Environmental Research 158 (2017) 105–117

Contents lists available at ScienceDirect

Environmental Research

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

A new air quality monitoring and early warning system: Air quality assessment and air pollutant concentration prediction

Zhongshan Yang, Jian Wang*

School of Statistics, Dongbei University of Finance and Economics, Dalian 116025, China

ARTICLE INFO

Keywords:
Air pollution
Fuzzy comprehensive evaluation
Complementary ensemble empirical mode decomposition
Cuckoo search
Differential evolution

ABSTRACT

Air pollution in many countries is worsening with industrialization and urbanization, resulting in climate change and affecting people's health, thus, making the work of policymakers more difficult. It is therefore both urgent and necessary to establish amore scientific air quality monitoring and early warning system to evaluate the degree of air pollution objectively, and predict pollutant concentrations accurately. However, the integration of air quality assessment and air pollutant concentration prediction to establish an air quality system is not common. In this paper, we propose a new air quality monitoring and early warning system, including an assessment module and forecasting module. In the air quality assessment module, fuzzy comprehensive evaluation is used to determine the main pollutants and evaluate the degree of air pollution more scientifically. In the air pollutant concentration prediction module, a novel hybridization model combining complementary ensemble empirical mode decomposition, a modified cuckoo search and differential evolution algorithm, and an Elman neural network, is proposed to improve the forecasting accuracy of six main air pollutant concentrations. To verify the effectiveness of this system, pollutant data for two cities in China are used. The result of the fuzzy comprehensive evaluation shows that the major air pollutants in Xi'an and Jinan are PM_{10} and $PM_{2.5}$ respectively, and that the air quality of Xi'an is better than that of Jinan. The forecasting results indicate that the proposed hybrid model is remarkably superior to all benchmark models on account of its higher prediction accuracy and stability.

1. Introduction

With the rapid expansion of China's economy and growth in the number of vehicles and industries, the problem of air pollution is becoming more serious. The economic take-off in a country of 1.4 billion people requires the consumption of more fossil energy and natural resources, and this has led to a change in the chemical composition of the atmosphere. The Environmental Performance Index (EPI) shows that the air quality in China is the second worst among 180 countries included in the index this year (EPI (Environmental Performance Index, 2016)). Fig. 1 shows the ten countries with the poorest air quality in 2016, and their corresponding EPI scores. Air pollution has important implications on the ecological environment (Wang et al., 2016), and could also destroy vegetation and monuments (Brook et al., 2004; Baker and Foley, 2011). In addition, air pollution can cause lung cancer and related diseases, and many epidemiological studies have consistently shown an association between particulate air pollution and cardiovascular and respiratory diseases (Gorai et al., 2014). To assist people in keeping fit and improving the quality of their lives, the development of a robust, accurate, yet simple, air quality monitoring and early warning system is highly desirable.

Environmental monitoring and early warning is the basic function of environmental protection work, not only in relation to scientific decision-making, but also long-term development. The people of China want to improve air quality and reduce haze weather. To achieve this, more than 2700 monitoring stations have been established with more than 268,000 sets of monitoring instruments, and approximately 60,000 monitoring personnel (MEP Ministry of Environmental Protection, 2015). Air quality systems are very complicated, and, although China has made great progress in reducing the haze and

Abbreviations: EPI, Environmental Performance Index; PSI, Pollutant Standard Index; AQI, Air Quality Index; MEP, Ministry of Environmental Protection; AHP, Analytic hierarchy process; CTMs, Chemical transport models; ARIMA, Autoregressive integrated moving average; MLR, Multi-linear regression; ANN, Artificial neural network; SVM, Support vector machine;; RBF, Radial basis function;; FL, Fuzzy logic; EEMD, Ensemble empirical mode decomposition;; CS, Cuckoo search; BP, Back-propagation artificial neutral networks; ENN, Elman neural network; CEEMD, Complementary ensemble empirical mode decomposition;; DE, Differential evolution algorithm; IMF, Intrinsic mode function;; D-M, Diebold-Mariano test; EMD, Empirical mode decomposition; GRNN, Generalized regression neural network

* Corresponding author.
E-mail address: jianwang0826@163.com (J. Wang).

http://dx.doi.org/10.1016/j.envres.2017.06.002
Received 4 December 2016; Received in revised form 31 May 2017; Accepted 2 June 2017

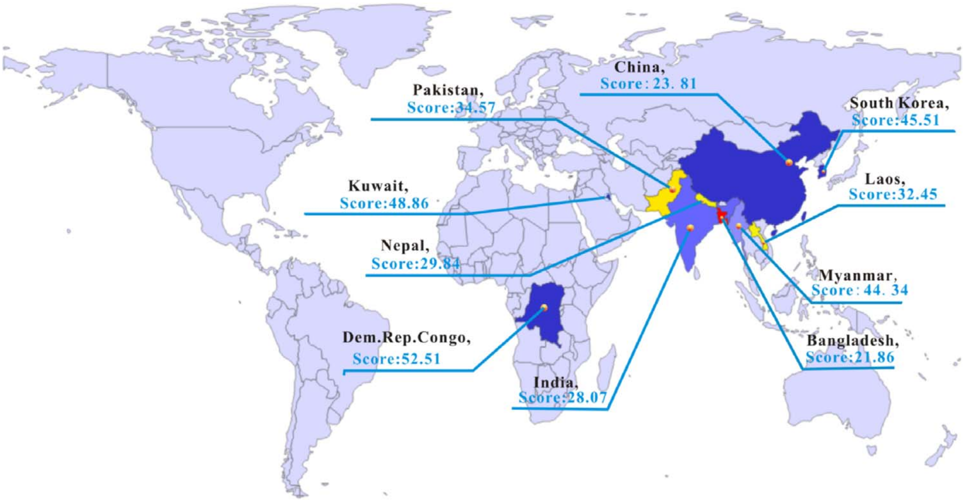


Fig. 1. Shows the ten countries with the poorest air quality in 2016.

improving air quality, the continuing challenges of reducing air pollutants still need to be addressed. Air quality problems are not unique to China; many countries, especially developing countries , suffer from air pollution. In recent years, significant research has been conducted on air quality monitoring and evaluation systems. The research can be divided into air quality assessment and air pollutants forecasting.

For air quality assessment, numerous indices have been proposed to evaluate the quality of the air. The first index was the “Pollutant Standard Index” (PSI), which was modified and replaced by the “Air Quality Index” (AQI), both of which were developed and introduced by United States Environmental Protection Agency (US EPA, 2009). The PSI considered five air pollutants: sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3), and particulate matter (PM_{10}). Compared to the PSI, $PM_{2.5}$ and 8-h average ozone concentrations were also included in the AQI (Cheng et al., 2007). The standards of countries using the AQI are not all same. China has its own air quality assessment system. The Ministry of Environmental Protection (MEP) of China promulgated new ambient air quality standards (GB3095-2012) to replace the previous standards (GB3095-1996) in 2012, whereby $PM_{2.5}$ and 8-h average ozone concentrations were included and the concentration of pollutants in the limits were redefined. Although it is used worldwide, the AQI system has certain limitations. Many other air quality indices have been proposed (Liu et al., 2012; Kyrkilis et al., 2007). Although traditional air quality assessment used a simple digital indicator as a dividing line, with each side divided into different levels, the classification standard is not objective. The evaluation of air quality, which is aimed at determining "the degree of pollution,” is a fuzzy concept, and it is difficult to find clear boundaries; thus, the evaluation classification standard of pollution levels should also be fuzzy. Therefore, fuzzy logic is a suitable tool for air quality assessment (Zadeh, 1965; Hájek and Olej, 2009; Abdullah and Khalid, 2012; Akkaya et al., 2015), and many fuzzy-based air quality indices have been proposed (Mohammad et al., 2011; Miguel and Ignacio, 2016).

Forecasting systems have been successfully used in many fields (Wang et al., 2014; Sun et al., 2013; Qin et al., 2015). Air pollution forecasting is essential, and air quality forecasting systems would allow for more efficient countermeasures to safeguard citizens' health (Yahya et al., 2014). The forecasting methods can be divided into two major categories: deterministic models and statistical models (Geoffrey Cobourn, 2010). The commonly used deterministic model is the chemical transport model (CTM). This model can forecast air pollution concentrations in places without monitoring the site and does not rely on a large quantity of historical data, but the accuracy is highly affected by the quality of the emission data and the scale used (Feng et al., 2015; Stern et al., 2008; Sun et al., 2013). As for the statistical models, the autoregressive integrated moving average (ARIMA) model (Konovalov et al., 2009), the multi-linear regression (MLR) model (Zhang et al., 2012), and the artificial neural network (ANN) model (Lin et al., 2010; Metaxiotis et al., 2003), have all been widely used for air pollutant concentration forecasting recently. The ARIMA and MLR models, both linear, have been adopted to forecast air quality (Kukkonen et al., 2003); however, a significant shortcoming is that the forecasting accuracy depends on the linear mapping ability in nonlinear processes (Song et al., 2015). To overcome this limitation, nonlinear models, such as the support vector machine (SVM) (Osowski and Garanty, 2007; Suarez Sanchez et al., 2011; Lin et al., 2011), the radial basis function (RBF) (Paschalidou et al., 2011), and fuzzy logic (FL) (Alhanafy et al., 2010), have been adopted to forecast air pollution. In addition, hybrid models have also been proposed to improve the forecasting accuracy of air pollutant concentration. Qin (2014) (Qin et al., 2014), proposed a hybrid model combing ensemble empirical mode decomposition (EEMD), cuckoo search (CS), and back-propagation (BP) artificial neural networks for forecasting PM concentrations.

This paper develops a new air quality monitoring and early warning system that integrates air quality assessment and air pollutant concentration forecasting. Fuzzy theory is used to evaluate the air quality more scientifically, and a new hybrid model MCSDE-CEEMD-ENN, combining complementary ensemble empirical mode decomposition (CEEMD), modified cuckoo search and differential evolution algorithm (MCSDE) and Elman neural network (ENN), is proposed to improve forecasting accuracy of air pollutant concentration. Due to the inherent complexity of pollutant concentration series, describing the moving trend of pollutants and accurate prediction is difficult. To overcome this, CEEMD is utilized to decompose the original time series and reconstruct a new series with the noise removed. ENN is used to forecast the air pollutant concentrations. An improved optimization algorithm (MCSDE) that combines CS and DE has been proposed to improve the stability of convergence by optimizing the initial weights and threshold of ENN. The main six air pollutants data recorded in Xi'an and Jinan, China, were used to verify the forecasting ability. The proposed air quality system has the following advantages:

1) Based on the analysis of existing air quality evaluations, imposing fuzzy theory to make fuzzy comprehensive evaluations about air quality can determine the most prevalent air pollutants in the city.

2) The proposed model can forecast the concentrations of the six main pollutants. There is no existing comprehensive research on the prediction of all the six main pollutants' concentrations, although many researchers focus on the prediction of one, or several, pollutants. The proposed MCSDE-CEEMD-ENN model has a good forecasting capability for air pollutant concentration.
3) The proposed optimization algorithm, MCSDE, combines CS and DE. It can provide better initial weights and thresholds to ENN, and improve the forecasting capability. Tests show that the model can avoid getting trapped into local optima, and the global searching capability is enhanced.
4) This new prediction model, valid for one day, will be updated by reconstructing and training new samples after providing 24 predictions. The model was continuously validated for one week.
5) More accurate metrics are applied to evaluate the forecasting performance of the proposed model. To estimate the forecasting performance, the Diebold-Mariano test, bias-variance framework, and three error criteria are adopted, including mean absolute error (MAE), mean absolute percentage error (MAPE), and mean square error (MSE).

2. Methodology

In this section, methods discussed in this paper are introduced briefly, including fuzzy comprehensive evaluation, CEEMD, Elman neural network, and the optimization algorithm.

2.1. Fuzzy comprehensive evaluation

The fuzzy comprehensive evaluation method appraises air quality by establishing and confirming gene subclass, estimation subclass, subject function, and weighted subclass. This method can solve the uncertainty and ambiguity among the factors in the air environment system, and evaluate the objective status of the air quality scientifically (Zhao et al., 2010).

The concrete steps of fuzzy comprehensive evaluation are as follows:

Step 1: Establish the gene subclass.

$$U = \{u_1, u_2, \ldots, u_6\} = \{SO_2, NO_2, CO, O_3, PM_{10}, PM_{2.5}\} \tag{1}$$

where u_i denote the six major air pollutants.

Step 2: Set up the estimation subclass.

$$V = v_{ij} = \{v_{i1}, v_{i2}, \ldots, v_{i5}\} \tag{2}$$

where $i = \{SO_2, NO_2, CO, O_3, PM_{10}, PM_{2.5}\}$, and j = {good, regular, bad, very bad, and extremely bad}.

Table 1
Membership function.

Air pollution level	Membership function	Variable domain
Good ($j = 1$)	1	$x_i \le s_{ij}$
	$(s_{i(j+1)} - x_i)/(s_{i(j+1)} - s_{ij})$	$s_{ij} \le x_i \le s_{i(j+1)}$
	0	$x_i \ge s_{i(j+1)}$
Regular, Bad, Very bad (j = 2,3,4)	$(x_i - s_{i(j-1)})/(s_{ij} - s_{i(j-1)})$	$s_{i(j-1)} \le x_i \le s_{ij}$
	$(s_{i(j+1)} - x_i)/(s_{i(j+1)} - s_{ij})$	$s_{ij} \le x_i \le s_{i(j+1)}$
	0	$x_i \ge s_{i(j+1)}$
Extremely bad ($j = 5$)	0	$x_i \le s_{i(j-1)}$
	$(x_i - s_{i(j-1)})/(s_{ij} - s_{i(j-1)})$	$s_{i(j-1)} \le x_i \le s_{ij}$
	1	$x_i > s_{ij}$

Step 3: Establish the fuzzy degree of the membership function. The membership function is defined as the halved trapezoidal distribution function. Details are presented in Table 1.

Step 4: Establish the fuzzy matrix. The observed value is brought into the membership function, and the multi-factor evaluation matrix is obtained.

$$R = r_{ij} = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1j} \\ r_{21} & r_{22} & \cdots & r_{2j} \\ \cdots & \cdots & \cdots & \cdots \\ r_{i1} & r_{i2} & \cdots & r_{ij} \end{bmatrix} \tag{3}$$

Step 5: Determine pollution factor weights. The weight of the pollution factor is a measure of the relative degree of the influence of pollutants on the air quality. Each pollution factor is assigned a different weight according to the influence of each factor.

$$\omega_i = [u_i^k/(\frac{1}{n}\sum_{j=1}^{n} v_{ij})]/\sum_{i=1}^{m}[u_i^k/(\frac{1}{n}\sum_{j=1}^{n} v_{ij})] \tag{4}$$

$$A = (\omega_1, \omega_2, \ldots, \omega_6) \tag{5}$$

Step 6: Compound calculation between weighted vectors and fuzzy relation matrix.

$$B = A(\bullet, \oplus)R = (\omega_1, \omega_2, \ldots, \omega_6)\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{15} \\ r_{21} & r_{22} & \cdots & r_{25} \\ \cdots & \cdots & \cdots & \cdots \\ r_{61} & r_{62} & \cdots & r_{65} \end{bmatrix} = [b_1, b_2, \ldots, b_5] \tag{6}$$

$$b_j = \min(1, \textstyle\sum_{i=1}^{m} w_i r_{ij}) \tag{7}$$

According to the principle of maximum membership degree, the maximum value of $\boldsymbol{B}$ is the result of fuzzy comprehensive evaluation of the air quality evaluation. In this paper, the concentrations of six major air pollutants are used as the input to evaluate the air quality.

2.2. Complementary ensemble empirical mode decomposition (CEEMD)

Complementary ensemble empirical mode decomposition (CEEMD) is a member of the empirical mode decomposition (EMD) family. The EMD family plays an important role in processing time series that are non-linear and non-stationary.

The EMD approach was proposed by Huang et al. Huang et al. (1998). With the EMD process, any complex time series can be decomposed into a finite number of intrinsic mode functions (IMF), and each IMF reflects the dynamic characteristics of the original signal.

The algorithm was considered a significant breakthrough for traditional methods, but the decomposition results often suffer from mode mixing. To solve this problem, Wu and Huang (2008) (Wu and Huang, 2008) proposed ensemble empirical mode decomposition (EEMD) that added Gaussian white noise to the original signal during the entire decomposition process. The EEMD method can suppress the mode mixing problem to some extent and improve the stability of the EMD algorithm significantly. However, the added white noise cannot be completely neutralized, and some residual white noise will remain in the components of IMF.

Yeh et al. (2010) introduced complementary ensemble empirical mode decomposition (CEEMD) to solve the problems of EMD and EEMD. Before EMD decomposition, they added opposite-sign white noise to the original time series. The CEEMD method not only effectively solves the mode mixing of EMD decomposition, but also almost completely eliminates the influence of residual white noise (Tang et al., 2015). To improve the forecasting accuracy and stability, CEEMD was utilized to remove the noise of the original data and decompose the original air pollutant concentration datasets into several IMFs. Each IMF will be an input for the Elman neural network to forecast air pollutant concentration.

2.3. Elman neural network (ENN)

The Elman neural network (ENN), a partial recurrent network model, was proposed by Elman in 1990 (Elman, 1990). The structure of ENN comprises an input layer, a hidden layer, an output layer, and a context layer. The characteristic of ENN is that the output of the hidden layer is connected to the input of the hidden layer by the delay and storage of the context layer. The addition of the context layer increases the ability of processing dynamic information, so as to achieve the goal of dynamic modeling.

The nonlinear state space expression is as follows:

$$\mathrm{y}(t) = g(\omega_1 \mathrm{x}(t)) \tag{8}$$

$$\mathrm{x}(t) = f(\omega_1 \mathrm{x_c}(t) + \omega_2(\mathrm{u}(t-1))) \tag{9}$$

$$\mathrm{x_c}(t) = \mathrm{x}(t-1) \tag{10}$$

where $\boldsymbol{y}$ and $\boldsymbol{x}$ are the output node vector and hidden layer node vector, respectively, $\boldsymbol{u}$ and x_c denote the input layer vector and feedback state vector, respectively, their corresponding weights are ω_1 and ω_2, and $g(\bullet)$ and $f(\bullet)$ denote their corresponding transfer functions, respectively.

Compared to traditional neural networks, ENN has a great advantage in the prediction of time series. However, ENN also has the shortcoming that the initialized weights and threshold values randomly generated can cause overfitting, easily fall into the local minimum value, and the convergence speed is slow (Lin and Hong, 2011). The intelligent optimization algorithm is essential for solving these problems.

2.4. Modified cuckoo search and differential evolution algorithm (MCSDE)

This section introduces the modified optimization algorithm. This algorithm is used to optimize the weights and thresholds for ENN to overcome the slow convergence and local maximum.

2.4.1. Cuckoo search (CS)

The cuckoo search algorithm, proposed by Yang and Deb (2009), is a metaheuristic algorithm based on the natural behavior of some cuckoo species along with their brood parasitism, and the Lévy flight behavior of birds (Yang and Deb, 2009). The CS algorithm is simple, has few parameters, and does not need to re-match a large number of parameters in solving special problems. Compared to other heuristic intelligent optimization algorithms, it performs better in solving many optimization problems (Gandomi et al., 2013; Marichelvam et al., 2014).

The CS algorithm can be summarized by three idealized conditions: a) each cuckoo lays only one egg at a time, and randomly selects a bird's nest to place it, b) the best nest will be preserved for the next generation, and c) the number of available host nests is fixed, and the probability of alien eggs being discovered by the nest's host is $p_a \in [0, 1]$. In this case, the host bird can either get rid of the egg, or simply abandon the nest and build a new one.

The algorithm combines the local stochastic process and the global search stochastic process effectively, both of which are controlled by the transformation parameter p_a. The local stochastic process can be described as:

$$\mathrm{x}_i^{(t+1)} = \mathrm{x}_i^{(t)} + \alpha s \oplus H(p_a - \varepsilon) \otimes (\mathrm{x}_j^t - \mathrm{x}_k^t) \tag{11}$$

where $H(\bullet)$ denotes the Heaviside function, x_j^t and x_k^t denote two different random sequences, ε is a random number from the random distribution, and s is the step size. The global stochastic process is carried out according to the Lévy flight process:

$$\mathrm{x}_i^{(t+1)} = \mathrm{x}_i^{(t)} + \alpha \frac{\lambda \Gamma(\lambda)\sin(\pi\lambda/2)}{\pi}\frac{1}{s^{1+\lambda}},\ s \gg s_0 > 0,\ 1 < \lambda \le 3 \tag{12}$$

where α is a scaling factor of the step size.

2.4.2. Differential evolution algorithm (DE)

As a simple and efficient global optimization method in continuous space, the differential evolution algorithm (DE) (Storn and Price, 1995) can be considered one of the most powerful random real-parameter optimization algorithms currently used. Compared to other algorithms for producing progeny by probability distribution, the DE uses the value obtained by multiplying the difference of the independent population members randomly selected, by a certain coefficient as the mutation operator to disturb the current population members.

Each iteration of the DE has four steps: parameter vector initialization, mutation, crossover, and selection. The initial population of DE is a randomly generated real parameter vector, consisting of NP individuals with d-dimension. The population is expressed as follows:

$$P^g = \{\vec{\mathrm{x}}_1^g, \vec{\mathrm{x}}_2^g, \ldots, \vec{\mathrm{x}}_2^g\} \tag{13}$$

$$\vec{\mathrm{x}}_i^g = [x_{1,i}^g, x_{2,i}^g, \ldots, x_{2,i}^g] \tag{14}$$

The donor vector $\vec{\mathrm{v}}_i^g = \{v_{1,i}^g, v_{2,i}^g, \ldots, v_{D,i}^g\}$ corresponding to the i-th target vector $\vec{\mathrm{x}}_i^g$ is created by the differential mutation strategy given below:

$$\vec{\mathrm{v}}_i^g = \vec{\mathrm{x}}_{r1}^g + F\cdot(\vec{\mathrm{x}}_{r2}^g - \vec{\mathrm{x}}_{r3}^g) \tag{15}$$

where $\vec{\mathrm{x}}_{r1}^g$, $\vec{\mathrm{x}}_{r2}^g$, and $\vec{\mathrm{x}}_{r3}^g$ are independent parameter vectors randomly selected from the current population, and $F \in [0.4, 1]$ is the proportional coefficient.

The trial vector corresponding to the i-th target vector $\vec{\mathrm{x}}_i^g$ is created using a binomial crossover as follows:

$$u_{j,i}^g = \begin{cases} v_{j,i}^g, & rand_{i,j}[0, 1] \le Cr \ or \ j = j_{rand} \\ x_{j,i}^g, & otherwise \end{cases} \tag{16}$$

where Cr denotes crossover rate, $rand_{i,j}[0, 1]$ is a random integer subject to uniform distribution, and $j_{rand} \in [1, 2, \ldots, D]$ is a randomly selected index value.

To keep the number of the offspring population constant, the next step of the algorithm decides whether the target vector or the test vector should be preserved for the next generation:

$$\vec{\mathrm{x}}_i^{g+1} = \begin{cases} \vec{\mathrm{u}}_i^g, & f(\vec{\mathrm{u}}_i^g) \le f(\vec{\mathrm{x}}_i^g) \\ \vec{\mathrm{x}}_i^g, & f(\vec{\mathrm{u}}_i^g) > f(\vec{\mathrm{x}}_i^g) \end{cases} \tag{17}$$

where $f(\bullet)$ denotes the minimized objective function.

2.4.3. Hybrid optimization algorithm

In this section, a modified-optimization algorithm, hybridizing the CS and DE, is proposed to optimize the initial weights and thresholds of ENN, avoid local optima, and improve prediction accuracy.

DE and CS algorithms are excellent meta-heuristic optimization methods. They have a faster convergence rate and strong global search capabilities, however, a single optimization model can easily fall into local optimization. Therefore, the MCSDE model is proposed to introduce a new exchange mechanism. This mechanism allows the information to be passed between the two populations to avoid falling into local optima, and improves the global and local search capabilities of the optimization algorithm.

A rudimentary MCSDE algorithm is outlined as follow:

***Algorithm:* MCSDE**

Input:
$x_t^{(0)} = (x_{(1)}^{(0)}, x_{(2)}^{(0)}, \ldots, x_{(t)}^{(0)})$—a series of training data
$x_v^{(0)} = (x_{(t+1)}^{(0)}, x_{(t+2)}^{(0)}, \ldots, x_{(t+d)}^{(0)})$—a series of verifying data
Output:
POP_{best}—the final solution with the best fitness value in population
Parameters:
$G_{\max}$—the maximum number of iterations
p—population size
d—the number of dimension
F, Cr—proportion of coefficient and mutation probability
Pa—discovery rate of alien eggs

/*set the parameters of CS and DE.*/
/*Generate initial population $x_i (i = 1, 2, \ldots, p)$ randomly.*/
FOR EACH *i*: $1 \le i \le p$ **DO**
Evaluate the corresponding fitness function f_i.
END FOR
The best fitness $f_{\min}$ and the best individual POP_{best}.
WHILE $g < G_{\max}$ **DO**
/*Cuckoo search process.*/
/*Generate new nests.*/
FOR EACH *i*: $1 \le i \le p$ **DO**
$x_i^{g+1} = x_i^g + \alpha L(s, \lambda)$
END FOR
/*Find the current best nest.*/
FOR EACH *i*: $1 \le i \le p$ **DO**
IF $f_i^{g+1} \le f_i^g$ **THEN** $nest_i^{best} = x_i^{g+1}$ **END IF**
END FOR
The best fitness f_{new} and the best nest $nest_{best}$.
/*Discovery and randomization.*/
FOR EACH *i*: $1 \le i \le p$ **DO**
$x_{i,new}^{g+1} = x_i^{g+1} + \alpha s \oplus H(p_a - rand()) \otimes (x_{randperm()}^{g+1} - x_{randperm()}^{g+1})$
END FOR
/* Evaluate this set of solutions.*/
FOR EACH *i*: $1 \le i \le p$ **DO**
IF $f_{i,new}^{g+1} \le f_i^{g+1}$ **THEN** $nest_{i,new}^{best} = nest_i^{best}$ **END IF**
END FOR
Find the best fitness f_{new} and the best nest $nest_{best}$
IF $f_{new} < f_{\min}^{CS}$ **THEN**
$f_{\min}^{CS} = f_{new}$; $POP_{best}^{CS} = nest_{best}$
END IF
/*Differential evolution.*/
FOR EACH *i*: $1 \le i \le p$ **DO**
/*Mutation.*/
$\vec{v}_i^g = \vec{x}_{rand()}^g + F \cdot (\vec{x}_{rand()}^g - \vec{x}_{rand()}^g)$
/*Crossover.*/
IF $rand_{i,j}[0, 1] \le Cr$ or $j = j_{rand}$ **THEN** $u_{j,i}^g = v_{j,i}^g$
ELSE $u_{j,i}^g = x_{j,i}^g$
END IF
/*Selection.*/
IF $f(\vec{u}_i^g) \le f(\vec{x}_i^g)$ **THEN** $\vec{x}_i^{g+1} = \vec{u}_i^g$
ELSE $\vec{x}_i^{g+1} = \vec{x}_i^g$
END IF
END FOR
Find the best fitness $f_{\min}^{DE}$ and the best individual POP_{best}^{DE}.
/*The global optimal solution.*/
IF $f_{\min}^{CS} < f_{\min}^{DE}$ **THEN** $POP_{best}^{DE} = POP_{best}^{CS}$
ELSE IF $f_{\min}^{CS} > f_{\min}^{DE}$ **THEN** $POP_{best}^{CS} = POP_{best}^{DE}$
END IF
END WHILE

2.4.4. Test of MCSDE

In this section, four benchmark functions are selected to test the optimization performance of CS, DE, and the improved algorithm optimization. The experiment is designed to find the global minimum of these functions. The mathematical expressions and the search ranges of these functions are presented in Table 2. The test conditions are set as follows: the maximum number of iterations, population size, and dimensions are 2000, 40 and 30, respectively; proportion of coefficient and mutation probability are 0.5 and 0.1, respectively, and the discovery rate of alien eggs is 0.25.

One hundred experiments were carried out independently. The experimental results, including the maximum, minimum, average, and standard deviation, are presented in Table 3, which shows the following:

(a) For the sphere function, the minimum values of DE, CS, and MCSDE are 2.79×10^{-30}, 4.44×10^{-13}, and 1.94×10^{-30}, the maximum values are 3.56×10^{-29}, 1.02×10^{-11}, and 3.40×10^{-29}, the average values are 1.22×10^{-29}, 2.88×10^{-12}, and 1.13×10^{-29}, and the standard deviation values are 7.17×10^{-30}, 2.14×10^{-12}, and 6.52×10^{-30}, respectively. The optimization effect of MCSDE is better than CS and DE.
(b) For the Rosenbrock function, the maximum, minimum, average, and deviation values of MCSDE are 21.29, 16.89, 19.62, and 0.86, respectively. All of the values are better than others.
(c) For the Rastrigin function, CS has a poor performance. The minimum values of DE and MCSDE are both 0, the maximum values are 7.59×10^{-13} and 1.10×10^{-13}, the average values are 1.47×10^{-14} and 1.23×10^{-14}, and the standard deviation values are 7.68×10^{-14} and 2.31×10^{-14}, respectively.
(d) For the Griewank function, the maximum, minimum, average, and deviation values of MCSDE are 1.11×10^{-16}, 0, 6.66×10^{-18}, and 2.65×10^{-17}, respectively.

Based on the results of the four test functions, MCSDE has the best optimization capability compared to CS and DE.

2.5. The proposed hybrid model

A new hybrid model, MCSDE-CEEMD-ENN, is proposed to forecast six major air pollutant concentrations. The details of the algorithm are described below, and the flow diagram is shown in Fig. 2.

Step 1: The CEEMD is employed to decompose the original air pollutant concentrations datasets into several IMFs. The new series is then reconstructed with the highest frequency IMF removed. The pollutant concentration data is usually a chaotic time series. The denoising technique can eliminate the influence of outliers and improve prediction accuracy.
Step 2: The optimization algorithm MCSDE, that combines CS and DE, is proposed to optimize the initial weights and thresholds of ENN.
Step 3: Construct ENN with optimized initial weights and thresholds. The optimized ENN model is utilized to forecast six major air pollutant concentrations of Xi'an and Jinan, China.
Step 4: The Diebold-Mariano test, bias-variance framework, and three error criteria are adopted to evaluate the predictive power of the model.

Table 2
The mathematical expressions and the search ranges of benchmark functions.

Function name	Test function	Variable domain	x*	f(x*)
Sphere	$f_1(x) = \sum_{i=1}^{d} x_i^2$	[−5.12,5.12]	[0,0,…,0]	0
Rosebrock	$f_2(x) = \sum_{i=1}^{d-1} (100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2)$	[−2.084,2.084]	[0,0,…,0]	0
Rastrigin	$f_3(x) = \sum_{i=1}^{d} [x_i^2 - 10\bullet \cos(2\pi x_i) + 10]$	[−5.12,5.12]	[0,0,…,0]	0
Griewank	$f_4(x) = \frac{1}{4000} \sum_{i=1}^{d} x_i^2 - \prod_{i=1}^{d} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	[−8,8]	[0,0,…,0]	0

Table 3
Test results of MCSDE.

Benchmark function	Algorithm	Min value	Max value	Average value	Standard deviation
Sphere	DE	2.79E-30	3.56E-29	1.22E-29	7.17E-30
	CS	4.44E-13	1.02E-11	2.88E-12	2.14E-12
	MCSDE	**1.94E-30**	**3.40E-29**	**1.13E-29**	**6.52E-30**
Rosebrock	DE	18.96	26.80	24.65	1.49
	CS	17.50	22.40	20.74	0.91
	MCSDE	**16.89**	**21.29**	**19.62**	**0.86**
Rastrigin	DE	0	7.59E-13	1.47E-14	7.68E-14
	CS	35.91	81.77	61.73	8.37
	MCSDE	**0**	**1.10E-13**	**1.23E-14**	**2.31E-14**
Griewank	DE	0	1.11E-16	8.88E-18	3.03E-17
	CS	3.33E-16	1.14E-14	2.57E-15	1.76E-15
	MCSDE	**0**	**1.11E-16**	**6.66E-18**	**2.65E-17**

To verify the validity and stability of the MCSDE-CEEMD-ENN, several single- and hybrid-models will be used as controlled trials.

3. Experimental simulation

This section introduces the results of a fuzzy comprehensive evaluation of air quality and the forecasting results of the six major air pollutant concentrations using the MCSDE-CEEMD-ENN model. To verify the forecasting ability of the proposed hybrid model, several single and combination forecasting models are utilized.

3.1. Performance metric

There is no fixed metric to measure the predictive power of these models. To estimate the forecasting performance from multiple perspectives, the Diebold-Mariano (D-M) test, bias-variance framework, and three error criteria are adopted, including MAE, MAPE, and MSE (Xiao et al., 2015a).

The three error criteria are expressed as:

$$MSE = \frac{1}{T}\sum_{t=1}^{T} (y_t - \hat{y}_t)^2 \tag{18}$$

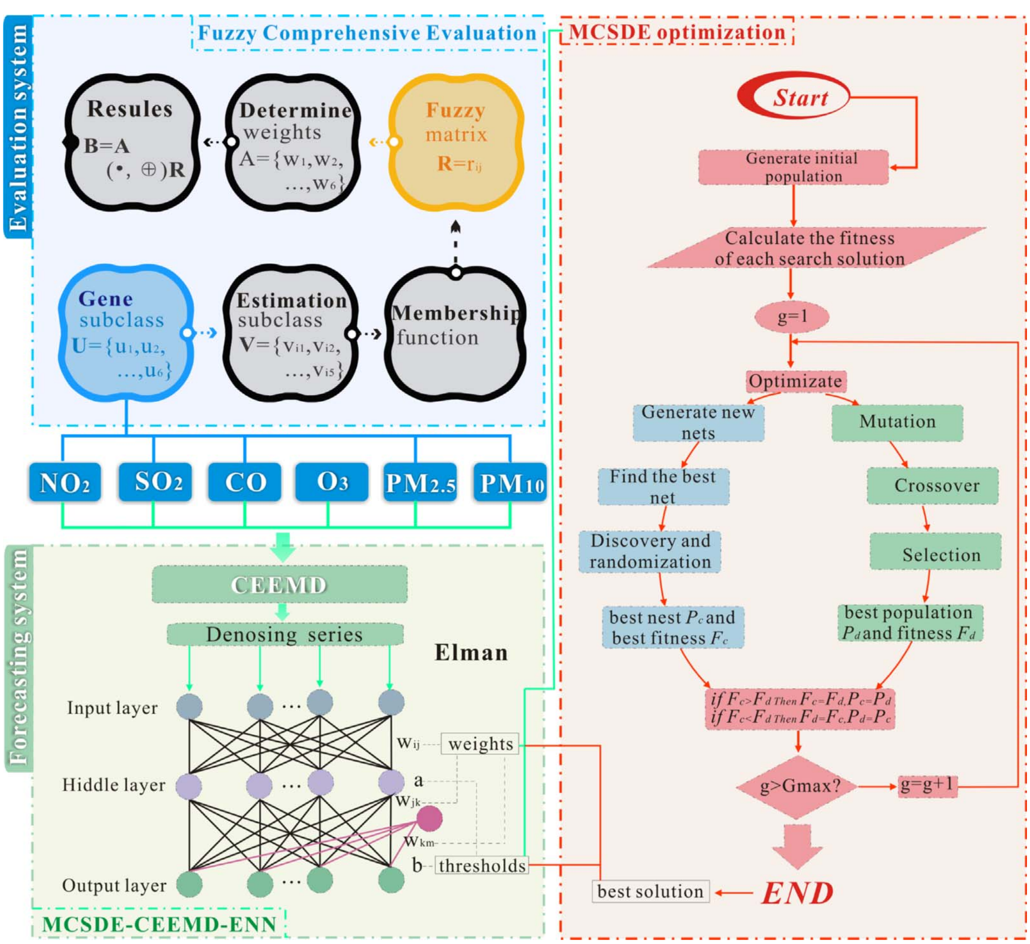


Fig. 2. Flowchart of air quality system.

Table 4
The statistical value of the data used for the simulation.

	City	Mean	Std.	Min	Max	Median
SO_2 (μg/m³)	Xian	43.67	17.58	9	122	42
	Jinan	76	34	14	251	71
NO_2 (μg/m³)	Xian	61.38	22.36	11	144	60
	Jinan	70	28	10	152	70
CO (mg/m³)	Xian	3.08	1.11	0.97	6.08	3.03
	Jinan	2	1	0.51	8.74	1.91
O_3 (μg/m³)	Xian	20.58	12.09	9	66	16
	Jinan	21	15	5	82	16
PM_{10} (μg/m³)	Xian	217.91	112.66	30	527	207.5
	Jinan	222	128	37	706	190
$PM_{2.5}$ (μg/m³)	Xian	119.22	73.26	12	316	115
	Jinan	150	99	11	608	128.5

$$MAE = \frac{1}{T}\sum_{t=1}^{T} |y_t - \hat{y}_t| \tag{19}$$

$$MAPE = \frac{1}{T}\sum_{t=1}^{T} \left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100\% \tag{20}$$

where y_t and $\hat{y}_t$ denote the real and predicted values at time t, respectively.

3.1.1. Diebold-Mariano (D-M) test

The Diebold-Mariano test, proposed by Diebold and Mariano (Diebold and Mariano, 1995), is a hypothesis test that focuses on the predictive accuracy. The hypothesis test is defined as:

$$H_0: E[L(\varepsilon_{n+t}^1)] = E[L(\varepsilon_{n+t}^2)] \tag{21}$$

$$H_1: E[L(\varepsilon_{n+t}^1)] \neq E[L(\varepsilon_{n+t}^2)] \tag{22}$$

where $L(\bullet)$ is the loss function, two popular versions of which are MAE and MSE, and ε_{n+t}^1 and ε_{n+t}^2 are the forecast errors of two forecasting models. The null hypothesis is that the two forecasts have the same accuracy.

The formula for the DM test is as follows:

$$DM = \frac{\sum_{t=1}^{T} (L(\varepsilon_{n+t}^1) - L(\varepsilon_{n+t}^2))/T}{\sqrt{s^2/T}} s^2 \tag{23}$$

where s^2 is an estimator of the variance of $d_t = L(\varepsilon_{n+t}^1) - L(\varepsilon_{n+t}^2)$. Under the null hypothesis, the test statistics DM are asymptotically N(0,1) distributed.

3.1.2. Bias-variance framework

The bias-variance framework (Xiao et al., 2015b) is used to estimate the accuracy and stability of the forecasting model, which is important when evaluating the effectiveness of the air pollutant concentrations prediction model. The absolute bias describes the average difference between the expected predicted value and the true value across all observed and predicted data. The greater the bias, the greater the deviation from the actual data. Variance describes the range of variation of the predicted value, and the degree of dispersion the distance from its expected value. The larger the variance, the more dispersed the data.

The bias-variance framework is decomposed as follows:

$$E(\hat{x}) = \frac{1}{N}\sum_{i=1}^{n} \hat{x}_i \tag{24}$$

$$E(x) = \frac{1}{N}\sum_{i=1}^{n} x_i \tag{25}$$

$$\begin{aligned} E(\hat{x} - x)^2 &= E[\hat{x} - E(\hat{x}) + E(\hat{x}) - x]^2 \\ &= E(\hat{x} - E(\hat{x}))^2 + [E(\hat{x}) - x]^2 \\ &\mathrm{Var}(\hat{x}) + \mathrm{Bias}^2(\hat{x}) \end{aligned} \tag{26}$$

where $\hat{x}_i$ and x_i denote the forecasting value and true value, respectively, and $E(\hat{x})$ $E(x)$ are the expected forecasted and actual values. A smaller $Bias(\hat{x})$ shows superior forecast accuracy. Similarly, a smaller $Var(\hat{x})$ indicates superior stability.

However, in this paper we need to validate the prediction ability of the proposed model for different pollutant concentrations. The statistical values of different air pollutants are very different, especially the mean values and variances. It is unscientific to simply use the variance of forecasting values to measure the stability of models. The average of variance of each pollutant is not sufficient to prove the stability of the prediction model. Therefore, we proposed a variance ratio that combines the variances of the predicted value and the true value to better illustrate the stability of the prediction model. The higher the variance ratio, the higher the prediction stability of the model. The variance ratio is described below:

$$VR = \min\{Var_{predicted}/Var_{real}, Var_{real}/Var_{predicted}\} \tag{27}$$

where VR is the variance ratio, and $Var_{predicted}$ and Var_{real} are the variances of the predicted value and the real value, respectively.

3.2. Data description and parameters of the related models

To verify the air quality fuzzy evaluation and air pollutant concentration prediction accuracy, the air pollutant concentration data (June 1, 2015 to May 31, 2016) from the cities of Xi'an and Jinan, in China, are used in this paper. In the air quality rankings of China's 74 major cities, Xi'an and Jinan both ranked in the bottom ten. It is thus imperative for the two cities to scientifically assess air quality and accurately predict air pollutant concentration.

The quarterly data for each city is used to demonstrate the fuzzy comprehensive evaluation. The one-hour datasets from December 4, 2015 to January 26, 2016 are used to verify the predictive ability of the proposed model. The total number of simulations is 1172, with 1004 used as training data, and 168 for verifying data. Table 4 presents the statistical values of the data used for the simulation. Fig. 3 shows the six major air pollutant concentration datasets for the two cities and how training data and verifying data are updated.

From Table 4 and Fig. 3, several features can be summarized:

(a) The data for the same air pollutants in the two cities are similar.
(b) The experimental data revealed the chaotic nature and inherent complexity of air pollutant concentration.
(c) The lifespan of the models is one day; the training set and test set will be updated after providing 24 hourly-forecasts. The model requires continuous validation for 7 days.
(d) For the mean value, standard deviation, and median of each air pollutant concentration, Jinan's values are greater than those of Xi'an, in addition to carbon monoxide.

In this paper, we employed five single-forecasting models and five hybrid models, including BP, double-layer back propagation neural network (2-BP), RBF, GRNN, Elman, CS-ENN, DE-ENN, MCSDE-ENN, MCSDE-SSA-ENN, and MCSDE-EEMD-ENN as benchmarks to evaluate and compare the proposed hybrid model. We need to set the parameters scientifically and rationally to ensure a fair experiment between the proposed model and the benchmark model. However, there is no clear theory for setting the number of neurons in network. We have conducted a large number of experiments and set the number of neurons per model according to the most accurate one. The parameters of several of the main models are presented in Table 5.

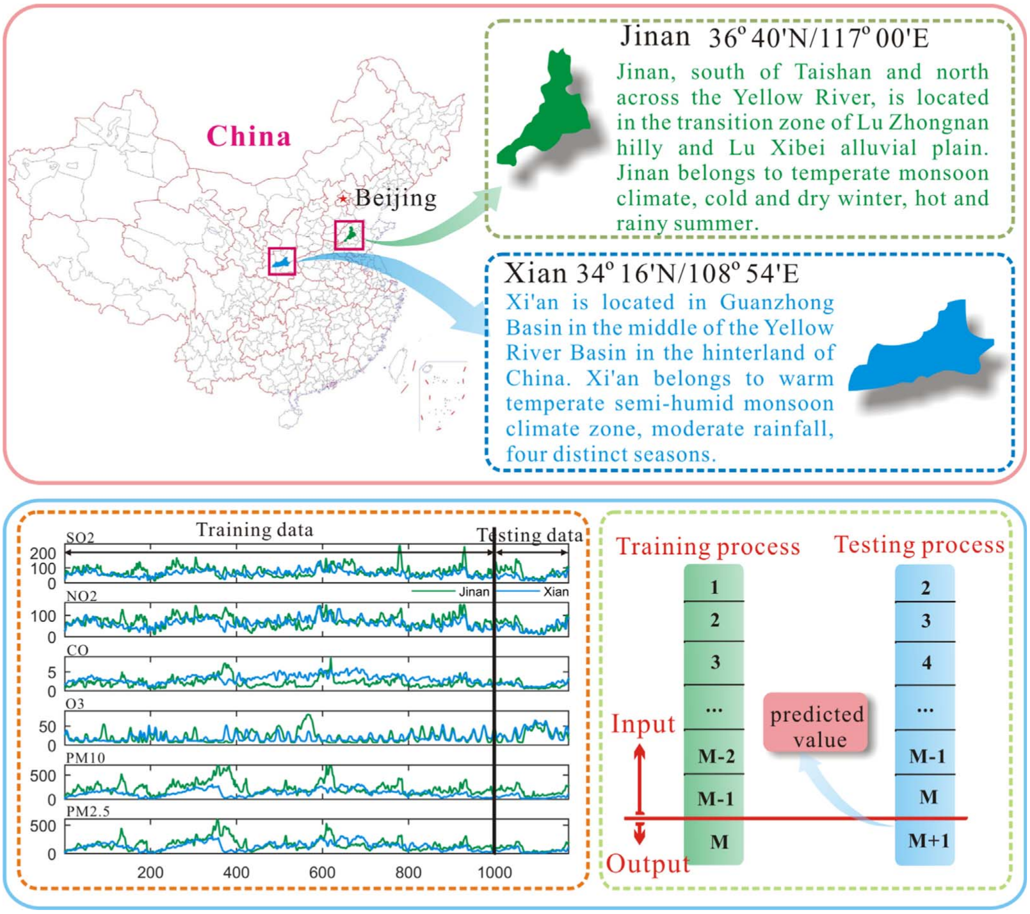


Fig. 3. Six major air pollutant concentration datasets from the two cities.

Table 5
The parameters of several main models.

Model	Experimental parameters	Default value
MCSDE	maximum number of iterations	100
	population size	40
	the number of dimension	56
	search boundaries	[−5,5]
	proportion of coefficient	0.5
	mutation probability	0.1
	discovery rate of alien eggs	0.25
CEEMD	noise standard deviation	0.2
	number of realizations	50
	maximum number of sifting iterations allowed	500
ENN	input layer	4
	middle layer	5
	output layer	1
	training requirement accuracy	0.000001
	learning rate	0.1
	iteration time	2000

Table 6
The air quality levels and the corresponding concentration limits.

Pollution level	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
Good	50	50	2	100	50	35
Regular	100	150	4	160	150	75
Bad	150	475	14	215	250	115
Very bad	200	800	24	265	350	150
Extremely bad	300	1600	36	800	420	250

Table 7
The weight set of each pollution factor.

City	Season	SO_2	NO_2	CO	O_3	PM_{10}	$PM_{2.5}$
	autumn	0.0534	0.0529	0.0627	0.2031	**0.3447**	0.2832
	winter	0.0923	0.0556	0.0759	0.0786	**0.3715**	0.3260
Xian	spring	0.1101	0.0426	0.0759	0.0382	0.3639	**0.3692**
	summer	0.0786	0.0589	0.0594	0.1163	**0.3823**	0.3045
	average	0.0836	0.0525	0.0685	0.1091	**0.3656**	0.3207
	autumn	0.0962	0.035	0.0377	0.1974	0.3039	**0.3298**
	winter	0.1121	0.0484	0.0471	0.0866	0.3263	**0.3794**
Jinan	spring	0.1718	0.0394	0.0457	0.0389	0.3107	**0.3936**
	summer	0.1275	0.0371	0.0334	0.1314	**0.3668**	0.3037
	average	0.1269	0.04	0.041	0.1136	0.3269	**0.3516**

3.3. Fuzzy comprehensive evaluation of air quality

In this section, quarterly averages are used to evaluate air quality and determine the most important air pollutants. Following the steps described in Section 2.1, we establish evaluation sets V = {good, regular, bad, very bad and extremely bad}, see Table 6. Then we calculate the membership degree of each evaluation factor to each evaluation criterion, using the membership degree calculation formula, and establish the fuzzy relation matrix.

The weight value of pollution factors is a measure of the impact of pollutants on ambient air quality, and directly affects the results of the comprehensive evaluation. Therefore, it is important to determine the weight of each evaluation parameter in the fuzzy comprehensive evaluation. The weights of six major air pollutants are presented in Table 7. From the average weights we can see that PM_{10} is the main pollutant in the ambient air of Xi'an during the evaluation period, followed by

Table 8
The air quality evaluation results of different cities.

city	season	Good	Regular	Bad	Very bad	Extremely bad	
	autumn	**0.82**	0.18	0.00	0.00	0.00	Good
Xian	Winter	**0.60**	0.40	0.00	0.00	0.00	Good
	Spring	0.24	**0.39**	0.37	0.00	0.00	Regular
	Summer	0.46	**0.54**	0.00	0.00	0.00	Regular
	autumn	0.45	**0.55**	0.00	0.00	0.00	Regular
Jinan	Winter	0.29	**0.55**	0.16	0.00	0.00	Regular
	Spring	0.23	0.25	**0.42**	0.10	0.00	Bad
	Summer	0.34	**0.57**	0.09	0.00	0.00	Regular

$PM_{2.5}$, O_3, SO_2, CO, and NO_2. In Jinan, the main pollutants were $PM_{2.5}$, PM_{10}, SO_2, O_3, CO, and NO_2. The main air pollutant in the two cities is different. For the same city, the main air pollutant differs seasonally as well. The two most important air pollutants are PM_{10} and $PM_{2.5}$.

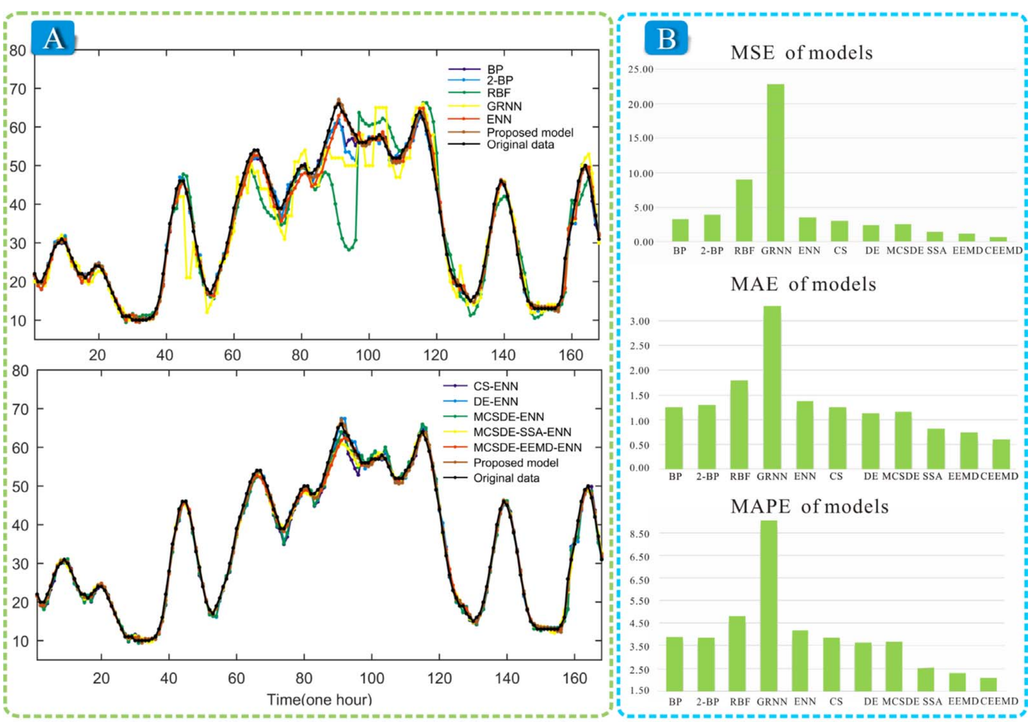
The weight matrix and the fuzzy matrix are combined to obtain the fuzzy matrix. Based on the principle of maximum membership, the maximum values of b1, b2, b3, b4, and b5 are the result of this ambient air quality evaluation, as seen in Table 8. For Xi'an, the pollution levels in autumn and winter are good, and the levels of spring and summer are regular. For Jinan, the levels of autumn, winter and summer are regular, and the level of spring is bad. Overall, the air quality of Xi'an is better than the air quality of Jinan.

3.4. Case studies

In this paper, one-hour data from two cities were selected to verify the validity of the proposed forecasting model. To validate the model's performance, five single models and five hybrid models were used to predict the six major air pollutant concentrations. The lifespan of these models was one day, and the training data and verifying data were updated after providing 24 predicted values. The model was continuously validated for one week, yielding 168 predicted values. Two experiments will now be conducted, the first one concerning the air pollution data from Xi'an, and the second from Jinan.

3.4.1. Case study one

In this section, one-hour air pollutant concentration series from



C

The results of DM test.

Model	DM value	Model	DM value
BP	4.53*	CS-ENN	5.17*
2-BP	3.97*	DE-ENN	4.23*
RBF	3.66*	MCSDE-ENN	4.78*
GRNN	5.33*	MCSDE-SSA-ENN	3.02*
ENN	6.03*	MCSDE-EEMD-ENN	2.26*

* is the 1% significance level; ** is the 10% significance level

Fig. 4. The O_3 concentration forecasting results of related models.

Table 9
Forecasting results obtained using the air pollution data from Xian.

		BP	2_BP	RBF	GRNN	ENN	CSENN	DEENN	MCSDEENN	MCSDESSAENN	MCSDEEEMDENN	Proposed model
SO_2	MSE	17.69	16.46	17.74	19.3152	16.37	16.97	16.37	15.80	5.24	5.22	**3.49**
	MAE	2.70	2.70	2.95	3.3158	2.64	2.63	2.61	2.56	1.53	1.43	**1.32**
	MAPE (%)	11.78	11.81	12.99	15.4771	11.46	11.40	11.03	10.74	6.61	6.40	**6.12**
NO_2	MSE	26.51	24.35	31.86	45.6012	19.78	19.00	19.45	18.96	7.66	3.83	**3.64**
	MAE	3.81	3.81	4.13	5.2362	3.40	3.34	3.37	3.30	2.21	1.51	**1.45**
	MAPE (%)	13.50	13.74	15.38	17.9798	11.58	11.49	11.43	10.94	7.53	5.16	**4.88**
CO	MSE	0.01	0.01	0.01	0.0402	0.01	0.01	0.01	0.01	0.01	0.00	**0.00**
	MAE	0.08	0.08	0.08	0.1506	0.08	0.07	0.08	0.07	0.06	0.04	**0.04**
	MAPE (%)	5.13	4.72	5.03	9.4754	4.59	4.49	4.62	4.34	3.53	**2.32**	2.67
O_3	MSE	3.18	3.86	8.99	22.7726	3.47	2.93	2.29	2.46	1.40	1.11	**0.60**
	MAE	1.25	1.29	1.78	3.3050	1.38	1.24	1.11	1.16	0.81	0.73	**0.59**
	MAPE (%)	3.86	3.84	4.79	9.0200	4.16	3.84	3.62	3.64	2.52	2.27	**2.05**
PM_{10}	MSE	147.83	135.51	147.09	389.6694	131.38	134.36	131.63	125.59	73.05	**33.92**	36.56
	MAE	9.17	8.84	8.95	14.2602	8.52	8.75	8.47	8.35	6.62	**4.48**	4.57
	MAPE (%)	9.45	9.14	9.29	14.8910	8.61	9.02	8.49	8.23	6.96	5.02	**4.76**
$PM_{2.5}$	MSE	27.86	26.46	28.43	77.8489	25.28	24.29	25.39	24.44	11.39	**6.25**	6.73
	MAE	3.83	3.62	3.44	5.8482	3.53	3.43	3.56	3.41	2.56	**1.88**	1.97
	MAPE (%)	10.51	9.39	8.57	15.4022	9.29	8.73	9.23	8.46	7.73	**5.45**	5.82
Average	MSE	37.18	34.44	39.02	92.54	32.72	32.93	32.52	31.21	16.46	**8.39**	8.50
	MAE	3.47	3.39	3.56	5.35	3.26	3.24	3.20	3.14	2.30	1.68	**1.66**
	MAPE (%)	9.04	8.77	9.34	13.71	8.28	8.16	8.07	7.73	5.81	4.44	**4.38**

Table 10
Results for DM test and bias-variance framework of Xian.

Model	Bias-variance framework		D-M value
	Bias	Variance ratio	
BP	1.2155	0.9372	5.28*
2_BP	1.1240	0.9257	5.37*
RBF	2.8492	0.7865	7.40*
GRNN	1.6042	0.9536	6.31*
ENN	0.9141	0.9366	5.52*
CS-ENN	1.0248	0.9445	5.26*
DE-ENN	0.8576	0.9484	5.30*
MCSDE-ENN	0.7724	0.9604	5.22*
MCSDE-SSA-ENN	0.2642	0.9394	3.66*
MCSDE-EEMD-ENN	0.3302	0.9530	1.67**
MCSDE-CEEMD-ENN	**0.2890**	**0.9607**	–

* is the 1% significance level;
** is the 10% significance level.

Xi'an is used to verify the forecasting ability of the proposed model. Five single forecasting models and five combined models are utilized in this section. To further demonstrate the predictive accuracy and stability of the model, the DM test and bias variance framework were used in this experiment, in addition to the traditional error representations. The O_3 concentration forecasting results of related models are presented in Fig. 4. The prediction curve is shown in part A, and the results of three error criteria and DM values are shown in parts B and C, respectively. Table 9 shows the six main air pollutant concentration forecasting results, measured by MAE, MAPE, and MSE, and forecasted by related models. The average results of the DM test and the bias variance framework are presented in Table 10.

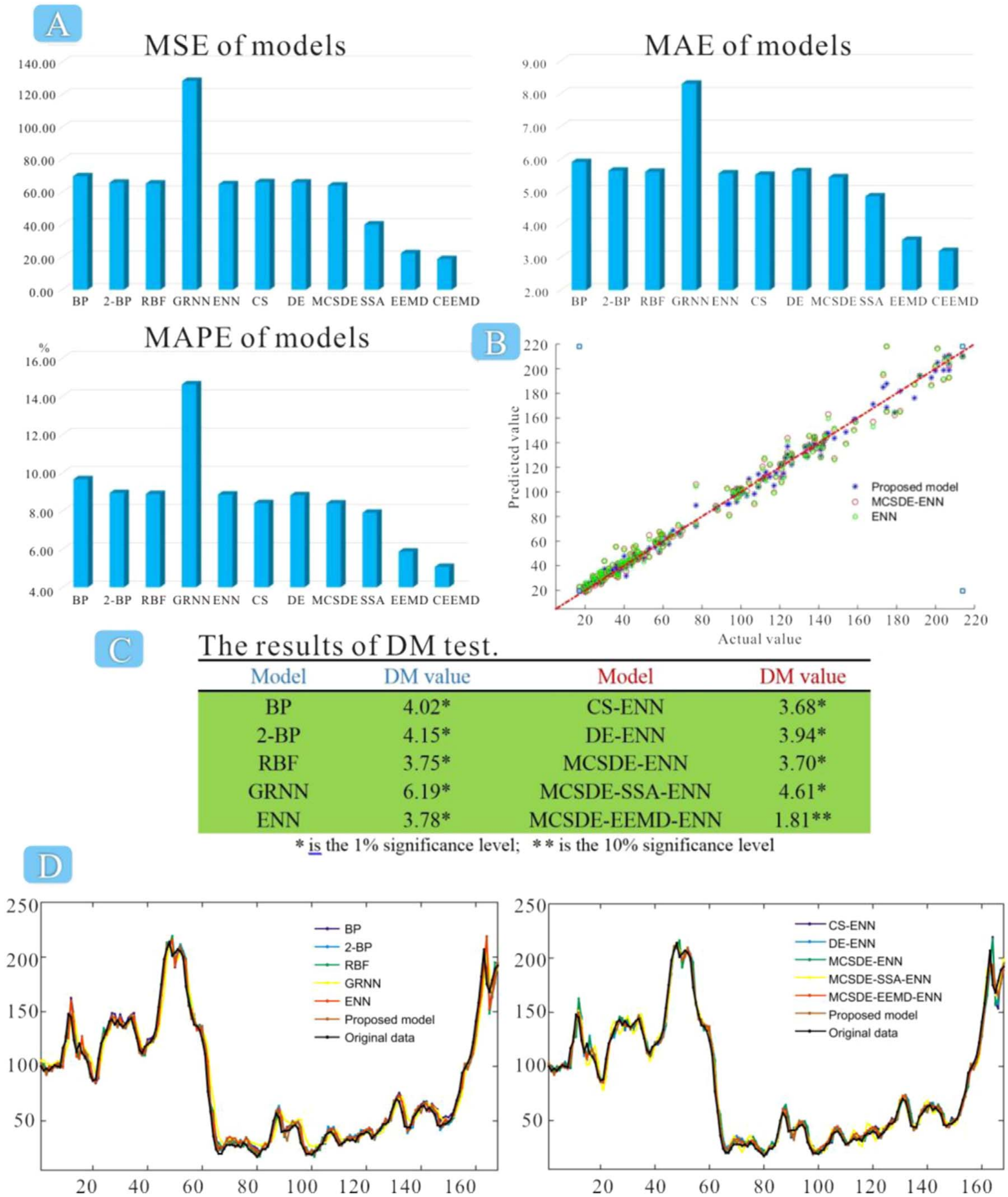
For these five models, the smallest statistical errors are consistently found from the ENN model by comparing the average of statistical errors. The average MSE, average MAE, and average MAPE of ENN are 32.72%, 3.26%, and 8.28%, respectively. However, not all the best values of MSE, MAE, and variance ratio were from the ENN model, the minimum values of MSE and MAE for PM2.5 were from the RBF model, and the best variance ratio from the GRNN model, but the ENN model has the best performance for the majority of the samples. The simulation results of the single models show that ENN is the most accurate model with the best performance in statistical errors. For the bias variance framework, ENN has the smallest bias, which also verifies the accuracy of ENN. ENN has the second best variance ratio, which shows good stability. The ENN model was therefore the best choice as the prediction model.

When MCSDE-ENN, CS-ENN, and DE-ENN, were compared, MCSDE-ENN had the smallest average MAPE, average MSE, and average MAE. For example, the MAPE of MCSDE-ENN increased by 0.43% and 0.34% compared to CS-ENN and DE-ENN, respectively. MCSDE-ENN has the best statistical error in addition to O_3, but the difference between the O_3 error and the minimum value is very small. In addition, the DE-ENN model performed well with O_3 forecasting, but both the bias and variance ratio of the model DE-ENN were worse than the MCSDE-ENN model. MCSDE-ENN offered the best bias and variance ratio, and the bias and variance ratio were 0.7724 and 0.9604, respectively. From the bias-variance and average error of MCSDE-ENN, we can see that the proposed optimization model MCSDE has the best optimization capabilities.

Compared to the proposed model MCSDE-CEEMD-ENN with MCSDE-SSA-ENN and MCSDE-EEMD-ENN, the proposed model has the smallest average MAE and average MAPE. Although the smallest average MSE belongs to MCSDE-EEMD-ENN, the bias and variance ratio of model MCSDE-EEMD-ENN are worse than the proposed. Among all the models, the smallest bias and the biggest variance ratio are from the proposed model. From the bias-variance framework, the proposed model has the best accuracy and stability. The DM value of MCSDE-CEEMD-ENN is larger than the upper limits at the 10% significance level, and the DM values of other models are larger than the upper limits at the 1% significance level. From the simulation results of several evaluated metrics, the proposed model has the best preference in bias-variance framework and DM test. Not every minimum error was from the proposed model, but the minimum and the predictive values of the proposed model are very close. In general, the proposed hybrid model has the best accuracy and stability.

3.4.2. *Case study two*

In this section, a one-hour air pollutants series from Jinan was used to evaluate the forecasting ability of the proposed model. The forecasting results of $PM_{2.5}$ are shown in Fig. 5. The forecasting results of six kinds of air pollutants by five single models and five combined models are shown in Table 11 and Table 12.



Model	DM value	Model	DM value
BP	4.02*	CS-ENN	3.68*
2-BP	4.15*	DE-ENN	3.94*
RBF	3.75*	MCSDE-ENN	3.70*
GRNN	6.19*	MCSDE-SSA-ENN	4.61*
ENN	3.78*	MCSDE-EEMD-ENN	1.81**

* is the 1% significance level; ** is the 10% significance level

Fig. 5. Forecasting results of $PM_{2.}$

Although not all of the error values of the proposed model are the smallest in the forecasting of the concentrations of the six major air pollutants, the average errors evaluated by three statistical metrics are the smallest. The average MSE, MAE, and MAPE are 19.04%, 2.42% and 4.87%, respectively. From the DM test, the value of MCSDE-CEEMD-ENN is larger than the upper limits at the 10% significance level, and the DM values from the other models are larger than the upper limits at the 1% significance level. To verify the accuracy and stability of related models, a bias variance framework is required. From the values of bias and variance, we can see that the smallest bias is from the proposed model, which verifies that the model has the best accuracy. The variance ratio of the proposed model is the second best, which shows good stability. From the above, the proposed model has the best accuracy and stability of all the models.

3.5. Summary

The average results of case studies one and two are presented in Table 13. The values in bold indices are the smallest MSE, MAE, and MAPE, and the best bias and variance ratio. The MSE, MAE, and MAPE of the proposed hybrid model are 13.77%, 2.04%, and 4.63%, respectively. The DM values of MCSDE-CEEMD-ENN are larger than the upper limits at the 10% significance level, and the DM values of other models are larger than the upper limits at the 1% significance level. The result indicates that the proposed hybrid model is significantly better than other models. For bias variance framework, the bias and variance ratio of the proposed model are 0.2828 and 0.9580, respectively. The smallest bias confirms the accuracy of the proposed model, and the largest variance ratio indicates that the model has the best stability.

Table 11
Forecasting results obtained using the air pollution data from Jian.

		BP	2_BP	RBF	GRNN	ENN	CSENN	DEENN	MCSDEENN	MCSDESSAENN	MCSDEEEMDENN	Proposed model
SO_2	MSE	73.16	68.27	72.57	86.31	72.72	70.33	72.25	71.25	33.21	18.13	**15.58**
	MAE	5.95	5.82	5.87	6.90	5.85	5.75	5.79	5.73	4.65	3.13	**2.87**
	MAPE (%)	10.29	10.27	9.94	13.16	9.94	9.88	9.79	9.60	8.82	5.71	**4.87**
NO_2	MSE	24.25	23.50	23.49	42.30	23.11	23.82	24.89	24.48	12.62	5.50	**4.99**
	MAE	3.90	3.82	3.81	4.85	3.77	3.77	3.85	3.79	2.87	1.81	**1.78**
	MAPE (%)	12.01	11.41	11.29	16.17	11.31	11.18	11.30	11.02	8.42	6.10	**5.97**
CO	MSE	0.04	0.04	0.04	0.0590	0.04	0.04	0.04	0.04	0.02	0.01	**0.01**
	MAE	0.13	0.13	0.13	0.1844	0.14	0.14	0.14	0.13	0.11	0.07	**0.07**
	MAPE (%)	10.15	9.52	9.14	15.08	10.90	10.44	10.39	9.80	9.81	5.50	**5.39**
O_3	MSE	2.57	6.31	6.59	6.40	2.39	2.42	2.40	2.50	1.07	0.69	**0.63**
	MAE	1.17	1.65	1.96	1.85	1.12	1.12	1.13	1.14	0.78	0.61	**0.61**
	MAPE (%)	5.36	5.02	8.30	6.95	5.00	4.94	4.95	4.94	3.81	**2.97**	2.98
PM_{10}	MSE	350.63	344.97	337.21	462.61	324.54	325.66	329.03	319.26	165.05	76.07	**74.15**
	MAE	13.08	12.83	12.69	15.03	12.62	12.69	12.60	12.43	9.79	6.41	**6.02**
	MAPE (%)	10.49	9.90	9.93	11.81	9.87	9.95	9.62	9.60	8.64	5.38	**4.96**
$PM_{2.5}$	MSE	69.78	65.79	65.33	128.38	64.90	66.13	65.95	64.07	40.14	22.43	**18.88**
	MAE	5.91	5.65	5.61	8.30	5.56	5.52	5.63	5.44	4.86	3.53	**3.19**
	MAPE (%)	9.65	8.93	8.88	14.61	8.85	8.40	8.81	8.39	7.90	5.86	**5.07**
Average	MSE	86.74	84.81	84.20	121.01	81.28	81.40	82.43	80.27	42.02	20.47	**19.04**
	MAE	5.02	4.98	5.01	6.19	4.84	4.83	4.86	4.78	3.85	2.59	**2.42**
	MAPE (%)	9.66	9.17	9.58	12.96	9.31	9.13	9.14	8.89	7.90	5.25	**4.87**

Table 12
Results for DM test and bias-variance framework of Jinan.

Model	Bias-variance framework		D-M value
	Bias	Variance ratio	
BP	1.2981	0.9347	5.30*
2_BP	0.9244	0.9417	5.42*
RBF	0.8558	0.9440	5.57*
GRNN	1.5235	0.8946	6.15*
ENN	1.0238	0.9295	5.17*
CS-ENN	0.9686	0.9282	5.05*
DE-ENN	0.8421	0.9357	5.08*
MCSDE-ENN	0.9355	0.9368	5.05*
MCSDE-SSA-ENN	0.3755	**0.9661**	4.79*
MCSDE-EEMD-ENN	0.5181	0.9513	1.76**
MCSDE-CEEMD-ENN	**0.2767**	0.9553	–

* is the 1% significance level;
** is the 10% significance level

Table 13
The average results of study case one and two.

Model	Error			Bias-Variance framework		D-M test
	MSE	MAE	MAPE	Bias	Variance ratio	
BP	61.96	4.25	9.35	1.2568	0.9359	5.29*
2_BP	59.63	4.19	8.97	1.0242	0.9337	5.39*
RBF	61.61	4.29	9.46	1.8525	0.8653	6.48*
GRNN	106.78	5.77	13.34	1.5639	0.9241	6.23*
ENN	57.00	4.05	8.80	0.9689	0.9331	5.34*
CS-ENN	57.17	4.04	8.65	0.9967	0.9364	5.15*
DE-ENN	57.48	4.03	8.61	0.8499	0.9421	5.19*
MCSDE-ENN	55.74	3.96	8.31	0.8540	0.9486	5.13*
MCSDE-SSA-ENN	29.24	3.08	6.86	0.3199	0.9527	4.22*
MCSDE-EEMD-ENN	14.43	2.14	4.85	0.4242	0.9521	1.72**
MCSDE-CEEMD-ENN	**13.77**	**2.04**	**4.63**	**0.2828**	**0.9580**	–

* is the 1% significance level;
** is the 10% significance level.

4. Conclusion

High levels of air pollution have obvious effects on human health, animals, plants and the environment, causing respiratory diseases and physiological dysfunction. Therefore, it is urgent and meaningful to establish an air quality monitoring and early warning system to evaluate the degree of air pollution scientifically, and forecast air pollutant concentrations more accurately. In this paper, fuzzy comprehensive evaluation was used to determine the main air pollutants and evaluate the level of air pollution. The evaluation results showed that the air quality of Xi'an is superior to the air quality of Jinan, and the main pollutant of Jinan and Xi'an are $PM_{2.5}$ and PM_{10}, respectively.

The prediction of air pollutant concentration is an important part of an air quality early warning system. More accurate forecasting could reduce the impact of air pollution on people's health and guide people's work and life. However, the chaotic nature and complexity of the air pollution data itself make prediction very difficult. In this paper, a novel hybrid model was proposed to forecast the concentration of six major air pollutants. CEEMD was used to capture the actual trend of the data and reduce the impact of noise on the prediction results, and the new optimization model combined the CS and DE models to optimize the initial weights and thresholds. To further verify the stability and accuracy of the model, a bias variance framework and DM test were used in an error evaluation system. The experimental results showed that the proposed model has the best accuracy and stability compared to the other five individual models and five combined models. For example, the errors measured by MAPE were reduced by 8.71%, 4.17%, 3.68%, and 0.22% compared to the GRNN, ENN, MCSDE-ENN and MCSDE-EEMD-ENN, respectively. This air quality monitoring and early warning system is highly effective for air quality assessment and air pollutant concentration prediction.

Acknowledgements

This work was supported by the National Social Science Foundation of China (Grant564 No. 13 & ZD171).

References

Abdullah, L., Khalid, N.D., 2012. Classification of air quality using fuzzy synthetic

multiplication. Environ. Monit. Assess. 184, 6957–6965.
Akkaya, G., Turanoğlu, B., Özta, S., 2015. An integrated fuzzy AHP and fuzzy MOORA approach to the problem of industrial engineering sector choosing. Expert Syst. Appl. 42 (24), 9565–9573.
Alhanafy, T.E., Zaghlool, F., El, A.S., Moustafa, D., 2010. Neuro fuzzy modeling scheme for the prediction of air pollution. J. Am. Sci. 6, 605–616.
Brook, R.D., Franklin, B., Cascio, W., Hong, Y., Howard, G., Lipsett, M., Luepker, R., Mittleman, M., Samet, J., Smith Jr., S.C., Tager, I., 2004. Air pollution and cardiovascular disease: a statement for health care professionals from the expert panel on population and prevention science of the American heart association. Circulation 109 (21), 2655–2671.
Baker, K.R., Foley, K.M., 2011. A nonlinear regression model estimating single source concentrations of primary and secondarily formed PM2.5. Atmos. Environ. 45, 3758–3767.
Cheng, W.L., Chen, Y.S., Zhang, J., Lyons, T.J., Pai, J.L., Chang, S.H., 2007. Comparison of the revised air quality index with the PSI and AQI indices. Sci. Total Environ. 382, 191–198.
Diebold, F.X., Mariano, R., 1995. Comparing predictive accuracy. J. Bus. Econ. Stat. 13, 253–263.
EPI (Environmental Performance Index), 2016. Global Metrics for the Environment-The Environmental Performance Index Ranks Countries Performance on High-priority Environmental Issues.〈http://environment.yale.edu/news/article/2016-yale-environmental-performance-index-released/〉.
US EPA, 2009. Air Quality Index: A Guide to Air Quality and Your Health.
Elman, J.L., 1990. Finding structure in time. Cogn. Sci. 14, 179–211.
Feng, Xiao, Li, Qi, Zhu, Yajie, Hou, Junxiong, Jin, Lingyan, Wang, Jingjie, 2015. Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation. Atmos. Environ. 107, 118–128.
Gorai, A.K., Tuluri, F., Tchounwou, P.B., 2014. A GIS based approach for assessing the association between air pollution and asthma in New York state, USA. Int. J. Environ. Res. Public Health 11, 4845–4869.
Geoffrey Cobourn, W., 2010. An enhanced PM2.5 air quality forecast model based on nonlinear regression and back-trajectory concentrations. Atmos. Environ. 44 (5), 3015–3023.
Gandomi, A.H., Talatahari, S., Yang, X.S., et al., 2013. Design optimization of truss structures using cuckoo search algorithm. Struct. Des. Tall Spec. Build. 22, 1330–1349.
Hájek, P., Olej, V., 2009. Air pollution assessment using hierarchical fuzzy inference systems. Sci. Pap. Univ. Pardubic. Univ. Pardubic. 52–62.
Huang, N.E., Shen, Z., Long, S.R., Wu, M.C., Shih, H.H., Zheng, Q., Yen, N.-C., Tung, C.C., Liu, H.H., 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proc. R. Soc. A: Math., Phys. Eng. Sci. 454, 903–995.
Kyrkilis, G., Chaloulakou, A., Kassomenos, P.A., 2007. Development of an aggregate Air Quality Index for an urban Mediterranean agglomeration: relation to potential health effects. Environ. Int. 33, 670–676.
Konovalov, I.B., Beekmann, M., Meleux, F., Dutot, A., Foret, G., 2009. Combining deterministic and statistical approaches for PM10 forecasting in Europe. Atmos. Environ. 43, 6425–6434.
Kukkonen, J., Partanen, L., Karppinen, A., Ruuskanen, J., Junninen, H., Kolehmainen, M., Niska, H., Dorling, S., Chatterton, T., Foxall, R., Cawley, G., 2003. Extensive evaluation of neural network models for the prediction of NO2 and PM10 concentrations, compared with a deterministic modeling system and measurements in central Helsinki. Atmos. Environ. 37, 4549–4550.
Liu, Chen, Lin, T., Chen, C., Chen, P., Wen, Sun, C., Juang, J., Jiang, J., 2012. An air quality monitoring system for urban areas based on the technology of wireless sensor networks. Int. J. Smart Sens. Intell. Syst. 5, 191–214.
Lin, W., Gow, H., Tsai, M., 2010. An enhanced radial basis function network for short-term electricity price forecasting. Appl. Energy 87 (10), 3226–3234.
Lin, K.P., Pai, P.F., Yang, S.L., 2011. Forecasting concentrations of air pollutants by logarithm support vector regression with immune algorithms. Appl. Math. Comput. 217, 5318–5327.
Lin, W.M., Hong, C.M., 2011. A new Elman neural network-based control algorithm for adjustable-pitch variable-speed wind-energy conversion systems. IEEE Trans. Power Electron. 26 (2), 473–481.
Mohammad, Hamed, Masud, Maryam T.M., Saeedeh, L., 2011. A novel, fuzzy-based air quality index (FAQI) for air quality assessment. Atmos. Environ. 45, 2050–2059.
Metaxiotis, K., Kagiannas, A., Askounis, D., Psarras, J., 2003. Artificial intelligence in short term electric load forecasting: a state-of-the-art survey for the researcher. Energy Convers. Manag. 44, 1525–1534.
Miguel, José J.C., Ignacio, Luis P.S., 2016. Air quality assessment using a weighted fuzzy inference system. Ecol. Inf. 33, 57–74.
MEP (Ministry of Environmental Protection), 2015. Enhancing Capability to Support Decision – A Summary of Environmental Monitoring in the First Half of 2015. 〈http://www.mep.gov.cn/xxgk/hjyw/201508/t20150804_307682.shtml〉.
Marichelvam, M.K., Prabaharan, T., Yang, X.S., 2014. Improved cuckoo search algorithm for hybrid flow shop scheduling problems to minimize makespan. Appl. Soft Comput. 19, 93–101.
Osowski, S., Garanty, K., 2007. Forecasting of the daily meteorological pollution using wavelets and support vector machine. Eng. Appl. Artif. Intell. 20, 745–755.
Paschalidou, A.K., Karakitsios, S., Kleanthous, S., Kassomenos, P.A., 2011. Forecasting hourly PM10 concentration in Cyprus through artificial neural networks and multiple regression models: implications to local environmental management. Environ. Sci. Pollut. Res. 18, 316–327.
Qin, S., Liu, F., Wang, J., Song, Y., 2015. Interval forecasts of a novelty hybrid model for wind speeds. Energy Rep. 1, 8–16.
Qin, S., Liu, F., Wang, J., Sun, B., 2014. Analysis and forecasting of the particulate matter (PM) concentration levels over four major cities of China using hybrid models. Atmos. Environ. 98, 665–675.
Sun, W., Zhang, H., Palazoglu, A., Singh, A., Zhang, W., Liu, S., 2013. Prediction of 24-hour average PM2.5 concentrations using a hidden Markov model with different emission distributions in Northern California. Total Environ. 443, 93–103.
Stern, R., Builtjes, P., Schaap, M., Timmermans, R., Vautard, R., Hodzic, A., Memmesheimer, M., Feldmann, H., Renner, E., Wolke, R., Kerschbaumer, A., 2008. A model inter-comparison study focusing on episodes with elevated PM10 concentrations. Atmos. Environ. 42, 4567–4588.
Sun, Wei, Zhang, Hao, Angadh Singh, Ahmet Palazoglub, Zhang, Weidong, 2013. Prediction of 24-hour-average PM2.5 concentration using a hidden Markov model with different emission distributions in Northern California. Sci. Total Environ. 443, 93–103.
Song, Y., Qin, S., Qu, J., Liu, F.,, 2015. The forecasting research of early warning systems for atmospheric pollutants: a case in Yangtze River Delta region. Atmos. Environ. 118, 58–69.
Suarez Sanchez, A., García Nieto, P.J., Riesgo Fernandez, P., Del Coz Díaz, J.J., Iglesias-Rodríguez, F.J., 2011. Application of a SVM-based regression model to the air quality study at local scale in the Aviles urban area (Spain). Math. Comput. Model 54, 1453–1466.
Storn, R., Price, K., 1995. Differential evolution-a simple and efficient adaptive scheme for global optimization over continuous spaces. 3 ICSI, Berkeley.
Tang, L., Dai, W., Yu, L., et al., 2015. A novel CEEMD-based EELM ensemble learning paradigm for crude oil price forecasting. Int. J. Inf. Technol. Decis. Mak. 14 (01), 141–169.
Wang, Huang, Cao, J., Shen, Tie, Han, Zhao, Li, G., Li, Z., Ni, H., Zhou, Ya, Wang, M., Su, Chen, 2016. Contribution of regional transport to the black carbon aerosol during winter haze period in Beijing. Atmos. Environ. 132, 11–18.
Wang, J. Wang, Li, J., Zhu, Y., Zhu, S., Zhao, J., 2014. Techniques of applying wavelet de-noising into a combined model for short-term load forecasting. Int. J. Electr. Power Energy Syst. 62, 816–824.
Wu, Z., Huang, N.E., 2008. Ensemble empirical mode decomposition: a noiseassisted data analysis method. Adv. Adapt. Data Anal. 1, 1–41.
Xiao, L., Wang, J., Yang, X., et al., 2015a. A hybrid model based on data preprocessing for electrical power forecasting. Int. J. Electr. Power Energy Syst. 64, 311–327.
Xiao, L., Wang, J., Hou, R., Wu, J., 2015b. A combined model based on data pre-analysis and weight coefficients optimization for electrical load forecasting. Energy 82, 524–549.
Yahya, K., Zhang, Y., Vukovich, J.M., 2014. Real-time air quality forecasting over the southeastern United States using WRF/Chem-MADRID: multiple-year assessment and sensitivity studies. Atmos. Environ. 92, 318–338.
Yeh, J.R., Shieh, J.S., Huang, N.E., 2010. Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. Adv. Adapt. Data Anal. 2, 135–156.
Yang, X., Deb, S., 2009. Cuckoo search via lévy flights. IEEE World Congr. Nat. Biol. Inspired Comput. 210–214.
Zhao, X., Qi, Q., Li, R., 2010. The establishment and application of fuzzy comprehensive model with weight based on entropy technology for air quality assessment. Procedia Eng. 7, 217–222.
Zadeh, L.A., 1965. Fuzzy sets. Inf. Control 8, 338–353.
Zhang, H., Zhang, W., Palazoglu, A., Sun, W., 2012. Prediction of ozone levels using a hidden markov model (HMM) with gamma distribution. Atmos. Environ. 62, 64–73.