), artificial intelligence models [19] and fuzzy time series forecasting methods.

Statistical approaches are famous for linear series forecasting. Among them, autoregressive (AR), ARIMA, multiple linear regression (MLR) and support vector regression (SVR) have been widely used in the prediction of pollutants. Zafra et al. applied the ARIMA model to analyze the PM_{10} concentration data and obtained good performance [20]. Wang et al. proposed a hybrid model based on the ARIMA model to forecast $PM_{2.5}$ concentrations with high accuracy [21]. However, statistical models are not suitable for long-term prediction and have their own limitations, as they cannot capture the non-linear patterns of the series [22].

CTMs, one of the most commonly used models for predicting pollutants, combined with statistical approaches can be applied to successfully forecast the PM concentrations [23]. However, at the same time, the shortcomings of the CTM model are also emerging in the application process. Stern et al. noted that there may be rather strong biases in the forecasting concentrations based on CTMs due to limited knowledge about pollutant sources and the incomplete representation of physicochemical processes [24].

In contrast, ANN models are adopted to forecast air pollutants. They can overcome the limitations of conventional models which can only deal with the linear problem based on hypothesis. Bai et al proposed that BPNN with wavelet transform model can significantly improve forecasting accuracy of daily air pollutants concentration [25]. Li et al put forward a novel long short-term memory neural network extended model that inherently considers spatiotemporal correlations for air pollutant concentration prediction and presents superior performance [26]. ANN models are also successfully applied in other fields such as: wind speed forecasting [27], electrical power system forecasting [28], oil price forecasting [29] and so on. Based on the above analysis, the ANN models, with the advantage of high forecasting accuracy in nonlinear series forecasting, require fewer assumptions and requirements for data series. However, many drawbacks may also occur with ANN models. For instance, owing to potential convergence to a local minimum and over-fitting, they may have insufficient accuracy [30].

Nevertheless, while the time series forecasting techniques mentioned above are widely used in the prediction of air pollutant concentrations, they also have unavoidable limitations, such as the following: a lack of knowledge of the data resources, uncertainty, vagueness, huge volatility in the data and so on. Fortunately, the fuzzy time series (FTS) forecasting technique first developed by Zadeh [31] can be successfully applied to forecasting when handling data series with imprecise and unidentifiable trends [32]. Jana et al found that it will get satisfying results when dealing with random variables with a certain probability distribution in a fuzzy environment [33]. Furthermore, several FTS forecasting approaches developed based on ANNs perform better than traditional FTS forecasting approaches such as the ensembles of prediction Models [34-35]. For high order fuzzy time series forecasting, the model based on fuzzy logic relations shows satisfactory forecasting results [36-37]. Moreover, an adaptive fuzzy inference system (ANFIS) has also been employed for forecasting fields [38-39]. In recent years, fuzzy logic showed significant advantages in air pollution prediction. D. Domańska proposed

a novel approach based fuzzy logic relations with high accuracy in pollutant concentration forecasting [40]. Nevin et al developed a fuzzy time series model based on robust clustering which can successfully deal with outliers and abnormal observations embed in air pollution [41]. On the other hand, by summarizing the literature, fuzzy time series forecasting mainly had the following three major drawbacks: (i) a lack of reliable interval lengths [42]; (ii) an excess of linguistic values [43]; and (iii) intervals that were set too short, which can result in some null sets [44]. Therefore, to optimize prediction methods for fuzzy time series, some authors applied an optimization algorithm to combine with FTS forecasting methods, which can overcome the shortcomings mentioned above to a certain degree, such as genetic algorithms [45], fuzzy C-means clustering [46], particle swarm optimization [47] and entropy-based discretization (EBD) [48].

The hybrid model is a widely used model and has the characteristics of high prediction accuracy and stability compared with the single model [49-51]. Hybrid models integrate superiority and overcome the drawbacks of single models by integrating two or more single models. In this way, considering that the pollutant concentration data is highly unstable and stochastic, data preprocessing is a crucial step to improving the forecasting accuracy in hybrid models. In recent years, a great many preprocessing methods have been used to address time series. Babu et al. proposed a fault classification algorithm based on empirical mode decomposition (EMD) [52]. In addition, Zhang et al. developed a new multidimensional k-nearest neighbor model based on the ensemble empirical mode decomposition (EEMD) method [53]. Furthermore, complementary ensemble empirical mode decomposition (CEEMD), improved the EMD and EEMD, which not only avoided the phenomenon of mode-mixing in the process of decomposition but was also capable of effectively removing the residual noise. Niu et al. found that CEEMD served as a decomposition method with good performance in data preprocessing [54].

Based on the above analysis, some drawbacks of the models discussed in previous studies can be summarized as follows: (1) single models have many shortcomings; for instance, statistical models forecast the linear series well but cannot address nonlinear series satisfactorily; ANN models can forecast highly nonlinear time series accurately, whereas it is easy to fall into over-fitting and a local minimum. Another major drawback is that a single model never cares about the significance of data preprocessing, thus it cannot satisfy the demand for time series forecasting. (2) time series forecasting technologies based on fuzzy logic in previous researches still need to be improved in partitioning discrete discourse adaptively. (3) considering the large fluctuations in the concentration of pollutants, most previous studies have focused on enhancing accuracy, while few have addressed the stability analysis, and this may lead to weak applicability. (4) researches always focus on the point forecast that ignored the uncertainty analysis about air pollutions which cannot provide sufficient and scientific early warning information.

Therefore, this paper developed a novel early warning system with both accuracy and stability. To better forecast the pollutant concentrations and evaluate the corresponding uncertainty of the forecasts, two strategies were used to conduct the

experiments: deterministic prediction and uncertainty analysis, which proved to be helpful in monitoring air quality and providing optimal advice to decision-makers. With regard to deterministic prediction, a hybrid model was proposed which combines the CEEMD and EBD algorithms to forecast three major air pollutant concentrations; furthermore, the results revealed the effectiveness of the model. For uncertainty analysis, the forecast interval was provided under several confidence levels, which should be effective for deterministic prediction. Furthermore, to verify the effectiveness of the proposed model, the assessment module was employed.

Therefore, the unique features of the early warning system and the main contributions of this study can be summarized as follows:

1）A novel early warning system, with both accuracy and stability, consisting of a deterministic prediction module, uncertainty analysis module and assessment module, was proposed.

2) A hybrid forecasting model based on fuzzy framework is developed for forecasting major pollutants. It solves the problem of poor accuracy and low stability in air pollutants forecasting. EBD algorithm is employed to partition the discrete discourse adaptively.

3) In the uncertainty analysis module, interval forecasting, which is capable to further mine and analyze the characteristics of air pollutants, is effectively implemented.

4) The proposed early warning system can also effectively assist decision makers in formulating preventive measures and provide useful guidance for people's daily lives.

The remainder of this article is organized as follows: Section 2 outlines the background and introduces the new proposed model in detail. Section 3 presents the experiments, and Section 4 analyzes the results of the experiments. The discussion is provided in Section 5, and Section 6 gives the conclusions.

2. Methodology

This section demonstrates two strategies for deterministic prediction and uncertainty analysis. The related approaches include FTS forecasting, EBD and CEEMD; these are described in brief.

2.1 Deterministic Prediction Module

This section introduced a novel hybrid model based on FTS with CEEMD decomposing technology. The basic theory components in hybrid models are described below.

2.1.1 Definition of Fuzzy Time Series

The fuzzy set theory was first proposed by Zadeh [55], and the FTS forecasting theory, which was developed by Song & Chissom [56], has a wide range of application in forecasting. The observed value of the fuzzy time series is the language value, whereas the traditional time series observation value involves real numbers, which is the most important difference between them. The general definitions of FTS are described briefly as follows [41]:

Definition 1. Define $U=\{u_1,u_2,\cdots,u_n\}$ as the universe of discourse. A fuzzy set A_i in U can be defined by its membership function:

$$A_i = \frac{f_{A_i}(u_1)}{u_1} + \frac{f_{A_i}(u_2)}{u_2} + \cdots + \frac{f_{A_i}(u_n)}{u_n} \quad (1)$$

where $f_{A_i} : U \mapsto [0,1]$ represents the membership function of the fuzzy set A_i and $f_{A_i}(u_j)$ represents the member degree of u_j to A_i.

Definition 2. Assume the time series $\{Y(t) \mid t = 0,1,\ldots\}$ is the universe which is a subset of $\mathbf{R}$, let $f_i(t)(i = 1,2,\ldots)$ be a fuzzy set in the universe $\{Y(t) \mid t = 0,1,\ldots\}$, and $F(t)$ is the set of $f_i(t)(i = 1,2,\ldots)$, then $F(t)$ is defined as a fuzzy time series on $\{Y(t) \mid t = 0,1,\ldots\}$.

Definition 3. If $F(t)$ is a fuzzy time series, then, a fuzzy relationship exists $R(t-p,t)$, such that

$$F(t) = F(t-p) \circ R(t-p,t) \quad (2)$$

where "$\circ$"is a max-min composition operator, and both $F(t)$ and $F(t-p)$ are fuzzy sets, then $F(t)$ is derived from $F(t-p)$, denoted by the fuzzy logical relation (FLR) "$F(t-p) \to F(t)$".

Definition 4. If $F(t)$ is a fuzzy time series, for $t = 0,1,2,\ldots$. If $F(t)$ is caused by $F(t-1), F(t-2), \ldots, F(t-p)$, then the *p-order* FLR can be represented by $F(t-p), F(t-p+1), \ldots, F(t-1) \to F(t)$. The relationship between $F(t)$ and $F(t-p)$ is denoted as $A_k \to A_j$, where A_k and A_j are called the left-hand side and the right hand side of the FLR, respectively. FLRs with the same left-hand side can be categorized into an ordered fuzzy logical group (FLG) [57].

Due to the advantages of fuzzy logic, it is widely applied in time series forecasting. To improve the stock index forecasts, Rubio et al proposed a new weighted fuzzy-trend time series method that proved more superior than other models [58]. Further, Stefanakos et al first applied fuzzy time series forecasting in wave field predictions which supposed to be a satisfying application for nonstationary series [59]. For wind speed series forecasting, fuzzy logic also has excellent performance [32]. Fuzzy time series forecasting also performs well in air quality forecasting, and this paper is a

successful application.

2.1.2 Entropy-based Discretization Algorithm

The EBD was developed by Shannon [60] in order to identify a set of breakpoints that can divide the original dataset into several small intervals. EBD performs better than conventional entropy-based method in label ranking problems [61] and it is also suitable for data streaming classification [62]. Therefore, it is considered as a very promising method in data identification and classification. According to Xe and Li [44, 48], the EBD algorithm can be defined by the following concepts:

Assuming $X \subseteq U$ and $|X|$ is the number of samples in X; $j(j=1,2,\ldots,k)$ is the decision attribute of X, then the information entropy of X can be defined as follows:

$$H(X) = -\sum_{j=1}^{k} p_j \log_2 p_j \tag{3}$$

$$p_j = \frac{k_j}{|X|} \tag{4}$$

The smaller the value of $H(X)$, the lower the disorder of the sequence in X. The minimum description length principle [63] is a well-known approach applied to discretize continuous attributes in classification tasks, which measures the information obtained by a given breakpoint by comparing the values of entropy before and after the partition. A breakpoint b_i^c divides X into two subsets, $l^X(b_i^c)$ denotes the numbers of samples whose decision attribute value on c is smaller than b_i^c; similarly, $r^X(b_i^c)$ denotes what decision attribute value on c is bigger than b_i^c.

$$l^X(b_i^c) = \sum_{j=1}^{k} l_j(b_i^c) \tag{5}$$

$$r^X(b_i^c) = \sum_{j=1}^{k} r_j(b_i^c) \tag{6}$$

X_l and X_r are two subsets of X, and their information entropy can be computed as follows:

$$H(X_l) = -\sum_{j=1}^{k} p_j \log_2 p_j, \quad p_j = \frac{l_j^X(b_i^c)}{l^X(b_i^c)} \tag{7}$$

$$H(X_r) = -\sum_{j=1}^{k} q_j \log_2 q_j, \quad q_j = \frac{r_j^X(b_i^c)}{r^X(b_i^c)} \tag{8}$$

Furthermore, the information entropy of b_i^c to X is

$$H^X\left(b_i^c\right)=\frac{|X_l|}{|U|}H\left(X_l\right)+\frac{|X_r|}{|U|}H\left(X_r\right) \tag{9}$$

This assumes that P and B are the set of selected breakpoints and candidate breakpoints, respectively, and $L=\{X_1,X_2,\cdots,X_m\}$ refers to an equivalent class set divided by P. Each time the candidate breakpoint is added to the selected set P, the information entropy can be calculated as follows:

$$H(b,L)=H^{X_1}(b)+H^{X_2}(b)+\cdots+H^{X_m}(b) \tag{10}$$

The initial value of H is set to $H(U)$ according to Eq.3, then the pseudo-code of the EBD algorithm is outlined as follow:

***Algorithm* 1: EBD**

Input: $L=(l(1),l(2),\cdots,l(n))$ -a sequence of pollutant concentrations

Output: $P=(p(1),p(2),\cdots,p(m))$-a set consists of breakpoints

Parameters:

P—Selected breakpoint set
B—Candidate breakpoint set
H—Information entropy
ρ—Threshold of Information entropy

```
/*Set the parameters of EBD.*/
/* Set the initial value of H.*/
/*Caculate the entropy of every point in B.*/
FOR EACH i: 1≤ i≤ n DO
    Caculate the H(b_i, L)
END FOR
/*Find the minmimum of H.*/
WHILE (H>min(b_i, L))
    Add b_min to P; Caculate the H(b_min, L)
    H=H(b_min, L); B=B-{b_min}; L={L_1,L_2};
    /*Update the entropy H.*/
    FOR EACH i=1:length(P)
        IF (min{H(b_i, L_1)}<min{H(b_i,L_2)}) THEN
            L=L_2; H=H(L_2);
        ELSE
            L=L_1; H=H(L_1);
        END IF
```

```
    END FOR
    /*Determine whether the entropy satisfies the threshold.*/
    IF (H<ρ) THEN
        Output P.
    END IF
END WHILE
RETURN P
```

2.1.3 Complementary Ensemble Empirical Mode Decomposition (CEEMD)

CEEMD performs well in decomposing unstable and nonlinear series compared with the traditional decomposition method. In fact, the traditional decomposition method concentrates on decomposing the time series using specific characteristics. In general, wavelet decomposition requires that the series be unstable, with linear characteristics, and Fourier decomposition defines the data as smooth and periodic [54]. In light of the above characteristics, complementary ensemble empirical mode decomposition is applied to preprocess the original pollution concentration series.

CEEMD, as a member of the empirical mode decomposition family, overcomes the shortcomings of EMD. The EMD method can decompose the original time series into a small and finite number of oscillating functions through the screening process. However, it easily falls into a mode-mixing phenomenon. Accordingly, Wu and Huang [64] proposed a new decomposition method that adds white noise into EMD, which can avoid the mode-mixing phenomenon and improve stability. However, some shortcomings still exist, including residual noise, time-consuming processing needs and other shortcomings. Therefore, Yeh [65] introduced CEEMD to improve the EMD and EEMD; CEEMD not only avoids the phenomenon of mode-mixing in the process of decomposition but is also capable of effectively removing residual noise, which more effectively improves the decomposition effect.

2.1.4 The Proposed Hybrid Model

Compared with previous models, the proposed model inherits the merits of the single model and improves the performance in forecasting the pollutant concentration. The following steps demonstrate the framework of the proposed model in detail, and they are also shown in Fig. 1.

Step 1 Data processing.

Due to the pollutant concentration data with great randomness and instability characteristics may lead to poor forecasting accuracy, CEEMD was applied to decompose the original pollution concentration series into several IMFs and to reconstruct the new series with the highest-frequency IMFs eliminated. In fact, CEEMD is a successful application that can eliminate the negative influence of noise and improve prediction accuracy.

Step 2 Define the universe of discourse, *U*.

Compute the maximum value $D_{\max}$ and the minimum value $D_{\min}$ of the pollution concentration data. Then, define the discrete discourse U as $\left[D_{\min}-D_1, D_{\max}+D_2\right]$,

where D_1 and D_2 are the appropriate positive real numbers. Finally, the discrete discourse is divided into several small intervals adaptively based on the EBD algorithm which can find out the best breakpoint of discourse by searching for the minimum value of information entropy in the iterative process so that the linguistic values close to a steady state belong to the fuzzy set.

Step 3 Define fuzzy sets.

Based on the sub intervals defined in step 2, fuzzy sets $A=\{A_1, A_2, \cdots, A_n\}$ can be defined as follows:

$$A_i = \frac{a_{i1}}{u_1} + \frac{a_{i2}}{u_2} + \cdots + \frac{a_{in}}{u_n} \tag{11}$$

where a_{ij} denotes the membership degree of the interval j $(1 \le j \le n)$ to the fuzzy set i, $a_{ij} \in [0,1]$, $u_i\ (i=1,2,\ldots,n)$ is the element of fuzzy set A_i. The value of a_{ij} can be defined as follows:

$$a_{ij} = \begin{cases} 1, & i = j \\ 0.5, & i = j \pm 1 \\ 0, & others \end{cases} \tag{12}$$

Step 4 Fuzzify the observed rules.

Herein, each observation will be fuzzified into a corresponding fuzzy set. The fuzzy set A_i that has the highest membership value of the defined sub-interval is determined where the membership degree is calculated as follows:

$$\mu_{A_i}(y_t) = \begin{cases} 1, & IF\ i = 1\ AND\ y_t \le m_1 \\ 1, & IF\ i = n\ AND\ y_t \ge m_n \\ \max\left\{0, 1 - \left|y_t - m_i\right| / (2 \times l_i)\right\}, & others \end{cases} \tag{13}$$

where $m_i, (i=1,2,\ldots,n)$ and $l_i, (i=1,2,\ldots,n)$ are the mid-value and the length of the ith interval, respectively.

Step 5 Establish FLRs and FLRGs.

Based on the **Definition 4,** FLRs are formed with the fuzzified observations of the pollution concentration, then the FLGs are established based on all FLRs.

Step 6 Build trend-weighted matrix.

The trend-weighted matrix, wherein each row denotes the occurrence frequency of the FLRs, is then generated for all FLRs. The trend-weighted is computed as follows:

$$w_i(t) = \left[w_1^{'}, w_2^{'}, \cdots, w_n^{'}\right] = \left[\frac{w_1}{\sum_{j=1}^{n} w_j}, \frac{w_2}{\sum_{j=1}^{n} w_j}, \cdots, \frac{w_n}{\sum_{j=1}^{n} w_j}\right] \tag{14}$$

where w_i is the weight for fuzzy set A_i, $1 \le i \le n$ and $1 \le j \le n$.

Step 7 Calculate the forecasted outputs.

In this step, the forecasted values are computed by multiplying the defuzzified matrix and weighting matrix as follows:

$$F(t) = M_{df}(t-1) \times w_i(t-1) \tag{15}$$

where $M_{df}(t-1)$ denotes the defuzzified matrix. The centroid defuzzification method is then used to derive the weighting matrix $w_i(t-1)$.

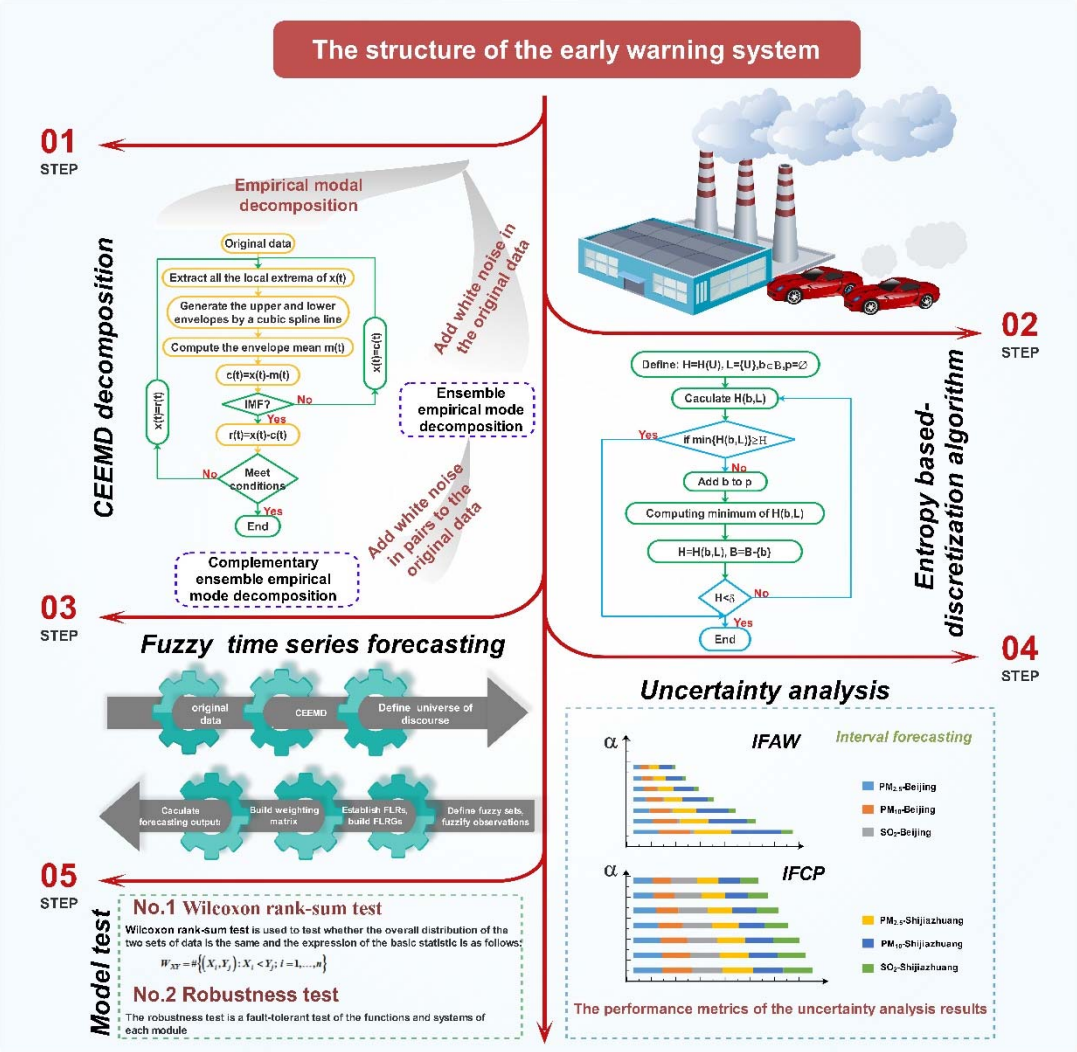

Fig.1. The general flowchart conducted in this paper

2.2 Uncertainty Analysis Module

To further forecast the pollution concentration and the uncertainty of the forecast, in this subsection, interval forecasting based on the deterministic prediction is applied to forecast the uncertainty of pollution concentrations. Interval forecasting is based on deterministic predictions and is often used to estimate the uncertainty trends of future

values [17-18, 66]. Each significance level will correspond to a forecasting interval, and the length of interval is not only related to the confidence level but also to the degree of volatility in the data. Therefore, the shorter the interval, the lower the uncertainty of the data, and the better the forecasting effect. At a confidence level α, the relationship between the confidence limit ($I_{\min}$ and $I_{\max}$) of the forecasting interval and the observed value Y_t can be expressed as follows:

$$P\left(I_{\min} \le Y_t \le I_{\min}\right) = 1 - 2\alpha \tag{16}$$

or

$$P\left\{I_{\min} \le Y_t \le I_{\max} \mid E\left(Y_t\right) = \hat{y}\right\} \square P\left\{E\left(Y_t\right) = \hat{y}\right\} = 1 - 2\alpha \tag{17}$$

2.3 Assessment Module

Although eight performance evaluation metrics are introduced above, in order to better prove the effectiveness of the model, this section introduces two kinds of testing methods: a Wilcoxon rank-sum test and a robustness test to further demonstrate the effect of the model from two aspects.

2.3.1 Wilcoxon Rank-sum Test

To verify the forecasting effectiveness of the two models and determine which is more effective, a Wilcoxon rank-sum test was applied in this study. Assuming *X* and *Y* are absolutely continuous random variables with distribution functions *F* and *G*, respectively, then, let $X_m = \left(X_1, \cdots, X_m\right)$ and $Y_n = \left(Y_1, \cdots, Y_n\right)$ be independent random samples from *F* and *G*, respectively. Assuming that the null hypothesis that *X* and *Y* are equal in distribution, the alternative hypothesis is that *Y* is stochastically strictly greater than *X*:

$$H_0 : Y \underline{d} X \quad vs. \quad H_1 : Y >_{st} X \tag{18}$$

The statistic (written in the Mann-Whitney form) is expressed as follows:

$$W_{XY} = \#\left\{\left(X_i, Y_j\right) : X_i < Y_j ; i = 1, \ldots, n\right\} \tag{19}$$

For a given level α, the Wilcoxon rank-sum test is

$$\Phi_\alpha\left(W_{XY}\right) = \begin{cases} 1, & if\ W_{XY} > w_{1-\alpha} \\ 0, & if\ W_{XY} \le w_{1-\alpha} \end{cases} \tag{20}$$

where w_q is the *q*-quantile of the null distribution of W_{XY}. The null distribution W_{XY} depends only on the sample sizes *m* and *n*. The approximated values of W_{XY} can be obtained through well-known normal approximation as follows:

$$\frac{W_{XY}-mn/2}{\left(mn\left(m+n+1\right)/12\right)^{1/2}} \square N\left(0,1\right) \tag{21}$$

2.3.2 The Robustness Test

The aim of the robustness test of the proposed model is to determine if it still performs well when the dataset has great fluctuation. A common method of robustness testing is to randomly increase or decrease the historical dataset by a few percentage points (to simulate the stochastic fluctuations of the data), then, to examine the forecasting accuracy of the model. If there is only small fluctuation of the forecasting accuracy, this indicates that the model is robust; however, the robustness of the proposed model can be denied.

3. Experimental Set Up

To better address air pollution problems and to understand the characteristics of pollutant concentrations, this section consists of three experiments. The experiments performed in our study were implemented on Matlab2016a, running on a Windows 8.1 Professional operating system. The specific hardware parameters were as follows: Intel (R) Core i5-4590 3.30 GHz CPU and 8 GB RAM.

3.1 Data Description

To verify the air quality and the effectiveness of the proposed hybrid model, two datasets from two cities (Beijing and Shijiazhuang) in the Jing-Jin-Ji region were evaluated in this study, as shown in Fig. 2. Jing-Jin-Ji is the economic center of China and has a problem of urban air pollution; the region consists of 13 cities: Beijing, Tianjin, Baoding, Tangshan, Langfang, Chengde, Zhangjiakou, Qinhuangdao, Hengshui, Cangzhou, Xingtai, Handan and Shijiazhuang [66]. There are many reasons for the air pollution caused by Beijing and Tianjin, including economic development but also geographical location. The reasons that we selected these two cities mainly include the following: (1) the air pollution problems in these two cities are notable, and they are very representative in and important for the treatment of air pollution. (2) At present, there have been many studies about air pollution in this area, and the two cities were selected to carry out comparative research.

The main air pollutants include $PM_{2.5}$, PM_{10}, SO_2, CO, NO_2, O_3 and so on; based on existing research, this study selected three pollutants ($PM_{2.5}$, PM_{10} and SO_2) that influence the air quality more significantly [17]. Furthermore, the sample data are hourly for the $PM_{2.5}$, PM_{10} and SO_2 pollutant concentrations from November 1, 2016 to July 31, 2017, and can be used to evaluate the performance of the proposed hybrid models. In this study, the sample datasets were divided into two parts: a training set and a testing set. There are 1000 observations in a testing set, and the remaining observations from the datasets were used as a training set.

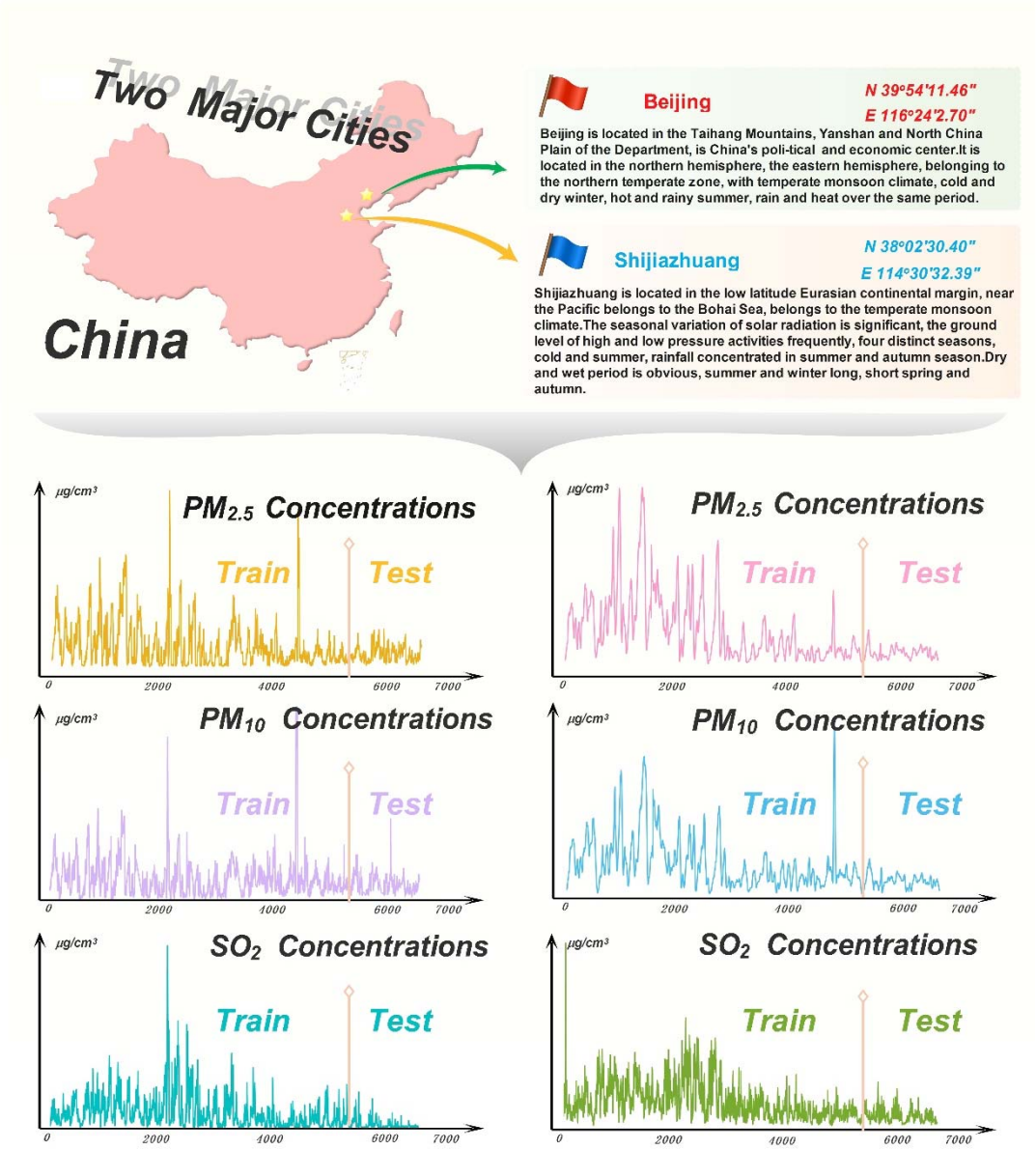

Fig.2. Specific locations of the two study cities as well as the climatic conditions.

3.2 The Performance Metric

In recent literature, there have been many metrics employed to evaluate forecasting models, but there is no clear rule about which specific metrics are standard. Multiple performance metrics can properly evaluate the performance of the model. Therefore, this study employed eight metrics to evaluate the forecasting accuracy in deterministic prediction; two metrics were introduced in this study to verify the performance of interval forecasting in the uncertainty analysis; the definitions and the expressions are detailed described in Table 1.

Table 1

The definitions and expressions of the metrics

Metric	Definition	Equation
MAPE	The average of N absolute percentage error	$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left\|\frac{A_i - F_i}{A_i}\right\| \times 100\%$

MAE	The mean absolute error of *N* forecasting results	$MAE = \frac{1}{N}\sum_{i=1}^{N}\lvert A_i - F_i\rvert$
RMSE	The square root of the average of the error squares	$RMSE = \sqrt{\frac{1}{N}\times\sum_{i=1}^{N}(A_i - F_i)^2}$
MdAPE	The median of *N* absolute percentage error	$MdAPE = median\left(\left\lvert\frac{A_i - F_i}{A_i}\right\rvert\times 100\%\right)$
DA	The direction accuracy of forecasting results	$DA=\frac{1}{l}\sum_{i=1}^{l}w_i,\ w_i=\begin{cases}1, if\ (A_{i+1}-A_i)(F_{i+1}-A_i)>0\\0, \quad otherwise\end{cases}$
FB	The fractional bias of *N* forecasting results	$FB=2\left(\bar{A}-\bar{F}\right)/\left(\bar{A}+\bar{F}\right)$
IA	The index of agreement of forecasting results	$IA=1-\sum_{i=1}^{N}(F_i-A_i)^2/\sum_{i=1}^{N}\left(\lvert F_i-\bar{A}\rvert+\lvert A_i+\bar{A}\rvert\right)^2$
R^2	Pearson's correlation coefficient	$R^2=\frac{\sum FA-\sum F\sum A/N}{\sqrt{\left(\sum F^2-(\sum F)^2/N\right)\left(\sum A^2-(\sum A)^2/N\right)}}$
IFCP	Interval forecasting coverage probability	$IFCP=\frac{1}{N}\sum_{i=1}^{N}c_i,\ c_i=\begin{cases}1, \quad if\ y_i\in[L_i,U_i]\\0, \quad otherwise\end{cases}$
IFAW	Interval forecasting average width	$IFAW = \frac{1}{N}\sum_{i=1}^{N}(U_i - L_i)$

In the Table 1, A_i and F_i represent the forecast value and the actual value of the pollutants concentrations respectively.

3.3 Aims of the Experiments

To verify the superiority of the proposed model and the performance of the model to analyze and monitor the air quality, three experiments were constructed. The experiments were carried out based on two datasets, as mentioned above. The details of the experiments are described as follows:

(1) Experiment I aims to forecast the target point data of the pollutant concentrations and two cities were selected to verify the performance of the proposed hybrid model. This study applied the FTS forecasting method and CEEMD to establish the hybrid model; moreover, the research on FTS forecasting mainly focuses on two aspects: division of the discrete domain and the method of weight distribution [46-47, 67]. The EBD algorithm was applied to define the numbers and the width of the interval; specifically, the pollutant concentration was divided into seven attribute classes according to the Air Quality Index (AQI) to divide the air pollution level method. The AQI corresponding to the air pollution level is shown in Table 2.

Table 2

Different classification standard

AQI	Level	Descriptions	Color	SO_2	$PM_{2.5}$	PM_{10}
0-50	I	Good	Green	≤50	≤35	≤50
51-100	II	Moderate	Yellow	50-150	35-75	50-150
101-150	III	Lightly Polluted	Orange	150-250	75-115	150-250
151-200	IV	Moderately Polluted	Red	250-475	115-150	250-350
201-300	V	Heavily Polluted	Purple	475-800	150-250	350-420
>300	VI	Severely Polluted	Maroon	≥800	≥250	≥420

(2) Experiment II aims to forecast the uncertainty of the pollutant concentrations based on interval forecasting. Interval forecasting can predict the range of pollutant concentration fluctuations under a confidence level α; as a result, it is of great help in the establishment of air early warning systems and the treatment of air pollution.

(3) Experiment III showed an additional experiment of comparing results with baselines based on two relative researches [2, 68]. Comparing with the benchmark models, the purpose of the experiment is to demonstrate the superiority of the proposed model and to prove that fuzzy logic is a better option for predicting air pollutants.

(4) Experiment IV established two kinds of testing methods to evaluate the proposed hybrid model. This study not only applies the performance metrics to verify the forecasting effectiveness of the model; in order to fully verify the validity of the model, the Wilcoxon rank-sum test and the robustness test were also employed to the text. Through the Wilcoxon rank-sum test, whether the forecasting effectiveness between the benchmark models and the proposed model had significant differences could be verified, which proves the validity of the model. In addition, the robustness test could verify whether the model is suitable for time series forecasting and to determine what data is unstable and stochastic.

4 Experiments and Analysis

In this section, three experiments are conducted based on the experimental aims mentioned above to predict and analyze the major air pollutants and to evaluate the system. The performance of the experiments was evaluated by three major pollutant concentrations in China, and the results and detailed analysis are illustrated below.

4.1 Experiment I: Forecasting models comparison

To verify the superiority of the proposed model in forecasting capability, some other popular forecasting models, ENN, BPNN and ARIMA, were constructed as benchmarks. To discuss the contribution of the CEEMD and EBD algorithm, EW*, CEW* and EBD* were constructed to compare with the proposed model. To evaluate the performance of the model, the metrics mentioned above were all calculated and are presented in Table 3 and Table 4. Moreover, Fig. 3 depicts the forecasting results of the two datasets and the model performance metrics value of the statistical model (ARIMA), ANN model (BPNN) and the proposed model. The detailed analysis of the forecasting results is as follows:

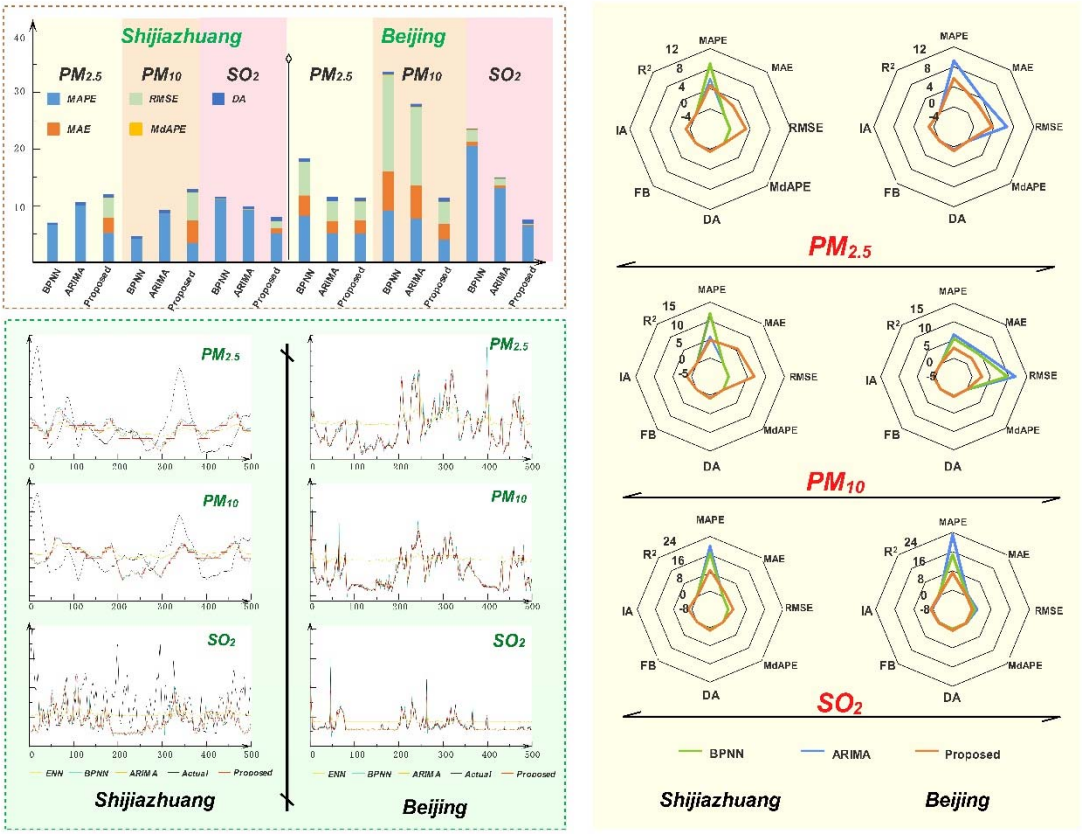

Fig.3. Comparison of the forecasting results obtained by experiment I

4.1.1 Case in Beijing

Beijing is the capital of China, whose air quality plays an important role in establishing an air quality early warning system. Thus, it is essential to analyze the primary pollutant concentrations in Beijing. In order to fully verify the validity of the proposed model, several benchmarks are employed to be compared and the validity of each part of the proposed model is also analyzed. Through the experimental results, the conclusions can be obtained as follows.

(1) The comparison results of proposed model and single models

As the forecasting performance is shown in Table 3, the forecasting accuracy of the proposed model is more precise than that of the other three models. More precisely, for $PM_{2.5}$ forecasting, the MAPE of ENN, BPNN and ARIMA are 93.4518%, 9.2301% and 5.7883%, respectively, whereas the MAPE of the proposed hybrid model is 5.6596%. For the PM_{10} forecasting, the proposed hybrid model has the smallest values of MAPE, and obtained a decrease of 78.8291%, 5.7909% and 4.1703% in the MAPE, whereas for SO_2 forecasting, the proposed model obtained a reduction of 85.2686%, 15.7666% and 7.4473% in the MAPE compared with ENN, BPNN and AERIMA. Furthermore, for the other metrics, the proposed model is almost more superior than that of the other compared models, which means that a single model cannot obtain satisfactory results.

In comparison with the two AI models in experiment I, the BPNN performs better than ENN with all metrics and provides great reduction. The results mean that ENN cannot satisfy the requirements for air quality forecasting. Furthermore, compared with

AI models (i.e., ENN, BPNN), the statistical model (i.e., ARIMA) has higher accuracy, but not all indicators are better than the AI models. Fig. 3 shows the metrics of the benchmarks models (BPNN and ARIMA); it can be concluded that the ARIMA and BPNN both have good forecasting results, but the proposed model is far better than that two single models.

In summary, the statistical model has better forecasting accuracy than the AI model, but not all performance metrics are superior to the AI model in the first case. However, the proposed hybrid model based on fuzzy logic is superior to both statistical model and AI model in forecasting performance. And from the experimental results, it is proved that fuzzy logic is a better option in air pollution forecasting.

Table 3

Results of the proposed model and the other single models

Pollutions	Model	MAPE	MAE	RMSE	MdAPE	DA	FB	IA	R^2
Beijing									
$PM_{2.5}$	**ENN**	93.4518	26.7606	30.9735	0.5604	0.5135	-0.3222	0.5297	0.4979
	BPNN	9.2301	3.9791	6.6215	0.0617	0.6256	-0.0032	0.9854	-0.0297
	ARIMA	5.7883	2.4315	3.8616	0.0394	0.8178	-0.0003	0.9949	0.0571
	Proposed	**5.6596**	**2.6760**	**3.6299**	**0.0479**	**0.8028**	**-0.0115**	**0.9955**	**-0.0008**
PM_{10}	**ENN**	83.3110	41.4097	48.4174	0.5725	0.4535	-0.3662	0.5034	0.1405
	BPNN	10.2728	7.7155	18.9002	0.0676	0.5195	-0.0106	0.9328	-0.0129
	ARIMA	8.6522	6.4973	15.4009	0.0589	0.6276	-0.0054	0.9545	0.0638
	Proposed	**4.4819**	**3.1849**	**4.2876**	**0.0376**	**0.7948**	**-0.0123**	**0.9966**	**-0.0030**
SO_2	**ENN**	92.4942	6.8868	7.1088	3.8088	0.1351	-1.0314	0.3093	-6.8499
	BPNN	22.9922	0.7577	2.2232	0.1599	0.1562	-0.1063	0.7895	-0.3766
	ARIMA	14.6729	0.4956	1.2898	0.0956	0.2372	-0.0240	0.9192	0.1254
	Proposed	**7.2256**	**0.1658**	**0.2477**	**0.0437**	**0.7958**	**-0.0145**	**0.9972**	**-0.0586**
Shijiazhuang									
$PM_{2.5}$	**ENN**	44.7254	0.0210	0.0236	0.3462	0.3554	-0.2954	0.5940	-0.5818
	BPNN	7.4357	0.0040	0.0042	0.0665	0.3554	-0.0649	0.9853	-0.0572
	ARIMA	11.1925	0.0063	0.0085	0.0085	0.6306	0.0000	0.9994	-0.0011
	Proposed	**5.8237**	**3.0605**	**3.9050**	**0.0451**	**0.6284**	**-0.0112**	**0.9910**	**-0.0237**
PM_{10}	**ENN**	32.0291	0.0259	0.0324	0.1837	0.3914	-0.1833	0.6442	0.4064
	BPNN	4.6425	0.0247	0.0297	0.0194	0.5345	-0.0200	0.9979	0.0320
	ARIMA	9.6000	0.0010	0.0014	0.0070	0.7267	-0.0006	0.9995	0.0033
	Proposed	**3.8466**	**4.3846**	**5.5049**	**0.0341**	**0.6551**	**-0.0015**	**0.9951**	**0.0084**
SO_2	**ENN**	91.3123	0.0243	0.0259	1.5679	0.3764	-0.7321	0.4074	-2.4857
	BPNN	12.4302	0.0023	0.0037	0.0915	0.4695	-0.0257	0.9776	0.0282
	ARIMA	10.4516	0.0020	0.0031	0.0725	0.5155	-0.0081	0.9826	0.1407
	Proposed	**5.8026**	**1.0090**	**1.3633**	**0.0393**	**0.8232**	**-0.0211**	**0.9974**	**-0.0026**

Note: The bold numbers in the table represent the results of the proposed model

(2) Compare CEEMD with other processing approaches

CEEMD, as one of the models with great decomposing capability, was applied to processing the original series of pollutants concentration. In this study, we set the ensemble member to 500 and the standard deviation of the added white noise in each ensemble member was set to 0.2. However, the large number of ensemble members will introduce model complexity and is time consuming. On the other hand, if the number of ensemble member sets is small, it is hard to obtain a satisfying performance of decomposition. To discuss the role that CEEMD played in the proposed hybrid model,

an extended comparison is illustrated below. From Table 4, in the selected forecasting of $PM_{2.5}$ for instance, the hybrid model with CEEMD achieved the smallest MAPE, at 5.6956%. It also reveals that CEW* had a decrease of 18.3029% in MAPE compared with EW*, and the proposed model had a reduction of 2.5494% in MAPE compared with the EBD*. Furthermore, the performance of the other seven metrics (i.e., MAPE, MAE, RMSE, MdAPE, DA, FB, IA, R^2) all had a different degree of improvement. Therefore, it can be concluded that the CEEMD can efficiently eliminate the noise and unstable elements of original series, and it could be a promising model to capture the primary components hidden in the original pollutant concentration time series.

Table 4

Results of the proposed model and the other hybrid models

Pollutions	Model	MAPE	MAE	RMSE	MdAPE	DA	FB	IA	R^2
Beijing									
$PM_{2.5}$	**EW***	64.5626	20.4301	23.3211	0.3632	0.5015	-0.1497	0.7465	0.2721
	CEW*	46.2097	17.4469	20.5871	0.3678	0.5415	-0.0127	0.8445	0.0010
	EBD*	8.2090	3.8403	5.2425	0.0644	0.7377	-0.0164	0.9905	0.0482
	Proposed	**5.6596**	**2.6760**	**3.6299**	**0.0479**	**0.8028**	**-0.0115**	**0.9955**	**-0.0008**
PM_{10}	**EW***	51.6665	28.3187	32.7669	0.3323	0.4615	-0.1172	0.6979	0.4485
	CEW*	37.7629	26.4086	31.8998	0.3379	0.5375	0.0810	0.7861	0.0608
	EBD*	6.5188	4.0095	5.3030	0.0403	0.7207	-0.0169	0.9948	0.0508
	Proposed	**4.4819**	**3.1849**	**4.2876**	**0.0376**	**0.7948**	**-0.0123**	**0.9966**	**-0.0030**
SO_2	**EW***	98.0464	9.3571	9.5824	5.2317	0.1301	-1.1897	0.2479	-13.7973
	CEW*	57.6256	5.6929	5.8965	2.9917	0.5005	-0.9251	0.3475	-4.7346
	EBD*	15.7718	0.3733	0.3968	0.2255	0.2633	-0.1063	0.9932	0.0243
	Proposed	**7.2256**	**0.1658**	**0.2477**	**0.0437**	**0.7958**	**-0.0145**	**0.9972**	**-0.0586**
Shijiazhuang									
$PM_{2.5}$	**EW***	30.3276	14.5505	17.9309	0.1797	0.3604	-0.0782	0.2988	0.9202
	CEW*	25.5265	14.3711	17.6581	0.1854	0.5015	-0.0596	0.2493	0.9546
	EBD*	7.5598	4.0231	4.9252	0.0670	0.3969	-0.0117	0.9856	-0.0222
	Proposed	**5.8237**	**3.0605**	**3.9050**	**0.0451**	**0.6284**	**-0.0112**	**0.9910**	**-0.0237**
PM_{10}	**EW***	30.9088	26.8718	33.6136	0.1812	0.4104	-0.0622	0.2554	0.9504
	CEW*	20.6018	26.7779	33.4330	0.1848	0.4985	-0.0542	0.2317	0.9626
	EBD*	4.9296	5.5139	6.8431	0.0413	0.5077	-0.0017	0.9925	-0.0285
	Proposed	**3.8466**	**4.3846**	**5.5049**	**0.0341**	**0.6551**	**-0.0015**	**0.9951**	**0.0084**
SO_2	**EW***	94.5054	16.4491	18.4185	1.0025	0.4024	-0.5474	0.5650	-0.7500
	CEW*	43.1343	9.4800	11.9580	0.4645	0.5435	0.1397	0.8439	-1.3799
	EBD*	6.3430	0.8002	1.0054	0.0398	0.7325	-0.0284	0.9986	0.0046
	Proposed	**5.8026**	**1.0090**	**1.3633**	**0.0393**	**0.8232**	**-0.0211**	**0.9974**	**-0.0026**

Note: EW, CEW and* represent EWP, CEEMD-EWP, FTS forecasting, respectively. For example, CEW* is CEEMD-EWP-FTS.

(3) Compare EBD algorithm with other partition approaches

Consider the suitable number of human short-term memory effects: the discourse universe is usually segregated into seven linguistic values [69], but in fact, seven linguistic values cannot completely divide the attributes of the data when the amount of data is large. Therefore, the EBD algorithm has been applied to segregate the discourse universe, and it terminates the iteration based on the threshold of entropy, such that the number of linguistic values varies from the iteration times, which are not specified. There are some commonly used approaches for dividing the universe: equal-

width pre-partitioning (EWP) and equal-depth (frequency) pre-partitioning (EDP). The EWP method is used to separate all of the linguistics with the same width, whereas the EDP method is used to separate all of the linguistics with the same frequencies [70]. Because the fluctuation range of the pollutant concentration data is relatively large, the forecasting error will increase based on the EDP. Therefore, the approach based on EWP serves as a benchmark model in this paper.

For fuzzy time series forecasting, the model based on CEEMD can heavily improve the accuracy, as mentioned above, and the EBD also provides significant improvement in forecasting accuracy. By observing Table 4, the following conclusions can be drawn. For $PM_{2.5}$ forecasting, the EBD* model decreases by 56.353% in MAPE compared with EW*, and the proposed model has a reduction of 40.5501% in MAPE compared with CEW*. Furthermore, the models with the EBD algorithm all provide huge improvements in the performance of the remaining seven metrics compared with the models with the EWP method. In addition, for PM_{10} and SO_2 forecasting, the models using the EBD method are all superior to the models with the EWP method; moreover, the proposed model almost performs best in all metrics. It reveals that the EBD plays an important role in FTS forecasting; furthermore, the proposed model is efficient in forecasting the pollutant concentration series.

Remark: From the above experimental results, compared with MAPE, MAE, RMSE, MdAPE, DA, FB, IA, R^2 and forecasting effectiveness, the proposed model almost performs best. Furthermore, the CEEMD and EBD approaches contribute much to the hybrid model, with significant improvements in accuracy. Compared with single models and other benchmark models, the proposed hybrid model performs the best in all cases. Thus, the experimental results indicate that the proposed hybrid model is a promising model for forecasting the primary air pollutant concentrations.

4.1.2 Case in Shijiazhuang

The proposed model and all benchmark models were applied to forecast the concentration of the pollutants in Shijiazhuang. The detailed analyses are described below.

(1) By comparing them with the single model, the forecasting accuracy of the ARIMA model and the BPNN model are similar, whereas the ENN model performs worse. The reason for the fluctuation in the forecasting accuracy may be due to the large volatility and the poor stability of the pollutant concentration data. However, under the same conditions, the proposed model performs better than any other model and has the best MAPE value of 5.8237%, 3.8466% and 5.8026% respectively in forecasting the three pollutant concentrations.

(2) From Table 4, using set $PM_{2.5}$ forecasting as an example, the models with the CEEMD method exhibit a reduction of 4.4735% and 1.7361%, respectively. It can be concluded that the CEEMD is a successful application for decomposing the original pollutant concentration series and makes a great contribution to increasing accuracy.

(3) In selecting $PM_{2.5}$ forecasting as an example, the EBD*model decreased by 22.7678% in MAPE compared with EW*, and the proposed model had a reduction of 19.7028% in MAPE compared with CEW*. Furthermore, the models with the EBD algorithm all exhibited significant improvement in the performance of the remaining

seven metrics compared with the models with the EWP method.

Remark: From the analysis of the experimental results mentioned above, the following conclusions could be made: (1) The proposed model based on fuzzy time series performs better than statistical models and AI models with better accuracy, it reveals that the fuzzy logic has a successful application in air pollution forecasting. (2) Compared with single models, the proposed model outperforms benchmarks in forecasting the concentration of pollutants, and it can also be applied to different environments with high accuracy. (3) All considered models based on the CEEMD method were superior to the models without a decomposition method, which demonstrates that the CEEMD can efficiently enhance the model performance. (4) The EBD algorithm outperforms the EWP method by properly pre-partitioning the discourse universe, and the experimental results demonstrate the effectiveness of adaptively partitioning the universe in forecasting fuzzy time series.

4.2 Experiment II: Uncertainty Analysis

The randomness and intermittence of pollutant concentration is the biggest challenge in the air quality monitor system. Accurate forecasting of pollutant concentration is a powerful tool to deal with such problem. Conventional pollutant concentration forecasting model usually produces a value at a time point in the future. However, any forecasting approach has its inherent and irreducible uncertainty. Compared with deterministic prediction model, interval forecasting integrated uncertainty analysis is a promising approach to providing a forecasting range, confidence level, and other uncertain information of future values, which can be a smart tool to assist decision-makers in analyzing and monitoring air quality. For the purpose of quickly and properly calculating the interval range, Gaussian distributions are employed here to estimate the bilateral limits of the interval. However, under different interval confidence levels, the length and range of the interval is different. To capture the information hidden in the interval forecasting, under different confidence levels, seven estimated intervals are listed in Table 5.

The average length of the interval is usually used to evaluate the interval prediction effect. It is well known that the shorter the average length of the prediction interval, the better the effect of the interval prediction. However, it is not reliable to only consider this rule to measure the effect of interval forecasting. If the observed data does not fall within the forecasting interval, the forecasting interval is meaningless. Thus, the forecasting interval should cover most of the observed data. Based on the analysis, this study uses double metrics to evaluate the effect of interval forecasting. IFCP represents the proportion of observed data falling within the forecasting interval, which is expected to be close to 1, as it must be in the range [0, 1]. Whereas the IFAW denotes the average length of the interval, which expected to be small.

By analyzing the results shown in Table 5, the following conclusions can be drawn: (1) different confidence levels bring about different IFCP and IFAW. For example, in $PM_{2.5}$ forecasting in Beijing, when $\alpha = 0.2$ and $\alpha = 0.3$, the values of IFCP and IFAW are 80.6%, 6.03 μg/m^3 and 73.4% and 4.9μg/m^3, respectively. The same situation in Shijiazhuang also arose in PM_{10} forecasting,, when $\alpha = 0.3$ and $\alpha = 0.4$, the values of IFCP and IFAW were 78.6%, 9.87 μg/m^3 and 74.3% and 7.97 μg/m^3, respectively. (2)

When the confidence level increased, the IFCP gradually increased, whereas the IFAW gradually decreased. For $PM_{2.5}$ forecasting in Beijing, when α expanded from 0.01 to 0.5, IFCP decreased by 32.9%, while IFAW exhibited a reduction of 8.86 μg/m^3. Similar things also occur in other pollutant concentration forecasting.

Remark: From the experimental results, it can be concluded that interval forecasts should meet the following conditions: under a proper α (choose the appropriate level of confidence based on the actual situation), the value of IFCP should be as large as possible; meanwhile, the value of IFAW should be as small as possible. Furthermore, the more important thing is that the prediction accuracy of the experiment must be high. In summary, this study selects a confidence level of 0.1, and the results of the interval forecasts are displayed in Table 5 in black bold.

Table 5

The interval forecasting results under different significant levels.

City	Level	$PM_{2.5}$		PM_{10}		SO_2	
Beijing	α	IFCP	IFAW	IFCP	IFAW	IFCP	IFAW
	0.01	93.80	12.01	95.80	15.32	99.00	1.92
	0.05	92.00	9.23	90.50	11.78	97.90	1.47
	0.10	**88.50**	**7.72**	**85.50**	**9.85**	**96.60**	**1.23**
	0.20	80.60	6.03	77.90	7.69	95.60	0.96
	0.30	73.40	4.90	72.10	6.25	93.70	0.78
	0.40	65.90	3.95	66.70	5.05	88.60	0.63
	0.50	60.90	3.15	61.00	4.03	83.70	0.50
Shijiazhuang	α	IFCP	IFAW	IFCP	IFAW	IFCP	IFAW
	0.01	98.10	17.77	98.00	24.20	95.90	5.59
	0.05	92.10	13.66	94.40	18.60	91.00	4.30
	0.10	**90.20**	**11.43**	**91.60**	**15.56**	**84.90**	**3.60**
	0.20	86.50	8.92	84.20	12.15	76.50	2.81
	0.30	81.10	7.25	78.60	9.87	71.50	2.28
	0.40	74.80	5.85	74.30	7.97	65.20	1.84
	0.50	69.80	4.67	69.80	6.36	58.50	1.47

4.3 Experiment III: Compare Results with Baselines

In order to clearly indicate the superiority of the proposed model and prove that fuzzy logic is a better option in air quality forecasting, this paper carried on an additional experiment based on two related researches [2, 68]. From the experimental results of the paper mentioned above, the neural networks structure with the best results is selected to carry on the experiment. Therefore, the benchmark models, FF and LR with the adaptive learning function of gradient descent weight and bias, ANFIS with the membership of Gaussian function and the training of the neural network based on hybrid algorithm are conducted in our paper in order to indicate the superiority of the proposed method. The number of hidden nodes has great effect on the forecasting accuracy of the neural network. Therefore, this paper selects different hidden layer nodes to construct the neural network and apply it to the prediction. The data sets from two sites are also selected to conduct the experiment, Table 6 shows the average results of two dataset.

As shown in Table 6, the best number of hidden layer nodes is different in different

dataset. Therefore, there is no definite theory to determine the optimal network structure. However, the proposed model performs better than FF, LR and ANFIS with the best MAPE of 5.7417%, 4.1643% and 6.5141% respectively. It also reveals that proposed model has a decrease of 3.8699% and 7.8481% in average MAPE compared with FF and LR in $PM_{2.5}$ forecasting, while it has a reduction of 7.6326% compared with ANFIS. Furthermore, the performance of the performance in other four metrics (i.e., MAE, RMSE, MdAPE, DA, FB, IA, R^2) almost performs better than other models.

Remark: Experimental results reveal that the proposed model based on fuzzy time series has the superiority to other benchmarks in air pollution forecasting and fuzzy logic is a better option in air quality forecasting.

4.4 Experiment IV: Comprehensive Test

To further evaluate the efficiency of the model, this section considers two kinds of tests to examine the forecasting performance. The Wilcoxon rank-sum test aims to determine whether the two models have significant differences, and the robustness test aims to examine whether the accuracy of the model fluctuates when the data set fluctuates. Moreover, the experimental results and the analysis of the results are detailed as follows.

4.4.1 The Wilcoxon Rank-sum Test of the Proposed Model

In this experiment, ENN, BPNN, ARIMA, EW*, CEW* and EBD* were employed as validation models, and the proposed hybrid model served as the tested model. To compare the significant differences of forecasting the effectiveness between the tested models with any of the validation models, the error series of the models were conducted to generate Wilcoxon rank-sum statistics. Since the experimental test sample is selected from the testing set; therefore, the sample size is large enough to generate Wilcoxon rank-sum statistics under a large sample size. The experimental results reveal the test results and the *P* value when rejecting the original hypothesis under a confidence level α. In addition, in the test results, 0 represents the accepted original hypothesis, which means that there is no significant difference between the two test samples; whereas 1 represents a rejection of the original hypothesis, which implies that there are significant differences between the two test samples. The *P* value when rejecting the original hypothesis under a confidence level of 5% of the experimental results is shown in Table 7.

Table 7

The results of the Wilcoxon rank-sum.

Beijing	**ENN**	**BPNN**	**ARIMA**	**EW***	**CEW***	**EBD***
$PM_{2.5}$	1.07×10^{-45}	3.20×10^{-3}	3.47×10^{-9}	9.69×10^{-33}	8.01×10^{-197}	1.38×10^{-2}
PM_{10}	8.46×10^{-158}	4.64×10^{-2}	2.40×10^{-2}	3.42×10^{-20}	1.65×10^{-2}	3.06×10^{-2}
SO_2	7.18×10^{-316}	7.21×10^{-97}	9.03×10^{-7}	3.81×10^{-365}	1.87×10^{-265}	1.37×10^{-192}
Shijiazhuang	**ENN**	**BPNN**	**ARIMA**	**EW***	**CEW***	**EBD***
$PM_{2.5}$	2.69×10^{-10}	1.93×10^{-10}	1.91×10^{-10}	3.18×10^{-13}	3.23×10^{-7}	4.34×10^{-2}
PM_{10}	1.65×10^{-2}	2.13×10^{-2}	2.15×10^{-2}	1.55×10^{-9}	7.98×10^{-7}	5.86×10^{-2}
SO_2	2.92×10^{-40}	2.00×10^{-46}	1.60×10^{-46}	5.70×10^{-241}	2.33×10^{-88}	2.03×10^{-6}

When the *P* value is close to zero, the original hypothesis can be rejected. Thus, the smaller the *P* value, the more significant the difference between the two samples. For the test results for Beijing, all of the *P* values are approximately zero, especially

with ENN, and the *P* values are 1.07×10^{-45}, 8.46×10^{-158}, 7.18×10^{-316} respectively, in $PM_{2.5}$, PM_{10} and SO_2 forecasting; it can concluded that the difference in the forecasting effectiveness between the model ENN and the proposed model is enormous. Similarly, the *P* value of the BPNN, ARIMA, EW*, CEW* and EBD* are 3.20×10^{-3}, 3.47×10^{-9}, 9.69×10^{-33}, 8.01×10^{-197} and 1.38×10^{-2}, respectively, in $PM_{2.5}$ forecasting. The same situation also appeared in the prediction of Shijiazhuang. As the forecasting accuracy of the proposed model is superior to the validation models, and there is significant difference between the proposed model and the validation models, this verified the effectiveness of the proposed model compared with the validation models.

Table 6

The experimental results of the additional experiment.

Metrics		Feed Forward Backpropagation (FF)					Layer Recurrent(LR)					ANFIS	Proposed
Hidden Nodes		4	5	6	7	Average	4	5	6	7	Average		
$PM_{2.5}$	MAPE	9.7568	10.8101	8.6662	9.2132	9.6116	7.5480	27.2750	7.8500	11.6860	13.5898	13.3743	5.7417
	MAE	5.8768	6.4319	4.9121	5.7849	5.7514	4.1007	13.0311	4.3593	7.1578	7.1622	5.3962	2.8683
	RMSE	8.3108	8.8174	6.7051	9.5666	8.3499	5.5949	15.6985	5.9345	10.3430	9.3927	13.8074	3.7675
	MdAPE	0.0796	0.0992	0.0681	0.0662	0.0782	0.0705	0.2197	0.0695	0.0953	0.1138	0.0764	0.0465
	DA	0.6441	0.5763	0.7119	0.7288	0.6653	0.7119	0.3390	0.7288	0.6949	0.6186	0.6352	0.7156
	FB	0.0730	0.0411	0.0513	0.0668	0.0580	0.0426	0.0047	0.0465	0.1003	0.0485	-0.0188	-0.0114
	IA	0.9599	0.9590	0.9748	0.9446	0.9596	0.9844	0.9313	0.9821	0.9388	0.9591	0.8546	0.9933
	R^2	0.3324	0.1600	0.2984	0.3756	0.2916	0.1029	-1.5718	0.1371	0.2709	-0.2652	-0.7927	-0.0123
PM_{10}	MAPE	10.1240	11.3161	9.1620	9.4004	10.0006	7.8468	11.6904	8.8077	9.9549	9.5750	13.1349	4.1643
	MAE	8.4359	10.4937	7.8507	7.9494	8.6824	6.4750	9.2697	7.4254	8.7069	7.9693	9.3899	3.7848
	RMSE	11.7710	15.6796	11.0894	11.0544	12.3986	10.0361	13.1831	10.6692	12.8006	11.6723	14.4563	4.8963
	MdAPE	0.0836	0.0875	0.0737	0.0709	0.0789	0.0640	0.0899	0.0748	0.0778	0.0766	0.0759	0.0359
	DA	0.5085	0.5593	0.5593	0.4915	0.5297	0.6102	0.3898	0.4407	0.6441	0.5212	0.6160	0.7250
	FB	-0.0094	0.1058	0.0539	0.0414	0.0479	0.0055	0.0187	0.0426	0.0505	0.0293	0.0957	-0.0069
	IA	0.9445	0.8646	0.9410	0.9455	0.9239	0.9567	0.9274	0.9496	0.9228	0.9391	0.8345	0.9959
	R^2	-0.1070	0.3880	0.2323	0.1126	0.1565	0.0598	-0.0199	0.1009	0.1763	0.0793	-0.3349	0.0027
SO_2	MAPE	7.8468	11.6904	8.8077	9.9549	9.5750	13.4083	12.5445	12.2840	11.1668	12.3509	12.9439	6.5141
	MAE	6.4750	9.2697	7.4254	8.7069	7.9693	0.3368	0.3346	0.3098	0.2954	0.3192	0.1560	0.5874
	RMSE	10.0361	13.1831	10.6692	12.8006	11.6723	0.4262	0.4702	0.4882	0.4352	0.4550	0.2460	0.8055
	MdAPE	0.0640	0.0899	0.0748	0.0778	0.0766	0.0894	0.0808	0.0589	0.0619	0.0727	0.0912	0.0415
	DA	0.6102	0.3898	0.4407	0.6441	0.5212	0.2034	0.2034	0.0847	0.2034	0.1737	0.1873	0.8095
	FB	0.0055	0.0187	0.0426	0.0505	0.0293	0.0150	0.0105	0.0486	0.0305	0.0261	-0.0483	-0.0178
	IA	0.9567	0.9274	0.9496	0.9228	0.9391	0.7426	0.7721	0.7930	0.7802	0.7720	0.7481	0.9973
	R^2	0.0598	-0.0199	0.1009	0.1763	0.0793	0.6255	0.1217	-0.1137	0.3550	0.2471	-0.2661	-0.0306

4.4.2 A Robustness Test of the Proposed Model

The purpose of the robustness test was to examine whether the forecasting accuracy of the model greatly changes when the historical datasets are nonstationary and not accurate. In this experiment, the data for the training sets randomly increased by 5%, which is considered to be from stochastic disturbances; then, a change of each performance metrics was observed; the comparison results are tabularized in Table 8.

Table 8

The results of the robustness test

	MAPE		MAE		RMSE		MdAPE	
	Random	**Proposed**	**Random**	**Proposed**	**Random**	**Proposed**	**Random**	**Proposed**
Beijing								
$PM_{2.5}$	5.2211	5.6596	2.4468	2.6760	3.4834	3.6299	0.0396	0.0479
PM_{10}	3.8308	4.4819	2.7501	3.1849	3.6045	4.2876	0.0338	0.0376
SO_2	7.3658	7.2256	0.2345	0.1658	0.2613	0.2477	0.0514	0.0437
Mean	**5.4726**	**5.7890**	**1.8105**	**2.0089**	**2.4497**	**2.7217**	**0.0416**	**0.0431**
Std	**1.4541**	**1.1239**	**1.1212**	**1.3197**	**1.5482**	**1.7699**	**0.0073**	**0.0042**
Shijiazhuang								
$PM_{2.5}$	5.7695	5.8237	3.1255	3.0605	3.9645	3.9050	0.0468	0.0451
PM_{10}	3.8338	3.8466	4.2153	4.3846	5.4636	5.5049	0.0325	0.0341
SO_2	5.6358	5.8026	0.9362	1.0090	1.2303	1.3633	0.0392	0.0393
Mean	**5.0797**	**5.1576**	**2.7590**	**2.8180**	**3.5528**	**3.5911**	**0.0395**	**0.0395**
Std	**0.8827**	**0.9271**	**1.3636**	**1.3887**	**1.7526**	**1.7053**	**0.0059**	**0.0045**
	DA		**FB**		**IA**		**R^2**	
	Random	**Proposed**	**Random**	**Proposed**	**Random**	**Proposed**	**Random**	**Proposed**
Beijing								
$PM_{2.5}$	0.8268	0.8028	-0.0151	-0.0115	0.9959	0.9955	0.0097	-0.0008
PM_{10}	0.8198	0.7948	-0.0106	-0.0123	0.9976	0.9966	0.0066	-0.0030
SO_2	0.7857	0.7958	-0.0241	-0.0145	0.9756	0.9972	-0.1369	-0.0586
Mean	**0.8108**	**0.7978**	**-0.0166**	**-0.0128**	**0.9897**	**0.9964**	**-0.0402**	**-0.0208**
Std	**0.0180**	**0.0036**	**0.0056**	**0.0012**	**0.0100**	**0.0007**	**0.0684**	**0.0268**
Shijiazhuang								
$PM_{2.5}$	0.6346	0.6284	-0.0115	-0.0112	0.9955	0.9910	-0.0248	-0.0237
PM_{10}	0.6550	0.6551	-0.0019	-0.0015	0.9976	0.9951	0.0076	0.0084
SO_2	0.8232	0.8232	-0.0211	-0.0211	0.9974	0.9974	-0.0024	-0.0026
Mean	**0.7043**	**0.7022**	**-0.0115**	**-0.0113**	**0.9968**	**0.9945**	**-0.0065**	**-0.0060**
Std	**0.0845**	**0.0862**	**0.0079**	**0.0080**	**0.0009**	**0.0026**	**0.0135**	**0.0133**

For the proposed model, the average MAPE is 5.7890% and 5.1576% respectively in two observation cities. With regard to the modified model, the values of MAPE decrease to 5.4726% and 5.0797% respectively, indicating that the stochastic disturbances do not affect the forecasting performance. Besides, for the proposed model, the average RMSE is 2.7217 with the standard deviation 1.7699 ranges of forecasting errors. For the perspective of modified model, the average RMSE decreases a little and the standard deviation decreases to 1.5482, revealing that the random disturbances are not significant. Moreover, it can be seen that the forecasting performance of the proposed model does not change significantly by observing the standard deviation fluctuations of other metrics. In the dataset for Shijiazhuang forecasting, the situation remained the same. When smaller instability occurred into model as compared to original model, it is very weak to deny the robustness of the proposed model [32]. Therefore, sufficient evidences have proved the robustness of the proposed model.

5 Discussion

One of the ultimate goals of every early warning system for air quality is to appraise the forecasting performance and stability as accurately as possible. Pollutant concentrations become more attractive to operators of economic systems and environmental monitoring systems, because model accuracy is improved, and better predictive techniques applications are introduced. In previous work, experiments proved that an early warning system for air quality can improve the accuracy of pollutant forecasting; at the same time, selecting the specific aspects of the system will be discussed in this section. First, correctly selecting the parameters is conducive to better performance in the CEEMD model and of the system, thus, the question of choosing parameters for the CEEMD model is discussed in this work. Furthermore, the selection of a partitioning method that is suitable and has an impact on FLR establishment and FTS forecasting performance is discussed in this section. Moreover, a high-precision and robust system is crucial for decision-making and analysis; therefore, the effectiveness and stability of an air quality early warning system will be further discussed and validated.

5.1 Discussion of the Parameter Ensemble Number in CEEMD

In the proposed early warning system, the first step is to use the CEEMD model to decompose the original series of three major pollutant concentrations into several independent IMFs. In this research, the standard deviation of the added white noise in each ensemble number was 0.2, and the ensemble number was set to 500. However, the variation in the parameters may affect the decomposition result of the model. Therefore, five values for the ensemble numbers were chosen in this section to determine the optimal parameters from the average error and calculation time of the experimental results. Table 9 shows the detailed results for three major pollutants, with different parameters applied in the CEEMD model.

Table 9

The results of the system with different parameters in CEEMD

EN	Metric	$PM_{2.5}$	PM_{10}	SO_2	Average
200	**FE**	0.6681	0.6223	0.1134	0.4679
	Time	63.8535	72.2888	68.8987	68.3470
300	**FE**	0.5946	1.0040	0.0983	0.5656
	Time	96.4800	103.1036	100.7629	100.1155
400	**FE**	0.3171	0.9887	0.1575	0.4878
	Time	124.3598	138.2738	136.2668	132.9668
500	**FE**	0.5756	0.4812	0.0463	0.3677
	Time	155.5923	171.8197	175.7477	167.7199
600	**FE**	0.5802	0.8686	0.0679	0.5056
	Time	184.5525	205.5541	191.6648	193.9238

Note: EN represents ensemble numbers and FE represents the forecasting error in table.

For the forecasting error, the smallest average forecasting error was 0.3677, when ensemble number was set to 500; furthermore, the forecasting error was minimal in the three pollutant forecasting attempts. As for computation time, it ranges from 68.3470 s to 193.9238 s, as the complexity of the model and the computation time increased with the addition of ensemble numbers.

Remark: From the analysis above, we can draw the conclusion that too many IMFs may lead to model complexity and computational cost. Furthermore, modeling too many IMFs cannot always generate satisfying final results because of the estimation error of each IMF, which can accumulate in the ensemble forecasting step. To avoid these problems, this research set the optimal ensemble number to 500 so that the

forecasting error is the smallest and the computational time is relatively short.

5.2 Discussion of the Partition Intervals Method

Partitioning discrete discourse is a significant step in FTS forecasting. Too many intervals will result in complex FLR and make it difficult to construct a weight matrix, whereas too few intervals will lead to poor forecasting accuracy. In addition, the accuracy of fuzzy time series model forecasting is invariably affected by interval length and it is difficult to formulate proper intervals. Determining the distance partitioning with the equal width can easily result in either excessive linguistic values or excessively short intervals which can lead to the generation of null sets among the FLRs [70]. Therefore, partitioning discrete discourse correctly plays a crucial role in improving forecasting accuracy. The EWP method uses semantic conventions to divide the universe of discourse into seven equal-width intervals. Due to the disadvantages of not being able to reasonably change the number and the length of the intervals, this easily lead to poor forecasting accuracy. In contrast, the EBD algorithm determines the length and number of intervals adaptively based on the principle of the smallest information entropy. The searching breakpoint process is applied recursively in each partition, and the process terminates when there is no need to search for the breakpoint so that the linguistic values close to a steady state belong to the fuzzy set. This work applied the EBD method to partition the discrete discourse in FTS forecasting compared with the EWP method. Table 10 presents two types of partitioning: EWP and EBD, and the forecasting accuracy.

Table 10

The forecasting results with different partition method

Pollutant	Model	MAPE	MAE	RMSE	MdAPE
$PM_{2.5}$	EWP	35.8681	15.9090	19.1226	0.2766
	EBD	5.7417	2.8683	3.7675	0.0465
PM_{10}	EWP	29.1824	26.5933	32.6664	0.2614
	EBD	4.1643	3.7848	4.8963	0.0359
SO_2	EWP	50.3800	7.5865	8.9273	1.7281
	EBD	6.5141	0.5874	0.8055	0.0415

From Table 9, the EBD algorithm had the best MAPE of 5.7417%, 4.1643% and 6.5141%, respectively, for $PM_{2.5}$, PM_{10} and SO_2 forecasting, and the accuracy was significantly improved compared with EWP. Furthermore, the EBD algorithm almost outperformed EWP in other metrics. Therefore, the EBD algorithm had successful application in partitioning the universe of discourse and in FTS forecasting.

Remark: For the forecasting of FTS with large data fluctuation, the application of the EBD algorithm to partitioning the universe of discourse can adaptively determine the number and length of the intervals, which is propitious for improving the forecasting accuracy.

5.3 The Forecasting Effectiveness and Stability of the System

Forecasting accuracy and good stability are two important factors for evaluating air quality early warning systems. An excellent early warning system for air quality with high forecasting precision and stability can provide strong informational support for the improvement of air quality and the treatment of air pollution. Forecasting stability often reflects the fluctuation of model forecasting accuracy. Variance, as an important measure of data fluctuation, can be used to demonstrate the forecasting stability of the model. Furthermore, forecasting error is a key indicator used to evaluate forecasting performance. Therefore, the variance in the forecasting error could be used to verify the forecasting stability of the model. This section focuses on forecasting accuracy and stability to verify the superiority of the proposed model compared with

other benchmark models, and the forecasting results can be seen in Table 11.

Table 11

The results about the forecasting accuracy and variance of the forecasting error

Pollutant	Metric	ENN	BPNN	ARIMA	EW*	CEW*	EBD*	Proposed
$PM_{2.5}$	MAPE	69.0886	8.3329	8.4904	47.4451	35.8681	7.8844	5.7417
	VAR	31.0237	19.3626	14.9268	479.6940	423.8471	26.8369	12.8580
PM_{10}	MAPE	57.6701	7.4577	9.1261	41.2877	29.1824	5.7242	4.1643
	VAR	1161.4078	356.9013	237.2520	983.3519	982.8329	26.4425	17.4904
SO_2	MAPE	91.9033	17.7112	12.5623	96.2759	50.3800	11.0574	6.5141
	VAR	4.3676	4.8201	0.5186	4.8979	4.8203	0.1305	0.0593

Based on the analysis of the previous experimental results, the accuracy of the proposed model was proven; to further verify the effectiveness of the system, all of the comparison models were combined in a unified analysis. This can be seen in Table 11, wherein the proposed model outperforms the other six comparison models with MAPE values of 5.7417%, 4.1643% and 6.5141%, respectively. For the forecasting stability, the variance values of the proposed model are smaller than those of the compared benchmark models, indicating that the proposed model is more stable than the other benchmark models.

5.4 The Real Application of the Proposed Early Warning System

The proposed early warning system possesses many practical applications, such as mining the characteristics of air pollutants, warning and guiding the public before the occurrence of hazardous air pollutants, etc. Additionally, it consists of two kinds prediction method: deterministic prediction and uncertainty prediction. And they both have their own practical applications.

1) The deterministic prediction provides accurate and reliable warning information by mining and forecasting air pollutants. The proposed hybrid model can be applied to forecast the future value of pollutant concentration, which can not only help environmental policy makers take effective protection measures before the occurrence of hazardous air pollutants but also provide useful guidance for people's daily lives combined with AQI index [71].
2) In the developed early warning system, an uncertainty analysis module is set up, which has capability to provide more effective and credible information than point forecasting through scientifically forecast the future range of pollutant concentration. The interval forecasting provides predictive ranges and confidence levels so that the speed and degree of diffusion of pollutants can be analyzed. When the concentration of pollutants exceeds the standard, the early warning system will make an alarm and the air quality supervision department can promptly make relevant prevention and control measures. At the same time, it also provides intuitive guidance for residents [72].

6 Conclusion and Future Work

Air pollution, which is a great threat to the economy, environment and human health, has become a major global problem. In recent years, many cities have not been able to get rid of the threat of air pollution, especially cities that have experienced rapid industrial development, such as Beijing, Shijiazhuang, Tianjin and so on. Consequently, it is worthwhile to scientifically forecast air pollutant concentrations to provide the public with sufficient information and time to respond to incoming air pollution.

This paper developed an effective and reliable hybrid air quality forecasting and early warning system to project the concentrations of three major pollutants. This proposed early warning system consists of three modules: a deterministic prediction module, an uncertainty analysis module and an assessment module. Specifically, the

experimental results of the deterministic module reveal that the proposed model, which served to perform target points forecasting, can remarkably enhance accuracy compared with benchmarks. Afterwards, in an analysis module, the experimental results illustrated the uncertainty information involved in future forecasts under different confidence levels. Finally, the assessment module provided comprehensive evaluation of the system and proved the effectiveness and robustness of the proposed hybrid model. In summary, the experimental results demonstrate that the proposed early warning system obtained the best performance, with high forecasting accuracy, robustness and stability, which suggests that it will be a useful tool for analyzing and monitoring air pollution. Its excellent performance reveals that it can also be applied to other fields, such as power-load forecasting, stock-price forecasting, wind-speed forecasting and traffic-flow forecasting.

Inspired by related literature, the pollutant data may have chaotic characteristic that leads to unsatisfactory performance [40]. Furthermore, if the fuzzy logic relationship can achieve adaptive clustering, membership functions will be easily formed to improve the prediction accuracy [73]. Therefore, solving the problem of forecasting time series with chaotic characteristics as well as the application of fuzzy logic adaptive clustering are researches focus for the future.

Acknowledgements

This work was supported by Major Program of National Social Science Foundation of China (Grant No.17ZDA093).

Appendix A.

A.1.

List of abbreviations			
PM	particulate matter	FTS	fuzzy time series
AQI	air quality index	FLR	fuzzy logical relation
Jing-Jin-Ji	Beijing-Tianjin-Hebei region	FLRG	fuzzy logical relation group
AR	autoregressive model	EBD	entropy-based discretization
ARMA	autoregressive moving average model	EWP	equal-width pre-partitioning
ARIMA	autoregressive integrated moving average model	EDP	equal-depth partitioning
MLR	multiple linear regression	ANFIS	adaptive fuzzy inference system
SVR	support vector regression	MAPE	mean absolute percentage error
CTMs	chemical transport models	MAE	mean absolute error
ANNs	artificial neural networks	RMSE	root mean square error
ENN	elman neural network	MdAPE	median absolute percentage error
BPNN	back propagation neural network	DA	direction accuracy
RBFNN	radial basis function neural network	FB	fractional bias
EMD	empirical mode decomposition	IA	index of agreement
EEMD	ensemble empirical mode decomposition	*r*	Pearson's correlation coefficient
CEEMD	complementary ensemble empirical mode decomposition	IFCP	interval forecasting coverage probability
IMFs	intrinsic mode functions	IFAW	interval forecasting average width

A.2.

FLR of $PM_{2.5}$ from the first site (Beijing)

	$PM_{2.5}$									
Day	**1st**	**2nd**	**3th**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**
0:00	A9	A15	A9	A10	A14	A16	A15	A15	A16	A21
1:00	A9	A15	A9	A10	A14	A15	A12	A14	A16	A23
2:00	A10	A14	A10	A14	A15	A14	A9	A15	A16	A24
3:00	A10	A10	A9	A15	A15	A13	A9	A14	A18	A24
4:00	A10	A10	A9	A15	A15	A12	A9	A14	A20	A24
5:00	A10	A10	A9	A15	A15	A13	A9	A14	A21	A24
6:00	A10	A10	A9	A15	A15	A14	A9	A15	A21	A26
7:00	A10	A14	A9	A15	A14	A14	A10	A15	A21	A26
8:00	A12	A15	A9	A15	A14	A15	A10	A16	A20	A29
9:00	A10	A15	A9	A15	A15	A15	A13	A18	A20	A29
10:00	A9	A14	A9	A14	A15	A16	A13	A20	A20	A28
11:00	A9	A10	A9	A12	A15	A16	A14	A21	A21	A26
12:00	A8	A8	A9	A10	A15	A18	A14	A21	A21	A26
13:00	A8	A9	A9	A10	A15	A18	A14	A21	A21	A24
14:00	A8	A9	A10	A10	A15	A18	A14	A21	A21	A24
15:00	A8	A9	A10	A10	A15	A16	A14	A21	A21	A21
16:00	A9	A9	A10	A12	A15	A16	A15	A21	A21	A21
17:00	A10	A10	A12	A14	A16	A18	A15	A21	A21	A21
18:00	A12	A9	A14	A15	A16	A16	A15	A21	A21	A21
19:00	A10	A9	A14	A15	A16	A18	A15	A21	A24	A21
20:00	A10	A9	A15	A15	A16	A16	A15	A21	A24	A21
21:00	A10	A9	A15	A15	A16	A16	A15	A21	A24	A21
22:00	A13	A9	A15	A14	A16	A15	A15	A21	A21	A24
23:00	A15	A9	A13	A14	A16	A15	A15	A21	A21	A24

A.3.

FLR of PM_{10} from the first site (Beijing)

	PM_{10}									
Day	**1st**	**2nd**	**3th**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**
0:00	A10	A16	A14	A12	A16	A12	A11	A14	A19	A22
1:00	A10	A16	A14	A14	A17	A12	A10	A12	A17	A22
2:00	A11	A14	A15	A16	A17	A12	A9	A12	A16	A22
3:00	A12	A14	A15	A17	A17	A12	A6	A12	A17	A22
4:00	A12	A16	A14	A19	A17	A10	A6	A12	A19	A22
5:00	A12	A16	A14	A19	A16	A10	A9	A16	A20	A23
6:00	A12	A19	A14	A19	A15	A10	A12	A17	A22	A24
7:00	A14	A23	A15	A17	A14	A12	A14	A19	A22	A27
8:00	A13	A31	A14	A16	A14	A14	A15	A19	A20	A31
9:00	A12	A32	A15	A15	A15	A16	A16	A20	A19	A31
10:00	A10	A24	A15	A16	A14	A16	A16	A20	A20	A27
11:00	A10	A18	A14	A16	A14	A15	A16	A19	A20	A24
12:00	A11	A12	A14	A15	A13	A14	A17	A19	A20	A22
13:00	A12	A12	A14	A14	A13	A14	A17	A20	A20	A20
14:00	A12	A12	A15	A14	A14	A14	A17	A20	A20	A19
15:00	A12	A15	A16	A14	A12	A14	A17	A20	A19	A19
16:00	A12	A16	A17	A14	A12	A13	A17	A20	A19	A20
17:00	A14	A16	A18	A16	A14	A14	A18	A19	A20	A20
18:00	A14	A15	A17	A16	A16	A15	A18	A19	A22	A20
19:00	A14	A14	A17	A16	A16	A15	A17	A20	A22	A19

20:00	A14	A14	A17	A16	A17	A15	A16	A22	A22	A19
21:00	A12	A14	A16	A16	A16	A14	A16	A24	A22	A20
22:00	A14	A14	A14	A16	A14	A14	A16	A24	A22	A22
23:00	A15	A14	A12	A16	A12	A12	A16	A22	A22	A22

A.4.
FLR of SO_2 from the first site (Beijing)

					SO_2					
Day	**1st**	**2nd**	**3th**	**4th**	**5th**	**6th**	**7th**	**8th**	**9th**	**10th**
0:00	A3	A12	A15	A12	A18	A8	A3	A4	A10	A18
1:00	A5	A11	A15	A13	A18	A6	A3	A4	A9	A18
2:00	A8	A9	A13	A15	A18	A6	A2	A3	A9	A18
3:00	A8	A8	A12	A15	A18	A6	A2	A2	A11	A18
4:00	A8	A8	A10	A15	A17	A6	A2	A3	A12	A17
5:00	A8	A9	A9	A13	A17	A5	A2	A3	A13	A15
6:00	A7	A11	A8	A12	A17	A3	A2	A4	A12	A15
7:00	A8	A12	A8	A9	A17	A3	A2	A6	A12	A15
8:00	A8	A12	A8	A8	A17	A3	A2	A9	A9	A15
9:00	A6	A13	A9	A8	A17	A4	A2	A10	A9	A15
10:00	A6	A13	A11	A9	A15	A5	A3	A12	A9	A13
11:00	A5	A15	A11	A10	A15	A4	A4	A12	A10	A12
12:00	A5	A17	A11	A12	A13	A5	A5	A12	A11	A11
13:00	A5	A19	A10	A12	A12	A6	A4	A12	A12	A12
14:00	A3	A20	A9	A12	A12	A5	A4	A12	A12	A15
15:00	A3	A20	A9	A13	A12	A4	A5	A12	A12	A17
16:00	A2	A20	A9	A18	A12	A4	A4	A10	A12	A18
17:00	A2	A19	A9	A20	A12	A4	A4	A9	A12	A18
18:00	A4	A18	A11	A22	A11	A4	A5	A11	A13	A18
19:00	A5	A18	A12	A22	A10	A5	A6	A12	A15	A18
20:00	A6	A17	A13	A22	A10	A6	A6	A13	A17	A18
21:00	A8	A15	A15	A21	A9	A6	A5	A13	A17	A18
22:00	A8	A13	A13	A20	A9	A5	A4	A12	A18	A20
23:00	A9	A15	A12	A19	A9	A4	A4	A12	A18	A20

A.5.
FLGs of $PM_{2.5}$ from the first site (Beijing)

PM2.5														
A1	→	A1	A10	→	A15	A18	→	A16	A24	→	A26	A31	→	A33
A1	→	A3	A10	→	A16	A18	→	A18	A24	→	A28	A31	→	A35
A1	→	A5	A11	→	A10	A18	→	A19	A25	→	A21	A32	→	A30
A1	→	A20	A12	→	A5	A18	→	A20	A25	→	A22	A32	→	A31
A2	→	A1	A12	→	A9	A18	→	A21	A25	→	A24	A32	→	A32
A2	→	A3	A12	→	A10	A18	→	A30	A25	→	A25	A32	→	A33
A3	→	A1	A12	→	A12	A19	→	A12	A25	→	A26	A32	→	A35
A3	→	A3	A12	→	A13	A19	→	A15	A26	→	A15	A32	→	A37
A3	→	A4	A12	→	A14	A19	→	A16	A26	→	A20	A33	→	A26
A3	→	A5	A12	→	A15	A19	→	A19	A26	→	A21	A33	→	A29
A4	→	A3	A13	→	A9	A19	→	A20	A26	→	A23	A33	→	A30
A4	→	A4	A13	→	A10	A19	→	A21	A26	→	A24	A33	→	A31
A4	→	A5	A13	→	A11	A20	→	A1	A26	→	A25	A33	→	A32
A5	→	A2	A13	→	A12	A20	→	A9	A26	→	A26	A33	→	A33
A5	→	A3	A13	→	A13	A20	→	A10	A26	→	A27	A33	→	A34
A5	→	A4	A13	→	A14	A20	→	A14	A26	→	A28	A33	→	A35

A5	→	A5	A13	→	A15	A20	→	A15	A26	→	A29	A33	→	A36
A5	→	A6	A13	→	A16	A20	→	A16	A26	→	A30	A34	→	A31
A5	→	A7	A13	→	A20	A20	→	A18	A26	→	A33	A34	→	A33
A5	→	A8	A14	→	A8	A20	→	A19	A27	→	A24	A34	→	A34
A5	→	A9	A14	→	A9	A20	→	A20	A27	→	A26	A34	→	A35
A5	→	A10	A14	→	A10	A20	→	A21	A27	→	A28	A34	→	A36
A6	→	A8	A14	→	A12	A20	→	A24	A27	→	A29	A35	→	A24
A7	→	A3	A14	→	A13	A20	→	A32	A27	→	A33	A35	→	A29
A7	→	A4	A14	→	A14	A21	→	A8	A28	→	A24	A35	→	A30
A7	→	A5	A14	→	A15	A21	→	A9	A28	→	A26	A35	→	A31
A7	→	A8	A14	→	A16	A21	→	A10	A28	→	A28	A35	→	A33
A7	→	A9	A14	→	A18	A21	→	A13	A28	→	A29	A35	→	A34
A8	→	A1	A15	→	A5	A21	→	A15	A28	→	A30	A35	→	A35
A8	→	A5	A15	→	A7	A21	→	A16	A29	→	A24	A35	→	A36
A8	→	A7	A15	→	A8	A21	→	A18	A29	→	A26	A36	→	A26
A8	→	A8	A15	→	A9	A21	→	A19	A29	→	A27	A36	→	A29
A8	→	A9	A15	→	A10	A21	→	A20	A29	→	A28	A36	→	A30
A8	→	A10	A15	→	A11	A21	→	A21	A29	→	A29	A36	→	A31
A8	→	A13	A15	→	A12	A21	→	A22	A29	→	A30	A36	→	A33
A8	→	A14	A15	→	A13	A21	→	A23	A29	→	A31	A36	→	A34
A9	→	A1	A15	→	A14	A21	→	A24	A29	→	A33	A36	→	A35
A9	→	A3	A15	→	A15	A21	→	A25	A29	→	A34	A36	→	A36
A9	→	A5	A15	→	A16	A21	→	A26	A29	→	A35	A36	→	A37
A9	→	A6	A15	→	A17	A21	→	A27	A30	→	A21	A36	→	A38
A9	→	A7	A15	→	A18	A22	→	A21	A30	→	A24	A37	→	A30
A9	→	A8	A15	→	A20	A23	→	A15	A30	→	A26	A37	→	A31
A9	→	A9	A15	→	A21	A23	→	A20	A30	→	A28	A37	→	A33
A9	→	A10	A16	→	A9	A23	→	A21	A30	→	A29	A37	→	A35
A9	→	A12	A16	→	A10	A23	→	A22	A30	→	A30	A37	→	A36
A9	→	A13	A16	→	A12	A23	→	A23	A30	→	A31	A37	→	A37
A9	→	A14	A16	→	A13	A23	→	A24	A30	→	A32	A37	→	A38
A9	→	A15	A16	→	A14	A23	→	A26	A30	→	A33	A37	→	A39
A9	→	A16	A16	→	A15	A24	→	A9	A30	→	A34	A38	→	A36
A9	→	A18	A16	→	A16	A24	→	A10	A30	→	A35	A38	→	A37
A10	→	A1	A16	→	A18	A24	→	A15	A30	→	A36	A38	→	A38
A10	→	A5	A16	→	A19	A24	→	A18	A30	→	A38	A38	→	A39
A10	→	A7	A16	→	A20	A24	→	A19	A31	→	A20	A38	→	A40
A10	→	A8	A16	→	A21	A24	→	A20	A31	→	A26	A39	→	A37
A10	→	A9	A16	→	A24	A24	→	A21	A31	→	A28	A39	→	A38
A10	→	A10	A17	→	A21	A24	→	A22	A31	→	A29	A39	→	A39
A10	→	A12	A18	→	A13	A24	→	A23	A31	→	A30	A39	→	A40
A10	→	A13	A18	→	A14	A24	→	A24	A31	→	A31	A40	→	A39
A10	→	A14	A18	→	A15	A24	→	A25	A31	→	A32	A40	→	A40

A.6.

FLGs of PM_{10} from the first site (Beijing)

PM_{10}											
A1	→	A1	A11	→	A9	A16	→	A19	A23	→	A21	A29	→	A27
A1	→	A12	A11	→	A10	A17	→	A12	A23	→	A22	A29	→	A31
A2	→	A2	A11	→	A11	A17	→	A14	A23	→	A23	A30	→	A27
A2	→	A3	A11	→	A12	A17	→	A15	A23	→	A24	A30	→	A28
A2	→	A4	A11	→	A13	A17	→	A16	A23	→	A25	A30	→	A31

A2	→	A5	A11	→	A14	A17	→	A17	A23	→	A27	A31	→	A19
A2	→	A14	A11	→	A15	A17	→	A18	A23	→	A31	A31	→	A24
A3	→	A1	A11	→	A16	A17	→	A19	A24	→	A13	A31	→	A25
A3	→	A2	A12	→	A6	A17	→	A20	A24	→	A15	A31	→	A26
A3	→	A3	A12	→	A9	A17	→	A23	A24	→	A16	A31	→	A27
A3	→	A4	A12	→	A10	A17	→	A27	A24	→	A18	A31	→	A28
A4	→	A2	A12	→	A11	A18	→	A12	A24	→	A19	A31	→	A29
A4	→	A3	A12	→	A12	A18	→	A14	A24	→	A20	A31	→	A30
A4	→	A4	A12	→	A14	A18	→	A15	A24	→	A22	A31	→	A31
A4	→	A5	A12	→	A15	A18	→	A16	A24	→	A23	A31	→	A32
A4	→	A6	A12	→	A16	A18	→	A17	A24	→	A24	A31	→	A33
A5	→	A4	A12	→	A17	A18	→	A18	A24	→	A25	A31	→	A35
A5	→	A5	A12	→	A34	A18	→	A19	A24	→	A26	A32	→	A24
A5	→	A6	A13	→	A2	A18	→	A20	A24	→	A27	A32	→	A27
A5	→	A7	A13	→	A10	A18	→	A22	A24	→	A31	A32	→	A31
A6	→	A2	A13	→	A12	A19	→	A12	A24	→	A33	A32	→	A32
A6	→	A3	A13	→	A13	A19	→	A14	A25	→	A19	A32	→	A33
A6	→	A4	A13	→	A14	A19	→	A15	A25	→	A22	A32	→	A35
A6	→	A5	A13	→	A15	A19	→	A16	A25	→	A23	A33	→	A24
A6	→	A6	A14	→	A6	A19	→	A17	A25	→	A24	A33	→	A27
A6	→	A7	A14	→	A7	A19	→	A18	A25	→	A25	A33	→	A28
A6	→	A8	A14	→	A10	A19	→	A19	A25	→	A26	A33	→	A31
A6	→	A9	A14	→	A11	A19	→	A20	A25	→	A27	A33	→	A32
A6	→	A10	A14	→	A12	A19	→	A22	A26	→	A20	A33	→	A33
A6	→	A11	A14	→	A13	A19	→	A23	A26	→	A23	A33	→	A34
A6	→	A12	A14	→	A14	A19	→	A24	A26	→	A24	A33	→	A35
A7	→	A6	A14	→	A15	A19	→	A26	A26	→	A25	A34	→	A31
A7	→	A7	A14	→	A16	A20	→	A17	A26	→	A26	A34	→	A33
A7	→	A9	A14	→	A17	A20	→	A18	A26	→	A27	A34	→	A34
A7	→	A10	A14	→	A18	A20	→	A19	A26	→	A28	A34	→	A38
A7	→	A11	A14	→	A19	A20	→	A20	A26	→	A35	A35	→	A24
A8	→	A6	A14	→	A20	A20	→	A22	A27	→	A22	A35	→	A27
A8	→	A9	A14	→	A24	A20	→	A23	A27	→	A23	A35	→	A31
A9	→	A6	A15	→	A2	A20	→	A24	A27	→	A24	A35	→	A32
A9	→	A7	A15	→	A11	A20	→	A27	A27	→	A25	A35	→	A33
A9	→	A8	A15	→	A12	A21	→	A18	A27	→	A26	A35	→	A34
A9	→	A9	A15	→	A14	A21	→	A20	A27	→	A27	A35	→	A35
A9	→	A10	A15	→	A15	A22	→	A17	A27	→	A28	A35	→	A36
A9	→	A11	A15	→	A16	A22	→	A18	A27	→	A29	A36	→	A35
A9	→	A12	A15	→	A17	A22	→	A19	A27	→	A30	A36	→	A36
A10	→	A3	A16	→	A6	A22	→	A20	A27	→	A31	A36	→	A37
A10	→	A6	A16	→	A10	A22	→	A21	A27	→	A32	A36	→	A38
A10	→	A7	A16	→	A11	A22	→	A22	A27	→	A35	A37	→	A36
A10	→	A9	A16	→	A12	A22	→	A23	A28	→	A24	A37	→	A37
A10	→	A10	A16	→	A13	A22	→	A24	A28	→	A27	A38	→	A36
A10	→	A11	A16	→	A14	A22	→	A26	A28	→	A28	A38	→	A38
A10	→	A12	A16	→	A15	A22	→	A27	A28	→	A30			
A10	→	A14	A16	→	A16	A22	→	A28	A28	→	A31			
A11	→	A6	A16	→	A17	A23	→	A19	A28	→	A33			
A11	→	A7	A16	→	A18	A23	→	A20	A28	→	A35			

A.7.
FLGs of SO_2 from the first site (Beijing)

SO_2												
A1	→	A1	A6	→	A8	A10	→	A8	A13	→	A20	A18	→	A21
A1	→	A2	A6	→	A9	A10	→	A9	A14	→	A12	A19	→	A12
A1	→	A3	A6	→	A10	A10	→	A10	A14	→	A13	A19	→	A15
A2	→	A1	A6	→	A12	A10	→	A11	A14	→	A14	A19	→	A17
A2	→	A2	A6	→	A13	A10	→	A12	A14	→	A15	A19	→	A18
A2	→	A3	A7	→	A3	A10	→	A13	A14	→	A17	A19	→	A19
A2	→	A4	A7	→	A5	A11	→	A2	A15	→	A9	A19	→	A20
A2	→	A5	A7	→	A6	A11	→	A6	A15	→	A10	A19	→	A21
A2	→	A6	A7	→	A8	A11	→	A8	A15	→	A12	A20	→	A12
A3	→	A1	A7	→	A9	A11	→	A9	A15	→	A13	A20	→	A13
A3	→	A2	A7	→	A10	A11	→	A10	A15	→	A14	A20	→	A15
A3	→	A3	A7	→	A11	A11	→	A11	A15	→	A15	A20	→	A17
A3	→	A4	A8	→	A2	A11	→	A12	A15	→	A16	A20	→	A18
A3	→	A5	A8	→	A3	A11	→	A13	A15	→	A17	A20	→	A19
A3	→	A6	A8	→	A4	A11	→	A15	A15	→	A18	A20	→	A20
A3	→	A8	A8	→	A5	A12	→	A1	A15	→	A19	A20	→	A21
A3	→	A12	A8	→	A6	A12	→	A3	A15	→	A20	A20	→	A22
A4	→	A2	A8	→	A7	A12	→	A4	A16	→	A15	A20	→	A23
A4	→	A3	A8	→	A8	A12	→	A6	A16	→	A16	A21	→	A13
A4	→	A4	A8	→	A9	A12	→	A7	A16	→	A17	A21	→	A17
A4	→	A5	A8	→	A10	A12	→	A8	A16	→	A18	A21	→	A19
A4	→	A6	A8	→	A11	A12	→	A9	A17	→	A8	A21	→	A20
A4	→	A7	A8	→	A12	A12	→	A10	A17	→	A9	A21	→	A21
A4	→	A8	A8	→	A13	A12	→	A11	A17	→	A11	A21	→	A22
A4	→	A9	A8	→	A15	A12	→	A12	A17	→	A12	A21	→	A23
A4	→	A12	A9	→	A2	A12	→	A13	A17	→	A13	A22	→	A19
A5	→	A2	A9	→	A3	A12	→	A14	A17	→	A14	A22	→	A20
A5	→	A3	A9	→	A4	A12	→	A15	A17	→	A15	A22	→	A21
A5	→	A4	A9	→	A5	A12	→	A17	A17	→	A16	A22	→	A22
A5	→	A5	A9	→	A6	A13	→	A1	A17	→	A17	A22	→	A23
A5	→	A6	A9	→	A7	A13	→	A5	A17	→	A18	A23	→	A20
A5	→	A7	A9	→	A8	A13	→	A6	A17	→	A19	A23	→	A21
A5	→	A8	A9	→	A9	A13	→	A11	A17	→	A20	A23	→	A22
A5	→	A9	A9	→	A10	A13	→	A12	A17	→	A21	A23	→	A23
A5	→	A12	A9	→	A11	A13	→	A13	A18	→	A13	A23	→	A24
A6	→	A3	A9	→	A12	A13	→	A14	A18	→	A15	A24	→	A23
A6	→	A4	A9	→	A13	A13	→	A15	A18	→	A17	A24	→	A24
A6	→	A5	A10	→	A5	A13	→	A16	A18	→	A18			
A6	→	A6	A10	→	A6	A13	→	A17	A18	→	A19			
A6	→	A7	A10	→	A7	A13	→	A18	A18	→	A20			

A.8. Forecasted output of the first site (Beijing)

Time	$PM_{2.5}$		PM_{10}		SO_2		$PM_{2.5}$		PM_{10}		SO_2	
	2017.07.22						**2017.07.27**					
0:00	A7	13.41	A12	60.11	A2	1.93	A9	21.61	A9	38.31	A2	1.93
1:00	A8	16.93	A10	45.93	A2	1.93	A9	21.61	A10	45.93	A2	1.93
2:00	A8	16.93	A10	45.93	A2	1.93	A9	21.61	A10	45.93	A3	2.55
3:00	A5	11.00	A7	33.98	A2	1.93	A10	26.81	A10	45.93	A3	2.55
4:00	A5	11.00	A6	28.19	A2	1.93	A10	26.81	A10	45.93	A3	2.55
5:00	A8	16.93	A4	18.70	A2	1.93	A12	29.95	A10	45.93	A3	2.55
6:00	A8	16.93	A6	28.19	A2	1.93	A13	31.49	A11	51.93	A2	1.93
7:00	A9	21.61	A6	28.19	A2	1.93	A14	34.93	A11	51.93	A2	1.93

8:00	A9	21.61	A6	28.19	A2	1.93	A14	34.93	A11	51.93	A2	1.93
9:00	A9	21.61	A6	28.19	A2	1.93	A14	34.93	A11	51.93	A2	1.93
10:00	A9	21.61	A6	28.19	A2	1.93	A14	34.93	A11	51.93	A2	1.93
11:00	A9	21.61	A6	28.19	A3	2.55	A15	39.82	A12	60.11	A2	1.93
12:00	A9	21.61	A6	28.19	A3	2.55	A15	39.82	A12	60.11	A2	1.93
13:00	A9	21.61	A6	28.19	A3	2.55	A15	39.82	A12	60.11	A3	2.55
14:00	A9	21.61	A6	28.19	A2	1.93	A15	39.82	A14	68.78	A3	2.55
15:00	A9	21.61	A6	28.19	A2	1.93	A15	39.82	A15	74.30	A4	3.05
16:00	A8	16.93	A6	28.19	A2	1.93	A15	39.82	A15	74.30	A4	3.05
17:00	A8	16.93	A6	28.19	A2	1.93	A16	47.45	A16	84.69	A4	3.05
18:00	A8	16.93	A9	38.31	A2	1.93	A16	47.45	A16	84.69	A4	3.05
19:00	A9	21.61	A9	38.31	A2	1.93	A16	47.45	A16	84.69	A4	3.05
20:00	A10	26.81	A11	51.93	A2	1.93	A16	47.45	A16	84.69	A4	3.05
21:00	A10	26.81	A12	60.11	A2	1.93	A18	53.83	A16	84.69	A4	3.05
22:00	A10	26.81	A12	60.11	A2	1.93	A18	53.83	A17	97.27	A4	3.05
23:00	A10	26.81	A10	45.93	A2	1.93	A18	53.83	A16	84.69	A4	3.05
	2017.07.23						**2017.07.28**					
0:00	A9	21.61	A9	38.31	A2	1.93	A16	47.45	A16	84.69	A4	3.05
1:00	A9	21.61	A9	38.31	A2	1.93	A16	47.45	A16	84.69	A4	3.05
2:00	A9	21.61	A9	38.31	A2	1.93	A15	39.82	A16	84.69	A3	2.55
3:00	A9	21.61	A9	38.31	A2	1.93	A15	39.82	A16	84.69	A3	2.55
4:00	A9	21.61	A9	38.31	A2	1.93	A15	39.82	A16	84.69	A2	1.93
5:00	A10	26.81	A10	45.93	A2	1.93	A15	39.82	A15	74.30	A2	1.93
6:00	A10	26.81	A10	45.93	A2	1.93	A14	34.93	A12	60.11	A2	1.93
7:00	A10	26.81	A10	45.93	A2	1.93	A10	26.81	A10	45.93	A2	1.93
8:00	A10	26.81	A10	45.93	A2	1.93	A10	26.81	A10	45.93	A2	1.93
9:00	A10	26.81	A10	45.93	A2	1.93	A9	21.61	A11	51.93	A2	1.93
10:00	A10	26.81	A10	45.93	A2	1.93	A9	21.61	A12	60.11	A2	1.93
11:00	A14	34.93	A11	51.93	A3	2.55	A10	26.81	A14	68.78	A2	1.93
12:00	A15	39.82	A12	60.11	A3	2.55	A10	26.81	A14	68.78	A3	2.55
13:00	A15	39.82	A14	68.78	A4	3.05	A10	26.81	A12	60.11	A4	3.05
14:00	A16	47.45	A14	68.78	A4	3.05	A9	21.61	A10	45.93	A3	2.55
15:00	A16	47.45	A14	68.78	A3	2.55	A9	21.61	A7	33.98	A3	2.55
16:00	A18	53.83	A14	68.78	A2	1.93	A8	16.93	A6	28.19	A2	1.93
17:00	A18	53.83	A12	60.11	A2	1.93	A8	16.93	A6	28.19	A2	1.93
18:00	A20	56.30	A12	60.11	A2	1.93	A8	16.93	A9	38.31	A2	1.93
19:00	A18	53.83	A12	60.11	A2	1.93	A9	21.61	A10	45.93	A2	1.93
20:00	A16	47.45	A12	60.11	A2	1.93	A9	21.61	A12	60.11	A2	1.93
21:00	A15	39.82	A12	60.11	A2	1.93	A10	26.81	A14	68.78	A2	1.93
22:00	A15	39.82	A12	60.11	A2	1.93	A10	26.81	A15	74.30	A2	1.93
23:00	A15	39.82	A12	60.11	A2	1.93	A14	34.93	A16	84.69	A3	2.55
	2017.07.24						**2017.07.29**					
0:00	A18	53.83	A14	68.78	A2	1.93	A15	39.82	A16	84.69	A3	2.55
1:00	A21	65.35	A16	84.69	A2	1.93	A14	34.93	A14	68.78	A3	2.55
2:00	A21	65.35	A17	97.27	A2	1.93	A14	34.93	A12	60.11	A2	1.93
3:00	A21	65.35	A17	97.27	A2	1.93	A12	29.95	A12	60.11	A2	1.93
4:00	A21	65.35	A17	97.27	A3	2.55	A10	26.81	A12	60.11	A2	1.93
5:00	A23	75.12	A17	97.27	A3	2.55	A10	26.81	A12	60.11	A2	1.93
6:00	A21	65.35	A16	84.69	A3	2.55	A10	26.81	A12	60.11	A2	1.93
7:00	A21	65.35	A16	84.69	A2	1.93	A10	26.81	A12	60.11	A2	1.93
8:00	A21	65.35	A15	74.30	A2	1.93	A10	26.81	A12	60.11	A2	1.93
9:00	A20	56.30	A14	68.78	A2	1.93	A10	26.81	A14	68.78	A2	1.93
10:00	A16	47.45	A12	60.11	A2	1.93	A14	34.93	A15	74.30	A2	1.93
11:00	A15	39.82	A12	60.11	A2	1.93	A15	39.82	A16	84.69	A2	1.93

12:00	A15	39.82	A12	60.11	A2	1.93	A16	47.45	A16	84.69	A3	2.55
13:00	A15	39.82	A14	68.78	A2	1.93	A18	53.83	A16	84.69	A4	3.05
14:00	A16	47.45	A14	68.78	A2	1.93	A18	53.83	A15	74.30	A4	3.05
15:00	A16	47.45	A14	68.78	A2	1.93	A15	39.82	A13	55.56	A3	2.55
16:00	A15	39.82	A12	60.11	A2	1.93	A15	39.82	A12	60.11	A3	2.55
17:00	A14	34.93	A10	45.93	A2	1.93	A13	31.49	A10	45.93	A2	1.93
18:00	A10	26.81	A10	45.93	A2	1.93	A10	26.81	A10	45.93	A2	1.93
19:00	A10	26.81	A9	38.31	A2	1.93	A10	26.81	A10	45.93	A2	1.93
20:00	A9	21.61	A10	45.93	A2	1.93	A9	21.61	A10	45.93	A2	1.93
21:00	A9	21.61	A9	38.31	A2	1.93	A9	21.61	A10	45.93	A2	1.93
22:00	A9	21.61	A9	38.31	A2	1.93	A9	21.61	A11	51.93	A2	1.93
23:00	A9	21.61	A9	38.31	A2	1.93	A10	26.81	A11	51.93	A3	2.55
	2017.07.25						**2017.07.30**					
0:00	A9	21.61	A11	51.93	A2	1.93	A10	26.81	A11	51.93	A3	2.55
1:00	A9	21.61	A12	60.11	A2	1.93	A10	26.81	A12	60.11	A4	3.05
2:00	A9	21.61	A11	51.93	A2	1.93	A13	31.49	A12	60.11	A5	3.43
3:00	A9	21.61	A10	45.93	A2	1.93	A14	34.93	A12	60.11	A5	3.43
4:00	A9	21.61	A10	45.93	A2	1.93	A14	34.93	A12	60.11	A5	3.43
5:00	A9	21.61	A10	45.93	A2	1.93	A15	39.82	A12	60.11	A4	3.05
6:00	A9	21.61	A10	45.93	A2	1.93	A15	39.82	A14	68.78	A3	2.55
7:00	A9	21.61	A9	38.31	A2	1.93	A15	39.82	A14	68.78	A3	2.55
8:00	A8	16.93	A9	38.31	A2	1.93	A15	39.82	A16	84.69	A3	2.55
9:00	A9	21.61	A10	45.93	A2	1.93	A16	47.45	A16	84.69	A3	2.55
10:00	A9	21.61	A10	45.93	A2	1.93	A18	53.83	A16	84.69	A4	3.05
11:00	A9	21.61	A11	51.93	A3	2.55	A18	53.83	A15	74.30	A4	3.05
12:00	A9	21.61	A12	60.11	A7	4.01	A16	47.45	A12	60.11	A4	3.05
13:00	A10	26.81	A12	60.11	A10	5.76	A15	39.82	A12	60.11	A4	3.05
14:00	A15	39.82	A15	74.30	A12	7.49	A14	34.93	A12	60.11	A4	3.05
15:00	A21	65.35	A16	84.69	A12	7.49	A15	39.82	A12	60.11	A4	3.05
16:00	A21	65.35	A16	84.69	A10	5.76	A15	39.82	A12	60.11	A4	3.05
17:00	A20	56.30	A15	74.30	A6	3.88	A14	34.93	A12	60.11	A4	3.05
18:00	A16	47.45	A14	68.78	A5	3.43	A14	34.93	A12	60.11	A4	3.05
19:00	A15	39.82	A12	60.11	A5	3.43	A12	29.95	A14	68.78	A5	3.43
20:00	A15	39.82	A12	60.11	A8	4.39	A12	29.95	A14	68.78	A3	2.55
21:00	A15	39.82	A11	51.93	A8	4.39	A14	34.93	A15	74.30	A3	2.55
22:00	A14	34.93	A11	51.93	A7	4.01	A15	39.82	A16	84.69	A2	1.93
23:00	A14	34.93	A10	45.93	A6	3.88	A15	39.82	A16	84.69	A2	1.93
	2017.07.26						**2017.07.31**					
0:00	A14	34.93	A10	45.93	A5	3.43	A15	39.82	A16	84.69	A2	1.93
1:00	A14	34.93	A12	60.11	A5	3.43	A16	47.45	A16	84.69	A2	1.93
2:00	A14	34.93	A12	60.11	A4	3.05	A20	56.30	A16	84.69	A2	1.93
3:00	A15	39.82	A12	60.11	A4	3.05	A21	65.35	A17	97.27	A2	1.93
4:00	A15	39.82	A12	60.11	A5	3.43	A21	65.35	A17	97.27	A2	1.93
5:00	A14	34.93	A12	60.11	A6	3.88	A21	65.35	A16	84.69	A2	1.93
6:00	A14	34.93	A12	60.11	A5	3.43	A21	65.35	A17	97.27	A2	1.93
7:00	A10	26.81	A11	51.93	A5	3.43	A24	80.72	A19	120.07	A3	2.55
8:00	A9	21.61	A10	45.93	A3	2.55	A26	98.02	A19	120.07	A3	2.55
9:00	A9	21.61	A10	45.93	A3	2.55	A26	98.02	A19	120.07	A4	3.05
10:00	A9	21.61	A9	38.31	A2	1.93	A26	98.02	A19	120.07	A4	3.05
11:00	A9	21.61	A9	38.31	A2	1.93	A26	98.02	A19	120.07	A3	2.55
12:00	A9	21.61	A9	38.31	A2	1.93	A26	98.02	A19	120.07	A3	2.55
13:00	A9	21.61	A9	38.31	A2	1.93	A25	88.70	A18	105.27	A4	3.05
14:00	A10	26.81	A6	28.19	A2	1.93	A24	80.72	A17	97.27	A4	3.05
15:00	A9	21.61	A6	28.19	A2	1.93	A24	80.72	A17	97.27	A3	2.55

16:00	A9	21.61	A6	28.19	A2	1.93	A24	80.72	A17	97.27	A3	2.55
17:00	A8	16.93	A6	28.19	A2	1.93	A24	80.72	A18	105.27	A2	1.93
18:00	A7	13.41	A5	22.45	A2	1.93	A24	80.72	A18	105.27	A2	1.93
19:00	A5	11.00	A4	18.70	A2	1.93	A24	80.72	A19	120.07	A2	1.93
20:00	A7	13.41	A4	18.70	A2	1.93	A24	80.72	A19	120.07	A2	1.93
21:00	A8	16.93	A6	28.19	A2	1.93	A24	80.72	A19	120.07	A2	1.93
22:00	A8	16.93	A6	28.19	A2	1.93	A24	80.72	A19	120.07	A3	2.55
23:00	A9	21.61	A7	33.98	A2	1.93	A24	80.72	A19	120.07	A3	2.55

913
914
915

References

[1] K. Prasad, A.K. Gorai, P. Goyal, Development of ANFIS models for air quality forecasting and input optimization for reducing the computational cost and time, Atmos. Environ. 128 (2016) 246–262. doi:10.1016/j.atmosenv.2016.01.007.

[2] A. Kurt, A.B. Oktay, Forecasting air pollutant indicator levels with geographic models 3 days in advance using neural networks, Expert Syst. Appl. 37 (2010) 7986–7992. doi:10.1016/j.eswa.2010.05.093.

[3] J. Cao, H. Xu, Q. Xu, B. Chen, H. Kan, Fine particulate matter constituents and cardiopulmonary mortality in a heavily polluted Chinese city, Environ. Health Perspect. 120 (2012) 373–378. doi:10.1289/ehp.1103671.

[4] J.J. Cao, Q.Y. Wang, J.C. Chow, J.G. Watson, X.X. Tie, Z.X. Shen, P. Wang, Z.S. An, Impacts of aerosol compositions on visibility impairment in Xi'an, China, Atmos. Environ. 59 (2012) 559–566. doi:10.1016/j.atmosenv.2012.05.036.

[5] S. Cai, Y. Wang, B. Zhao, S. Wang, X. Chang, J. Hao, The impact of the "Air Pollution Prevention and Control Action Plan" on PM 2.5 concentrations in Jing-Jin-Ji region during 2012–2020, Sci. Total Environ. 580 (2017) 197–209. doi:10.1016/j.scitotenv.2016.11.188.

[6] D. Sun, J. Fang, J. Sun, Health-related benefits of air quality improvement from coal control in China: Evidence from the Jing-Jin-Ji region, Resour. Conserv. Recycl. 129 (2018) 416–423. doi:10.1016/j.resconrec.2016.09.021.

[7] Y. Wang, H. Liu, G. Mao, J. Zuo, J. Ma, Inter-regional and sectoral linkage analysis of air pollution in Beijing–Tianjin–Hebei (Jing-Jin-Ji) urban agglomeration of China, J. Clean. Prod. 165 (2017) 1436–1444. doi:10.1016/j.jclepro.2017.07.210.

[8] J. Haas, Y. Ban, Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, River Delta and the Pearl River Delta, Int. J. Appl. Earth Obs. Geoinf. 30 (2014) 42–55. doi:10.1016/j.jag.2013.12.012.

[9] Z. Qi, W. Zhenshu, S. Yu, W. Yu, L. Zhanjie, The Power System Environmental Optimal Dispatch Containing Air Quality Forecast, in: Energy Procedia, 2017: pp. 3623–3628. doi:10.1016/j.egypro.2017.03.1053.

[10] N. Gouveia, T. Fletcher, Time series analysis of air pollution and mortality: Effects by cause, age and socioeconomic status, J. Epidemiol. Community Health. 54 (2000) 750–755. doi:10.1136/jech.54.10.750.

[11] C.M. Vong, W.F. Ip, P.K. Wong, C.C. Chiu, Predicting minority class for suspended particulate matters level by extreme learning machine, Neurocomputing. 128 (2014) 136–144. doi:10.1016/j.neucom.2012.11.056.

[12] G. Kiesewetter, W. Schoepp, C. Heyes, M. Amann, Modelling PM2.5impact indicators in Europe: Health effects and legal compliance, Environ. Model. Softw. 74 (2015) 201–211. doi:10.1016/j.envsoft.2015.02.022.

[13] The Central People's Government of the People's Republic of China. http://www.gov.cn/zwgk/2013-09/12/content_2486773.htm.

[14] D.W. van der Meer, J. Widén, J. Munkhammar, Review on probabilistic forecasting of photovoltaic power production and electricity consumption, Renew. Sustain. Energy Rev. 81 (2018) 1484–1512. doi:10.1016/j.rser.2017.05.212.

[15] Z. Yang, J. Wang, A new air quality monitoring and early warning system: Air quality assessment and air pollutant concentration prediction, Environ. Res. 158 (2017) 105–117. doi:10.1016/j.envres.2017.06.002.

[16] J. Wang, X. Zhang, Z. Guo, H. Lu, Developing an early-warning system for air quality prediction and assessment of cities in China, Expert Syst. Appl. 84 (2017) 102–116. doi:10.1016/j.eswa.2017.04.059.

[17] Y. Xu, P. Du, J. Wang, Research and application of a hybrid model based on

dynamic fuzzy synthetic evaluation for establishing air quality forecasting and early warning system: A case study in China, Environ. Pollut. 223 (2017) 435–448. doi:10.1016/j.envpol.2017.01.043.
[18] Y. Song, S. Qin, J. Qu, F. Liu, The forecasting research of early warning systems for atmospheric pollutants: A case in Yangtze River Delta region, Atmos. Environ. 118 (2015) 58–69. doi:10.1016/j.atmosenv.2015.06.032.
[19] P. Jiang, Q. Dong, P. Li, A novel hybrid strategy for PM2.5 concentration analysis and prediction, J. Environ. Manage. 196 (2017) 443–457. doi:10.1016/j.jenvman.2017.03.046.
[20] C. Zafra, Y. Ángel, E. Torres, ARIMA analysis of the effect of land surface coverage on PM10 concentrations in a high-altitude megacity, Atmos. Pollut. Res. 8 (2017) 660–668. doi:10.1016/j.apr.2017.01.002.
[21] P. Wang, H. Zhang, Z. Qin, G. Zhang, A novel hybrid-Garch model based on ARIMA and SVM for PM2.5concentrations forecasting, Atmos. Pollut. Res. 8 (2017) 850–860. doi:10.1016/j.apr.2017.01.003.
[22] M. Qin, Z. Li, Z. Du, Red tide time series forecasting by combining ARIMA and deep belief network, Knowledge-Based Syst. 125 (2017) 39–52. doi:10.1016/j.knosys.2017.03.027.
[23] I.B. Konovalov, M. Beekmann, F. Meleux, A. Dutot, G. Foret, Combining deterministic and statistical approaches for PM10 forecasting in Europe, Atmos. Environ. 43 (2009) 6425–6434. doi:10.1016/j.atmosenv.2009.06.039.
[24] R. Stern, P. Builtjes, M. Schaap, R. Timmermans, R. Vautard, A. Hodzic, M. Memmesheimer, H. Feldmann, E. Renner, R. Wolke, A. Kerschbaumer, A model inter-comparison study focussing on episodes with elevated PM10 concentrations, Atmos. Environ. 42 (2008) 4567–4588. doi:10.1016/j.atmosenv.2008.01.068.
[25] Y. Bai, Y. Li, X. Wang, J. Xie, C. Li, Air pollutants concentrations forecasting using back propagation neural network based on wavelet decomposition with meteorological conditions, Atmos. Pollut. Res. 7 (2016) 557–566. doi:10.1016/j.apr.2016.01.004.
[26] X. Li, L. Peng, X. Yao, S. Cui, Y. Hu, C. You, T. Chi, Long short-term memory neural network for air pollutant concentration predictions: Method development and evaluation, Environ. Pollut. 231 (2017) 997–1004. doi:10.1016/j.envpol.2017.08.114.
[27] P. Du, J. Wang, Z. Guo, W. Yang, Research and application of a novel hybrid forecasting system based on multi-objective optimization for wind speed forecasting, Energy Convers. Manag. 150 (2017) 90–107. doi:10.1016/j.enconman.2017.07.065.
[28] J. Wang, W. Yang, P. Du, Y. Li, Research and application of a hybrid forecasting framework based on multi-objective optimization for electrical power system, Energy. 148 (2018) 59–78. doi:10.1016/j.energy.2018.01.112.
[29] A. Safari, M. Davallou, Oil price forecasting using a hybrid model, Energy. 148 (2018) 49–58. doi:10.1016/j.energy.2018.01.007.
[30] P. Du, J. Wang, W. Yang, T. Niu, Multi-step ahead forecasting in electrical power system using a hybrid forecasting system, Renew. Energy. 122 (2018) 533–550. doi:10.1016/j.renene.2018.01.113.
[31] L.A. Zadeh, The concept of a linguistic variable and its application to approximate reasoning-I, Inf. Sci. (Ny). 8 (1975) 199–249. doi:10.1016/0020-0255(75)90036-5.
[32] P. Jiang, Q. Dong, P. Li, L. Lian, A novel high-order weighted fuzzy time series model and its application in nonlinear time series prediction, Appl. Soft Comput. J.

55 (2017) 44–62. doi:10.1016/j.asoc.2017.01.043.
[33] D.K. Jana, B. Das, M. Maiti, Multi-item partial backlogging inventory models over random planninghorizon in random fuzzy environment, Appl. Soft Comput. J. 21 (2014) 12–27. doi:10.1016/j.asoc.2014.02.021.
[34] M. Pulido, P. Melin, O. Castillo, Particle swarm optimization of ensemble neural networks with fuzzy aggregation for time series prediction of the Mexican Stock Exchange, Inf. Sci. (Ny). 280 (2014) 188–204. doi:10.1016/j.ins.2014.05.006.
[35] J. Soto, P. Melin, O. Castillo, A New Approach for Time Series Prediction Using Ensembles of IT2FNN Models with Optimization of Fuzzy Integrators, Int. J. Fuzzy Syst. 20 (2018). doi:10.1007/s40815-017-0443-6.
[36] C. Kocak, ARMA(p,q) type high order fuzzy time series forecast method based on fuzzy logic relations, Appl. Soft Comput. 58 (2017) 92–103. doi:10.1016/j.asoc.2017.04.021.
[37] M. Bose, K. Mali, A novel data partitioning and rule selection technique for modeling high-order fuzzy time series, Appl. Soft Comput. 63 (2017) 87–96. doi:10.1016/j.asoc.2017.11.011.
[38] L. Tan, S. Wang, K. Wang, A new adaptive network-based fuzzy inference system with adaptive adjustment rules for stock market volatility forecasting, Inf. Process. Lett. 127 (2017) 32–36. doi:10.1016/j.ipl.2017.06.012.
[39] A.A.M. Ahmed, S.M.A. Shah, Application of adaptive neuro-fuzzy inference system (ANFIS) to estimate the biochemical oxygen demand (BOD) of Surma River, J. King Saud Univ. - Eng. Sci. 29 (2017) 237–243. doi:10.1016/j.jksues.2015.02.001.
[40] D. Domańska, M. Wojtylak, Application of fuzzy time series models for forecasting pollution concentrations, Expert Syst. Appl. 39 (2012) 7673–7679. doi:10.1016/j.eswa.2012.01.023.
[41] N. Güler Dincer, Ö. Akkuş, A new fuzzy time series model based on robust clustering for forecasting of air pollution, Ecol. Inform. 43 (2018) 157–164. doi:10.1016/j.ecoinf.2017.12.001.
[42] C.H. Cheng, T.L. Chen, H.J. Teoh, C.H. Chiang, Fuzzy time-series based on adaptive expectation model for TAIEX forecasting, Expert Syst. Appl. 34 (2008) 1126–1132. doi:10.1016/j.eswa.2006.12.021.
[43] T.H.-K. Yu, K.-H. Huarng, A bivariate fuzzy time series model to forecast the TAIEX, Expert Syst. Appl. 34 (2008) 2945–2952. doi:10.1016/j.eswa.2007.05.016.
[44] M.-Y. Chen, B.-T. Chen, Online fuzzy time series analysis based on entropy discretization and a Fast Fourier Transform, Appl. Soft Comput. 14 (2014) 156–166. doi:10.1016/j.asoc.2013.07.024.
[45] E. Bas, · Vedide, R. Uslu, U. Yolcu, E. Egrioglu, E. Bas, V.R. Uslu, · E Egrioglu, A modified genetic algorithm for forecasting fuzzy time series, Appl Intell. 41 (2014) 453–463. doi:10.1007/s10489-014-0529-x.
[46] O. Cagcag Yolcu, H.K. Lam, A combined robust fuzzy time series method for prediction of time series, Neurocomputing. 247 (2017) 87–101. doi:10.1016/j.neucom.2017.03.037.
[47] S.-H. Cheng, S.-M. Chen, W.-S. Jian, Fuzzy time series forecasting based on fuzzy logical relationships and similarity measures, Inf. Sci. (Ny). 327 (2016) 272–287. doi:10.1016/j.ins.2015.08.024.
[48] L. Xie, G. Li, M. Xiao, L. Peng, Novel classification method for remote sensing images based on information entropy discretization algorithm and vector space model, Comput. Geosci. 89 (2016) 252–259. doi:10.1016/j.cageo.2015.12.015.
[49] H. Liu, H. Tian, X. Liang, Y. Li, Wind speed forecasting approach using secondary

decomposition algorithm and Elman neural networks, Appl. Energy. 157 (2015) 183–194. doi:10.1016/j.apenergy.2015.08.014.
[50] S. Wang, N. Zhang, L. Wu, Y. Wang, Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method, Renew. Energy. 94 (2016) 629–636. doi:10.1016/j.renene.2016.03.103.
[51] M.G. De Giorgi, M. Malvoni, P.M. Congedo, Comparison of strategies for multi-step ahead photovoltaic power forecasting models based on hybrid group method of data handling networks and least square support vector machine, Energy. 107 (2016) 360–373. doi:10.1016/j.energy.2016.04.020.
[52] N. Ramesh Babu, B. Jagan Mohan, Fault classification in power systems using EMD and SVM, Ain Shams Eng. J. (2015) 1–9. doi:10.1016/j.asej.2015.08.005.
[53] N. Zhang, A. Lin, P. Shang, Multidimensional k-nearest neighbor model based on EEMD for financial time series forecasting, Phys. A Stat. Mech. Its Appl. 477 (2017) 161–173. doi:10.1016/j.physa.2017.02.072.
[54] M. Niu, Y. Wang, S. Sun, Y. Li, A novel hybrid decomposition-and-ensemble model based on CEEMD and GWO for short-term PM2.5concentration forecasting, Atmos. Environ. 134 (2016) 168–180. doi:10.1016/j.atmosenv.2016.03.056.
[55] L. a. Zadeh, Fuzzy sets, Inf. Control. 8 (1965) 338–353. doi:10.1016/S0019-9958(65)90241-X.
[56] Q. Song, B.S. Chissorn, Forecasting enrollments with fuzzy time series-part II, Fuzzy Sets Syst. 62 (1994) 1–8. doi:10.1016/0165-0114(94)90067-1.
[57] H.K. Yu, Weighted fuzzy time series models for TAIEX forecasting, Phys. A Stat. Mech. Its Appl. 349 (2005) 609–624. doi:10.1016/j.physa.2004.11.006.
[58] A. Rubio, J.D. Bermúdez, E. Vercher, Improving stock index forecasts by using a new weighted fuzzy-trend time series method, Expert Syst. Appl. 76 (2017) 12–20. doi:10.1016/j.eswa.2017.01.049.
[59] C. Stefanakos, Fuzzy time series forecasting of nonstationary wind and wave data, Ocean Eng. 121 (2016) 1–12. doi:10.1016/j.oceaneng.2016.05.018.
[60] C.E. Shannon, A mathematical theory of communication, Bell Syst. Tech. J. 27 (1948) 379–423. doi:10.1145/584091.584093.
[61] C.R. De Sá, C. Soares, A. Knobbe, Entropy-based discretization methods for ranking data, Inf. Sci. (Ny). 329 (2016) 921–936. doi:10.1016/j.ins.2015.04.022.
[62] S. Ramírez-Gallego, S. García, F. Herrera, Online entropy-based discretization for data streaming classification, Futur. Gener. Comput. Syst. 86 (2018) 59–70. doi:10.1016/j.future.2018.03.008.
[63] K. Irani, U. Fayyad, Multi-lnterval Discretization of Continuous-Valued Attributes for Classification learning, Proc. Natl. Acad. Sci. U. S. A. (1993) 1022–1027. doi:10.1109/TKDE.2011.181.
[64] N.E. Wu, Zhaohua and Huang, Ensemble Empirical Mode Decomposition : A Noise Assisted Data Analysis Method, Adv. Adapt. Data Anal. 1 (2009) 1–41. doi:10.1142/S1793536909000047.
[65] J.-R. YEH, J.-S. SHIEH, N.E. HUANG, COMPLEMENTARY ENSEMBLE EMPIRICAL MODE DECOMPOSITION: A NOVEL NOISE ENHANCED DATA ANALYSIS METHOD, Adv. Adapt. Data Anal. 02 (2010) 135–156. doi:10.1142/S1793536910000422.
[66] S. Qin, F. Liu, J. Wang, B. Sun, Analysis and forecasting of the particulate matter (PM) concentration levels over four major cities of China using hybrid models, Atmos. Environ. 98 (2014) 665–675. doi:10.1016/j.atmosenv.2014.09.046.
[67] S.M. Chen, B.D.H. Phuong, Fuzzy time series forecasting based on optimal partitions of intervals and optimal weighting vectors, Knowledge-Based Syst. 118

(2017) 204–216. doi:10.1016/j.knosys.2016.11.019.
[68] M. Oprea, S. F. Mihalache, M. Popescu, A comparative study of computational intelligence techniques applied to PM2. 5 air pollution forecasting, (2016) 103-108. IEEE. In Computers Communications and Control (ICCCC). doi: 10.1109/ICCCC.2016.7496746
[69] A. Ockelford, The magical number two, plus or minus one: Some limits on our capacity for processing musical information, Music. Sci. 6 (2002) 185–219. doi:10.1177/102986490200600205.
[70] M.-Y. Chen, B.-T. Chen, A hybrid fuzzy time series model based on granular computing for stock price forecasting, Inf. Sci. (Ny). 294 (2015) 227–241. doi:10.1016/j.ins.2014.09.038.
[71] Y. Xu, W. Yang, J. Wang, Air quality early-warning system for cities in China, Atmos. Environ. 148 (2017) 239–257. doi:10.1016/j.atmosenv.2016.10.046.
[72] C. Li, & Z. Zhu, Research and application of a novel hybrid air quality early-warning system: A case study in China, Sci. Total Environ, 626 (2018) 1421–1438. doi:10.1016/j.scitotenv.2018.01.195
[73] C. Gupta, A. Jain, D. K. Tayal, ClusFuDE: Forecasting low dimensional numerical data using an Improved Method based on Automatic Clustering, Fuzzy Relationships and Differential Evolution, Eng. Appl. Artif. Intell. 71 (2018) 175-189. doi:10.1016/j.engappai.2018.02.015.