Atmospheric Environment 200 (2019) 264–279

Contents lists available at ScienceDirect

Atmospheric Environment

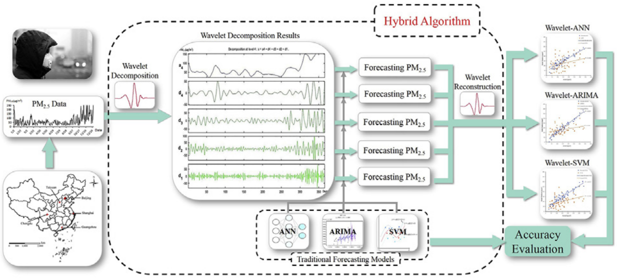
journal homepage: www.elsevier.com/locate/atmosenv

Hybrid algorithm for short-term forecasting of $PM_{2.5}$ in China

Yong Cheng[a], Hong Zhang[a,*], Zhenhai Liu[a], Longfei Chen[a], Ping Wang[b]

[a] College of Environmental & Resource Sciences, Shanxi University, Taiyuan, China
[b] School of Environmental Economics, Shanxi University of Finance & Economics, Taiyuan, China

GRAPHICAL ABSTRACT



ARTICLE INFO

Keywords:
Hybrid models
$PM_{2.5}$
Short-term prediction
Wavelet-ANN
Wavelet-ARIMA
Wavelet-SVM

ABSTRACT

In recent years, the forecasting of particles with a diameter of 2.5 μm or less ($PM_{2.5}$) has been a popular research topic, and involves multiple sources of pollution, making it difficult to determine all of the contributing meteorological and environmental factors. When only the $PM_{2.5}$ concentration time series is considered without other exogenous information, accurate forecasting is important and should be efficient. To address this problem, this paper proposes a hybrid algorithm consisting of multiple models to improve prediction accuracy. The innovation of the proposed hybrid algorithm is to decompose the original single one-dimensional (1D) $PM_{2.5}$ data into multi-dimensional information which effectively mines information hidden in the 1D data. Then, it uses traditional prediction methods to forecast each sequence and to reconstruct its forecasting results to obtain the final forecasting results. Three hybrid models, Wavelet-ANN, Wavelet-ARIMA and Wavelet-SVM, are developed to forecast the 2016 $PM_{2.5}$ trends in 5 Cities in China. The results showed that: (1) Hybrid models (Wavelet-ANN, Wavelet-ARIMA and Wavelet-SVM) can forecast short-term $PM_{2.5}$ concentrations in China. Compared with the traditional Autoregressive Integrated Moving Average (ARIMA), Artificial Neural Network (ANN) and Support Vector Machine (SVM) models, hybrid models can significantly improve prediction accuracy. (2) The Wavelet-ARIMA model has higher accuracy with respect to predicting $PM_{2.5}$ concentrations. In particular, it can more accurately capture the mutational points of $PM_{2.5}$ concentrations, which can provide effective information support for generating warnings about atmospheric pollution. The hybrid algorithm proposed in this paper can be effectively applied to the short-term forecasting of $PM_{2.5}$ concentrations and can significantly improve the accuracy of prediction.

1. Introduction

$PM_{2.5}$ (particles with a diameter of 2.5 μm or less) are mainly composed of water-soluble ions, organic carbon matter (OCM), elemental carbon (EC), and other inorganic compounds (Zhou et al., 2016), which are very harmful to human health (Yan et al., 2017).

[*] Corresponding author.
E-mail address: zhanghong@sxu.edu.cn (H. Zhang).

https://doi.org/10.1016/j.atmosenv.2018.12.025
Received 11 July 2018; Received in revised form 3 December 2018; Accepted 12 December 2018
Available online 21 December 2018

Owing to increased public concern, research on the prediction of $PM_{2.5}$ concentration has been the ongoing focus of air-quality research (Wu et al., 2016), and the accurate prediction of $PM_{2.5}$ concentration is one of the key goals of the 2018 Air Pollution Prevention and Control Action Plan (APPCAP). Cai et al. (2017) proved that APPCAP provided an effective approach to alleviate $PM_{2.5}$ pollution levels, and there is a need for more attention on controlling $PM_{2.5}$ in future. At the same time, by achieving accurate predictions of $PM_{2.5}$ concentrations, the goals of the Air Quality Guidelines (AQG) and the three interim targets (ITs) can be achieved, realizing several potential health benefits. (Maji et al., 2017). Because $PM_{2.5}$ is harmful to humans, it is very important to establish a high-precision $PM_{2.5}$ concentration prediction model for control and monitoring (Ke et al., 2018). In addition, government departments can generate more scientific warnings of potentially severe air pollution incidents based on accurate prediction results (Ping et al., 2017; Lv et al., 2016). Research on the prediction of $PM_{2.5}$ concentrations and the generation of more accurate scientific warnings of potentially severe air pollution incidents based on accurate prediction results are the current focus of ongoing air-quality research, and are areas of significant public concern (Yan et al., 2018).

Many methods have been proposed to analyse and forecast the $PM_{2.5}$ concentration. The commonly employed method uses mainly the Weather Research and Forecasting (WRF) model or the WRF-CMAQ air-quality modelling system for the analysis and prediction (Cao et al., 2018; Weimiao et al., 2017); however, its main functions are simulation and analysis, which require a large amount of meteorological and environmental data, and which are not suitable for predicting future pollutant concentration changes (He et al., 2017). Some scholars (Long et al., 2014; Sun and Sun, 2016) have proposed methods based on principal component analysis (PCA) to predict $PM_{2.5}$ concentrations, but the accuracy of these models cannot be guaranteed. Another kind of model has been proposed based on the relationship between $PM_{2.5}$ concentrations and other meteorological variables based on multi-source data (Ni et al., 2017), but for most situations, accurate pollutant source data and meteorological variables are unavailable (Whiteman et al., 2014; Tai et al., 2010).

Over the past few years, the Autoregressive Integrated Moving Average(ARIMA) model has been widely used to forecast $PM_{2.5}$ concentrations because it is very suitable for univariate time-series predictions (Wang et al., 2018; Jian et al., 2012; Ni et al., 2017). Zafra et al. (2017) used the ARIMA model to perform daily temporal analyses of the effect of land surface coverage (LSC) on PM_{10} concentrations in a high-altitude megacity. García et al. (2018) applied the ARIMA model to calculate daily PM_{10} concentrations based on a simulation in Northern Spain, which effectively improved the prediction accuracy of the model. Jian et al. (2012) used ARIMA to build a framework that could predict the impact of meteorological factors on submicron particle concentrations in cities. However, it assumes that pollutant data are stationarities, and the opportunity to capture non-linearities and non-stationarities in time-series data is limited (Choubin and Malekian, 2017; Ni et al., 2017).

Artificial neural networks (ANNs) are also applied to forecasting in engineering, in which an unknown relationship exists between a set of input factors and an output (Li et al., 2012). ANNs have become a valuable tool for simulating non-linear phenomena, such as $PM_{2.5}$ and PM_{10} concentration prediction (Zhan et al., 2017), SO_2 prediction (Xue and Liu, 2014), surface ozone prediction (Faris et al., 2014), and temperature forecasting (Tao et al., 2013). Perez and Gramsch (2015) presented an hourly $PM_{2.5}$ concentration forecasting model which is based on feed-forward neural networks in Santiago, Chile, and the accuracy of the forecasting model was significantly better than that of traditional forms, which may make it a useful tool for $PM_{2.5}$ forecasting. De et al. (2013) developed an ANN, and tested it to forecast daily PM_{10} concentrations in two different environments in NE Spain, a regional background site, and an urban background site, and the measurement was significantly influenced by vehicular emissions. Park et al. (2017) predicted the indoor PM concentration using the information of outdoor PM, the number of subway trains running, and information on ventilation operation by the ANN model.

Support Vector Machine (SVM) has good robustness and good effect on classification and regression. Many scholars applied SVM model to simulation and prediction in atmospheric particulate matter(PM) (García et al., 2018; Sun and Sun, 2016). García et al. (2018) applied VARIMA, ARIMA, MLPNN, SVM to prediction in PM_{10} concentration, and results have shown that the SVM model was better than the other models to forecast PM_{10} concentration. Sun and Sun (2016) established model which is based on PCA and improved SVM appears to be very attractive and shows a great ability of generalization.

Although ARIMA, ANN and SVM showed excellent performance in predicting $PM_{2.5}$ concentrations, the model's generalisation ability is significantly affected when the predicted data only consist of 1D time series (Garg et al., 2016; Ping et al., 2015). Thus, the focus of this paper is to decompose 1D data into multi-dimensional data as well as to mine as much information hidden in the data as possible to improve the prediction accuracy. For this purpose, we proposed a hybrid algorithm consisting of multiple models to improve prediction accuracy. This hybrid algorithm was applied to the short-term forecasting of $PM_{2.5}$ in 5 cities in China. The objectives of the study were: (1) to propose a logical, simple and widely adaptable hybrid algorithm for the short-term forecasting of $PM_{2.5}$, (2) based on this hybrid algorithm, construct hybrid models which combine wavelet decomposition with benchmark models (ANN, ARIMA and SVM), which are called the Wavelet-ANN model Wavelet-ARIMA model and Wavelet-SVM model, to forecast the $PM_{2.5}$ concentrations, and (3) evaluate the prediction accuracy of hybrid models compared with traditional benchmark models. We hope the hybrid algorithm proposed in this paper provide effective information support for atmospheric pollution warnings.

2. Materials and methods

2.1. Area description

Poor air quality in urban areas is considered to be one of the most serious toxic pollution problems in the world. In China, air pollution has also become a major issue and poses a huge threat to Chinese public health in the period of rapid economic development. In 2016, only 84 out of 338 prefecture-level (administrative division of the People's Republic of China (PRC), ranking below a province and above a county) or higher cities met the national standard for air quality. At present, China's urban-centric air pollution is intensifying and is spreading to the countryside. In some areas that the economy is developed and the population density is high, air pollution matter is particularly acute.

This paper selects five representative cities in China. Beijing is the capital of China and the third most populous city in the world; Chengdu is a sub-provincial city which serves as the capital of Sichuan province and can represents the inland cities of southern China; Guangzhou also known as Canton, is the capital and most populous city of the province of Guangdong in the southern part of China, which able to represents the southern cities of China; Shanghai is one of the four municipalities under the direct administration of the central government of the China, the largest city in China by population, which represents the coastal cities of China, and Taiyuan can represents the inland cities of northern China, which is one of the most important heavy industrial and mining cities in China. The geographical locations of these five cities are shown in Fig. 1.

2.2. Data collection

The $PM_{2.5}$ pollutant data used in this paper were downloaded and recorded through the website http://www.stateair.net and http://hbj.taiyuan.gov.cn. The $PM_{2.5}$ concentration was measured using an RP1400a atmospheric autosampler (fully automatic operation), which

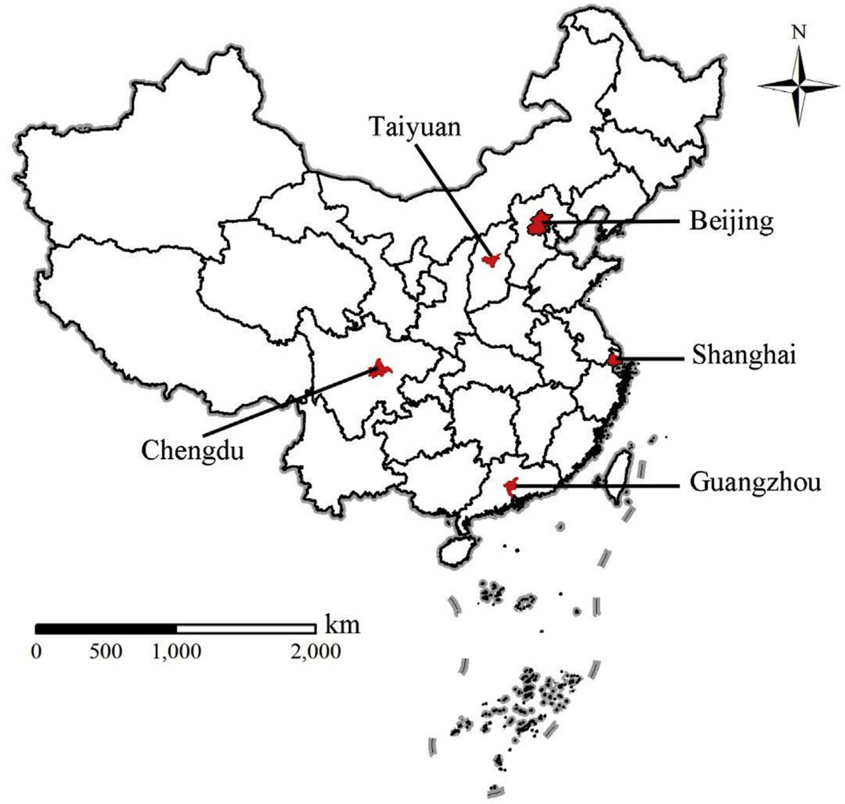


Fig. 1. Five cities in China.

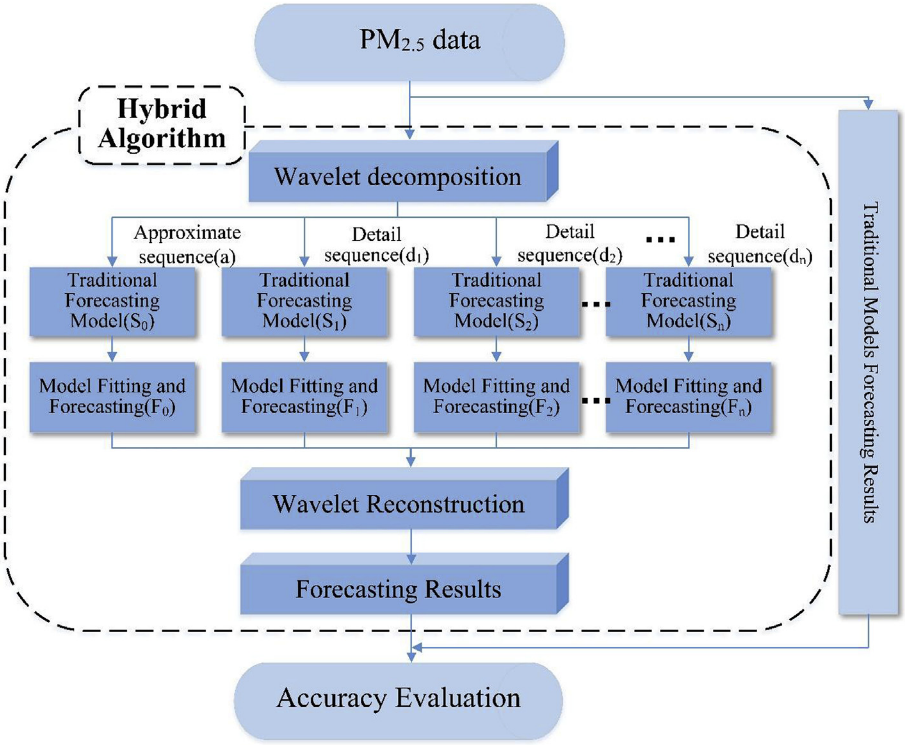

PM2.5 data
Hybrid Algorithm
Wavelet decomposition
Approximate sequence(a)
Detail sequence(d1)
Detail sequence(d2)
Detail sequence(dn)
Traditional Forecasting Model(S0)
Traditional Forecasting Model(S1)
Traditional Forecasting Model(S2)
Traditional Forecasting Model(Sn)
Model Fitting and Forecasting(F0)
Model Fitting and Forecasting(F1)
Model Fitting and Forecasting(F2)
Model Fitting and Forecasting(Fn)
Wavelet Reconstruction
Forecasting Results
Traditional Models Forecasting Results
Accuracy Evaluation

Fig. 2. Technical flowchart of $PM_{2.5}$ forecasting.

can realise real-time automatic sampling, automatic analysis, and data return in real-time. The calibration of zero-gas and standard gas can be used to achieve quality control of the $PM_{2.5}$ concentration data, and these data can be corrected using manual monitoring analysis methods. The time scale of the measured $PM_{2.5}$ data was based on the daily average concentrations from Jan. 1, 2016 to Dec. 31, 2016.

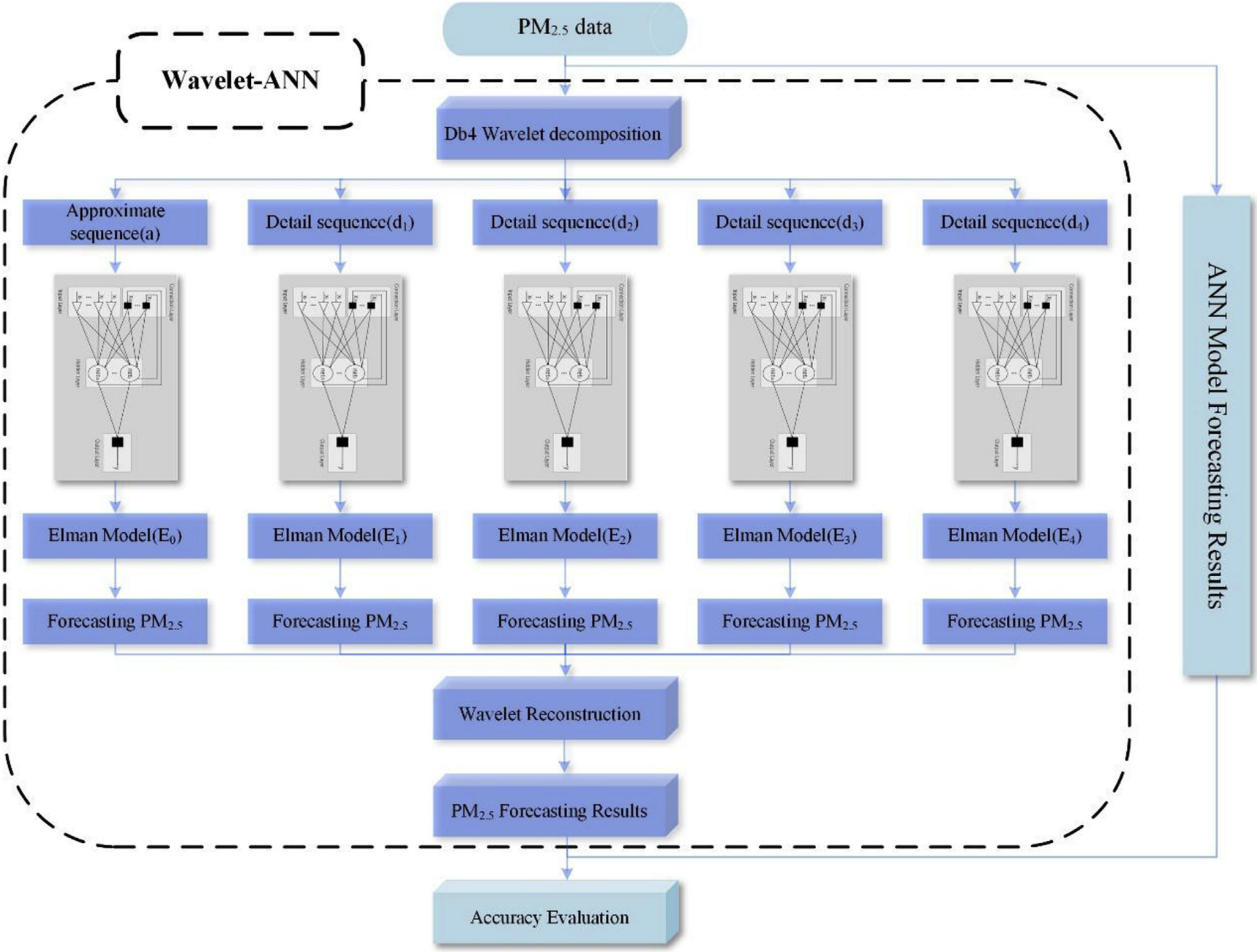


Fig. 3. Framework of the Wavelet-ANN model and ANN model.

2.3. Hybrid forecasting algorithm

In this paper, a hybrid algorithm combined with a benchmark model with wavelet decomposition was proposed to solve the problem of the static structure and poor prediction accuracy of 1D time series. The key steps of the hybrid algorithm are to decompose the original data by wavelet decomposition, convert the original single 1D data into multi-dimensional information, mine the hidden information of the original data, use traditional prediction methods (such as ANN, ARIMA and SVM) to forecast data for each sequence, and finally, to reconstruct the forecasting results of each sequence using wavelets to compose the final forecasting results. The framework of this hybrid algorithm is shown in Fig. 2, and the key procedures of the algorithm are as follows:

Input: Daily $PM_{2.5}$ concentration time series.
Output: Prediction of the daily average concentrations of $PM_{2.5}$ data and evaluation of the accuracy of the models.

Step 1: Standardise the original data (the time series of $PM_{2.5}$ daily average concentration data).
Step 2: Select a suitable wavelet function.
Step 3: Decompose the original 1D data into approximation and detail sequences.
Step 4: Construct a traditional prediction model (e.g. ANN, ARIMA and SVM) to fit and forecast data for each sequence separately.
Step 5: Reconstruct the forecasting results of each sequence, and then compose the final forecasting results.
Step 6: Evaluate the forecasting accuracies of the hybrid model and traditional models.

2.4. Wavelet-ANN method

Based on the key concepts of the hybrid algorithm, we constructed three hybrid models: A Wavelet-ANN model, a Wavelet-ARIMA model and a Wavelet-SVM. The Wavelet-ANN model is a combination of wavelet composition and Elman ANNs.

The first step is to choose the wavelet type which was used to decompose the original $PM_{2.5}$ time series and a level of decomposition *N*. The principle of wavelet decomposition is that the original time series is high-pass filtered to produce several levels of detail in the data, each describing a detailed local change in the original data. It is then low-pass filtered to produce approximate data. Thus, the key concept of wavelet analysis is to select the suitable wavelet basis function. In this paper, the db4 wavelet function was chosen to detect and filter noise. Specifically, this db4 wavelet decomposed the original data into approximation sequences of one layer, and detailed sequences of four layers each.

Following the decomposition of the original data, the next step is to use traditional prediction methods to forecast data for each sequence. In this study, the Elman ANN was chosen to perform $PM_{2.5}$ forecasting. It is a typical local regression neural network which is a type of feedback neural network, and is very similar to forward neural networks,

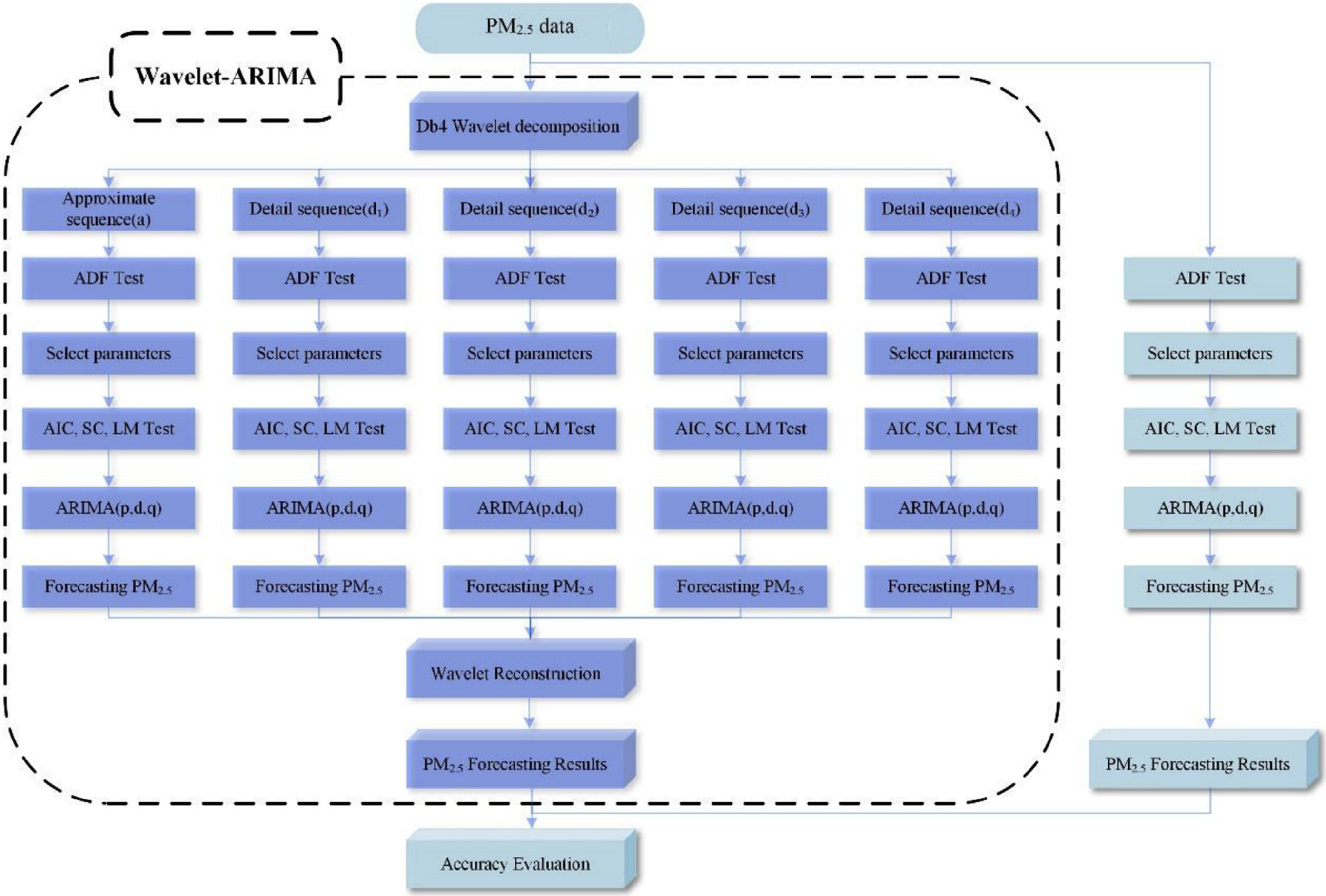


Fig. 4. Framework of the Wavelet-ARIMA model and ARIMA model.

but with superior computing capabilities. Its prime advantages are strong optimal calculations and associative memory functions.

The basic Elman neural network architecture in this hybrid model consists of one input layer, one hidden layer, one connection layer, and one output layer. Compared with BP networks, Elman neural networks have one more connection layer for the construction of local feedback. The transfer function of the connection layer is a linear function, but there is a delay unit. Thus, the connection layer can remember the past state together with the input of the network as the input of the hidden layer at the next moment (Fann et al., 2012). This gives the network a function that occupies dynamic memory. Therefore, Elman neural networks are relatively suitable for time-series forecasting. Prior to creating the Elman neural network and initialising the network, we need to specify the delay parameters, the number of neurons in the hidden layer, the training function, and the determination conditions. Then, we import the normalised training sample set into the Elman network for training. When the network is trained well, we import the normalised test sample set for testing. The output is the forecasting data of each sequence.

The final step is to reconstruct the forecasting results from each of sequence. The forecast accuracies of the Wavelet-ANN model and the ANN model were evaluated. The technology framework of this modelling process is shown in Fig. 3. MATLAB was used to import and forecast the data.

2.5. Wavelet-ARIMA method

Based on the key concepts of the hybrid algorithm, the other hybrid model which we constructed was the Wavelet-ARIMA model, which was a combination of wavelet composition and the ARIMA model. The wavelet method is the same as that in the previous section, and the db4 wavelet function was selected to decompose the $PM_{2.5}$ time series into one approximation and four detailed sequences. Then, with regard to the decomposed sequences, the traditional benchmark ARIMA prediction method was used to forecast data for each sequence.

The autoregressive moving average (ARMA) model considers the time series as a random process, and considers the statistical properties of the time series. Although the composition sequence value of the time series is uncertain, the overall sequence change has a certain regularity. Simply put, it is a concrete description of the autocorrelation of a time series with its own dynamic memory. The process of modelling is the process of quantifying dynamic memory.

The ARMA method is a model which requires data stationarity. If the data is stationary, we can use it to build the ARMA model directly. However, the obtained data are generally not stationary; if we perform differential calculations on the data and make it stationary, then we can build the ARMA model. This whole process is called the ARIMA method. In summary, the ARIMA model is an algorithm which builds an ARMA model using differentially stationary data.

The equation for the ARMA (p, q) model is:

$$y_t = \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \cdots + \varphi_p y_{t-p} + e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \quad (1)$$

where e_t is a random perturbation term.

Stationarity is an assumption of the ARIMA method, and non-stationary data need to be transformed to become stationary. Therefore, prior to using the ARMA algorithm to build the model, we need to perform an Augmented dickey-fuller (ADF) unit root test to justify the stationarity of each sequence. A non-stationary process can easily be transformed into a stationary process by differentiation, which helps to

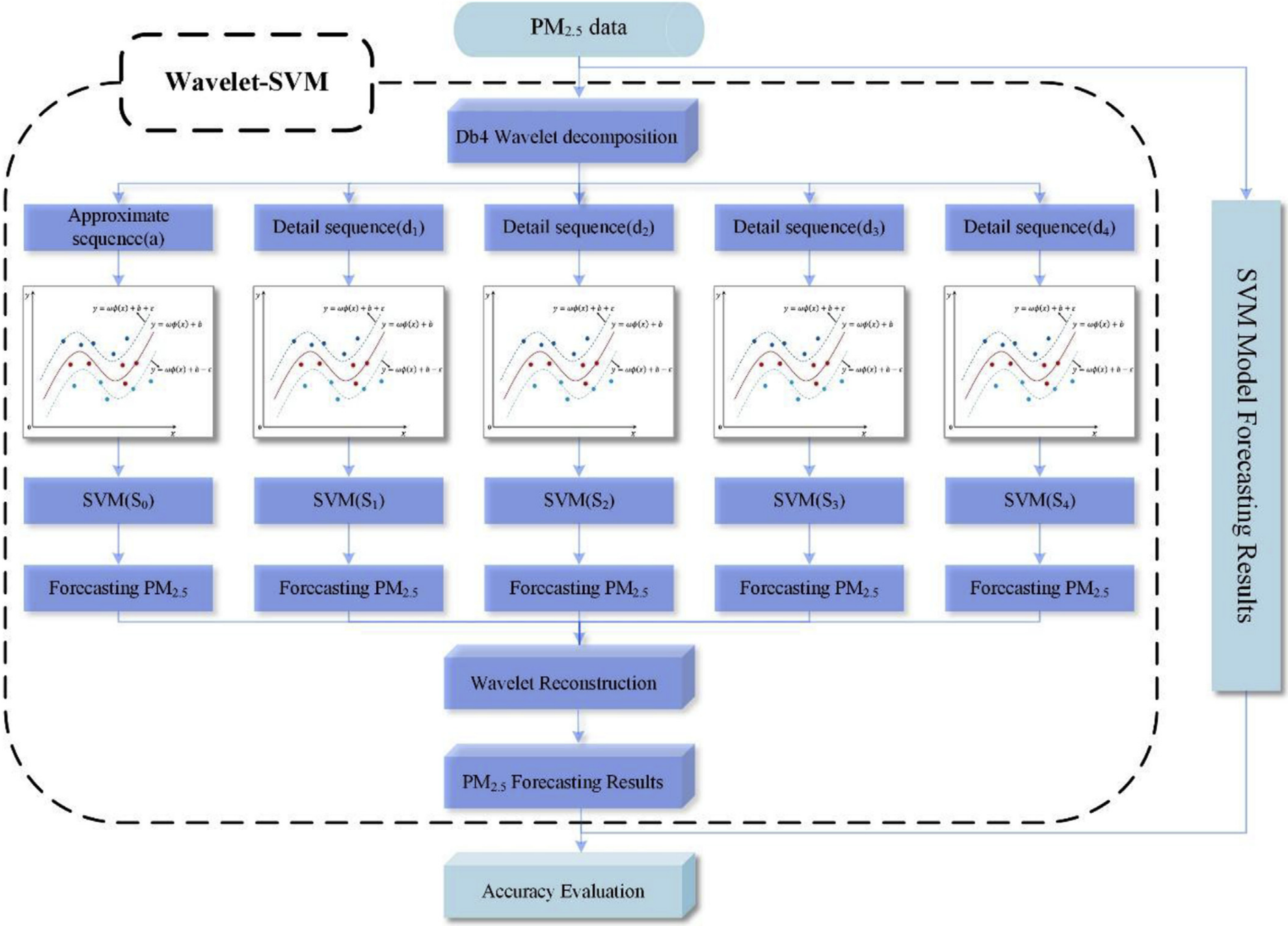


Fig. 5. Framework of the Wavelet-SVM model and SVM model.

extract the underlying trends.

The next step is to identify and determine the model parameters for creating the ARIMA model and Wavelet-ARIMA model. The range of parameter selection for the ARIMA model needs to be determined based on two kinds of statistics: the autocorrelation coefficient (AC) and the partial autocorrelation coefficient (PAC). Then, the specific parameters of p and q are determined by the Akaike information criterion value (AIC) criterion and the Schwarz criterion (SC). The specific principles are as follows:

The parameters p and q can be determined by the truncation and smearing features of AC and PAC. If the AC or PAC decays exponentially or the sine wave attenuates and approaches zero as the lag period increases, then it is said to have smearing. If the AC or PAC quickly approaches zero from a certain lag period as the lag period increases, it is said to be truncated. With this method, the parameters can only be preliminarily determined. If the optimal model cannot be identified, the p and q values can only be tried from lowest to highest. In this process, the model can be adjusted using AIC and SC to detect its quality. The optimal model is one that minimises the variance of AIC and SC.

Finally, the established ARIMA model and Wavelet-ARIMA model should be tested. This step is to check whether the residual sequence of the established model is a self-noise sequence. The Lagrange multiplier (LM) test is generally used to detect random probability. Under normal circumstances, if the random probability is more than 0.05, the residual sequence can be considered as white noise. However, if it is less than 0.05, the residual sequence cannot be considered as a white noise sequence, and the model needs to be re-established; otherwise, the accuracy of the prediction will be reduced.

The technology framework of the Wavelet-ARIMA and ARMA models are shown in Fig. 4. MATLAB and EViews were used to construct the ARIMA model.

2.6. Wavelet-SVM method

Support Vector Machine (SVM) is an effective machine learning method based on statistical learning theory, which is based on the VC dimension theory (Riondato and Upfal, 2015) and the Structural Risk Minimization Inductive Principle of statistical theory. Compared with some traditional machine learning methods, this method has strong robustness and can solve practical problems such as nonlinearity (Balabin and Lomakina, 2011), small sample, high dimensionality and local minimum point, and can effectively avoid “over-harmony". This method has been successfully applied to classification, function approximation and time series forecasting.

In support vector regression (SVR), the learning machine will use the nonlinear mapping to map the prepared training data to the high-dimensional space. Then, in the high-dimensional feature space, the hyperplane (contain slack variables) can form nonlinear relationship between training data and output data, which is called the SVR function, and can also be expressed as a convex optimization problem.:

$$\min_{\alpha} \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}(\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j)K(x_i, y_i) + \varepsilon\sum_{i=1}^{l}(\alpha_i^* + \alpha_i) - \sum_{i=1}^{l} y_i(\alpha_i^* - \alpha_i) \tag{2}$$

Table 1
Descriptive statistics of average daily $PM_{2.5}$ concentrations in 2016 in five cities.

City	Statistics	s	a_4	d_1	d_2	d_3	d_4
Beijing	Mean	72.89	74.54	−0.02	0.06	−0.08	−1.61
	Median	56.60	62.96	0.05	−0.41	−0.17	−0.04
	Standard deviation	65.37	37.79	23.23	30.43	35.43	28.49
	Kurtosis	4.31	6.83	5.04	1.42	3.12	7.17
	Skewness	1.91	2.28	−0.19	0.09	0.20	−1.26
	Minimum	5.50	21.99	−122.38	−100.19	−123.39	−157.91
	Maximum	381.67	245.32	110.68	99.70	154.68	100.24
Chengdu	Mean	72.83	73.05	0.00	0.01	0.07	−0.29
	Median	62.23	64.46	−0.34	0.82	1.13	−0.07
	Standard deviation	38.40	25.91	11.73	13.56	18.18	11.41
	Kurtosis	0.96	0.45	3.08	0.58	1.27	0.99
	Skewness	1.14	0.94	0.06	−0.29	−0.32	0.16
	Minimum	14.38	35.03	−48.21	−42.49	−67.96	−30.21
	Maximum	221.71	149.02	48.47	41.10	56.58	36.06
Guangzhou	Mean	32.99	32.92	0.00	−0.02	0.06	0.03
	Median	29.75	30.40	−0.05	−0.25	−0.27	−0.09
	Standard deviation	20.51	11.63	7.48	7.28	10.60	7.25
	Kurtosis	5.28	0.01	2.22	1.10	4.46	0.46
	Skewness	1.67	0.73	−0.25	0.00	0.62	0.25
	Minimum	2.14	14.94	−27.15	−30.11	−36.15	−22.23
	Maximum	159.00	67.39	25.70	21.47	56.91	21.61
Shanghai	Mean	45.43	45.50	0.00	0.00	0.05	−0.12
	Median	38.98	49.83	0.00	0.46	0.26	0.01
	Standard deviation	28.63	18.03	12.38	12.89	10.67	8.13
	Kurtosis	1.97	0.13	2.72	0.94	4.16	4.32
	Skewness	1.34	0.25	−0.20	−0.07	−0.51	0.26
	Minimum	4.55	12.17	−52.49	−43.61	−58.50	−28.53
	Maximum	160.25	95.60	45.19	36.97	40.96	36.80
Taiyuan	Mean	66.21	67.49	−0.01	−0.01	−0.08	−1.32
	Median	51.00	52.34	−0.25	−0.23	−0.43	−0.39
	Standard deviation	53.30	40.98	16.93	21.65	20.67	20.56
	Kurtosis	6.25	1.41	4.62	3.73	3.09	12.36
	Skewness	2.20	1.55	0.10	−0.02	0.16	−1.75
	Minimum	5.00	21.92	−82.97	−87.89	−83.16	−126.38
	Maximum	341.00	192.00	86.00	98.47	73.84	90.57

s refers to the original time series of $PM_{2.5}$ concentrations; a_4 refers to the approximation sequence; and d_1, d_2, d_3, and d_4 are the four detail sequences.

$$\text{s.t.} \quad \sum_{i=1}^{l}(\alpha_i^* - \alpha_i) = 0, \qquad 0 \le \alpha_i, \quad \alpha_i^* \le \frac{C}{l}, \qquad i = 1,2, \ldots, l,$$

where $K(x_i, y_i)$ is kernel function and α_i, α_i^* are Lagrange multipliers. The resulting regression function $f(x)$ is as follows:

$$f(x) = \sum_{i=1}^{l}(\alpha_i^* - \alpha_i)K(x, x_i) + b \tag{3}$$

Because SVR forms a regression relationship between input data and output data, the quality of the input data greatly affects the prediction effect of the model. Based on the hybrid algorithm, Wavelet decomposition is performed on the input data, which decompose the $PM_{2.5}$ time series into one approximation and four detailed sequences, and then divide the data of each sequence into training data and test data. The SVM model is used to train and predict each layer of data after decomposition. The next is to reconstruct the forecasting results from each of sequence. And the last is to evaluate the performance of Wavelet-SVM. The technology framework of this modelling process is shown in Fig. 5. MATLAB was applied to finish the process.

2.7. Accuracy evaluation

The average absolute error (MAE), root mean-squared error (RMSE) and coefficient of determination (R^2) were used to evaluate the accuracy of the prediction model. Formulae (4), (5) and (6) are as follows:

$$MAE = \frac{\sum_{i=1}^{n}|y_i - \hat{y}_i|}{n} \tag{4}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n}} \tag{5}$$

$$R^2 = \frac{ESS}{TSS} = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6}$$

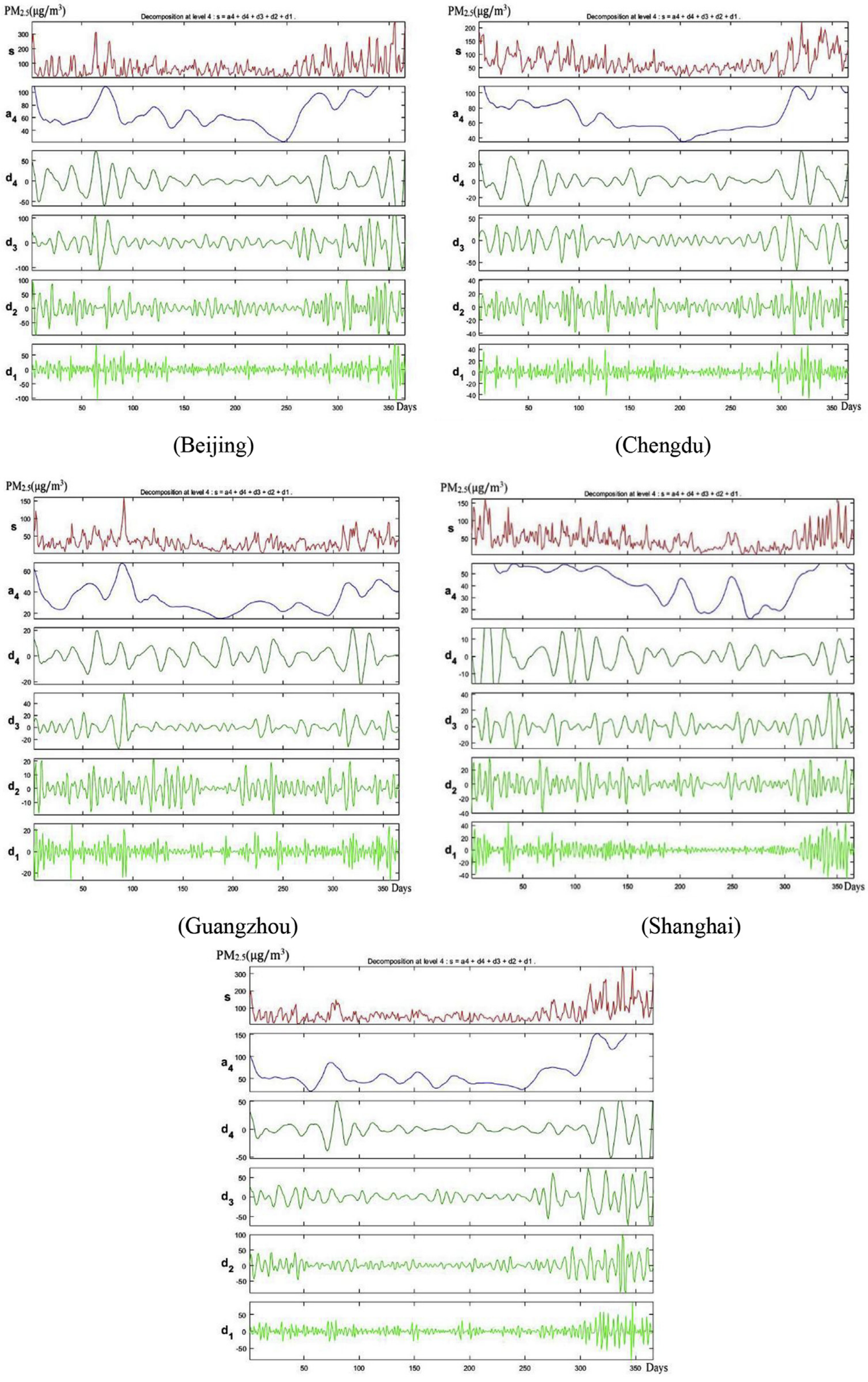
where y_i represents the actual value of sample i, $\hat{y}_i$ represents the predicted value of sample i, $\bar{y}$ is the average of the actual value of sample i, and n is the number of samples.

The smaller the values of MAE and RMSE, the smaller will be the prediction error of the model and the better the prediction model. The value of R^2 indicates the proportion of the explained sum of squares (ESS) in the total sum of squares (TSS), and is between 0 and 1. R^2 can be used as a measure of the goodness of fit of the regression equation. The larger the value of R^2, the better is the fitting effect.

3. Result analysis

3.1. Wavelet decomposition and descriptive statistics

Wavelet decomposition of the original $PM_{2.5}$ data using the db4 wavelet was performed, and the original data of these five cities (Beijing, Chengdu, Guangzhou, Shanghai, Taiyuan) were decomposed into one approximation sequence (a_4), which reflects the overall trend



(Beijing)

(Chengdu)

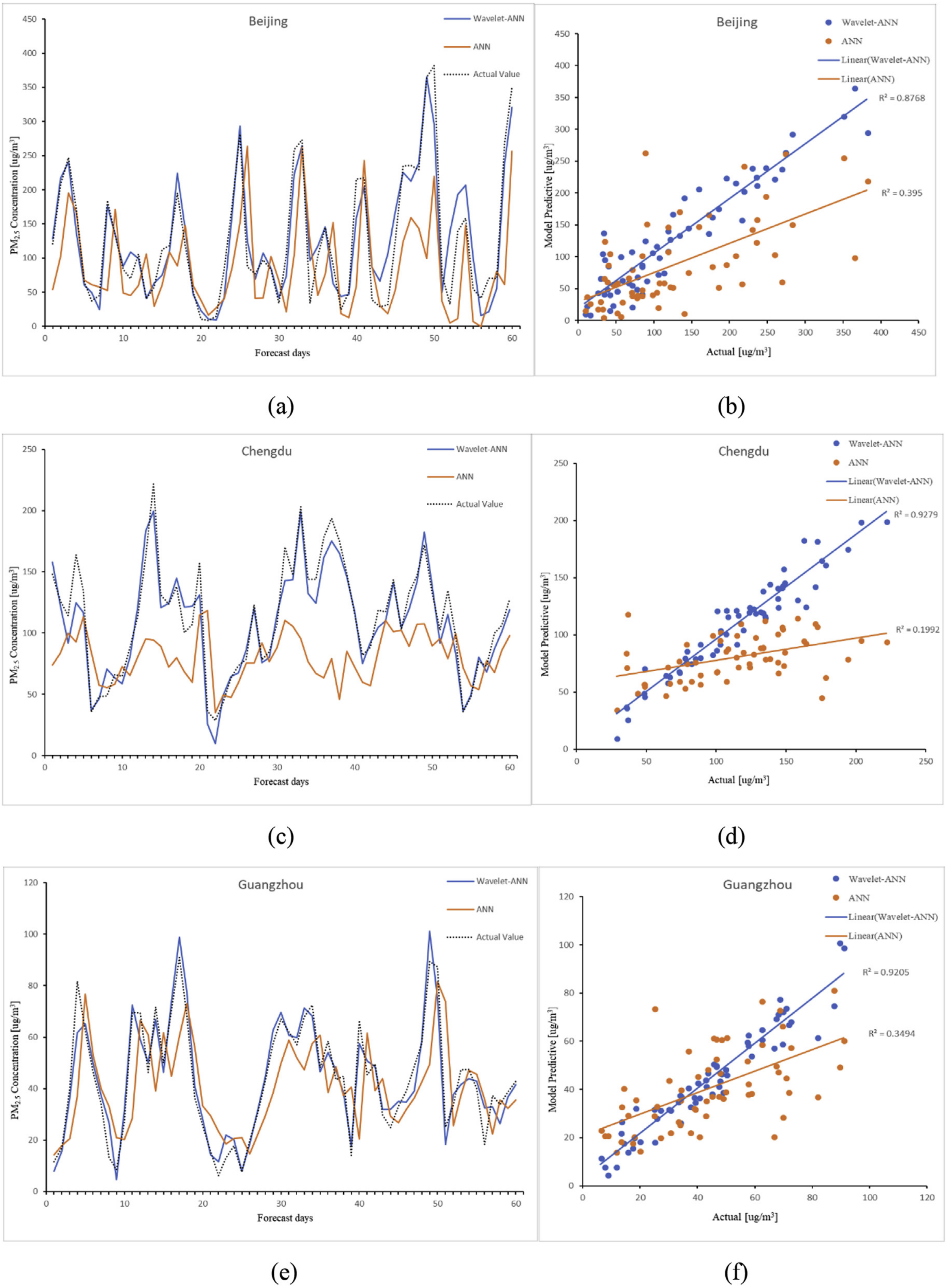
(Guangzhou)

(Shanghai)

(Taiyuan)

Fig. 6. Wavelet decomposition of daily average $PM_{2.5}$ concentrations in 2016 in five cities.

of the original time series of $PM_{2.5}$ concentrations and four detail sequences (d_1, d_2, d_3, and d_4) which reflect the noise of the original time series of $PM_{2.5}$ concentrations. Descriptive statistics, including the mean, median, maximum, minimum, standard deviation, skewness, and kurtosis were used to quantitatively summarise the characteristics of average daily $PM_{2.5}$ concentrations of each decomposed sequence to further explain their significance. The mean and median are statistics which reflect the central tendencies in the data set. The minimum and maximum values show the amplitude of the time series. The standard deviation is a measure of the degree of dispersion of the data distribution, and is used to measure the degree to which the data deviate from the arithmetic mean. Skewness and kurtosis are used to assess



(a)

(b)

(c)

(d)

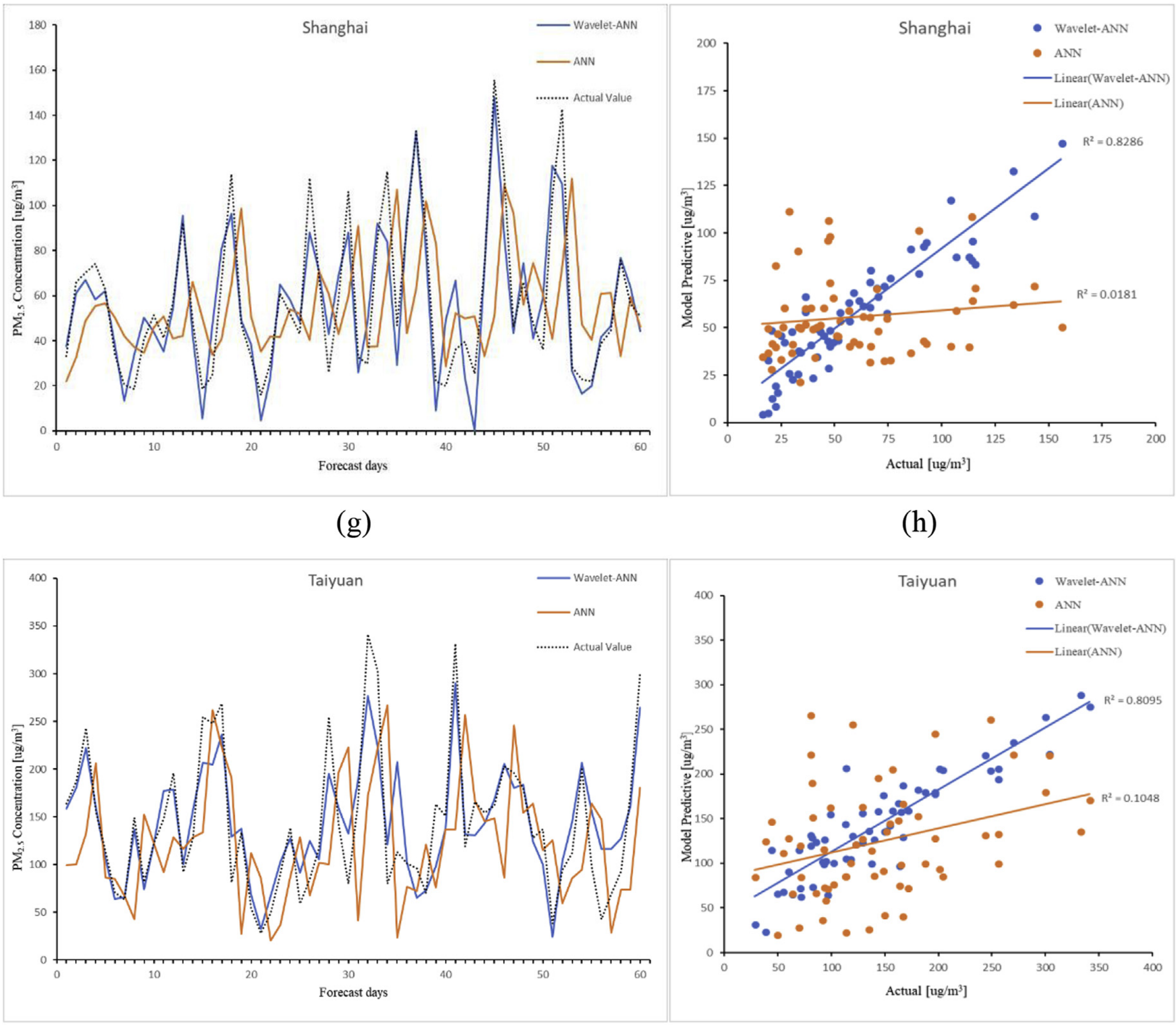
(e)

(f)

Fig. 7. Forecasting results of the ANN model and Wavelet-ANN model.

whether the sampling distribution is normal. The descriptive statistics of these five cities are shown in Table 1 and Fig. 6.

The statistical results revealed the following: (1) The mean and standard deviation of the original time series were 72.89 and 65.37 in Beijing, 72.83 and 38.40 in Chengdu, 32.99 and 20.51 in Guangzhou, 45.43 and 28.63 in Shanghai, 66.21 and 53.30 in Taiyuan respectively. The skewness and kurtosis of the five cities were much greater than 1, which indicates that the original time series did not follow the normal distribution roughly. (2) The approximation sequence (a_4) reflects the overall trend of the original time series, and the detail sequences (d_1, d_2, d_3, and d_4) reflect the noise of the original time series. From the approximation sequence (a_4) in Fig. 6, $PM_{2.5}$ concentrations in these



(g)

(h)

(i)

(j)

Fig. 7. (*continued*)

five cities are generally high in spring and winter, and low in summer and autumn. (3) In most cases of these five cities, comparing the original time series with the decomposed sequences, the skewness and kurtosis of the approximation sequence and the four detail sequences were smaller than that of the original $PM_{2.5}$ data. This indicates that compared with the original $PM_{2.5}$ sequence, the decomposed sequences using wavelets were more stable, and tended to follow a normal distribution.

3.2. Forecasting of $PM_{2.5}$ with ANN and Wavelet-ANN

The Wavelet-ANN model was established to forecast the $PM_{2.5}$ concentrations in five cities (Beijing, Chengdu, Guangzhou, Shanghai, Taiyuan) in China. To justify the use of the Wavelet-ANN model, it is important to know whether the hybrid Wavelet-ANN model leads to an improvement relative to traditional ANN. Therefore, comparisons based on the forecasting performance of $PM_{2.5}$ were made between the traditional ANN and the hybrid Wavelet-ANN in five cities in China.

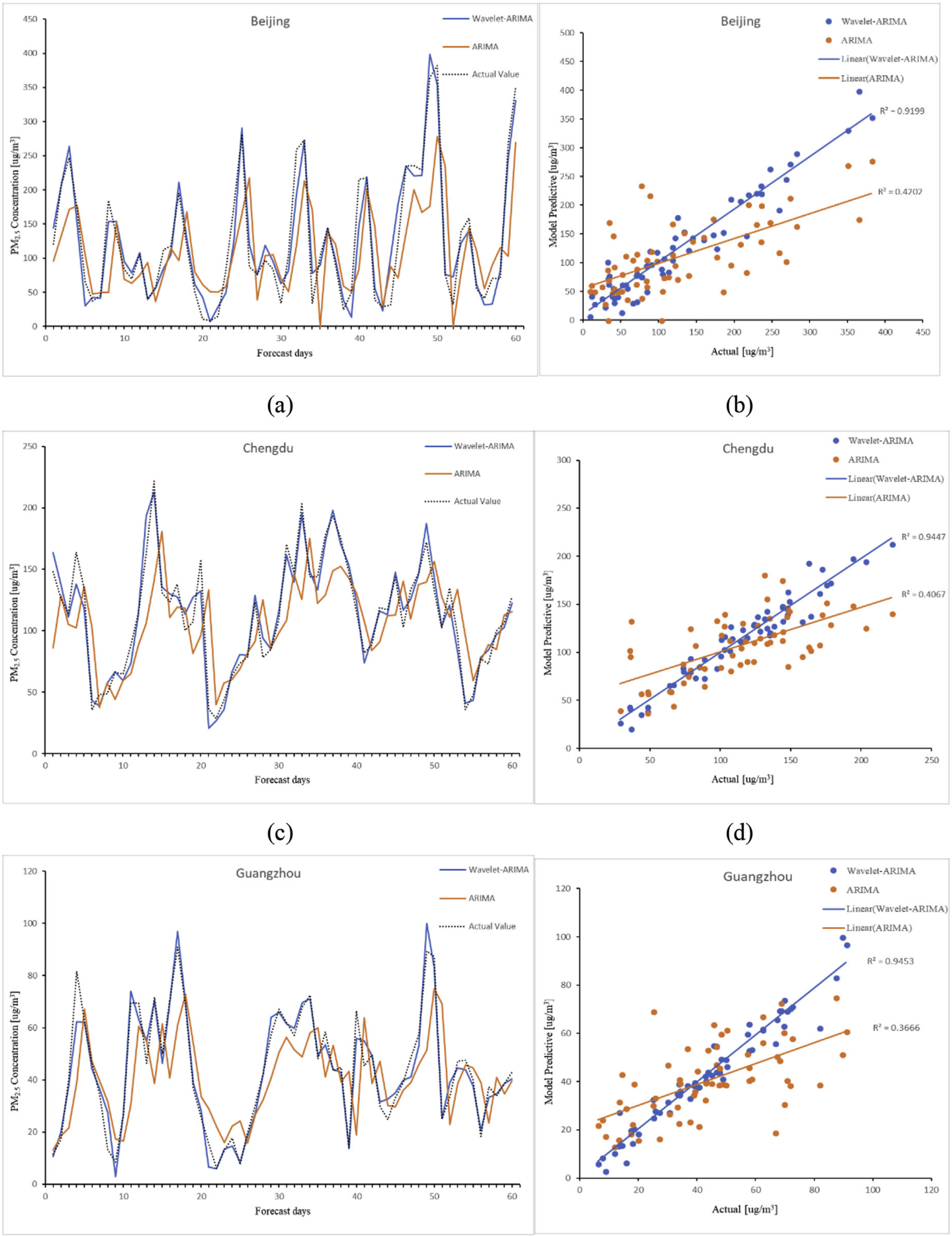
The $PM_{2.5}$ time series of the first 300 days of 2016 was used as a source for the training samples to fit the model, and the remainder of the 66 data points of 2016 were used as the test samples to verify the model. The Elman ANN has one input layer with six variables x, one hidden layer with fifteen neurons, one connection layer with fifteen independent variables, and one output layer with a dependent variable. When six independent variables are input, that is, the $PM_{2.5}$ concentrations of the current day are to be calculated using data from six days earlier, we used the data from the second day to the seventh day to calculate the data of the eighth day. We can repeatedly calculate and obtain the last 60 days of forecasting data.

For traditional ANN, the network was applied only to the original $PM_{2.5}$ time series; the output contained the final forecasting results. For Wavelet-ANN, the network was applied to four detail sequences and one approximation sequence of $PM_{2.5}$ time series; the final forecasting results integrated the output of the approximation sequence and the four detail sequences. The performances of the two models are shown in Fig. 7(a,c,e,g,i). The correlation between daily observed and modelled $PM_{2.5}$ concentrations is shown in Fig. 7(b,d,f,h,j).

It can be seen from Fig. 7(a,c,e,g,i) that both the Wavelet-ANN and ANN models produced results which are essentially in agreement with the actual values; however, the $PM_{2.5}$ concentrations modelled by the ANN were consistently underestimated compared to those modelled by the Wavelet-ANN. In particular, for peak values, the Wavelet-ANN can provide much more accurate and usable information than the ANN.

As can be seen from Fig. 7(b,d,f,h,j), the scatter plot indicates the strength of a linear relationship between the modelled value (y) and the observed value (x).

Take Beijing as an example, it is obvious that the wavelet-ANN (blue dot) appears to be distributed more normally than the ANN (orange



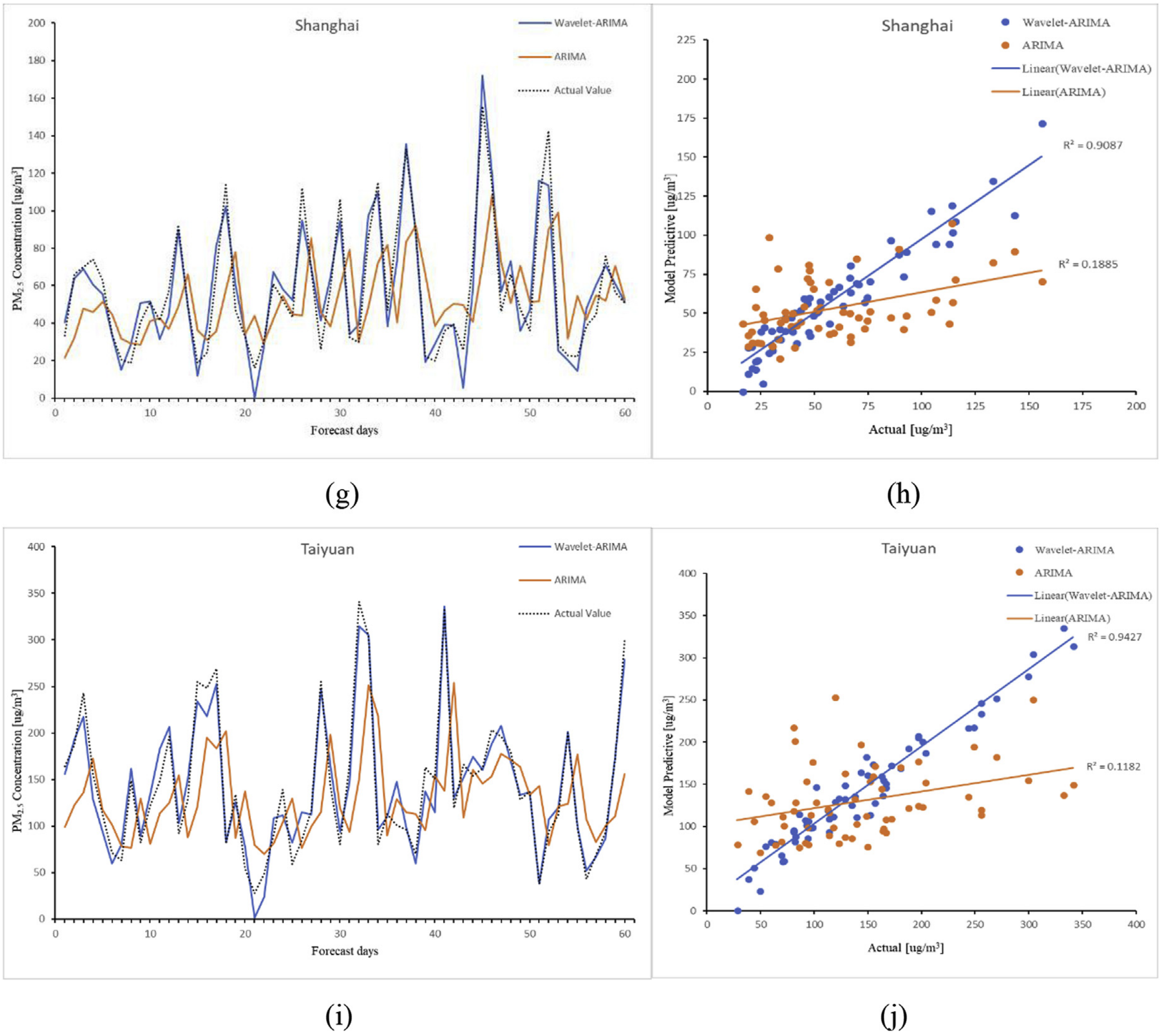
(a) (b)

(c) (d)

(e) (f)

Fig. 8. ARIMA model and Wavelet-ARIMA model forecasting results.

dot), and the coefficient of determination of the Wavelet-ANN was 0.8768, which is higher than 0.3950 of the ANN. The comparison results of other cities are basically the same as that in Beijing. Meanwhile, we can find that the best forecasting result of the Wavelet-ANN model is Chengdu, and the coefficient of determination is 0.9279. Using the Wavelet-ANN model revealed that the wavelet decomposed the single signals of the original $PM_{2.5}$ time series to multi-layer signals, which contain useful knowledge and can improve the forecasting performance. Overall, the Wavelet-ANN model outperforms the ANN model.



(g) (h)

(i) (j)

Fig. 8. (*continued*)

3.3. Forecasting of $PM_{2.5}$ with ARIMA and Wavelet-ARIMA

The hybrid Wavelet-ARIMA and the traditional ARIMA methods were established to forecast future $PM_{2.5}$ concentrations, and comparisons between the $PM_{2.5}$ forecasting performances were also made for the two models. $PM_{2.5}$ data of the first 300 days of 2016 were used as a source for training samples to fit the model, and the rest of the 66 data points of 2016 were used as the test samples to verify the model.

The training samples were decomposed by the wavelet function into four detail sequences and one approximation sequence. The ADF unit root test method was used on the original $PM_{2.5}$ data and the five wavelet-decomposed sequences to test whether the time-series variable is non-stationary. The results showed that the first-order differential $PM_{2.5}$ data and five wavelet-decomposed sequences all pass the stationarity test ($P < 0.001$), which indicated that these time series data points were stationary.

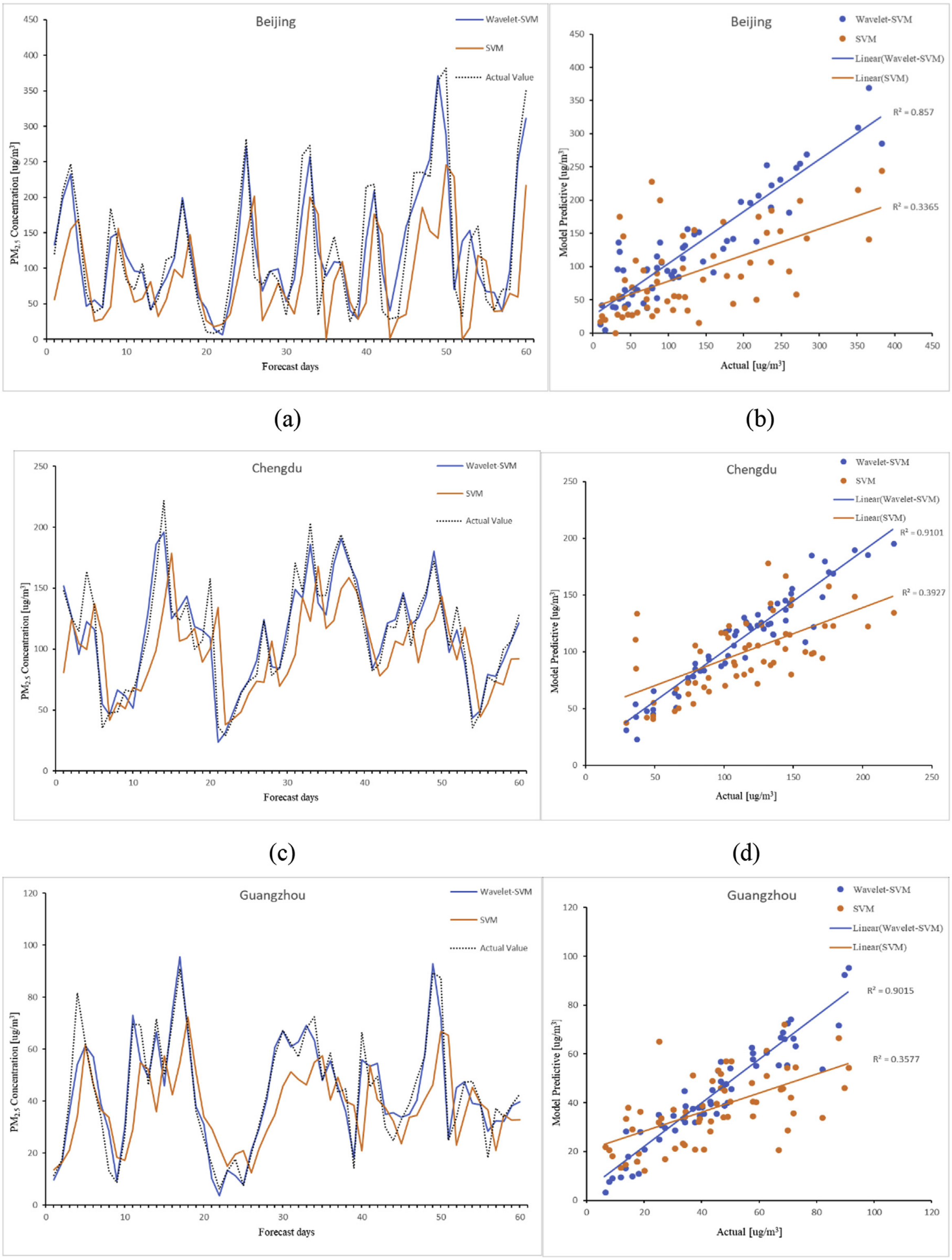
Where the data show evidence of stationarity, the identification and specification of appropriate ARIMA parameters is an important step. The range of ARIMA parameters was determined by AC and PAC; specifically, the values of p and q are determined by AIC and SC. The results showed that the p and q values of the original $PM_{2.5}$ data and the five wavelet-decomposed sequence data are all different. Thus, we established six different ARIMA models for the original $PM_{2.5}$ data and the five wavelet-decomposed sequences in each city.

The LM test was performed on the rest of the established ARIMA model to test whether the noise values constituted a white noise sequence. The results showed that the residue sequences of the six ARIMA models are all white noise sequences, which means that these models can be used for further forecasting. Thus, based on the six trained ARIMA models, we forecasted the rest of the 66 $PM_{2.5}$ concentrations, and obtained the final forecast results.

For traditional ARIMA, the model was used only on the original $PM_{2.5}$ time series, and the output contained the final forecasting results. For Wavelet-ARIMA, the model was applied to four detail sequences and one approximation sequence of $PM_{2.5}$ time series, and the final forecasting results integrated the output of the approximation sequence and the four detail sequences. The performances of the two models are shown in Fig. 8(a,c,e,g,i). The correlation between the daily observed and modelled $PM_{2.5}$ concentrations is shown in Fig. 8(b,d,f,h,j).

Fig. 8(a,c,e,g,i) shows that both the Wavelet-ARIMA and ARIMA models can predict future points in the series in agreement with the actual values. By comparing the two models, we can see that the $PM_{2.5}$ concentrations modelled by ARIMA were consistently underestimated compared to those modelled by Wavelet-ARIMA, both for the peaks and the valleys.

Fig. 8(b,d,f,h,j) shows the strength of a linear relationship between the modelled value (y) and observed value (x). Take Beijing as an example, it is obvious that Wavelet-ARIMA (blue dot) appear to be closer to a linear curve than the ARIMA (orange dot), and the coefficient of determination of the Wavelet-ARIMA was 0.9199, which was higher



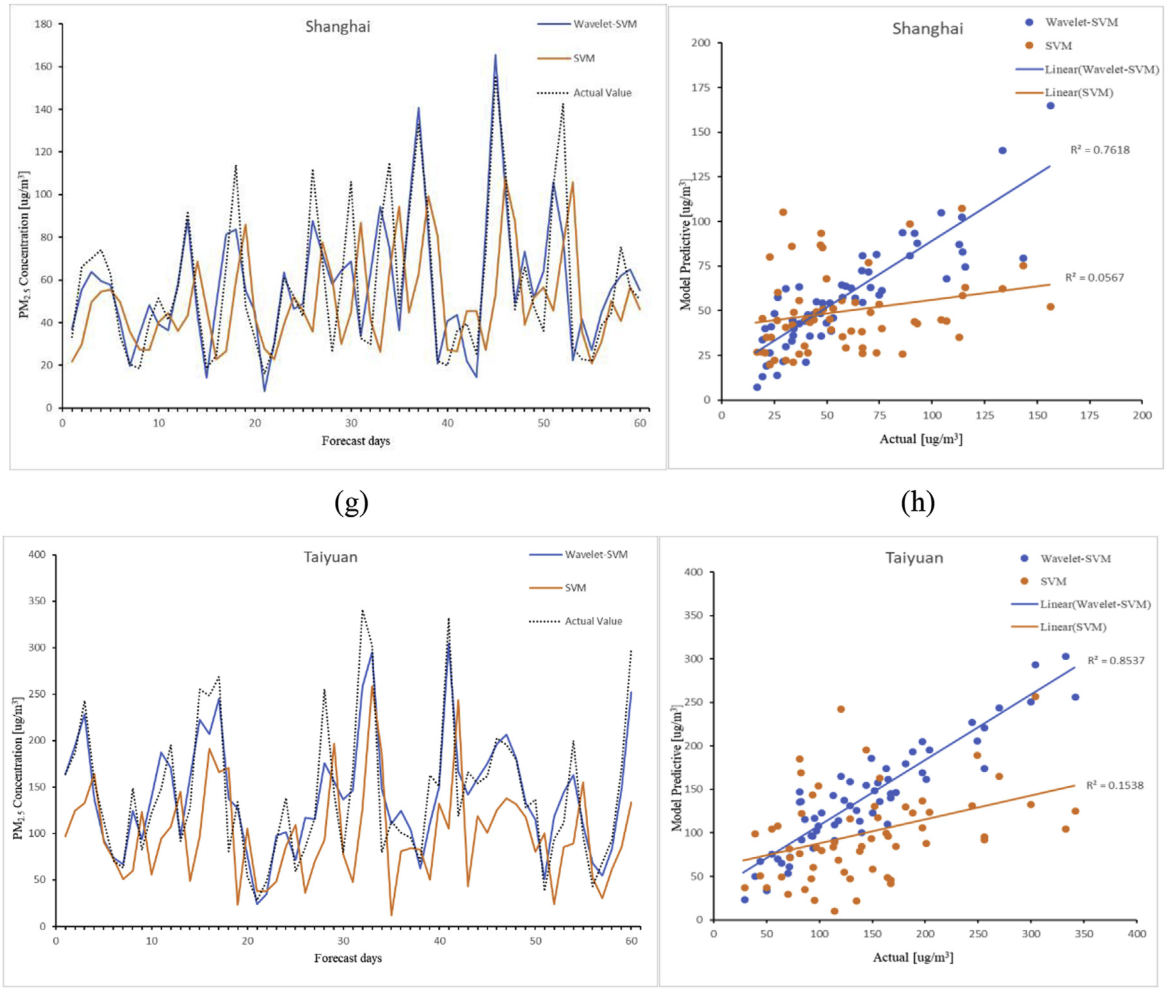
(a) (b)

(c) (d)

(e) (f)

Fig. 9. SVM model and Wavelet-SVM model forecasting results.

than 0.4202. The comparison results of other cities are basically the same as that in Beijing. Meanwhile, we can see from Fig. 8 that the Wavelet-ARIMA model has the best forecasting effect in Chengdu, and the predicted data is basically consistent with the trend of the actual data, and the coefficient of determination is the highest of 0.9453. However, the forecasting effect is slightly worse in Shanghai, with the coefficient of determination is 0.9087, but also exceeds 0.9. The use of the Wavelet-ARIMA model revealed that the wavelet decomposed the single signals of the original $PM_{2.5}$ time series into multi-layer signals, either to better understand the data or to improve the forecasting performance.



(g)

(h)

(i)

(j)

Fig. 9. (*continued*)

3.4. Forecasting of $PM_{2.5}$ with SVM and Wavelet-SVM

The traditional SVM method and hybrid Wavelet-SVM were established. The five representative cities (Beijing, Chengdu, Guangzhou, Shanghai and Taiyuan) in China were predicted respectively, comparisons $PM_{2.5}$ forecasting performances between SVM and Wavelet-SVM models. The $PM_{2.5}$ time series of the first 300 days of 2016 was used as a source for the training samples to fit the model, and the remainder of the 66 data points of 2016 were used as the test samples to verify the model.

In the Wavelet-SVM model, we separately decomposed the $PM_{2.5}$ data of each city, and then decompose each city's $PM_{2.5}$ data into five layers, including one approximation and four detailed sequences, and then divide each layer into training data and test data. The choice was to use the six-day as the independent variable and the seventh day as the dependent variable, which was the same as the data processing method selected by the Wavelet-ANN model.

The performances of the two models (SVM and Wavelet-SVM) are shown in Fig. 9(a,c,e,g,i). The correlation between daily observed and modelled $PM_{2.5}$ concentrations is shown in Fig. 9(b,d,f,h,j).

From Fig. 9(a,c,e,g,i) of the five cities, we can see that the prediction data of the two models can reflect the fluctuation trend of the original data, but overall, the Wavelet-SVM (blue dot) model is closer to the actual value (black dot), while the SVM (orange dot) model does not perform well without the Wavelet-SVM model. Take Beijing as an example, the coefficient of determination of the Wavelet-SVM were 0.8570, which was higher than 0.3365. Meanwhile, the comparison results of other cities are basically the same as that in Beijing. Overall, the Wavelet-SVM model based on the hybrid algorithm in this paper has greatly improved the prediction accuracy compared with the traditional benchmark model, especially in Shanghai and Taiyuan, with an increase of 0.7051 and 0.699 respectively. From the above, the hybrid algorithm model has high precision and superiority.

3.5. Comparison of forecasting accuracy

An excellent prediction method can yield uncorrelated residuals with an average of zero. If there is a correlation between the residual values, the information left in the residuals should be used in the calculation of the prediction. If the mean of the residuals is not zero, this indicates that the forecast is biased. The forecast error metrics of the six models chosen in this study were MAE and RMSE, as shown in Table 2. The values of R^2, which reflect the forecasting accuracy, are also shown in Table 2.

The forecasting performance of the five cities showed the same trend. Take Beijing as an example, among the three benchmark models, the MAE and RMSE of the ARIMA model were the smallest and the R^2 was the largest. Among the three hybrid models, the Wavelet-ARIMA model has the best prediction results, and the MAE and RMSE are the smallest among the three hybrid models, as well as the R^2 is 0.9199, which is higher than the Wavelet-ANN model (0.8768) and Wavelet-SVM model (0.8570).

Table 2
Prediction of the six models' MAE, RMSE and R^2 in five cities.

City	Model		MAE	RMSE	R^2
Beijing	Traditional forecasting models	ANN	61.95	82.82	0.3950
		ARIMA	56.12	72.94	0.4202
		SVM	65.70	85.55	0.3365
	Hybrid forecasting models	Wavelet-ANN	24.20	32.87	0.8768
		Wavelet-ARIMA	19.82	26.43	0.9199
		Wavelet-SVM	27.02	36.11	0.8570
Chengdu	Traditional forecasting models	ANN	40.24	51.47	0.1992
		ARIMA	26.79	35.19	0.4067
		SVM	29.08	37.72	0.3927
	Hybrid forecasting models	Wavelet-ANN	10.07	12.93	0.9279
		Wavelet-ARIMA	8.48	10.62	0.9447
		Wavelet-SVM	9.91	13.50	0.9101
Guangzhou	Traditional forecasting models	ANN	14.25	18.31	0.3494
		ARIMA	13.72	17.83	0.3666
		SVM	14.27	18.57	0.3577
	Hybrid forecasting models	Wavelet-ANN	4.87	6.17	0.9205
		Wavelet-ARIMA	3.55	5.14	0.9453
		Wavelet-SVM	4.82	6.87	0.9015
Shanghai	Traditional forecasting models	ANN	29.12	37.14	0.0181
		ARIMA	24.30	31.00	0.1885
		SVM	27.42	36.33	0.0567
	Hybrid forecasting models	Wavelet-ANN	10.68	14.00	0.8286
		Wavelet-ARIMA	8.28	10.21	0.9087
		Wavelet-SVM	11.70	16.38	0.7618
Taiyuan	Traditional forecasting models	ANN	65.68	81.69	0.1048
		ARIMA	56.30	72.45	0.1182
		SVM	67.42	83.60	0.1538
	Hybrid forecasting models	Wavelet-ANN	25.20	33.54	0.8095
		Wavelet-ARIMA	14.49	17.78	0.9427
		Wavelet-SVM	22.31	29.39	0.8537

Similarly, we analysed the forecasting results of Chengdu, Guangzhou, Shanghai, and Taiyuan. We can find that the prediction results of several models are basically the same as those in Beijing. Based on our proposed hybrid algorithm, MAE and RMSE have a significant reduction compared to the value of the benchmark model, and the R^2 has been significantly improved. The hybrid model prediction results are all superior to the benchmark model. The prediction results of the Wavelet-ARIMA model in these five cities are always the best of the six models, and the values of R^2 are all higher than 0.9, showing a good prediction effect. As the Wavelet-ANN model and the Wavelet-SVM model, we can find that in the four cities of Beijing, Chengdu, Guangzhou, and Shanghai, the Wavelet-ANN model is better than the Wavelet-SVM model. Only the Taiyuan City, Wavelet-SVM model prediction results are better than the Wavelet-ANN model.

In general, based on our proposed hybrid algorithm, the prediction verification is carried out in five representative cities in China, and the proposed hybrid model can significantly improve the prediction accuracy. Overall, the Wavelet-ARIMA model has the best prediction effect, followed by Wavelet-ANN model and Wavelet-SVM model.

4. Conclusion and discussion

This paper proposed a hybrid algorithm for short-term air pollution forecasting, and two hybrid models were developed (Wavelet-ANN, Wavelet-ARIMA and Wavelet-SVM) based on this idea. The basic principles and modelling steps were described, and an experiment was carried out based on 2016 data of $PM_{2.5}$ concentrations in 5 Cities in China. The forecasting accuracy of the hybrid models was compared with that of traditional models respectively.

The results indicated that (1) the proposed hybrid algorithm in this paper combined the traditional model with wavelet decomposition, were superior to traditional benchmark models, which the result is available from five cities in China. (2) Compared with traditional benchmark models, the forecasting precision of the hybrid algorithm was greatly improved. In particular, the Wavelet-ARIMA model was the best of the six models in five cities. (3) The hybrid algorithm is a good alternative for the short-term forecasting of $PM_{2.5}$ concentrations, and it can be applied to realistic predictions.

However, there remain problems in the study: (1) the $PM_{2.5}$ concentration is highly dependent on the spatial and temporal scale. To better explain the changes in the $PM_{2.5}$ concentration, we propose incorporating the hourly value of the $PM_{2.5}$ concentrations into future predictions because hourly changes in $PM_{2.5}$ concentrations are closely related to our daily lives, which is important to protecting the health of residents (Samiksha et al., 2017). (2) This model can only accurately reflect the short-term change of $PM_{2.5}$ concentrations, and cannot capture sudden changes in the concentration of pollutants caused by meteorological and emission sources, which is not suitable for long-term prediction problems. If we apply this model to long-term prediction, forecasting accuracy cannot be guaranteed. (3) The results of the study showed that the Wavelet-ARIMA model is superior to the Wavelet-ANN model in forecasting results in China, which is contrary to our previous knowledge from existing studies. According to the analysis by Ni et al. (2017), the ANN model was applicable to systems with large-scale data and complex structures. The larger the amount of original data and the more types of data, the more accurate will be the prediction. For the 1D $PM_{2.5}$ time series processed in this paper, the data volume and types are small. However, the ARIMA model considers more the influence of residual sequences on the prediction results, which is an advantage for single 1D data prediction. (4) When well designed, the hybrid algorithm proposed in this paper is a good alternative for the short-term forecasting of other air pollutants and other cities.

Declaration of interest statement

The authors declared that they have no conflicts of interest to this work. We declare that we do not have any commercial or associative interest that represents a conflict of interest in connection with the work submitted.

Acknowledgements

This work was supported by the National Natural Science Foundation of China (41871193) and Natural Science Foundation of Shanxi Province, China (201601D102055). Meanwhile,We are grateful to the reviewers for providing us with the reference $PM_{2.5}$ data and many useful suggestions.

References

Balabin, R.M., Lomakina, E.I., 2011. Support vector machine regression (svr/ls-svm)—an alternative to neural networks (ann) for analytical chemistry? comparison of nonlinear methods on near infrared (nir) spectroscopy data. Analyst 136 (8), 1703.

Cai, S., Wang, Y., Zhao, B., Wang, S., Chang, X., Hao, J., 2017. The impact of the "air pollution prevention and control action plan" on pm2.5 concentrations in jing-jin-ji region during 2012-2020. Sci. Total Environ. 580, 197–209.

Cao, Q., Shen, L., Chen, S.C., Pui, D., 2018. Wrf modeling of pm2.5 remediation by salscs and its clean air flow over beijing terrain. Sci. Total Environ. 626, 134.

Choubin, B., Malekian, A., 2017. Combined gamma and m-test-based ann and arima models for groundwater fluctuation forecasting in semiarid regions. Environmental Earth Sciences 76 (15), 538.

De, G.G., Trizio, L., Di, G.A., Pey, J., Pérez, N., Cusack, M., et al., 2013. Neural network model for the prediction of pm10 daily concentrations in two sites in the western mediterranean. Sci. Total Environ. 463–464 (5), 875–883.

Fann, N., Lamson, A.D., Anenberg, S.C., Wesson, K., Risley, D., Hubbell, B.J., 2012. Estimating the national public health burden associated with exposure to ambient pm2.5 and ozone. Risk Anal. 32 (1), 81–95.

Faris, H., Alkasassbeh, M., Rodan, A., 2014. Artificial neural networks for surface ozone prediction: models and analysis. Pol. J. Environ. Stud. 23 (2).

García Nieto, P.J., Sánchez, L.F., García-Gonzalo, E., Fj, D.C.J., 2018. Pm10 concentration forecasting in the metropolitan area of oviedo (northern Spain) using models based on svm, mlp, varma and arima: a case study. Sci. Total Environ. 621, 753–761.

Garg, N., Sharma, M.K., Parmar, K.S., Soni, K., Singh, R.K., Maji, S., 2016. Comparison of arima and ann approaches in time-series predictions of traffic noise. Noise Control Eng. J. 64 (4), 522–531.

Ke, G., Junfei, Q., Xiaoli, L., 2018. Highly efficient picture-based prediction of pm2.5 concentration. In: IEEE Transactions on Industrial Electronics, pp. 1.

He, J., Zhang, Y., Wang, K., Chen, Y., Leung, L.R., Fan, J., et al., 2017. Multi-year application of wrf-cam5 over east asia-part i: comprehensive evaluation and formation regimes of o3 and pm2.5. Atmos. Environ. 165.

Jian, L., Zhao, Y., Zhu, Y.P., Zhang, M.B., Bertolatti, D., 2012. An application of arima model to predict submicron particle concentrations from meteorological factors at a busy roadside in hangzhou, China. Sci. Total Environ. 426 (2), 336–345.

Li, W., Raskin, R., Goodchild, M., 2012. Semantic similarity measurement based on knowledge mining: an artificial neural net approach. Int. J. Geogr. Inf. Sci. 26 (8), 1415–1435.

Long, L.I., Lei, M.A., Jianfeng, H.E., Shao, D., Sanli, Y.I., Xiang, Y., et al., 2014. Pm2. 5 concentration prediction model of least squares support vector machine based on feature vector. J. Comput. Appl. 34 (8), 2212–2216.

Lv, B., Hu, Y., Chang, H.H., Russell, A.G., Bai, Y., 2016. Improving the accuracy of daily pm2.5 distributions derived from the fusion of ground-level measurements with aerosol optical depth observations, a case study in north China. Environ. Sci. Technol. 50 (9), 4752.

Maji, K.J., Dikshit, A.K., Arora, M., Deshpande, A., 2017. Estimating premature mortality attributable to pm2.5 exposure and benefit of air pollution control policies in China for 2020. Sci. Total Environ. 612, 683.

Ni, X.Y., Huang, H., Du, W.P., 2017. Relevance analysis and short-term prediction of pm2.5 concentrations in beijing based on multi-source data. Atmos. Environ. 150, 146–161.

Park, S., Kim, M., Kim, M., Namgung, H.G., Kim, K.T., Cho, K.H., et al., 2017. Predicting pm10 concentration in seoul metropolitan subway stations using artificial neural network (ann). J. Hazard Mater. 341, 75–82.

Perez, P., Gramsch, E., 2015. Forecasting hourly pm2.5 in santiago de Chile with emphasis on night episodes. Atmos. Environ. 124, 22–27.

Ping, J., Dong, Q., Li, P., 2017. A novel hybrid strategy for pm 2.5, concentration analysis and prediction. J. Environ. Manag. 196, 443–457.

Ping, W., Yong, L., Qin, Z., Zhang, G., 2015. A novel hybrid forecasting model for pm 10, and so 2, daily concentrations. Sci. Total Environ. 505 (505C), 1202–1212.

Weimiao, Q., Jing, C., Juncai, H., Xinghong, C., Xiaomin, W., 2017. Research on a non-linear forecast method based on wrf-cmaq model. Environ. Sci. Technol.

Riondato, M., Upfal, E., 2015. VC-dimension and rademacher averages: from statistical learning theory to sampling algorithms. In: Acm Sigkdd International Conference on Knowledge Discovery & Data Mining. ACM.

Samiksha, S., Sunder, R.R., Nirmalkar, J., Kumar, S., Sirvaiya, R., 2017. Pm10 and pm2.5 chemical source profiles with optical attenuation and health risk indicators of paved and unpaved road dust in bhopal, India. Environ. Pollut. 222, 477–485.

Sun, W., Sun, J., 2016. Daily pm2.5 concentration prediction based on principal component analysis and lssvm optimized by cuckoo search algorithm. J. Environ. Manag. 188, 144.

Tai, A.P.K., Mickley, L.J., Jacob, D.J., 2010. Correlations between fine particulate matter (pm2.5) and meteorological variables in the United States: implications for the sensitivity of pm2.5 to climate change. Atmos. Environ. 44 (32), 3976–3984.

Tao, W., Yang, K.L., Guo, X.L., Hui, F.U., 2013. Comparative study of anfis and ann applied to freeze-up water temperature forecasting. J. Hydraul. Eng. 44 (07), 842–847.

Wang, Y., Wang, C., Shi, C., Xiao, B., 2018. Short-term cloud coverage prediction using the arima time series model. Remote Sensing Letters 9 (3), 275–284.

Whiteman, C.D., Hoch, S.W., Horel, J.D., Charland, A., 2014. Relationship between particulate air pollution and meteorological variables in Utah's salt lake valley. Atmos. Environ. 94, 742–753.

Wu, J., Zhang, P., Yi, H., Qin, Z., 2016. What causes haze pollution? an empirical study of pm2.5 concentrations in Chinese cities. Sustainability 8 (2), 132.

Xue, D., Liu, Q., 2014. Prediction of surface so2 concentration in shanghai using artificial neural network. Appl. Mech. Mater. 522–524, 44–47.

Yan, D., Lei, Y., Shi, Y., Zhu, Q., Li, L., Zhang, Z., 2018. Evolution of the spatiotemporal pattern of pm2.5 concentrations in China – a case study from the beijing-tianjin-hebei region. Atmos. Environ. 183.

Yan, J., Lai, C.H., Lung, S.C., Chen, C., Wang, W.C., Huang, P.I., et al., 2017. Industrial pm2.5 cause pulmonary adverse effect through rhoa/rock pathway. Sci. Total Environ. 599–600, 1658–1666.

Zafra, C., Ángel, Y., Torres, E., 2017. Arima analysis of the effect of land surface coverage on pm 10, concentrations in a high-altitude megacity. Atmospheric Pollution Research 8 (1), 1–12.

Zhan, Y., Luo, Y., Deng, X., Grieneisen, M.L., Zhang, M., Di, B., 2017. Spatiotemporal prediction of daily ambient ozone levels across China using random forest for human exposure assessment. Environ. Pollut. 233, 464.

Zhou, J., Xing, Z., Deng, J., Du, K., 2016. Characterizing and sourcing ambient pm 2.5, over key emission regions in China i: water-soluble ions and carbonaceous fractions. Atmos. Environ. 135, 20–30.