RESEARCH ARTICLE

Short-term air quality forecasting model based on hybrid RF-IACA-BPNN algorithm

De-wen Qiao[1] · Jian Yao[1] · Ji-wen Zhang[1] · Xin-long Li[1] · Tan Mi[1] · Wen Zeng[2]

Abstract

Despite the apparent improvement in air quality in recent years through a series of effective measures, the concentration of $PM_{2.5}$ and O_3 in Chengdu city remains high. And both the two pollutants can cause serious damage to human health and property; consequently, it is imperative to accurately forecast hourly concentration of $PM_{2.5}$ and O_3 in advance. In this study, an air quality forecasting method based on random forest (RF) method and improved ant colony algorithm coupled with back-propagation neural network (IACA-BPNN) are proposed. RF method was used to screen out highly correlated input variables, and the improved ant colony algorithm (IACA) was adopted to combine with BPNN to improve the convergence performance. Two datasets based on two different kinds of monitoring stations along with meteorological data were applied to verify the performance of this proposed model and compared with another five plain models. The results showed that the RF-IACA-BPNN model has the minimum statistical error of the mean absolute error, root mean square error, and mean absolute percentage error, and the values of R^2 consistently outperform other models. Thus, it is concluded that the proposed model is suitable for air quality prediction. It was also detected that the performance of the models for the forecasting of the hourly concentrations of $PM_{2.5}$ were more acceptable at suburban station than downtown station, while the case is just the opposite for O_3, on account of the low variability dataset at suburban station.

Keywords Hourly $PM_{2.5}$ and O_3 concentration prediction · Random forest · Improved ant colony algorithm · BP neural network

Introduction

Air pollution has become a major concern around the world since it is highly correlated with a variety of adverse effects on public health, especially in some developing countries. Nearly 91% of the world's population inhales polluted air, and about 4.2 million deaths occur every year because of ambient air pollution, according to air pollution program of the World Health Organization (WHO) (Krishan et al. 2019). Among the six conventional air pollutants, $PM_{2.5}$ and O_3 constitute the greatest harm to the human bodies. Long-term exposure to $PM_{2.5}$ (fine particles measuring less than 2.5 μm in diameter) air pollution is an important environmental risk factor for cardiopulmonary and lung cancer mortality (Arden Pope III 2002). O_3 (formed as a secondary pollutant through photochemical reactions of NOx and VOCs) is known to significantly decrease crop yield and a key ingredient of smog that is potentially toxic to animal and plant life (Wang et al. 2019). High concentrations of O_3 and everlasting haze pollution are also the main concerns for China at present, especially for Chengdu, Southwest China (Wang et al. 2017a, b). Hence, considering the serious harm and damage of air pollution to human health and agricultural economy, accurate air quality forecasting is of great significance, so as to assist and support the government in city atmospheric early warning, crisis response, and emergency planning and contribute to making travel plan in advance for citizens (Corani and Scanagatta 2016, Fan et al. 2018, Lauret et al. 2016, Prasad et al. 2016, Yang et al. 2018, Yang and Christakos 2015). Especially, short-term air pollution forecast (with lead time

✉ Jian Yao
   676400867@qq.com

1   College of Architecture and Environment, Sichuan University, Chengdu 610065, China

2   Institute for Disaster Management and Reconstruction, Sichuan University-the Hong Kong Polytechnic University, Chengdu, Sichuan, China

of 1 or 2 days) plays an important role in the daily operation of Local Emergency Management Agency.

Generally, the air quality forecasting models can be classified into two main types: the numerical prediction models and the statistical prediction models (Liu et al. 2018; Yang and Wang 2017). The numerical prediction models traditionally consist of the simulation of dispersion and transform mechanisms using emission source data and the knowledge of the transformations in the atmosphere (Zhang et al. 2012; Cobourn 2010; Hoshyaripour et al. 2016). Chemical transport model (CTM) refers to a type of atmospheric composition models relevant for air quality forecasting, which has been successfully used for decades since middle 1990s in the USA, China, Europe (Jakobs et al. 2002), and Canada (Pudykiewicz et al. 1997). This model can forecast the concentrations of air pollutants under both typical and atypical scenarios but will be independent of a large quantity of measurement data (Stern et al. 2008, Sun et al. 2012). Meanwhile, it is physically based and provides scientific insights of pollutant formation processes; thus, it can address issues that cannot be handled by other forecasting models such as long-range transport of air pollutants, emissions, and changes in air quality under different meteorological and emission scenarios (Vautard et al. 2001; Mchenry et al. 2010). By contrast, the statistical approaches usually require a large quantity of historical measured data under a variety of atmospheric conditions, which can produce good estimation results in the short-term forecasting (Wang et al. 2015a, b, Konovalov et al. 2009). They are easy to set up and fast to compute to handle nonlinear and chaotic chemical system at a site (Pires and Martins 2011). The statistical models have several common drawbacks. First, the nature of statistical modeling does not enable better understanding of chemical and physical processes (Gao et al. 2018). Second, they cannot predict concentrations under extreme air pollution conditions that deviate significantly from the historical records (Zhang et al. 2012). Third, they are usually confined to a given site and cannot be generalized to other sites (Stockwell et al. 2002). Throughout the course of air quality prediction research, scores of statistical prediction models were used, such as linear or nonlinear regression model, time series model, Markov model, gray model, artificial neural networks (ANNs), and so on.

Among most statistical prediction methods, artificial neural networks (ANNs) are easier and faster to establish and have a more flexible nonlinear modeling capability, adding to its strong adaptability and massive parallel computing abilities, and it proved effective to predict environmental parameters, especially for wind speed, water temperature, and air quality (Feng et al. 2015; Taylan 2017; Wang et al. 2015a, b). For instance, adaptive neuro-fuzzy inference system (ANFIS), Elman neural network (ENN), long short-term memory networks (LSTMs), multi-layer perceptron (MLP), the backpropagation neural networks (BPNN), and the support vector machine (SVM) have been extensively used for modelling air quality (Noori et al. 2009; Krishan et al. 2019; Li et al. 2018; Taylan 2017; Voukantsis et al. 2011; Wang et al. 2014, 2016; Wen and Yuan 2020; Zhou et al. 2020). Considering the atmospheric environment is an extremely complex huge and nonlinear system, which is under dynamic changing. And the concentration of air pollutants is influenced by a number of complex factors such as human activity, atmospheric pressure, wind direction, wind speed, temperature, humidity, temperature inversion, and rainfall (Wang et al. 2019). And even some physicochemical processes have significant impacts; as for chemical process, ozone is formed through the reaction of nitrogen oxides and volatile organic chemicals when there is strong sunshine and high temperatures (Wei et al. 2021), and nitrogen-oxide and sulfur oxide can react with other pollutants in the air to generate particles such as nitrate and sulfate, thus turning gaseous pollutants into solid pollutants and increasing PM2.5 concentration in the air (Luo et al. 2020). As for physical process, amines combine with sulfuric acid to form highly stable aerosol particles under typical atmospheric concentrations (Almeida et al. 2013). Moreover, most of the interrelationship between the various factors is uncharted. This is consistent with the multilayer mapping and nonlinear modeling ability of back propagation neural network (BPNN), a fairly common used ANNs (Guo et al. 2011; Wang et al. 2006), also referred to as error back propagation network that minimizes an error backward while information is propagated forward (Wang et al. 2015a, b). However, one apparent shortcoming of BPNN is that it easily gets in the local minimum, due to its randomly allocated initial connection weights and thresholds, leading to an inaccurate result.

To overcome the aforementioned shortcomings, many researchers have proposed various of improved methods to improve precision. There are generally three sub-categories, firstly, single intelligent evolution algorithms, such as particle swarm optimization (PSO) (Huang et al. 2020; Jin et al. 2012; Qiu et al. 2020; Ren et al. 2014; Wen and Yuan 2020) and genetic algorithm (GA) (Feng et al. 2011; Wang et al. 2016; Zhang et al. 2020), were used to select the initial connection weights and thresholds of BPNN. Secondly, the application of two joint intelligent evolution algorithms, Hu et al. (2019) proposed a forecasting model based on the hybrid GA-PSO-BPNN algorithm to avoid the defect that the prediction result easily falling into local optimum. Thirdly, the combination of BPNN with another algorithm, such as Bayesian regularization (Tang et al. 2020), support vector machine (SVM) (Sun et al. 2020), and empirical mode decomposition (EMD) (Liu et al. 2016). Results demonstrate that the above proposed models are superior to traditional BPNN models on the basis of convergence rate or prediction precision.

Nevertheless, to our knowledge, the optimization of BPNNs using the ant colony algorithm which possess the property of fast global searching and strong robustness has not been attempted in air quality forecasting. Last but not least to mention is that only limited attention appears to have been given to input parameters that largely determines the level of prediction precision (Hu et al. 2019). Too many input variables can not only prolong the training time but irrelevant or noisy variables may have adverse effects on the training process, leading to an unacceptable convergence speed and poor generalization power. Consequently, A quantifiable random forest method (RF) is used to select input parameters carefully to achieve desired results.

In summary, this paper proposes a novel air quality forecasting method based on a hybrid RF and IACA-BPNN neural network method. The RF method was used to remove the irrelevant factors after the original data was processed. And then the chosen input variables were loaded into the improved IACA-BPNN model for air quality forecasting both in two sites. Furthermore, horizontal and longitudinal contrast were carried out to assess the forecasting performance of the proposed model based on a series of models which contains BPNN, IACA-BPNN, ACA-BPNN, PSO-BPNN, GA-BPNN, and RF-IACA-BPNN.

The specific structure of the paper is organized as follows. The second section are a general introduction of the three concerned algorithms, two improvements to the ant colony algorithm, and procedure of IACA-BPNN prediction model. The third section are data pre-processing and description and then screening results of input variables. The fourth section are results and discussion of predictive simulation experiments both in two sites (the traffic site and the park site) in Chengdu, as well as the results under four evaluation indicators between observed and predicted value. The last section is the conclusion of the whole paper.

Method

The hybrid method of RF-IACA-BPNN plus the other five kinds of artificial neural networks (ANNs) were adopted for air quality forecasting in this paper. The models mentioned above were programmed in MATLAB version R2018a (Mathworks Inc., Natick, USA). The random forest (RF) algorithm involved in this study was implemented by python 3.6 to calculate the importance evaluation values.

Related theoretical basis

BPNN

By mimicking the behavior characteristics of biological neural networks, ANN is an algorithmic mathematical model

in essence endowed with parallel and distributed information processing ability (Wang et al. 2016). As one of the most widely used ANNs, BPNN is an adaptive and nonlinear dynamic system composed of interconnected neurons, known for its forward propagation of information and error backward propagation, proposed by Rumelhart in 1986. Generally, the steepest descent method was adopted during the error propagation process to minimize the error of the network; meanwhile, the weights and thresholds adjust successively to form an expected model to properly reflect the mapping relationship between input and output values. BPNN usually consists of three layers (input layer, hidden layer, and output layer) (Park et al. 2017); theoretically, one single hidden layer BPNN is able to approximate any nonlinear function with satisfactory accuracy (Aslanargun et al. 2007). The brief structure of BPNN is shown in Fig. 1.

The random forest

As an ensemble supervised learning method from machine learning based on bagging algorithm, random forest (RF) composes of multiple mutually independent decision trees which are the combination of classification and regression tree together (Wei et al. 2018). In virtue of the characteristic of fast calculation speed and effectively applied to a wide range of problems that are nonlinear and involving complex high-order interaction effects, coupled with reliably identifying relevant predictors from a large set of candidate variables, random forests have been widely applied to various fields relate to bioinformatics (Strobl et al. 2007) and environmental sciences (Wen and Yuan 2020) for prediction and variable selection. Whereupon, the variable importance score (VIS) derived from RF is adopted in this research to measure relative importance of factors affecting $PM_{2.5}$ and O_3. Using $PM_{2.5}$ as an example, D represents the overall
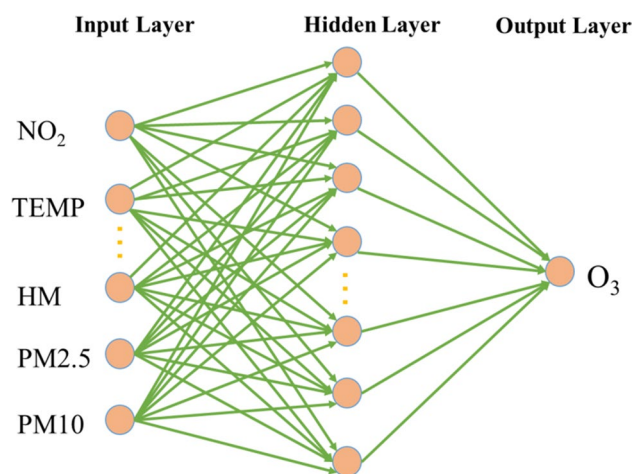

Fig. 1 Structure of one possible BPNN model used in this study

training sets, and vector X represents the sets of 10 factors affecting $PM_{2.5}$. $X = X_1, X_2, \ldots X_j, \ldots X_{10} | j = \{1, 2, \ldots, 10\}$.

The bootstrap method was adopted to randomly sample k training subsets from the overall training sets; thus, the k-th training subset represented as $D_K | k = 1, 2, \ldots K$.

$$VIS = VIS_1, VIS_2, \ldots, VIS_j, \ldots, VIS_{10} \mid j = j = \{1, 2, \ldots, 10\}$$

ACA

Ant colony algorithm (ACA) was originally theorized by Dorigo proposed in 1991 (Liu et al. 2019) (Dorigo and Stützle 2003), through observing the ants foraging behavior, which is actually a random searching algorithm in nature that successfully applying to many optimization problems (Liu et al. 2007). On account of the ability that ACA is able to successfully find the optimal path through the positive feedback and distributed cooperation, which has been widely combined with BPNN to improve its convergence speed and avoid trapping into the local minimum.

ACA improvements

Improved content

Generally, the better a solution is, the more likely it is to find the optimal solution around it (Yu and Zhou 2006). Hence, the basic idea of this algorithm rests with enhancing excellent solutions and weakening inferior solutions. With the increasing differences of pheromone between excellent and inferior solutions, it is more likely that the searching path of ants concentrates in the vicinity of the optimal solution.

(a) Improvement on pheromone updating rule

Ants are sorted according to the length of their paths. The more excellent ants pass by, the stronger pheromone they leave behind. The pheromone is adjusted according to Eq. (1):

$$\tau(r, s) = (1 - \rho)\tau(r, s) + \varepsilon_0 \frac{L_{worst}}{L_n} \tag{1}$$

where $\tau(r, s)$ represents the pheromone intensity between city r and s; $\varepsilon_0 \in (0, 1]$ is the pheromone penalty factor; L_{worst} is the length of the tour of the worst ant; L_n is the path length of the nth ant.

(b) Adaptive improvement on evaporation rate of pheromone

Pheromone volatile factor (ρ) measures the reduction extent during the evaporation process, which greatly affects global searching ability and convergence speed of the algorithm. To avoid trapping into local optimization, the pheromone volatile factor was improved by

$$\rho(nc) = \frac{\varphi}{1 + \lambda \frac{nc}{e^{nc_{max}}}} \tag{2}$$

where φ and λ both are constant coefficient; nc remains iteration times; nc_{max} represents maximum iterations. The analysis reveals that the convergence speed of the algorithm accelerating due to the large value of $\rho(nc)$ at the beginning. Hereafter, the accumulation of $\Delta\tau_{ij}(t)$ leads to the algorithm falling into local optimum. As can be seen from Eq. (8) above, the $\rho(nc)$ gradually decreases, making it easy for the algorithm to jump out of local optimization.

IACA-BPNN prediction model

Suppose the BPNN possesses m weights and thresholds in total, and each weight or threshold has n values to choose from, which are randomly generated within $[0,1]$ and thus forming set $A_i (1 \leq i \leq M)$.

Based on the above improvement of ACA, the process structure of the developed RF-IACA-BPNN prediction model is shown as Fig. 2. The specific steps of ACA optimizing the initial weight and threshold of BPNN are as follows:

Step 1: Set the initial pheromone value $\tau(A_i)(0)$, ant number (m), maximum iteration number (nc_{max}), and other parameters (including the parameters of BPNN).

Step 2: Elements from each set were selected by m ants according to Eq. (2), and all the elements selected by each ant constitutes a set of initial weights and thresholds of the neural networks.

Step 3: When m ants complete one cycle, m sets of initial weights and thresholds selected in Step 2 are used to train the BPNN model, and the output error of the network is calculated simultaneously according to Eq. (2). Record the set of weights and thresholds with the smallest error, and compare the error with the expected error ε. If it is less than the expected error ε. Then output algorithm results and enter Step 6. Otherwise, entering step 4.

Step 4: Pheromone of each element in the set $A_i (1 \leq i \leq M)$ is updated according to Eqs. (4) and (5).

Step 5: Repeat the steps (2) and (3) until all ants converge to the same path or reach the maximum iteration number nc_{max}.

Step 6: Take the optimal initial solutions selected by IACA as the weights and thresholds of BPNN. The neural network is further trained until the exit status is reached.
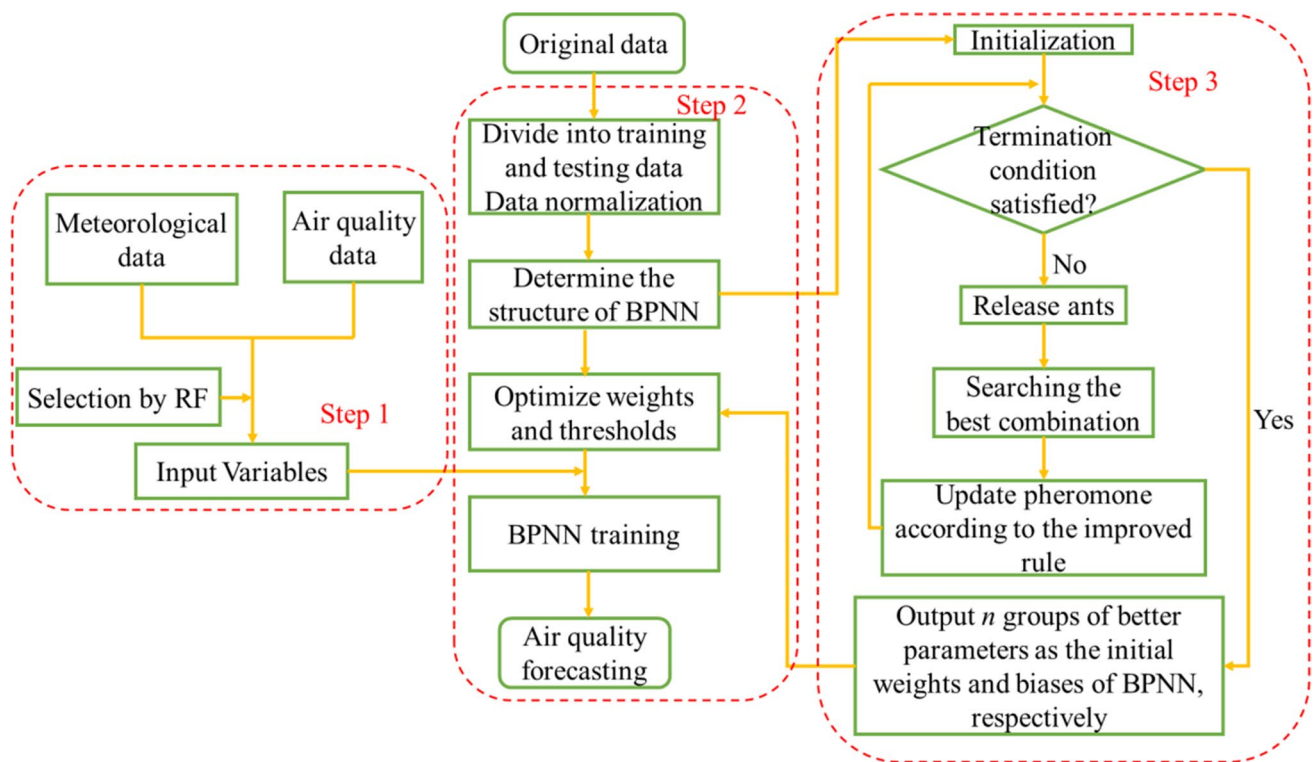
Fig. 2 Flow chart of the proposed RF-IACA-BPNN based forecasting model

Data

Data source

Chengdu, being the capital city of Sichuan province and one of the biggest cities in west China, has been regarded as commercial and cultural center for its strong economic growth and numerous cultural heritages. As an inland city located in the west of Sichuan Basin, with diversified landforms, the area covering 14335 km^2 is mainly composed of fertile plains that are surrounded by mountains. Independent of effective measures that have been taken by local government, air pollutants remain to be problematic in certain seasons especially for $PM_{2.5}$ and O_3. Generally, there are plenty of distinctive factors which impact the air pollution levels in Chengdu such as multitudinous populations (20 million in 2020), a large number of industrial enterprises and ever-increasing motor vehicles within the city, unfavorable local meteorological conditions due to distinct geographical location and topographic condition such as a high frequency of quiet wind throughout the year, and noteworthy temperature inversion in autumn and winter which result in negative atmospheric dispersion and transport mechanisms (Alimissis et al. 2018). Specifically, as the site of the FISU World University Games in 2022, it is indispensable and beneficial to accurately forecast of $PM_{2.5}$ and O_3 concentrations

such that regional air quality can be managed and controlled appropriately.

Fifteen citywide air quality stations were established by Ministry of Ecological Environment of China to monitor air quality trends, coupled with twelve meteorological monitoring stations (Fig. 3). In this study, two different air quality monitoring stations were selected, station A1 (Dashi Road West, a downtown station located in areas of heavy traffic) and station A2 (Long-Quan, a suburban station near the forest park), to validate the stability of the prediction model. The A1 station is located in the inner part of the Chengdu city, the density of population around this station is substantially higher as well. While the A2 station is far away from the downtown area, which is located in the Chengdu Forest Park, involving a country lane and higher percentage of open and green spaces. The concentration of air pollutants at A1 station are obviously higher than A2 station, except for O_3, which is 1.5 times less than A2 station (Table 1). The hourly data of six air quality factors ($PM_{2.5}$, PM_{10}, NO_2, SO_2, O_3, CO) are acquired automatically by each monitoring station to form a dataset ranging from 8 January 2019 to 8 November 2021, accompanied by five meteorological parameters including wind speed (WS), wind direction (WD), relative humidity (RH), temperature (TEMP), and atmospheric pressure (ATM) for the same period from the China Weather Website Platform, which is maintained by

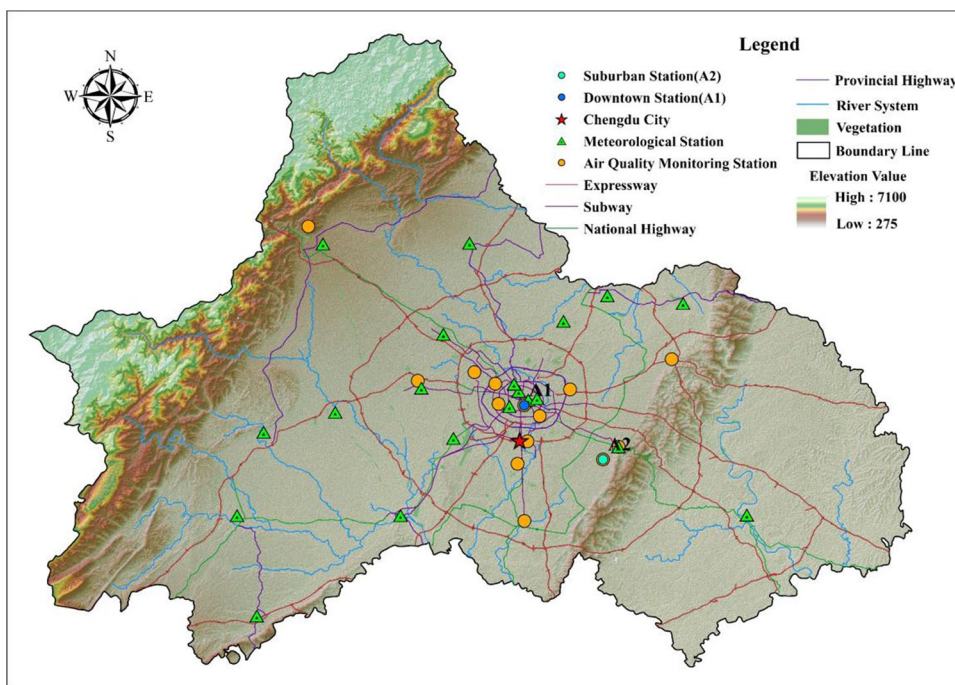**Fig. 3** Location of the air quality and meteorological measurement sites in Chengdu

Table 1 Basic statistics for the air quality and meteorological parameters of the two stations

A1 Jan, 2019–Nov, 2021 (data rows: 24,200)					A2 Jan, 2019–Nov, 2021 (data rows: 24,200)				
Pollutant	Mean	Std	Min	Max	Pollutant	Mean	Std	Min	Max
$PM_{2.5}$ ($\mu g/m^3$)	38.3	29.9	0	271	$PM_{2.5}$ ($\mu g/m^3$)	28.1	19.5	0	148
PM_{10} ($\mu g/m^3$)	66.0	44.0	0	366	PM_{10} ($\mu g/m^3$)	44.0	28.7	0	372
SO_2 ($\mu g/m^3$)	8.3	4.6	1	471	SO_2 ($\mu g/m^3$)	7.0	2.2	1	34
NO_2 ($\mu g/m^3$)	48.1	25.7	2	183	NO_2 ($\mu g/m^3$)	17.4	10.9	0	115
CO (mg/m^3)	0.7	0.3	0	23.7	CO (mg/m^3)	0.4	3.6	0	2.8
O_3 ($\mu g/m^3$)	48.4	48.0	0	346	O_3 ($\mu g/m^3$)	72.0	44.7	0	297
WS (m/s)	1.3	0.7	0	7	WS (m/s)	1.3	0.8	0	7
RH (%)	72.0	16.5	13	100	RH (%)	73.3	16.8	12	100
TEMP (°C)	14.5	7.6	−3	38	TEMP (°C)	16.5	7.6	−3	37
ATM (hPa)	954.5	7.8	935	977	ATM (hPa)	949.1	11.7	926	984

China Meteorological Bureau. The air quality data were matched corresponding to the closest meteorological parameters (Feng et al. 2015). The forecasting models under investigation will forecast $PM_{2.5}$ and O_3 concentrations over the next 24 h.

Data preprocessing

Due to the power failures or equipment breakdown, some data were missing. These missing data need to be properly processed to develop a well model. Rows with consecutive hourly gaps of more than 4 h of missing data were discarded; other missing data were supplemented. Given the datasets developed are sequences of random variables indexed by time, it is feasible to employ nearest neighbor interpolation method to re-fill the datasets, which is a method that interpolates the gray value with the nearest point (Hu et al. 2019). After that, a total of 24200 and 24200 hourly datasets left for A1 and A2 stations, respectively. The air quality and meteorological data involved in each station were randomly divided into training sets, validation sets, and test sets according to a ratio of approximately 6:2:2. Subsequently, random forest method was used to calculate the relative importance between input variables (meteorological and air quality factors) and output variables ($PM_{2.5}$, O_3).

To eliminate the adverse effects caused by the singular data and improve the convergence speed of the model, the datasets are linearly scaled to within the range [0,1] by

adopting linear transformation through the normal standardization formula (3):

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \tag{3}$$

where X_{norm} is the normal standardization data, X is the original value, and X_{max} and X_{min} are the maximum and minimum values of the series, respectively.

Data description

Based on the datasets aforementioned of the two sites, the basic statistics about the meteorological and air quality data were studied (Table 1).

The suburban station demonstrated significantly lower level of concentrations of pollutants, except for O_3, for which concentration was 1.5 times higher than the downtown station. A couple of factors help to explain, one is the consumption of O_3 in the oxidation process with nitrogen oxides emitted by vehicles, ozone precursors are carried by the wind to the suburbs is another. Mean values of $PM_{2.5}$ and PM_{10} concentrations at the downtown station were 1.3 and 1.5 times higher than the suburban station, respectively, on account of the mass emission sources, while the standard deviation was similar between the two stations. Besides, both of the 4 indicators for meteorological parameters of the two sites were nearly the same.

Model performance metrics

Till now, many researchers have employed various statistical indices to verify the predictive performance of the models, according to the literature (Alimissis et al. 2018; Li et al. 2018; Yildirim and Bayramoglu 2006). In this study, four statistical indices were adopted: the mean absolute error (*MAE*), the mean absolute percentage error, the coefficient of determination (*R*-square), and the root mean-square error (*RMSE*) defined as

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_o - y_p}{y_o} \right| * 100\% \tag{4}$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} \left| y_o - y_p \right| \tag{5}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(y_o - y_p \right)^2} \tag{6}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_o - y_p \right)^2}{\sum_{i=1}^{n} \left(y_o - \bar{y}_p \right)^2} \tag{7}$$

where y_o and y_p represent the observed and predicted value, respectively. $\bar{y}_p$ is the average of observed value, and n represents the number of observed data. The superiority of the model performance was measured by smaller values of *MAPE*, *MAE*, *RMSE*, and values of R^2 closest to one, all of the indices aforementioned manifests a model with high forecasting accuracy be related to the observed values in the testing data.

Screening of prediction indicators

The visual map of each input variable's importance is shown in Fig. 4. It is important to note that this test was repeated for several times in order to assure that their selection was not biased. With regard to the results of RF, the first eight indicators were sorted out as the input variables of each model in this study, as the importance of each single indicator was greater than 0.005 and the sum of the eight indicators was accounting for nearly 99%.

The reason why PM_{10} had a strong positive correlation with $PM_{2.5}$ is that $PM_{2.5}$ is included in PM_{10} and they can transform to each other under specified conditions. Higher RH is favorable for particulate matter to adhere to water vapor, which increases the mass concentration of particles (Zhang et al. 2017). Meanwhile, increases in RH favor nitric acid partitioning to the aerosol phase and therefore can lead to nitrate concentration increases (Dawson et al. 2007). Increasing TEMP can lead to elevated sulfate concentrations due to the increased rate of SO_2 oxidation (Jacob and Winner 2009). TEMP also has a significant indirect effect on secondary organic aerosol (SOA) concentrations (Megaritis et al. 2014). CO concentrations had significant positive effects on $PM_{2.5}$, and this positive correlation was likely due to industrial emissions and exhaust fumes, which produce large amounts of $PM_{2.5}$ as well as CO. On the contrary, given that $PM_{2.5}$ can reduce the radiation flux during photochemical reactions, there was a significant negative nonlinear relationship between O_3 and $PM_{2.5}$ (Cheng et al. 2021).

O_3 is a secondary product of the oxidation of hydrocarbons (CH_4 and NMHCs) and CO via reactions catalyzed by HO_X and NO_X radicals (Jacob 2000), which helps to explain the high correlation of O_3 with NO_X and also weak correlation with CO. While the correlation of O_3 with NO_X at A2 station is relatively low compared with A1 station, it can be contributed to low NO_X concentration. PM can affect the atmospheric photochemistry by scattering the solar and terrestrial radiation, indirectly altering the air temperature and subsequently affecting the formation process of O_3 (Sharma et al. 2016). The relationship between O_3 and temperature is indirect, which is realized through higher downward solar radiation, high temperature
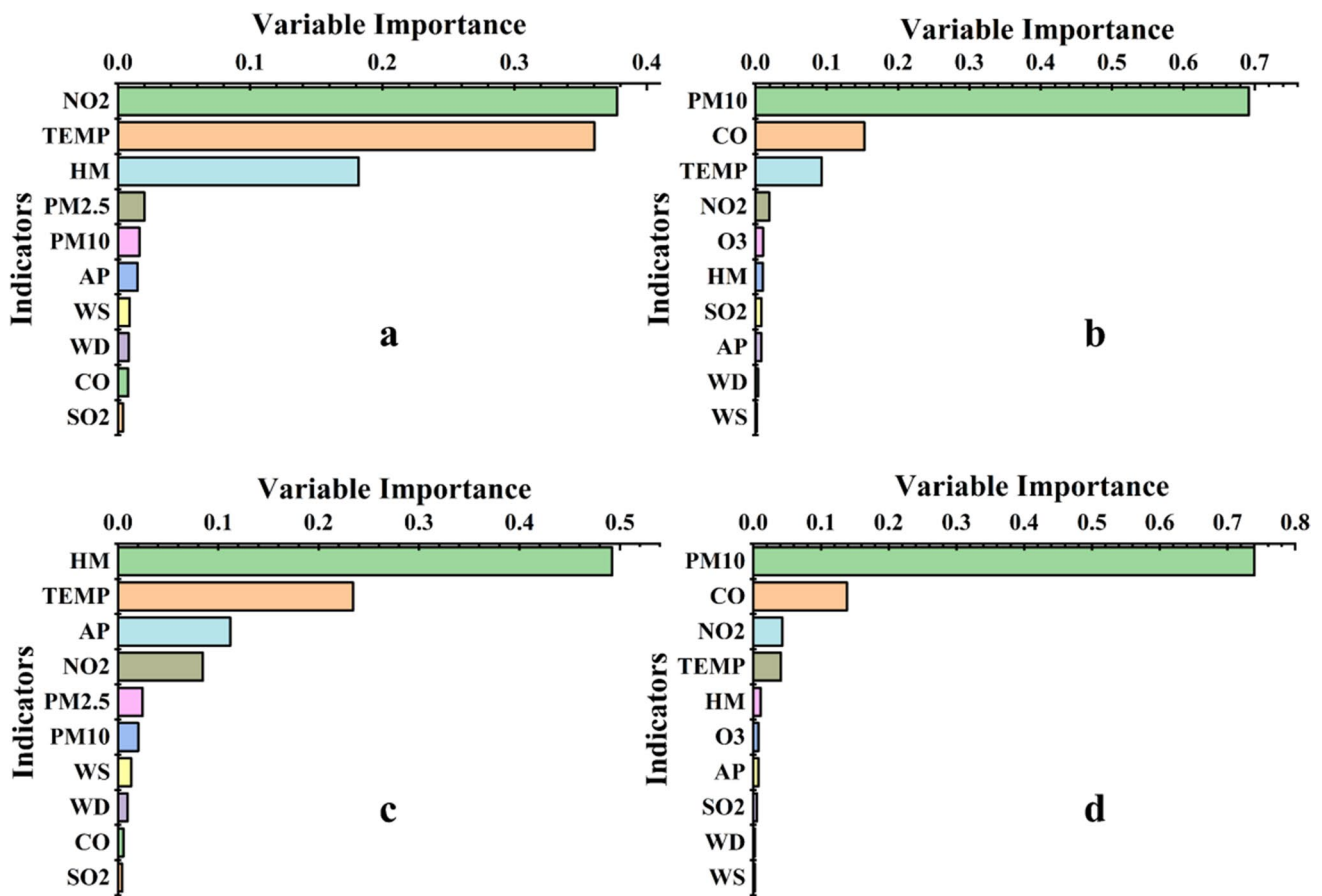
Fig. 4 The visual graph of input variables' importance. **a** O_3 at A1 station. **b** $PM_{2.5}$ at A1 station. **c** O_3 at A2 station. **d** $PM_{2.5}$ at A2 station

promotes the propagation rate of the radical chains and the formation of O_3 (Tu et al. 2007; Martins et al. 2012). Furthermore, RH is also a vital factor for O_3 formation; an appropriate RH can promote the formation of O_3 (Xu and Zhu 1994).

Generally, O_3 and $PM_{2.5}$ showed little correlation with the WD, WS at both two sites. One reason may explain this phenomenon is that the variation range of the WD and WS during the study period was too limited (the annual average wind speed of Chengdu is less than 1.2m/s) to show distinctive effects on O3 and $PM_{2.5}$ (Ahmad et al. 2019). AP affects the diffusion of O_3 and $PM_{2.5}$; if surface pressure field was mainly controlled by the huge clod high pressure, downdraft appears in the center, which inhibits the upward diffusion of O_3 and $PM_{2.5}$.

Results and discussion

Six models based on different artificial intelligence algorithms and different input variables were adopted to predict $PM_{2.5}$ and O_3 in two stations, respectively. There are four

Table 2 Performances of different models for O_3 during the testing phases at A1 station

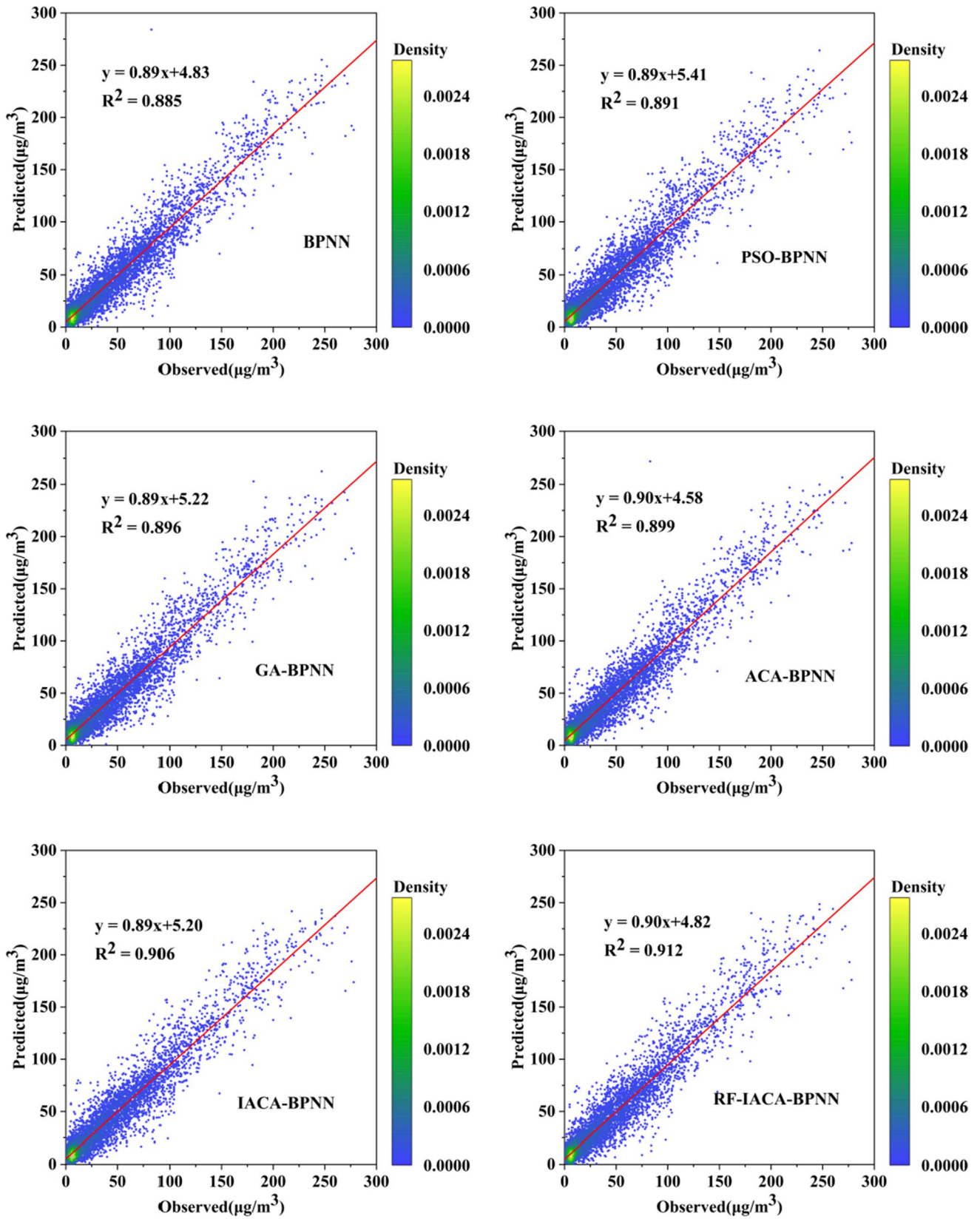| Air pollutants(A1 station)	Model	Testing			
		MAPE	*MAE*	*RMSE*	R^2
O_3	BPNN	0.714	10.499	16.380	0.885
	PSO-BPNN	0.423	10.928	15.924	0.891
	GA-BPNN	0.419	10.775	15.607	0.896
	ACA-BPNN	0.647	10.568	15.159	0.899
	IACA-BPNN	0.498	10.264	15.553	0.906
	RF-IACA-BPNN	0.453	10.448	15.336	0.912

Fig. 5 Comparison between predicted and observed O_3 concentration using different models at A1 station during testing periods

Table 3 Performances of different models for PM$_{2.5}$ during the testing phases at A1 station

Air pollutants(A1 station)	Model	Testing			
		MAPE	MAE	RMSE	R^2
PM$_{2.5}$	BPNN	0.185	4.765	8.522	0.919
	PSO-BPNN	0.208	5.092	7.412	0.935
	GA-BPNN	0.165	4.857	7.719	0.932
	ACA-BPNN	0.182	4.694	7.254	0.934
	IACA-BPNN	0.182	4.708	7.308	0.941
	RF-IACA-BPNN	0.184	4.667	7.242	0.942

experiments in total; each experiment contains six different models. The input variables of each experiment were solely selected by RF method and applied to six different models. It should be mentioned that each experiment in this study was independently carried out 15 times and the performance metrics (MAPE, MAE, RMSE, and R^2) were calculated on account of predicted and observed values subsequently.

Case study 1

Accuracy of various models for O$_3$ forecasts at A1 station

The input variables screened out by RF were NO$_2$ (0.378), TEMP (0.360), HM (0.182), PM$_{2.5}$ (0.020), PM$_{10}$ (0.016), AP (0.014), WS (0.009), and WD (0.008), respectively, for predicting hourly O$_3$ concentration at A1 station. The results showed that NO$_2$ made the greatest contribution, as one important predecessor of O$_3$, which is an important material to produce O$_3$ along with VOCs in the presence of heat and sunlight (Abdul-Wahab and Al-Alawi 2002, Heo and Heo and Kim 2004). Furthermore, temperature and humidity were also two main meteorological factors that affect atmospheric photochemical reaction of O$_3$.

Table 2 shows the performances of different models; the proposed RF-IACA-BPNN model provides significantly better forecasts when compared with the other 5 employed models for O$_3$ prediction in case study 1. The order of R^2 among different models from highest to lowest were RF-IACA-BPNN, IACA-BPNN, ACA-BPNN, GA-BPNN, PSO-BPNN, BPNN. GA-BPNN, and PSO-BPNN models yielded roughly similar results in terms of the four metrics. Despite RMSE, the rest of the three statistical criteria of IACA-BPNN outperform those of ACA-BPNN, which demonstrates that improvements on pheromone updating rule and evaporation rate of pheromone could enhance the fitting ability of ACA-BPNN. Moreover, the prediction results of 6 models were presented in Fig. 5. It can be seen that the scatters of the RF-IACA-BPNN model are closest and most concentrated on the regression line.

Performance of various models for PM$_{2.5}$ forecasts at A1 station

As can be seen from Fig. 4b, the most influential factors of PM$_{2.5}$ were PM$_{10}$ (0.692), CO (0.152), TEMP (0.093), NO$_2$ (0.020), O$_3$ (0.011), HM (0.010), SO$_2$ (0.008), and AP (0.008), respectively. PM$_{10}$ contains PM$_{2.5}$, the difference between the two is the aerodynamic diameter, which exactly explains the highly correlation of PM$_{10}$ to PM$_{2.5}$ (Biancofiore et al. 2017). As an important component of automobile exhaust, CO also has an impact on the concentration of PM$_{2.5}$, meanwhile, temperature can also influence PM$_{2.5}$ concentration by affecting boundary layer height.

The statistical metrics are listed in Table 3 and it can be seen more clearly in Fig. 6, the optimal indicator is marked in bold. The 3 metrics of RF-IACA-BPNN outperform the other 5 models. In the case of MAE, the RF-IACA-BPNN model is decreased by 2.06% compared with BPNN and by 0.87% compared with IACA-BPNN. It can be noted that the performance of this model with two optimizations for ACA along with screening process was significantly improved.

Case study 2

Performance of various models for O$_3$ forecasts at A2 station

Factors highly associated with O$_3$ included HM (0.492), TEMP (0.234), AP (0.112), NO$_2$ (0.084), PM$_{2.5}$ (0.025), PM$_{10}$ (0.020), WS (0.013), and WD (0.009), which is nearly the same with O$_3$ at A1 station but in a different order (Fig. 4c). Although the smallest MAE (11.534) appeared in the ACA-BPNN, the RF-IACA-BPNN model performs 4.82%, 22.26%, and 2.66% superior in determining RMSE values, MAPE values, and R^2 values, respectively, when compared with BPNN model (Table 4).

The prediction results of the RF-IACA-BPNN model are the closest to the actual values along with that even as peaks or valleys where the concentration fluctuates greatly, denoting that the RF-IACA-BPNN model performs the best (Fig. 7).
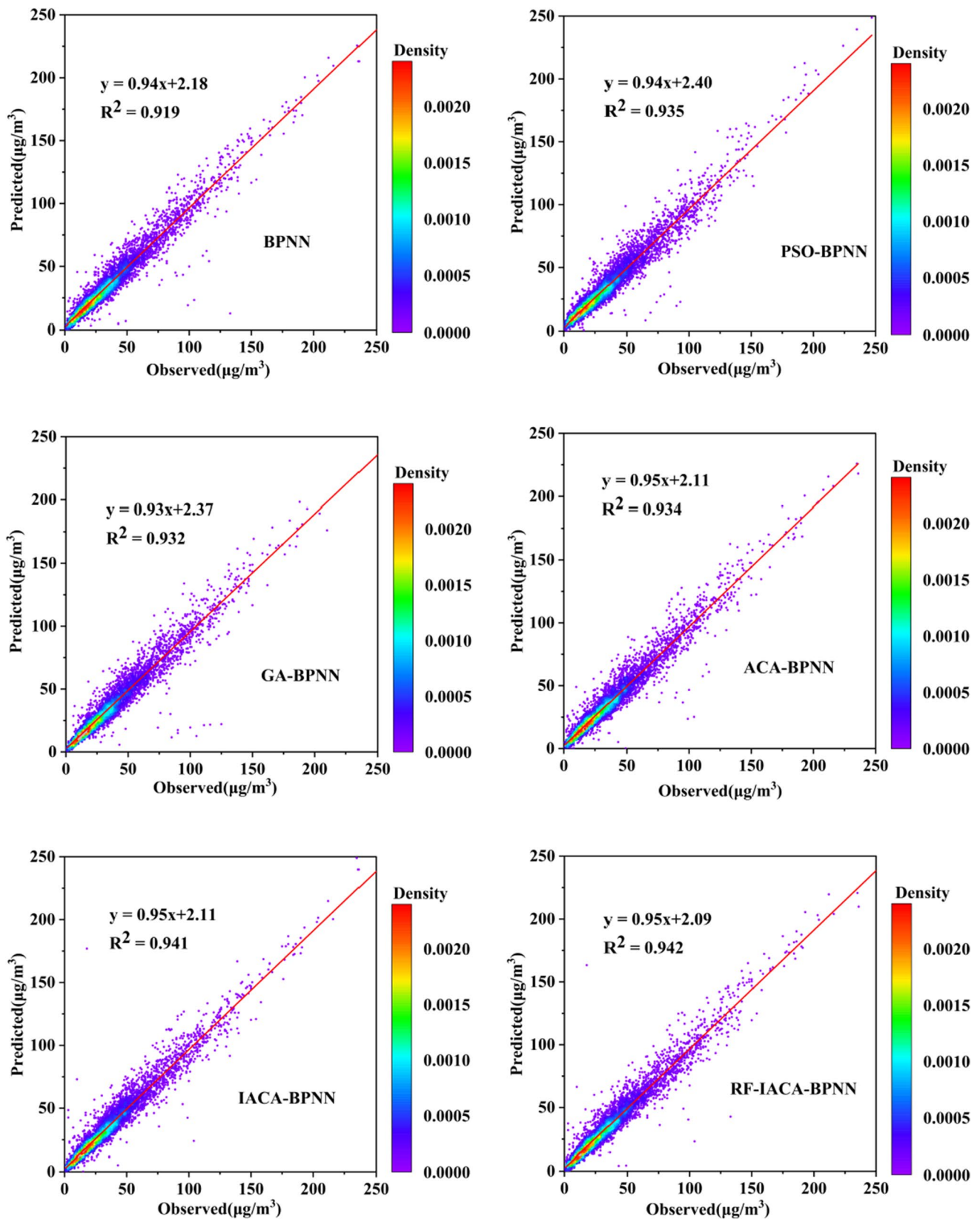
Fig. 6 Comparison between predicted and observed PM$_{2.5}$ concentration using different models at A1 station during testing periods

Table 4 Performances of different models for O_3 during the testing phases at A2 station

Air pollutants (A2 station)	Model	Testing			
		MAPE	*MAE*	*RMSE*	R^2
O_3	BPNN	0.319	11.683	15.179	0.864
	PSO-BPNN	0.347	12.286	15.915	0.875
	GA-BPNN	0.252	12.151	15.997	0.873
	ACA-BPNN	0.491	11.534	14.864	0.871
	IACA-BPNN	0.335	12.055	15.764	0.875
	RF-IACA-BPNN	0.248	11.826	14.448	0.887

In allusion to the scatter correlation figure, the regression line indicates the observed value is exactly the same as the predicted value; therefore, the closer the scatters are to this line, the better the performance (Sun and Li 2020a, b). It can be observed that the scatters of the RF-IACA-BPNN model are hithermost and most centralized on the regression line.

Performance of various models for PM$_{2.5}$ forecasts at A2 station

In this suburban station, the factors most closely related to PM$_{2.5}$ were PM$_{10}$ (0.740), CO (0.138), NO$_2$ (0.043), TEMP (0.041), HM (0.011), O$_3$ (0.008), AP (0.008), and SO$_2$ (0.006), respectively. The components and order of its relative importance are very similar to that A1 station.

The observed and predicted PM$_{2.5}$ values during testing phase at A2 station were presented in Fig. 8, coupled with the 4 performance metrics among 6 models (Table 5). It was found that models with improvement had better performance and were more easily and faster to converge to a better solution than a plain BPNN model; the screening process of input variables is also helpful in improving prediction performance. For example, the RF-IACA-BPNN model exhibits a 36.99% and 29.80% decrease in *MAPE* compared with the basis models IACA-BPNN and BPNN, respectively. Although the best metrics of *MAE* (3.280) did not appear in RF-IACA-BPNN model, however, the values of R^2 exhibit 1.39%, 1.61%, 1.61%, 4.06%, and 5.68% increase compared with the models IACA-BPNN, ACA-BPNN, GA-BPNN, and PSO-BPNN, BPNN, respectively.

Comparison of the same pollutant between two stations

To verify the performance of the same model against data based on different kinds of monitoring sites, here, the performance of PM$_{2.5}$ and O$_3$ are compared at the two stations, respectively.

With regard to the forecasting results of O$_3$, the *MAPE* of each model at A2 station (0.319, 0.347, 0.252, 0.491, 0.335, and 0.248) are smaller than that of corresponding 6 models that made it up at A1 station; it is the same with R^2. The R^2 of RF-IACA-BPNN at A1 station (0.912) is obviously higher than that of A2 station (0.887). As for PM$_{2.5}$, the *MAE* and *RMSE* values at A1 station significantly higher ($p < 0.05$) than those corresponding models at A2 station. While R^2 values at A2 station are generally higher than that of A1, furthermore, the R^2 of RF-IACA-BPNN increased by 0.74% compared with A1 station.

Overall, for PM$_{2.5}$, the prediction results of the models at A2 station are more acceptable than those at A1 station, while the prediction results of the models showed a better performance at A1 station for O$_3$. Part of the reason lies in variability of data. The mean value and standard deviation of PM$_{2.5}$ at A1 station are much higher than that of A2 station (Table 1). And the situation for O$_3$ is just the opposite.

Stability of the proposed model

In order to comprehensively test the robustness of the new air quality forecasting model, 10 additional sites were selected to verify the robustness of the model. Five of them are downtown sites (*Di, i = 1, 2, 3, 4, 5*) which close to the city center, and the others are suburban sites (*Si, i = 1, 2, 3, 4, 5*) far from the highways and population centers. The stability test is implemented in this part according to Eq. (8). It is known that the stability of the forecasting performance can be indicated by the variance (Var) of the forecasting error (Sun et al. 2020, Hao et al. 2019, Wang et al. 2017a, b). Generally, a smaller variance represents a more stable model. The stability test results are shown in Table 6. As for PM$_{2.5}$, the Var values of the proposed model at the 5 downtown sites are obviously higher than the Var values at the 5 suburban sites. The average Var value of downtown sites is 0.142, which is 73.17% more than that of suburban sites. As for O$_3$, conversely, the average Var value of downtown sites is 1.320, which is 4.90% less than that of suburban sites. In conclusion, the model can be more stable when it was used to forecast hourly PM$_{2.5}$ concentration at suburban sites. It's just the opposite for O$_3$, which had a higher stability at downtown sites.

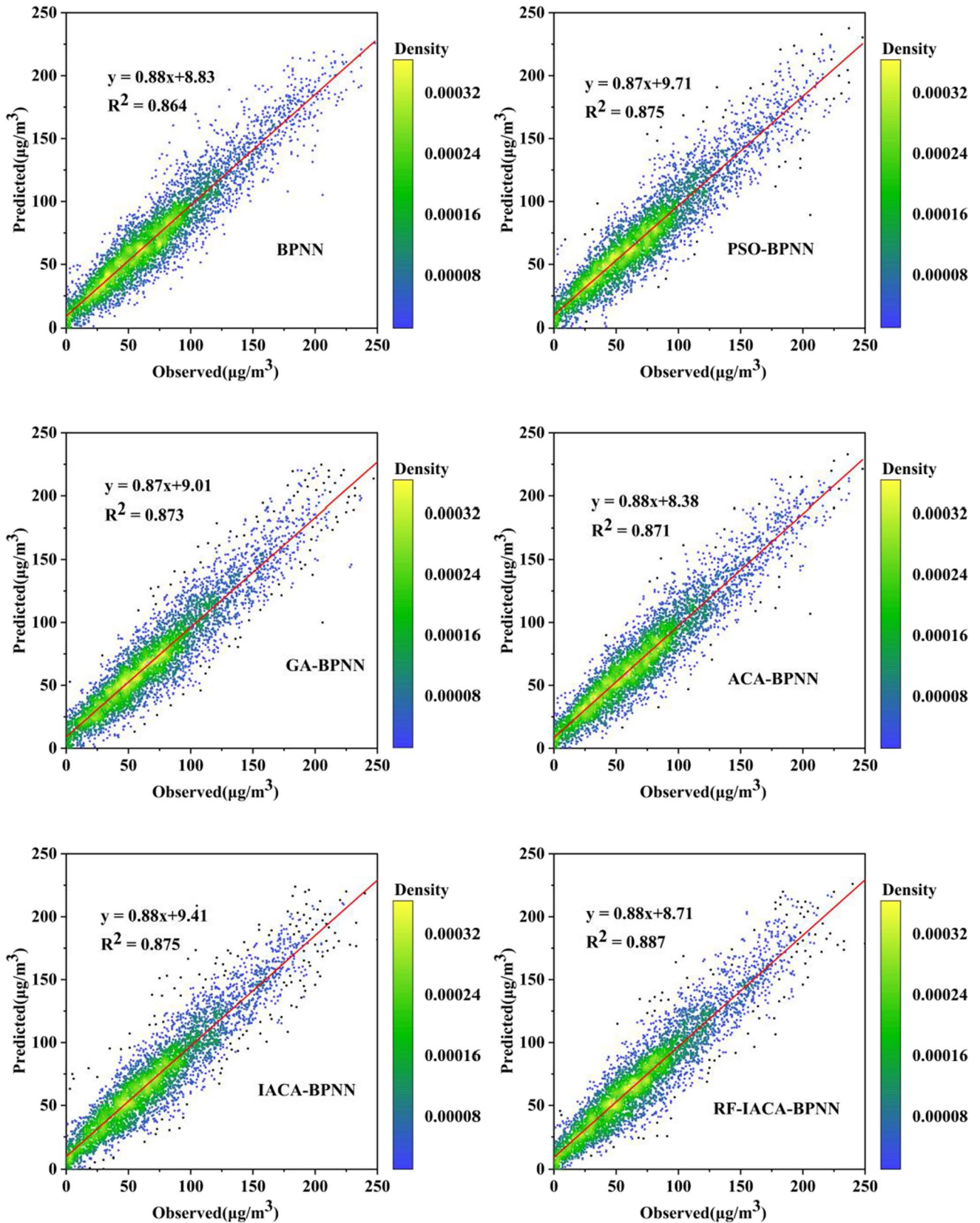$$S_{Var} = \text{var}\left(\left| \frac{y_o - y_p}{y_o} \right| \right) \qquad (8)$$

Fig. 7 Comparison between predicted and observed O_3 concentration using different models at A2 station during testing periods
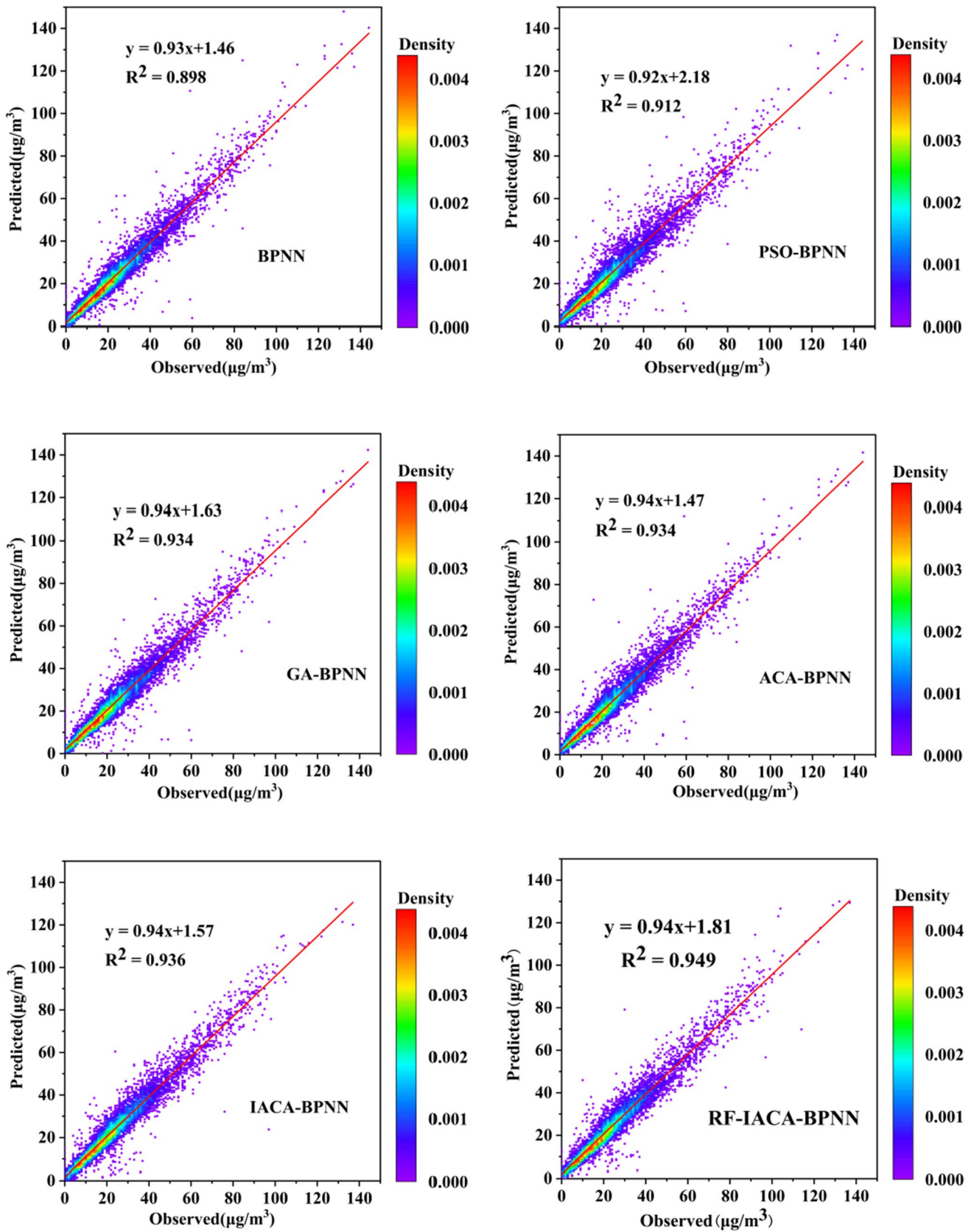
Fig. 8 Comparison between predicted and observed PM$_{2.5}$ concentration using different models at A2 station during testing periods

Table 5 Performances of different models for $PM_{2.5}$ during the testing phases at A2 station

Air pollutants (A2 station)	Model	Testing			
		MAPE	*MAE*	*RMSE*	R^2
$PM_{2.5}$	BPNN	0.245	3.351	4.996	0.898
	PSO-BPNN	0.196	3.884	5.798	0.912
	GA-BPNN	0.214	3.344	5.030	0.934
	ACA-BPNN	0.189	3.351	5.030	0.934
	IACA-BPNN	0.273	3.393	4.954	0.936
	RF-IACA-BPNN	0.172	3.409	4.941	0.949

Table 6 Stability test results at other 10 sites

Sites	Var (PM2.5)	Ave	Var(O3)	Ave
Stability test results of RF-IACA-BPNN model at 5 downtown sites.				
D1	0.167	0.142	1.062	1.320
D2	0.126		1.275	
D3	0.147		1.465	
D4	0.116		1.515	
D5	0.154		1.284	
Stability test results of RF-IACA-BPNN model at 5 suburban sites.				
S1	0.076	0.082	1.354	1.388
S2	0.082		1.110	
S3	0.078		1.334	
S4	0.084		1.552	
S5	0.092		1.588	

Conclusions

Many studies have put their efforts to unilaterally improve the accuracy of air quality forecasting through various optimization algorithms. Nevertheless, few did it from the perspective of addressing original datasets, i.e., screening out highly related input variables, which have a significant impact on the forecasting precision and the training process of ANNs. In this study, a new hybrid method based on improved ACA and BPNN model was proposed, with the combination of screening ability of RF method to forecast hourly concentration of $PM_{2.5}$ and O_3. Two datasets based on two different types of monitoring stations were used to compare the forecasting performance of RF-IACA-BPNN with those of other models and to verify the feasibility and effectiveness of ACA optimization and screening process. The following conclusions are drawn.

(1) With the determination of relative importance of input variables using RF method, results showed that the factors that most affected O_3 were similar at both downtown and suburban stations, so is the case with $PM_{2.5}$.

As for O_3, the five factors that have the greatest influence on ozone concentration were NO_2, TEMP, HM, $PM_{2.5}$, PM_{10}, respectively, while the top five highly corelated factors to $PM_{2.5}$ were PM_{10}, CO, TEMP, NO_2, O_3, respectively.

(2) In the case study 1, *MAE* and R^2 of the RF-IACA-BPNN model for O_3 were 10.448 and 0.912. As for $PM_{2.5}$ at downtown station, 3 statistical criteria of RF-IACA-BPNN outperform other 5 models. And in the case study 2, *MAPE*, *RMSE*, and R^2 of the proposed method for O_3 were 0.248, 15.448, and 0.887, in allusion to $PM_{2.5}$ at suburban station; the value of *MAPE* and R^2 were 0.172 and 0.949, respectively. It is concluded that the proposed model is the ideal one compared with the other plain models.

(3) On the whole, for $PM_{2.5}$, the prediction results of the models at A2 station are more acceptable than those at A1 station, while the prediction results of the models showed a better performance at A1 station for O_3, especially for RF-IACA-BPNN model, on account of the low variability at suburban station.

(4) The model can be more stable when it was used to forecast hourly $PM_{2.5}$ concentration at suburban sites. It's just the opposite for O_3, which had a higher stability at downtown sites.

Declarations

Ethics approval and consent to participate Not applicable

Consent for publication Not applicable

Competing interests The authors declare no competing interests.

References

Abdul-Wahab SA, Al-Alawi SM (2002) Assessment and prediction of tropospheric ozone concentration levels using artificial neural networks. Environ Modell Software 17(3):219–228. https://doi.org/10.1016/S1364-8152(01)00077-9

Ahmad M, Cheng S, Yu Q et al (2019) Chemical and source characterization of PM 2.5 in summertime in severely polluted Lahore,

Pakistan. Atmospheric Res 234:104715. https://doi.org/10.1016/j.atmosres.2019.104715

Alimissis A, Philippopoulos K, Tzanis CG et al (2018) Spatial estimation of urban air pollution with the use of artificial neural network models. Atmospheric Environ 191:205–213. https://doi.org/10.1016/j.atmosenv.2018.07.058

Almeida J, Schobesberger S, Kürten A et al (2013) Molecular understanding of sulphuric acid–amine particle nucleation in the atmosphere. Nature 502:359–363. https://doi.org/10.1038/nature12663

Arden Pope C III (2002) Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution. Jama 287(9):1132–1141. https://doi.org/10.1001/jama.287.9.1132

Aslanargun A, Mammadov M, Yazici B et al (2007) Comparison of ARIMA, neural networks and hybrid models in time series: tourist arrival forecasting. J Stat Comput Simul 77(1/2):29–53. https://doi.org/10.1080/10629360600564874

Biancofiore F, Busilacchio M, Verdecchia M et al (2017) Recursive neural network model for analysis and forecast of PM10 and PM2.5. Atmospheric. Pollution Res 8(4):652–659. https://doi.org/10.1016/j.apr.2016.12.014

Cheng B, Ma Y, Feng F et al (2021) Influence of weather and air pollution on concentration change of $PM_{2.5}$ using a generalized additive model and gradient boosting machine. Atmospheric Environ 255(D12):118437. https://doi.org/10.1016/j.atmosenv.2021.118437

Corani G, Scanagatta M (2016) Air pollution prediction via multi-label classification. Environ Modell Software 80:259–264. https://doi.org/10.1016/j.envsoft.2016.02.030

Cobourn WG (2010) An enhanced $PM_{2.5}$ air quality forecast model based on nonlinear regression and back-trajectory concentrations. Atmospheric Environ 44(25):3015–3023. https://doi.org/10.1016/j.atmosenv.2010.05.009

Dawson JP, Adams PJ, Pandis SN (2007) Sensitivity of PM2.5 to climate in the Eastern US: a modeling case study. Atmospheric Chem Phys 7:4295–4309. https://doi.org/10.5194/acp-7-4295-2007

Dorigo M, Stützle T (2003) The ant colony optimization metaheuristic: algorithms, applications, and advances. New Ideas Optimiz 57:250–285. https://doi.org/10.1007/0-306-48056-5_9

Fan YV, Perry S, Klemeš JJ, Lee CT (2018) A review on air emissions assessment: transportation. J Cleaner Product 194:673–684. https://doi.org/10.1016/j.jclepro.2018.05.151

Feng X, Li Q, Zhu Y, Hou J, Jin L, Wang J (2015) Artificial neural networks forecasting of PM2.5 pollution using air mass trajectory based geographic model and wavelet transformation. Atmospheric Environ 107:118–128. https://doi.org/10.1016/j.atmosenv.2015.02.030

Feng Y, Zhang W, Sun D, Zhang L (2011) Ozone concentration forecast method based on genetic algorithm optimized back propagation neural networks and support vector machine data classification. Atmospheric Environ 45(11):1979–1985. https://doi.org/10.1016/j.atmosenv.2011.01.022

Gao M, Yin L, Ning J (2018) Artificial neural network model for ozone concentration estimation and Monte Carlo analysis. Atmospheric Environ 184:129–139. https://doi.org/10.1016/j.atmosenv.2018.03.027

Guo ZH, Jie W, Lu HY et al (2011) A case study on a hybrid wind speed forecasting method using BP neural network. Knowledge-Based Syst 24(7):1048–1056. https://doi.org/10.1016/j.scitotenv.2003.11.009

Hao Y, Tian C, Wu C (2019) Modelling of carbon price in two real carbon trading markets. J Cleaner Product 244:118556. https://doi.org/10.1016/j.jclepro.2019.118556

Heo JS, Kim DS (2004) A new method of ozone forecasting using fuzzy expert and neural network systems. Sci Total Environ 325(1-3):221–237. https://doi.org/10.1016/j.scitotenv.2003.11.009

Hoshyaripour G, Brasseur G, Andrade MF et al (2016) Prediction of ground-level ozone concentration in São Paulo, Brazil: Deterministic versus statistic models. Atmospheric Environ 145:365–375. https://doi.org/10.1016/j.atmosenv.2016.09.061

Hu Y, Li J, Hong M et al (2019) Short term electric load forecasting model and its verification for process industrial enterprises based on hybrid GA-PSO-BPNN algorithm—A case study of papermaking process. Energy 170:1215–1227. https://doi.org/10.1016/j.energy.2018.12.208

Huang Y, Xiang Y, Zhao R, Cheng Z (2020) Air quality prediction using improved PSO-BP neural network. IEEE Access 8:99346–99353. https://doi.org/10.1109/ACCESS.2020.2998145

Jacob DJ (2000) Heterogeneous chemistry and tropospheric ozone. Atmospheric. Environ 34(12):2131–2159. https://doi.org/10.1016/S1352-2310(99)00462-8

Jacob DJ, Winner DA (2009) Effect of climate change on air quality. Atmospheric Environ 43:51–63. https://doi.org/10.1016/j.atmosenv.2008.09.051

Jakobs HJ, Tilmes S, Heidegger A et al (2002) Short-term ozone forecasting with a network model system during summer 1999. J Atmospheric Chem 42:23–40. https://doi.org/10.1023/A:1015767207688

Jin C, Jin SW, Qin LN (2012) Attribute selection method based on a hybrid BPNN and PSO algorithms. Appl Soft Comput 12(8):2147–2155. https://doi.org/10.1016/j.asoc.2012.03.015

Konovalov IB, Beekmann M, Meleux F et al (2009) Combining deterministic and statistical approaches for PM10 forecasting in Europe. Atmospheric Environ 43(40):6425–6434. https://doi.org/10.1016/j.atmosenv.2009.06.039

Krishan M, Jha S, Das J et al (2019) Air quality modelling using long short-term memory (LSTM) over NCT-Delhi, India. Air Quality, Atmosphere & Health 12(8):899–908. https://doi.org/10.1007/s11869-019-00696-7

Lauret P, Heymes F, Aprin L, Johannet A (2016) Atmospheric dispersion modeling using artificial neural network based cellular automata. Environ Modell Software 85:56–69. https://doi.org/10.1007/s11869-019-00696-7

Li H, Wang J, Lu H, Guo Z (2018) Research and application of a combined model based on variable weight for short term wind speed forecasting. Renewable Energy 116:669–684. https://doi.org/10.1016/j.renene.2017.09.089

Liu F, Gong H, Cai L, Xu K (2019) Prediction of ammunition storage reliability based on improved ant colony algorithm and BP neural network. Complexity 2019:1–13. https://doi.org/10.1155/2019/5039097

Liu H, Mi XW, Li YF (2018) Wind speed forecasting method based on deep learning strategy using empirical wavelet transform, long short term memory neural network and Elman neural network. Energy Convers Manag 156:498–514. https://doi.org/10.1016/j.enconman.2017.11.053

Liu S, Xu L, Li D (2016) Multi-scale prediction of water temperature using empirical mode decomposition with back-propagation neural networks. Comput Electrical Eng 49:1–8. https://doi.org/10.1016/j.enconman.2017.11.053

Liu YP, Wu MG, Qian JX (2007) Predicting coal ash fusion temperature based on its chemical composition using ACO-BP neural network. Thermochimica Acta 454(2007):64–68. https://doi.org/10.1016/j.tca.2006.10.026

Luo L, Zhu RG, Song CB et al (2020) Changes in nitrate accumulation mechanisms as PM2.5 levels increase on the North China Plain: a perspective from the dual isotopic compositions of nitrate. Chemosphere. 263(10):127915. https://doi.org/10.1016/j.chemosphere.2020.127915

Martins DK, Stauffer RM, Thompson AM et al (2012) Surface ozone at a coastal suburban site in 2009 and 2010: relationships to chemical and meteorological processes. J Geophys Res Atmospheres 117(D5). https://doi.org/10.1029/2011JD016828

Mchenry JN, Ryan WF, Seaman NL et al (2010) A real-time Eulerian photochemical model forecast system: overview and initial ozone forecast performance in the northeast U.S. corridor. Bull Am Meteorol Soc 85(4):525–548. https://doi.org/10.1175/BAMS-85-4-525

Megaritis AG, Fountoukis C, Charalampidis PE et al (2014) Linking climate and air quality over Europe: effects of meteorology on $PM_{2.5}$ concentrations. Atmospheric Chem Phys 14(18):10283–10298. https://doi.org/10.5194/acp-14-10283-2014

Noori R, Hoshyaripour G, Ashrafi K et al (2009) Uncertainty analysis of developed ANN and ANFIS models in prediction of carbon monoxide daily concentration. Atmospheric Environ 44(4):476–482. https://doi.org/10.1016/j.atmosenv.2009.11.005

Park S, Kim M, Kim M et al (2017) Predicting PM_{10} concentration in Seoul metropolitan subway stations using artificial neural network (ANN). J Hazardous Mater 341:75–82. https://doi.org/10.1016/j.jhazmat.2017.07.050

Pires JC, Martins FG (2011) Correction methods for statistical models in tropospheric ozone forecasting. Atmospheric Environ 45(14):2413–2417. https://doi.org/10.1016/j.atmosenv.2011.02.011

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

Pudykiewicz JA, Kallaur A, Smolarkiewicz PK (1997) Semi-Lagrangian modelling of tropospheric ozone. Tellus B 49(3). https://doi.org/10.3402/tellusb.v49i3.15964

Qiu R, Wang Y, Wang D et al (2020) Water temperature forecasting based on modified artificial neural network methods: two cases of the Yangtze River. Sci Total Environ. 737:139729. https://doi.org/10.1016/j.scitotenv.2020.139729

Ren C, An N, Wang J et al (2014) Optimal parameters selection for BP neural network based on particle swarm optimization: a case study of wind speed forecasting. Knowledge-Based Syst 56:226–239. https://doi.org/10.1016/j.knosys.2013.11.015

Sharma A, Mandal TK, Sharma SK et al (2016) Relationships of surface ozone with its precursors, particulate matter and meteorology over Delhi. J Atmospheric Chem 74:451–474. https://doi.org/10.1007/s10874-016-9351-7

Stockwell WR, Artz RS, Meagher, JF, et al. (2002) The scientific basis of NOAA's air quality forecasting program. EM: Air and Waste Management Association's Magazine for Environmental Managers. December,20-27

Strobl C, Boulesteix AL, Zeileis A, Hothorn T (2007) Bias in random forest variable importance measures: illustrations, sources and a solution. BMC Bioinformatics 8:25. https://doi.org/10.1186/1471-2105-8-25

Stern R, Builtjes P, Schaap M et al (2008) A model inter-comparison study focusing on episodes with elevated PM10 concentrations. Atmospheric Environ 42(19):4567–4588. https://doi.org/10.1016/j.atmosenv.2008.01.068

Sun Q, Tan Z, Zhou X (2020) Workload prediction of cloud computing based on SVM and BP neural networks. J Intell Fuzzy Syst 39(3):2861–2867. https://doi.org/10.1186/1471-2105-8-25

Sun W, Li Z (2020a) Hourly $PM_{2.5}$ concentration forecasting based on feature extraction and stacking-driven ensemble model for the winter of the Beijing-Tianjin-Hebei area. Atmospheric. Pollution Res 11(6):110–121

Sun W, Li Z (2020b) Hourly PM2.5 concentration forecasting based on mode decomposition-recombination technique and ensemble learning approach in severe haze episodes of China. J Cleaner Product 263:121442. https://doi.org/10.1016/j.jclepro.2020.121442

Sun W, Zhang H, Palazoglu A et al (2012) Prediction of 24-hour-average $PM_{2.5}$ concentrations using a hidden Markov model with different emission distributions in Northern California. Sci Total Environ 443(2013):93–103. https://doi.org/10.1016/j.scitotenv.2012.10.070

Tang Z, Wang M, Chen Z et al (2020) Design of multi-stage gear modification for new energy vehicle based on optimized BP neural network. IEEE Access 8:199034–199050. https://doi.org/10.1109/ACCESS.2020.3035570

Taylan O (2017) Modelling and analysis of ozone concentration by artificial intelligent techniques for estimating air quality. Atmospheric Environ 150:356–365. https://doi.org/10.1016/j.atmosenv.2016.11.030

Tu J, Xia ZG, Wang H et al (2007) Temporal variations in surface ozone and its precursors and meteorological effects at an urban site in China. Atmospheric Res 85(3):310–337. https://doi.org/10.1016/j.atmosres.2007.02.003

Vautard R, Beekmann M, Roux J et al (2001) Validation of a hybrid forecasting system for the ozone concentrations over the Paris area. Atmospheric Environ 35(14):2449–2461. https://doi.org/10.1016/S1352-2310(00)00466-0

Voukantsis D, Karatzas K, Kukkonen J et al (2011) Intercomparison of air quality data using principal component analysis, and forecasting of PM_{10} and $PM_{2.5}$ concentrations using artificial neural networks, in Thessaloniki and Helsinki. Sci Total Environ 409(7):1266–1276. https://doi.org/10.1016/j.scitotenv.2010.12.039

Wang L, Zeng Y, Zhang J, et al. (2006) The criticality of spare parts evaluating model using artificial neural network approach. International Conference on Computational Science. pp. 728 – 735

Wang J, Zhang W, Li Y et al (2014) Forecasting wind speed using empirical mode decomposition and Elman neural network. Appl Soft Comput 23:452–459. https://doi.org/10.1016/j.asoc.2014.06.027

Wang JZ, Wang Y, Jiang P (2015a) The study and application of a novel hybrid forecasting model–a case study of wind speed forecasting in China. Appl Energy 143:472–488. https://doi.org/10.1016/j.apenergy.2015.01.038

Wang L, Zeng Y, Chen T (2015b) Back propagation neural network with adaptive differential evolution algorithm for time series forecasting. Expert Syst Appl 42(2):855–863. https://doi.org/10.1016/j.eswa.2014.08.018

Wang J, Heng J, Xiao L et al (2017b) Research and application of a combined model based on multi-objective optimization for multi-step ahead wind speed forecasting. Energy 125:591–613. https://doi.org/10.1016/j.energy.2017.02.150

Wang P, Guo H, Hu J, Kota SH, Ying Q, Zhang H (2019) Responses of PM2.5 and O3 concentrations to changes of meteorology and emissions in China. Sci Total Environ 662:297–306. https://doi.org/10.1016/j.scitotenv.2019.01.227

Wang S, Zhang N, Wu L, Wang Y (2016) Wind speed forecasting based on the hybrid ensemble empirical mode decomposition and GA-BP neural network method. Renewable Energy 94:629–636. https://doi.org/10.1016/j.renene.2016.03.103

Wang WN, Cheng TH, Gu XF et al (2017a) Assessing spatial and temporal patterns of observed ground-level ozone in China. Sci Rep 7(1):3651. https://doi.org/10.1038/s41598-017-03929-w

Wei S, Wang YW, Zhang CC (2018) Forecasting CO2 emissions in Hebei, China, through moth-flame optimization based on the random forest and extreme learning machine. Environ Sci Pollut R 25(29):28985–28997. https://doi.org/10.1007/s11356-018-2738-z

Wei L, Yu C, Yang K et al (2021) Recent advances in VOCs and CO removal via photothermal synergistic catalysis. Chin J Catalysis

42(7):1078–1095. https://doi.org/10.1016/S1872-2067(20)63721-4

Wen L, Yuan X (2020) Forecasting CO2 emissions in Chinas commercial department, through BP neural network based on random forest and PSO. Sci Total Environ 718:137194. https://doi.org/10.1016/j.scitotenv.2020.137194

Xu J, Zhu Y (1994) Some characteristics of ozone concentrations and their relations with meteorological factors in Shanghai. Atmospheric Environ 28(20):3387–3392. https://doi.org/10.1016/1352-2310(94)00154-D

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

Yang G, Huang J, Li X (2018) Mining sequential patterns of PM 2.5 pollution in three zones in China. J Cleaner Product 170(jan.1):388–398. https://doi.org/10.1016/j.jclepro.2017.09.162

Yang Y, Christakos G (2015) Spatiotemporal characterization of ambient PM2.5 concentrations in Shandong Province (China). Environ Sci Technol 49(22):13431–13438. https://doi.org/10.1021/acs.est.5b03614

Yildirim Y, Bayramoglu M (2006) Adaptive neuro-fuzzy based modelling for prediction of air pollution daily levels in city of Zonguldak. Chemosphere 63(9):1575–1582. https://doi.org/10.1016/j.chemosphere.2005.08.070

Yu Y, Zhou ZH (2006) A new approach to estimating the expected first hitting time of evolutionary algorithms. Artificial Intell 172(15):1809–1832. https://doi.org/10.1016/j.artint.2008.07.001

Zhang YP, Chen J, Yang HN et al (2017) Seasonal variation and potential source regions of PM2.5-bound PAHs in the megacity Beijing, China: impact of regional transport. Environ Pollut 231:329–338

Zhang K, Lv G, Guo S et al (2020) Evaluation of subsurface defects in metallic structures using laser ultrasonic technique and genetic algorithm-back propagation neural network. NDT & E Int 116:102339. https://doi.org/10.1016/j.ndteint.2020.102339

Zhang Y, Bocquet M, Mallet V et al (2012) Real-time air quality forecasting, part I: history, techniques, and current status. Atmospheric Environ 60(2012):632–655. https://doi.org/10.1016/j.atmosenv.2012.06.031

Zhou G, Moayedi H, Bahiraei M, Lyu Z (2020) Employing artificial bee colony and particle swarm techniques for optimizing a neural network in prediction of heating and cooling loads of residential buildings. J Cleaner Product 254. https://doi.org/10.1016/j.jclepro.2020.120082