Research Article

Atmospheric PM$_{2.5}$ Concentration Prediction Based on Time Series and Interactive Multiple Model Approach

Jihan Li,[1] Xiaoli Li [ID],[1,2] and Kang Wang [ID][1]

[1]*Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China*
[2]*Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing Laboratory for Urban Mass Transit, Beijing 100124, China*

Correspondence should be addressed to Xiaoli Li; lixiaolibjut@bjut.edu.cn

Urbanization, industrialization, and regional economic integration have developed rapidly in China in recent years. Air pollution has attracted more and more attention. However, PM$_{2.5}$ is the main particulate matter in air pollution. Therefore, how to predict PM$_{2.5}$ accurately and effectively has become a concern of experts and scholars. For the problem, atmosphere PM$_{2.5}$ concentration prediction algorithm is proposed based on time series and interactive multiple model in this paper. PM$_{2.5}$ concentration is collected by using the monitor at different air quality levels. The time series models are established by historical PM$_{2.5}$ concentration data, which were given by the autoregressive model (AR). In the paper, three PM$_{2.5}$ time series models are established for three different air quality levels. Then, the three models are converted to state equation, respectively, by autoregressive integrated with Kalman filter (AR-Kalman) approaches. Besides, the proposed interactive multiple model (IMM) algorithm is, respectively, compared with autoregressive (AR) model algorithm and AR-Kalman prediction algorithm. It is turned out the proposed IMM algorithm is more accurate than the other two approaches for PM$_{2.5}$ prediction, and it is effective.

1. Introduction

Air pollution is composed of harmful gases and particulate matter. PM$_{2.5}$ is one kind of the particulate matter, and it is one of the main indicators affecting air quality. The air pollution statuses are not only the key issues in the scientific research but also the hot social issues of the public's life. Therefore, many experts and scholars at home and abroad have done a lot of research on PM$_{2.5}$ concentration prediction. A number of prediction methods [1] have been widely developed, such as principal component analysis [2], regression analysis [3, 4], neural network [5–8], genetic algorithm, time series [9, 10], and Kalman filter [11]. Azid et al. used principal component analysis (PCA) [12] to analyse the major components affecting air quality and to predict the air pollutant concentration by the predictive ability of neural network [13]. However, the disadvantage of this method is easily trapped in the local minimum, so the accuracy of air quality prediction will be affected. Zhou

et al. propose a hybrid model method of combining empirical model decomposition with neural network for predicting PM$_{2.5}$ concentration [14]. But the EEMD method [15] needs a large amount of calculation and time and can retain residual noise in decomposition process. Ping et al. found a hybrid strategy used to predict the PM$_{2.5}$ concentration in Beijing, Tianjin, and Hebei [16]. However, the selection of parameters and the training time may be inaccurate. Wang et al. used the autoregressive integrated moving average model to establish the time series of the PM$_{2.5}$ concentration, and the noise deviation is trained by SVM [17]. But this algorithm is simple and robust and it has some disadvantages such as needing a lot of computing time and equipment requirements. Voukantsis and Dagoumas used the principal component analysis and neural network [8] for comparing the air quality in Thessaloniki and Helsinki [18]. Michano Michanowicz et al. used the AERM-OD (AMS/EPA regulatory model) model to predict the PM$_{2.5}$ concentration [19]. Many experts and scholars

have proposed the Kalman filtering algorithm [20, 21] to predict $PM_{2.5}$ concentration. In the abovementioned algorithms, there are still great limitations in accuracy and effectiveness because these algorithms predict $PM_{2.5}$ concentration mostly for a certain period of time or a single model. In different weather conditions, atmospheric $PM_{2.5}$ concentration levels are different. The single model is only suitable for $PM_{2.5}$ concentration prediction of a certain weather quality. Once the weather changes, the $PM_{2.5}$ concentration level is also changed. The single model will produce errors in $PM_{2.5}$ prediction and cannot be relied on for accurate prediction. And the multiple model algorithm is widely used in the field of control [22, 23] and other fields [24–26], but not in the field of weather prediction. Thus, the multiple model algorithm under different weather conditions is selected to reduce the errors and make the prediction more accurate.

Different weather conditions are considered to predict $PM_{2.5}$ for the above method. Therefore, based on multiple model theory [27–29] and interactive multiple model algorithm, in this paper, the main objective of this study is to accurately predict $PM_{2.5}$ concentration according to different air quality levels. A feasible and effective method is provided for air pollution prediction.

2. Materials and Methods

2.1. Data Collection.
In order to prove the prediction accuracy of the proposed method, the $PM_{2.5}$ concentration data were provided by the monitoring station in Beijing. The monitoring station is shown in Figure 1. The data were collected at different air quality levels from Sept. 29th to Oct. 1st, Mar. 21st to 23rd, and Oct. 14th to 16th in 2018. The information sequence, which includes $PM_{2.5}$, NO_2, SO_2, CO, PM_{10}, and O_3 concentrations ($\mu g/m^3$), temperature, and humidity, is used to predict the performance of the proposed methods. Then, the time series model is established and combined with the proposed method to predict the $PM_{2.5}$ concentration.

Figure 2 shows the monitoring area. The red circle represents a monitoring point in the figure.

Figure 3 shows the air quality in three months, and different colors represent different air quality levels. The green color represents air quality level is excellent, yellow represents air quality level is good, and orange represents air quality level is slightly polluted. Specific air quality assessment standard is shown in Table 1.

2.2. Time Series Model.
Prediction methods of $PM_{2.5}$ concentration mainly include numerical model method and statistical method. Time series analysis method belongs to the statistical category, and it can be used to extract the relevant information. Thus, the structure of time series and pattern can be used to deduce the future change trend of the system. The time series model includes autoregressive (AR) model, moving average (MA) model, and autoregressive and moving average (ARMA) model. Since the 20th century, various parameterization methods such as linear regression,

FIGURE 1: Air pollution monitor (1 is the device monitoring probe, 2 is the equipment power supply box, and 3 is the solar electronic board).
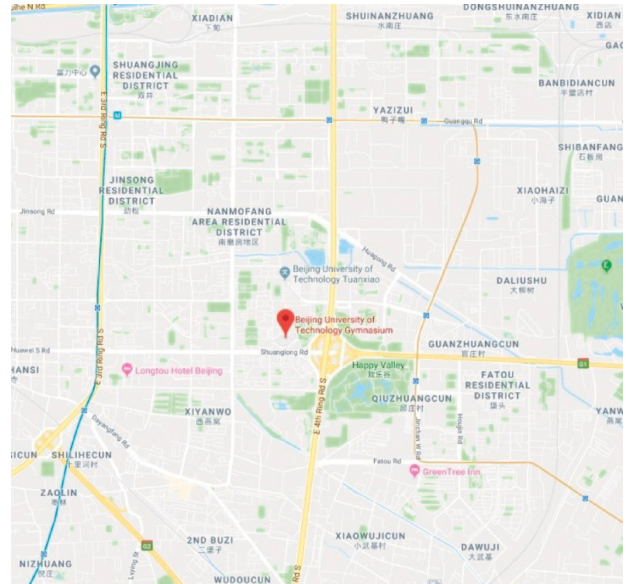

FIGURE 2: $PM_{2.5}$ monitoring area.

smoothing method, and autoregressive process have been widely used in weather forecasting and have achieved good results. In this paper, AR(p) model is given as follows:

$$X_n = \varphi_1 X_{n-1} + \varphi_2 X_{n-2} + \cdots + \varphi_p X_{n-p} + \varepsilon_n, \qquad (1)$$

where $\varphi_1, \varphi_2, \ldots, \varphi_p$ are the autoregressive coefficient and ε_n is the independent sequence of identically distributed random variable, and it is white noise. $\{X_n, n = 0, \pm1, \pm2, \ldots\}$ is the p-order autoregressive model AR(p).
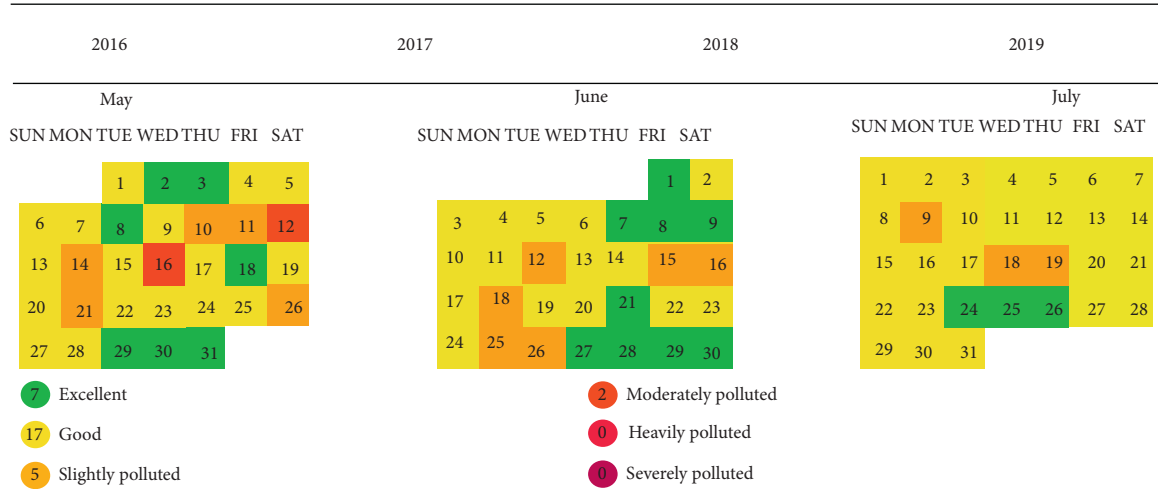
FIGURE 3: Air quality monitoring.

TABLE 1: Air quality standard based on ambient air quality index (AQI).

AQI	Air quality index level	Air quality index category	Health implications
0~50	Level 1	Excellent	No air pollution
50~100	Level 2	Good	Air quality is acceptable
100~150	Level 3	Slight pollution	Symptoms of irritation appear
150~200	Level 4	Moderate pollution	Respiratory system is affected
200~300	Level 5	Heavy pollution	Symptoms of patients increase
>300	Level 6	Severe pollution	Diseases appear

2.3. Autoregressive-Kalman Filter Method. For complex atmospheric dynamic systems, traditional methods are difficult to measure every predicted variable. But the Kalman filter method can directly use the limited and indirect measurement information to estimate the missing information and predict the future change trend of the atmospheric dynamic system. Because the collected $PM_{2.5}$ concentration data are time series, $PM_{2.5}$ concentration time series model and Kalman filtering approach are combined in the paper. The AR model was firstly established, and then, the AR model is transformed into state equation for Kalman filtering. The state space model of $PM_{2.5}$ concentration at different air quality levels is written as follows:

$$X(k+1) = A(k+1,k)X(k) + \Gamma(k+1,k)W(k),$$
$$Y(k+1) = H(k+1)X(k+1) + V(k+1), \quad (k = 0, 1, \ldots),$$
$$(2)$$

where $X(k+1)$ is the n-dimensional state vector, $Y(k+1)$ is the m-dimensional observation vector, A and H, are, respectively, $m \times n$-dimensional state transition matrix and observation matrix, and $W(k)$ and $V(k)$ represent process and measurement noises, respectively. The state equation in the AR-Kalman method [26] is given as follows:

$$
\begin{bmatrix}
X_1(k+1) \\
X_2(k+1) \\
X_3(k+1) \\
\vdots \\
X_{n-1}(k+1) \\
X_n(k+1)
\end{bmatrix}
=
\begin{bmatrix}
\varphi_1 & \varphi_2 & \varphi_3 & \cdots & \varphi_{p-2} & \varphi_{p-1} & \varphi_p \\
1 & 0 & 0 & \cdots & 0 & 0 & 0 \\
0 & 1 & 0 & \cdots & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & \cdots & 1 & 0 & 0 \\
0 & 0 & 0 & \cdots & 0 & 1 & 0
\end{bmatrix}
\begin{bmatrix}
X_1(k) \\
X_2(k) \\
X_3(k) \\
\vdots \\
X_{n-1}(k) \\
X_n(k)
\end{bmatrix}
+
\begin{bmatrix}
1 \\
0 \\
0 \\
\vdots \\
0 \\
0
\end{bmatrix}.
\tag{3}
$$

Meanwhile, the measurement equation is expressed as

$$Y(k+1) = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \end{bmatrix} \begin{bmatrix} X_1(k+1) \\ X_2(k+1) \\ X_3(k+1) \\ \vdots \\ X_n(k+1) \end{bmatrix} + V(k+1). \quad (4)$$

According to the Technical Regulation on Ambient Air Quality Index, the air pollution index is divided into different levels (excellent, good, mild pollution, and so on). In this paper, the three kinds of models at different air quality levels are only used for research and simulation.

2.4. Interactive Multiple Model Approach.

In different weather conditions, atmospheric $PM_{2.5}$ concentration levels are different. The interactive multiple model (IMM), which uses two or more models to describe the possible states in the process of work, is a soft switching algorithm and estimates the state of the system through effective weighted fusion. When $PM_{2.5}$ concentration changes, the different models can switch to the corresponding $PM_{2.5}$ concentration model by the interactive multiple model method in this paper. Therefore, the IMM model has strong robustness. IMM is used to achieve the initial condition of a specific model matching filter by mixing all the state estimations generated from the filter at the current moment under the assumption condition. Each model is performed by a standard Kalman Filter [30]; then, all filters generate updated state estimates weighted combination. At last, the results of state and co-variance estimates are finally obtained. At time k, when the target state estimation is calculated by the interactive multiple model method [31–33], all filters generate updated state estimates weighted combination. Thus, the time series models of $PM_{2.5}$ concentration are established at different air quality levels in this paper. The model at different air quality levels can be represented as follows:

$$M = \{M_1, M_2, \ldots, M_j\}, \quad j = 1, 2, 3. \quad (5)$$

Each model M_j has a prior probability. It is given as follows:

$$\mu_j(0) = P(M_j(0)). \quad (6)$$

The transformation probability from model i to model j is denoted as follows:

$$\pi_{ij} = \Pr\{M_j(k+1) \mid M_i(k)\}. \quad (7)$$

Next, we need to calculate π_{ij}. In the random process,

$$P(X_{n+1=j} \mid X_n = i_n, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = P(X_{n+1} = j \mid X_n = i),$$
$$i, j \in E, \ E = \{0, 1, \ldots\}, \quad (8)$$

we set $E = \{1, 2, 3\}$. E is a set of three $PM_{2.5}$ concentration levels. The levels 1, 2, and 3 represent excellent, good, and slight pollution. $\{X_n, n = 0, 1, 2, \ldots\}$ is a Markova chain of

discrete time parameters of state space transition probability, and it is given as follows:

$$P(X_{m+n} = j \mid X_m = i), \quad (i, j = 1, 2, \ldots, m). \quad (9)$$

From (9), the air quality level can be approximately considered as Markov chain. In this paper, the Markov state transition probability matrix can be obtained as follows:

$$p = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}. \quad (10)$$

The flow chart of IMM is shown in Figure 4:

In Figure 4, $\hat{X}_n(k)$ represents the initial value of the state of each model before input interaction. $\hat{P}_n(k)$ represents the initial covariance of each model after input interaction. $\hat{X}_n(k)^*$ represents the initial value of the state of each model. $\hat{P}_n(k)^*$ represents the initial covariance of each model. $\hat{X}_n(k+1)$ and $\hat{P}_n(k+1)$ are the updated state vector and covariance matrix of each model. $\hat{X}(k+1)$ is the prediction value. $\hat{P}(k+1)$ is estimation covariance matrix; $n = 1, 2, 3, \ldots$.

In order to apply the IMM algorithm, for different air quality levels, we establish different $PM_{2.5}$ time series models. When air quality grades are excellent, good, and slightly polluted, respectively, the corresponding $PM_{2.5}$ concentration models are model 1, model 2, and model 3, respectively. When the time series model is combined with Kalman filter, according to the three different air quality levels that are mentioned above, model 4, model 5, and model 6 are established, respectively. The state space of the models is written as follows:

$$X_j(k+1) = \Phi_j(k)X_j(k) + G_j(k)W_j(k), \quad (11)$$

where X_j is the state vector. The target state equation is represented by model j, $i, j \in [1, 2, 3]$. The measurement equation is given as follows:

$$Y_j(k+1) = H_j(k+1)X_j(k+1) + V_j(k+1), \quad (12)$$

where $\Phi_j(k)$ is the state transition matrix of model M_j, $W_j(k)$ is the white noise, $W_j \in (0, Q_j)$, $V_j(k)$ is the observed noise sequence with zero mean and covariance matrix is Q_j, and $V \in N(0, Q_j)$. $H_j(k+1)$ is the observation matrix of model M_j. G_j is the noise correlation matrix of model M_j. There are four steps in the IMM algorithm, which are reinitialization, model filtering, probability updation, and prediction. In this paper, we use homogeneous Markov chain to achieve the transformation between different models and calculate the transition probability matrix using submodel predictions. The specific steps of the IMM algorithm are as follows:

Step 1 (reinitialization). The prediction probability of the model from k time to time $k+1$ is performed as follows:

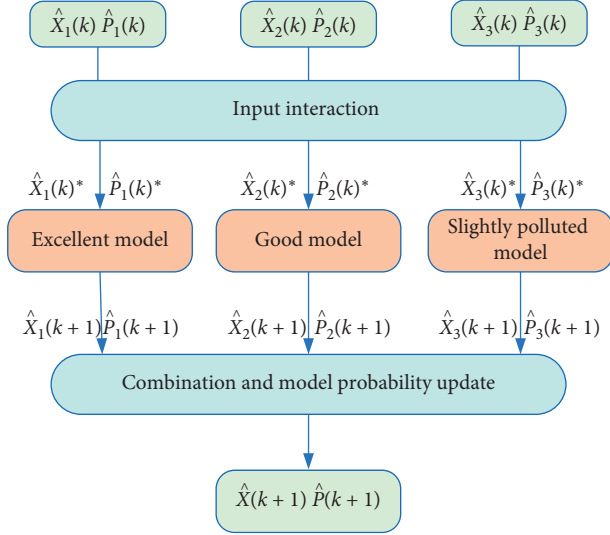$$\hat{\mu}_j = P(M_j(k) \mid Y(k)) = \sum_{i=1}^{r} \pi_{ij}\mu_i(k), \quad r = 3, \quad (13)$$

FIGURE 4: Flow chart of IMM.

where $M_j(k)$ is the selected model at time k and $Y(k)$ is the measured value at time k. Based on the model probability, the estimated results of each filter are weighted and merged. The probability of model j is expressed as follows:

$$\mu_j(k+1) = \frac{\hat{\mu} \cdot \Lambda_j(k+1)}{\sum_{j=1}^{r} \hat{\mu}_j \Lambda_j(k+1)}. \tag{14}$$

Each Kalman filter is initialized as follows:

$$\hat{X}_j(k \mid k) = E[X(k) \mid M_i(k+1), Y(k+1)]$$
$$= \sum_{i=1}^{r} \hat{X}_i(k \mid k)\mu_{ij}(k),$$
$$\hat{P}_j(k \mid k) = \sum_{i}^{3} \left[P_i(k \mid k) + \left(\hat{X}_i(k \mid k) - \hat{X}_j(k \mid k) \right) \right. \tag{15}$$
$$\left. \cdot \left(\hat{X}_i(k \mid k) - \hat{X}_j(k \mid k) \right)^T \right] \cdot \mu_{ij}(k),$$

where $\hat{P}_j(k \mid k)$ is the reinitialized state vector at time $k+1$, μ_j is the normalized constant for model j, $\hat{X}_i(k \mid k)$ is the corresponding covariance state estimation of the j filter, $i, j = 1, 2, \ldots, r$, and $\mu_{ij}(k)$ is the initial weight of each filter.

Step 2 (model filtering)

$$\hat{X}_j(k+1 \mid k) = \Phi_j(k)\hat{X}_j(k \mid k), \tag{16}$$

where $\hat{X}_j(k \mid k)$ is the reinitialization state at time k. The covariance matrix is as follows:

$$P_j(k+1 \mid k) = \Phi_j P_j(k \mid k)\Phi_j(k) + G_j(k)Q_j(k)G_j(k). \tag{17}$$

The Kalman gain is as follows:

$$K_j(k+1) = P_j(k+1 \mid k) \cdot \left(H_j(k+1) \right)^T \cdot \left(S_j(k+1) \right)^{-1}. \tag{18}$$

Covariance matrix can be obtained as follows:

$$S_j(k+1) = H_j(k+1) \cdot P_j(k+1 \mid k) \cdot \left(H_j(k+1) \right)^T + Q_j(k+1). \tag{19}$$

State estimation updates $\hat{X}_j(k+1 \mid k+1)$ and $P_j(k+1 \mid k+1)$ are given as

$$\hat{X}_j(k+1 \mid k+1) = \hat{X}_j(k+1 \mid k+1) + K_j(k+1) \cdot v_j(k+1),$$
$$P_j(k+1 \mid k+1) = P_j(k+1 \mid k) - K_j(k+1) \cdot S_j(k+1)$$
$$\cdot \left(K_j(k+1) \right)^T. \tag{20}$$

The residual is as follows:

$$v_j(k+1) = Y(k+1) - H_j(k+1) \cdot \hat{X}_j(k+1 \mid k). \tag{21}$$

Step 3 (probability updating). Likelihood function $\Lambda_j(k+1)$ is designed as follows:

$$\Lambda_j(k+1) = \frac{1}{\left| 2\pi S_j(k+1) \right|}$$
$$\cdot \exp\left\{ -\frac{1}{2} v_j(k+1)^T \cdot S_j^{-1}(k+1) v_j(k+1) \right\}. \tag{22}$$

Initialization of the weight of each Kalman filter in IMM algorithm is denoted as follows:

$$\mu_{ij}(k) = P\left(M_i(k) \mid M_j(k+1), Y(k) \right) = \frac{\pi_{ij}\mu_i(k)}{\hat{\mu}_j}, \tag{23}$$

where π_{ij} denotes the transition probability from model i to model j. $\mu_i(k)$ is model probability vector.

Step 4 (prediction). Prediction results are obtained by the update probability of each model and hybrid prediction model can be computed as follows:

$$\hat{X}(k+1 \mid k+1) = \sum_{j=1}^{r} \hat{X}_j(k+1 \mid k+1) \cdot \mu_j(k+1), \quad r = 3. \tag{24}$$

Estimation covariance matrix is given as follows:

$$P^{\text{IMM}}(k+1 \mid k+1) = \sum_{j=1}^{r} \left[P_j(k+1 \mid k+1) \right.$$
$$+ \left(\hat{X}(k+1 \mid k+1) - \hat{X}_j(k+1 \mid k+1) \right)$$
$$\cdot \left(\hat{X}(k+1 \mid k+1) - \hat{X}_j(k+1 \mid k+1) \right)^T \right]$$
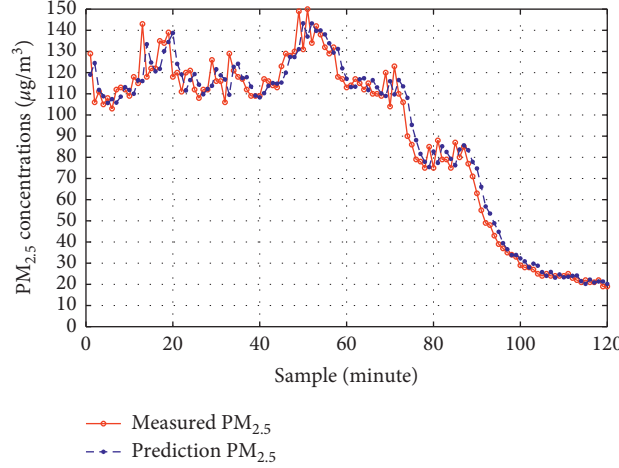$$\left. \cdot \mu_j(k+1), \quad r = 3. \right. \tag{25}$$

FIGURE 5: PM$_{2.5}$ concentration prediction result of model 1 (AR).

3. Simulation Analysis and Prediction

To verify the feasibility and effectiveness of the proposed IMM algorithm, some simulations are made to predict PM$_{2.5}$ concentration. The PM$_{2.5}$ concentration data from Sept. 29th 00 : 00 to Oct. 1st 00 : 00, Mar. 21st 00 : 00 to 23th 00 : 00, and Oct. 14th 00 : 00 to 16th 00 : 00 in 2018 (sampled every ten minutes) are selected as the experimental objects.

PM$_{2.5}$ concentration in the atmosphere always changes with the change of environment. For this problem, different models are transferred by the Markov chain. Thus, in the experimental study, the transfer is determined by the Markov probability transfer matrix. The Markov transition probability is given as follows:

$$\pi = \begin{bmatrix} 0.963 & 0.037 & 0 \\ 0 & 0.9444 & 0.0556 \\ 0 & 0 & 1 \end{bmatrix}. \tag{26}$$

Initial state probability is $\mu(0) = 0.33$. Initial value $X(0) = E[X(0)]$. Initial covariance $P(0) = E[(X(0) - E[X(0)])(X(0) - E[X(0)])^T]$.

The state space equation of model 1 is given as follows:

$$\begin{bmatrix} X_1(k+1) \\ X_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.6946 & 0.3054 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X_1(k) \\ X_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e(k+1). \tag{27}$$

The state space equation of model 2 is given as follows:

$$\begin{bmatrix} X_1(k+1) \\ X_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.931 & 0.069 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X_1(k) \\ X_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e(k+1). \tag{28}$$

The state space equation of model 3 is given as follows:

$$\begin{bmatrix} X_1(k+1) \\ X_2(k+1) \end{bmatrix} = \begin{bmatrix} 0.551 & 0.449 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} X_1(k) \\ X_2(k) \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} e(k+1). \tag{29}$$

The model 1 can predict PM$_{2.5}$ concentration at 0~50; the predictive effect is shown in Figure 5. Since model 1 is established based on data with PM$_{2.5}$ concentrations between 0 and 50, when the PM$_{2.5}$ concentration is more than 50, the predictive effect of the model 1 cannot truly reflect prediction result at other air quality levels.

In Figure 6, model 2 is used to predict PM$_{2.5}$ concentration. It shows that the model has a small error when the PM$_{2.5}$ concentration is between 50 and 100, and it cannot predict PM$_{2.5}$ concentration at other air quality levels. It is similar to model 1: since the model 2 is established based on data with PM$_{2.5}$ concentrations between 50 and 100, the predictive effect of the model 2 cannot truly reflect prediction result at other air quality levels.

In Figure 7, model 3 is used to predict PM$_{2.5}$ concentrations. From Figure 7, we can see that when PM$_{2.5}$ concentrations are between 100 and 150, the model can well predict trend of PM$_{2.5}$ concentration. It is similar to model 1 and model 2: model 3 can only be used to predict PM$_{2.5}$ concentrations from 100 to 150.

In Figure 8, PM$_{2.5}$ concentration is predicted by model 4 (AR-Kalman). When PM$_{2.5}$ concentration is between 0 and 50, prediction error is obviously small. Compared with the predictive effect of model 1, the predictive effect of model 4 is better. It is similar to model 1: model 4 can only be used to predict PM$_{2.5}$ concentrations from 0 to 50.

In Figure 9, model 5 is used to predict PM$_{2.5}$ concentrations. From Figure 9, we can see that the prediction error of PM$_{2.5}$ concentration is smaller than that of model 2 when the PM$_{2.5}$ concentration is between 50 and 100.

In Figure 10, the PM$_{2.5}$ concentration is predicted by model 6. It can be clearly seen that compared with model 3, model 6 has a higher prediction accuracy than model 3.

In Figure 11, we adopt the multiple model method to predict PM$_{2.5}$ concentration. From Figure 11, we can see that the IMM method can be used to predict PM$_{2.5}$ concentration values at all air quality levels effectively. Besides, compared with the single model, the IMM method has a better predictive effect. In order to more intuitively compare the prediction effects among various models, the performance
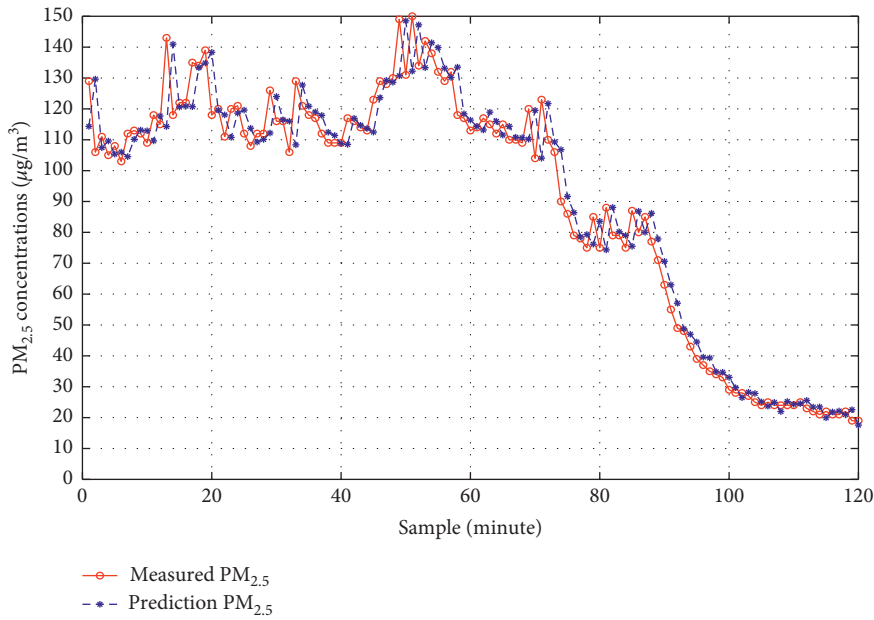
FIGURE 6: PM$_{2.5}$ concentration prediction result of model 2 (AR).
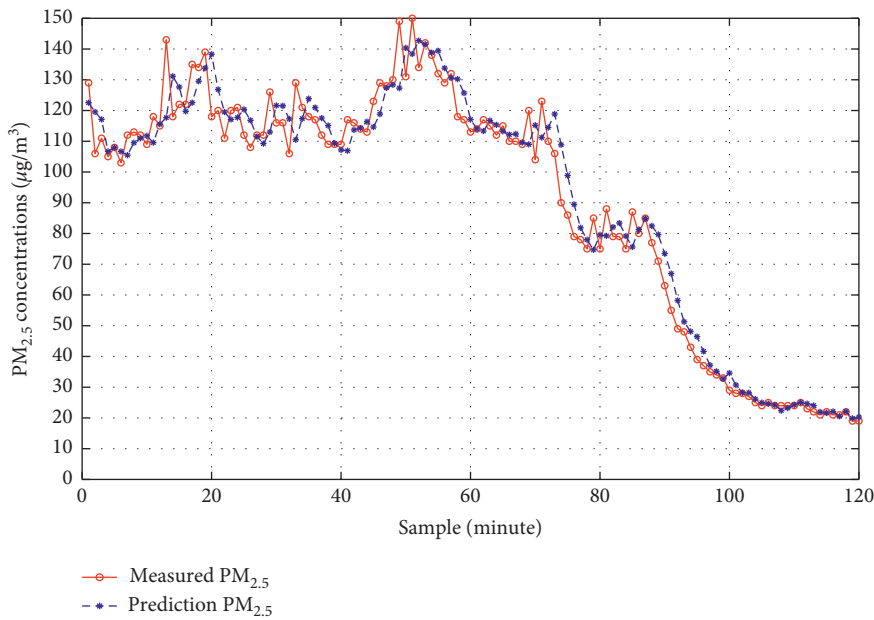

FIGURE 7: PM$_{2.5}$ concentration prediction result of model 3 (AR).

indicators of each prediction model are given in Table 2. Table 2 shows that the prediction error of AR model is significantly larger than that of the IMM method. In addition, Figures 5–7 also indicate that the predictive result of the single model is effective. The prediction performance analysis of AR-Kalman model and AR model is given in Table 2. From Table 2, we can see that the AR-Kalman model is better than the AR method. However, from Figures 8 to 10, the AR-Kalman model is only applicable to predict PM$_{2.5}$ concentration at the corresponding air quality level and it cannot be used to predict accurately PM$_{2.5}$ concentration at different air quality levels. In Table 2, when the air quality level is excellent, the prediction error of IMM is more accurate than the single model (AR-Kalman). The single model can only obtain higher prediction accuracy under the corresponding air quality conditions. However, IMM can be used to accurately predict PM$_{2.5}$ concentration at different air quality levels.
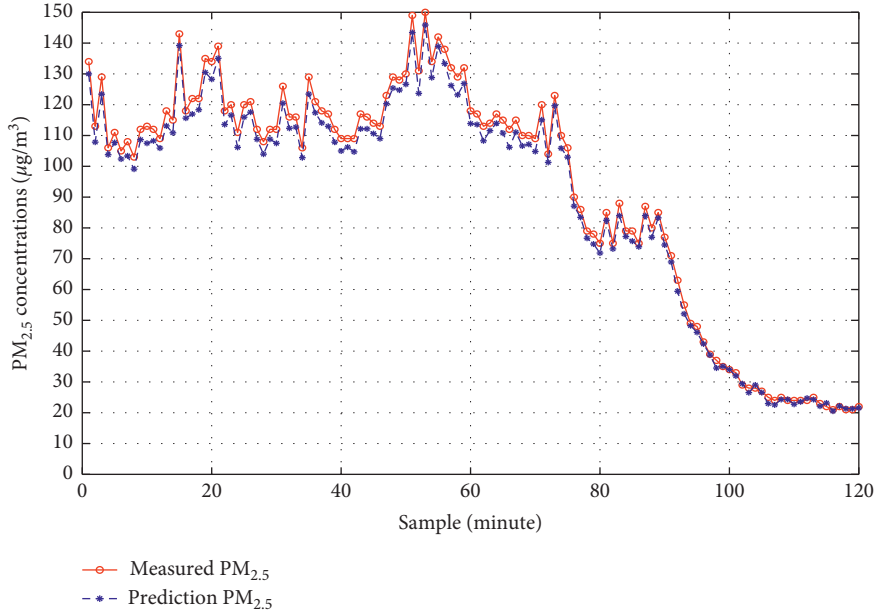
FIGURE 8: PM$_{2.5}$ concentration prediction result of model 4 (AR-Kalman).
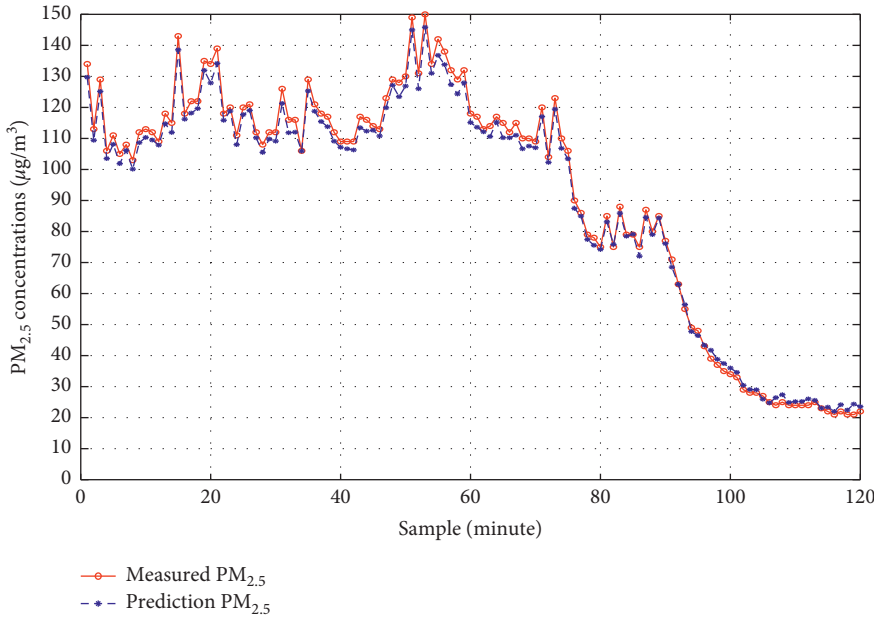

FIGURE 9: PM$_{2.5}$ concentration prediction result of model 5 (AR-Kalman).

The performance indicators in the above table can be given by the following metrics. The mean absolute error (MAE) is defined as follows:

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} \left| \hat{X}_i - \bar{X} \right|, \qquad (30)$$

where X_i is the observation, $\hat{X}_i$ is the predictive value, $\bar{X}$ is the mean value of observed PM$_{2.5}$ concentration, N is the number of points to sample data, and the second indicator is MAPE, which is used to calculate mean absolute percentage error:

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^{N} \frac{\left| \hat{X}_i - X_i \right|}{X_i} \times 100\%. \qquad (31)$$

The root mean square error (RMSE) is defined as follows:

$$\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} \left(\hat{X}_i - X_i \right)^2}. \qquad (32)$$

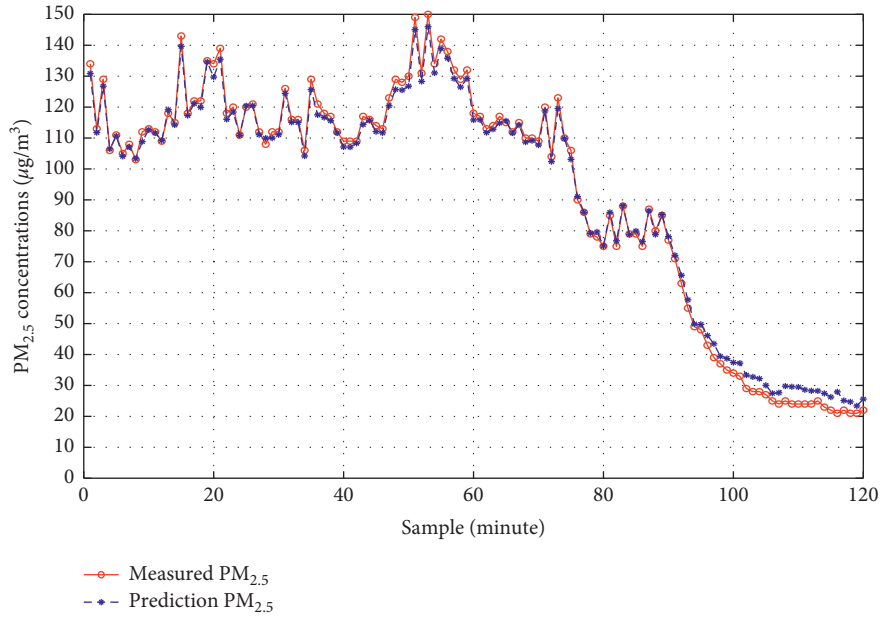However, the mean square error (MSE) is defined as follows:

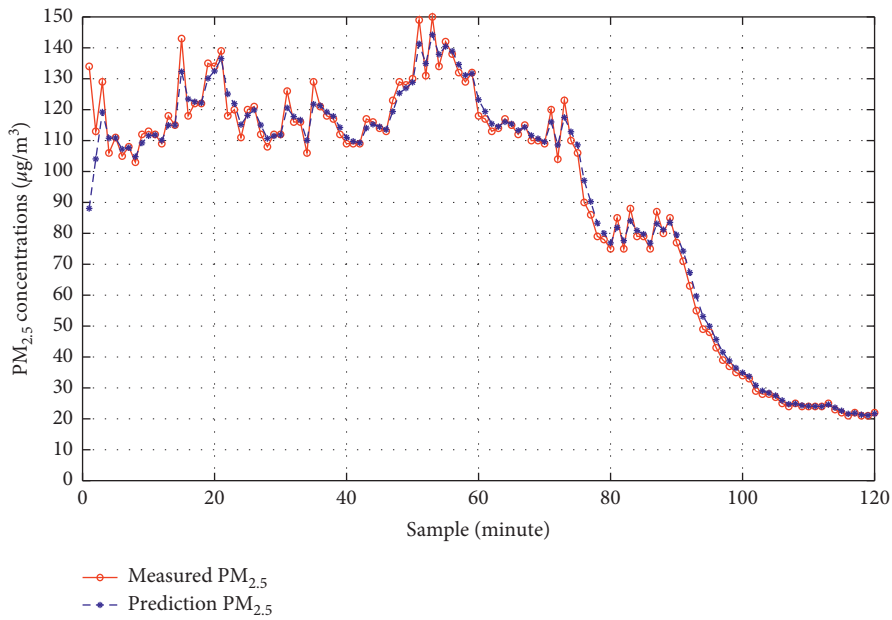FIGURE 10: PM$_{2.5}$ concentration prediction result of model 6 (AR-Kalman).

FIGURE 11: PM$_{2.5}$ concentration prediction result of IMM.

TABLE 2: AR-Kalman model, AR model prediction, and IMM results for PM$_{2.5}$ concentration.

Forecasting model	MAE	MAPE	RMSE	MSE
Excellent (AR)	5.48	6.46	7.47	41.8
Good (AR)	5.88	6.50	8.19	48.5
Slight pollution (AR)	5.46	6.31	7.37	40.7
Excellent (AR-Kalman)	2.87	3.18	3.25	17.81
Good (AR-Kalman)	2.49	3.04	2.82	15.45
Slight pollution (AR-Kalman)	2.16	4.72	2.70	14.82
IMM	0.54	0.55	0.95	5.23

$$\text{MSE} = \frac{1}{N} \sum_{i=1}^{N} \left(\hat{X}_i - X_i \right)^2. \tag{33}$$

The validity and rationality of the proposed method can be proved by analysing and comparing various indicators.

4. Conclusions

In this paper, the time series model (AR) of PM$_{2.5}$ concentration is established at different air quality levels and these models are used to predict PM$_{2.5}$ concentration. The Kalman filter method is introduced by transforming the AR model into the form of equation of state, and the AR-Kalman hybrid prediction method was obtained. The method was used to predict PM$_{2.5}$ concentration at different air quality levels. Then, the method of AR model and AR-Kalman is compared for PM$_{2.5}$ concentration prediction. It is indicated that the AR-Kalman model could predict more accurately than the AR model. Finally, the proposed interactive multiple model method is applied and analysed to predict PM$_{2.5}$ concentration by historical data at different air quality levels. The mean average prediction error is proposed as the evaluation index for the prediction models. The comparison indicates that the interactive multiple model (IMM) could predict more accurately than the single AR model and single AR-Kalman model. For PM$_{2.5}$ concentration, this method has lower prediction error at different air quality levels.

In future work, temperature, pressure, humidity, and other factors in the weather environment are to be considered. Thus, nonlinear the PM$_{2.5}$ model will be built and the dynamic update transition matrix may be considered to predict PM$_{2.5}$ concentration.

Data Availability

The data used to support the findings of this study have not been made available because data sharing is not applicable to this article as no new data were created or analysed in this study.

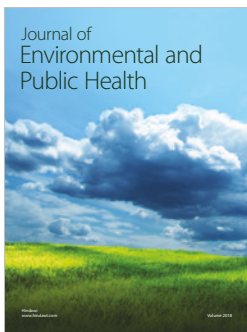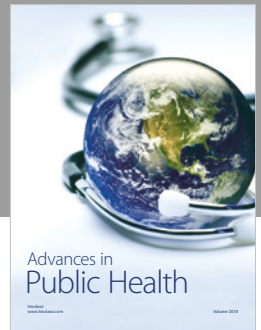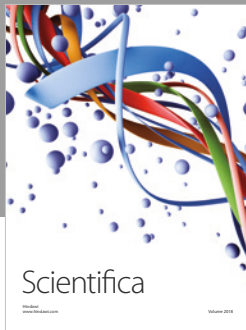## Conflicts of Interest
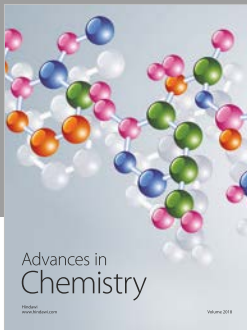
The authors declare no conflicts of interest.

Acknowledgments

References

[1] S.-Y. Kim, C. Olives, L. Sheppard et al., "Historical prediction modeling approach for estimating long-term concentrations of PM$_{2.5}$ in cohort studies before the 1999 implementation of widespread monitoring," *Environmental Health Perspectives*, vol. 125, no. 1, pp. 38–46, 2017.

[2] W. Sun and J. Sun, "Daily PM$_{2.5}$ concentration prediction based on principal component analysis and LSSVM optimized by cuckoo search algorithm," *Journal of Environmental Management*, vol. 188, pp. 144–152, 2017.

[3] Q. Di, P. Koutrakis, and J. Schwartz, "A hybrid prediction model for PM$_{2.5}$ mass and components using a chemical transport model and land use regression," *Atmospheric Environment*, vol. 131, pp. 390–399, 2016.

[4] Y. Guo, Q. Tang, D.-Y. Gong, and Z. Zhang, "Estimating ground-level PM$_{2.5}$ concentrations in Beijing using a satellite-based geographically and temporally weighted regression model," *Remote Sensing of Environment*, vol. 198, pp. 140–149, 2017.

[5] H. Liu, H.-Q. Tian, Y.-F. Li, and L. Zhang, "Comparison of four Adaboost algorithm based artificial neural networks in wind speed predictions," *Energy Conversion and Management*, vol. 92, pp. 67–81, 2015.

[6] A. Kumar, Z.-Z. Hu, B. Jha, and P. Peng, "Estimating ENSO predictability based on multi-model hindcasts," *Climate Dynamics*, vol. 48, no. 1-2, pp. 39–51, 2017.

[7] A. M. Torkashvand, A. Ahmadi, and N. L. Nikravesh, "Prediction of kiwifruit firmness using fruit mineral nutrient concentration by artificial neural network (ANN) and multiple linear regressions (MLR)," *Journal of Integrative Agriculture*, vol. 16, no. 7, pp. 1634–1644, 2017.

[8] I. P. Panapakidis and A. S. Dagoumas, "Day-ahead electricity price forecasting via the application of artificial neural network based models," *Applied Energy*, vol. 172, pp. 132–151, 2016.

[9] R. Adhikari, "A neural network based linear ensemble framework for time series forecasting," *Neurocomputing*, vol. 157, pp. 231–242, 2015.

[10] X. Qiu, Y. Ren, P. N. Suganthan, and G. A. J. Amaratunga, "Empirical mode decomposition based ensemble deep learning for load demand time series forecasting," *Applied Soft Computing*, vol. 54, pp. 246–255, 2017.

[11] L. Zanni, J.-Y. Le Boudec, R. Cherkaoui, and M. Paolone, "A prediction-error covariance estimator for adaptive Kalman filtering in step-varying processes: application to power-system state estimation," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 5, pp. 1683–1697, 2017.

[12] S. Ausati and J. Amanollahi, "Assessing the accuracy of ANFIS, EEMD-GRNN, PCR, and MLR models in predicting PM$_{2.5}$," *Atmospheric Environment*, vol. 142, pp. 465–474, 2016.

[13] A. Azid, H. Juahir, M. E. Toriman et al., "Prediction of the level of air pollution using principal component analysis and artificial neural network techniques: a case study in Malaysia," *Water, Air, & Soil Pollution*, vol. 225, p. 2063, 2014.

[14] Q. Zhou, H. Jiang, J. Wang, and J. Zhou, "A hybrid model for PM$_{2.5}$ forecasting based on ensemble empirical mode decomposition and a general regression neural network," *Science of the Total Environment*, vol. 496, pp. 264–274, 2014.

[15] W.-C. Wang, K.-W. Chau, D.-M. Xu, and X.-Y. Chen, "Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition," *Water Resources Management*, vol. 29, no. 8, pp. 2655–2675, 2015.

[16] P. Jiang, Q. Dong, and P. Li, "A novel hybrid strategy for PM$_{2.5}$ concentration analysis and prediction," *Journal of Environmental Management*, vol. 196, pp. 443–457, 2017.

[17] P. Wang, H. Zhang, Z. Qin, and G. Zhang, "A novel hybrid-Garch model based on ARIMA and SVM for PM$_{2.5}$ concentrations forecasting," *Atmospheric Pollution Research*, vol. 8, no. 5, pp. 850–860, 2017.

[18] D. Voukantsis, K. Karatzas, J. Kukkonen, T. Räsäbnene, A. Karppinen, and M. Kolehmainen, "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," *Science of The Total Environment*, vol. 409, no. 7, pp. 1266–11276, 2011.

[19] D. R. Michanowicz, J. L. C. Shmool, B. J. Tunno et al., "A hybrid land use regression/AERMOD model for predicting intra-urban variation in $PM_{2.5}$," *Atmospheric Environment*, vol. 131, pp. 307–315, 2016.

[20] K. Granstrom, P. Willett, and Y. Bar-Shalom, "Approximate multi-hypothesis multi-Bernoulli multi-object filtering made multi-easy," *IEEE Transactions on Signal Processing*, vol. 64, no. 7, pp. 1784–1797, 2016.

[21] I. Djalalova, L. Delle Monache, and J. Wilczak, "$PM_{2.5}$ analog forecast and Kalman filter post-processing for the Community Multiscale Air Quality (CMAQ) model," *Atmospheric Environment*, vol. 108, pp. 76–87, 2015.

[22] X. L. Li, W. Wang, and W. Sun, "Multi-model adaptive control," *Control and Decision*, vol. 15, pp. 390–394, 2000.

[23] X.-L. Li, D.-X. Liu, C. Jia, and X.-Z. Chen, "Multi-model control of blast furnace burden surface based on fuzzy SVM," *Neurocomputing*, vol. 148, pp. 209–215, 2015.

[24] X.-L. Li, C. Jia, D.-X. Liu, and D. W. Ding, "Nonlinear adaptive control using multiple models and dynamic neural networks," *Neurocomputing*, vol. 136, pp. 190–200, 2014.

[25] X. L. Li, K. Wang, and D. X. Liu, "An improved result of multiple model iterative learning control," *IEEE/CAA Journal of Automatica Sinica*, vol. 1, no. 3, pp. 315–322, 2014.

[26] H. Liu, H.-Q. Tian, and Y.-F. Li, "Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction," *Applied Energy*, vol. 98, pp. 415–424, 2012.

[27] P. Marco, B. Francesco, M. De Felice et al., "Multi-model ensemble for day ahead prediction of photovoltaic power generation," *Solar Energy*, vol. 134, pp. 132–146, 2016.

[28] T. Kim and H.-Y. Jeong, "A novel algorithm for crash detection under general road scenes using crash probabilities and an interactive multiple model particle filter," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2480–2490, 2014.

[29] G. Xie, H. Gao, L. Qian, B. Huang, K. Li, and J. Wang, "Vehicle trajectory prediction by integrating physics- and maneuver-based approaches using interactive multiple models," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 7, pp. 5999–6008, 2018.

[30] X. L. Li, X. F. Zhang, C. Jia, and D. X. Liu, "Multi-model adaptive control based on fuzzy neural networks," *Journal of Intelligent & Fuzzy Systems*, vol. 27, no. 2, pp. 965–975, 2014.

[31] B. Pourbabaee, N. Meskin, and K. Khorasani, "Sensor fault detection, isolation, and identification using multiple-model-based hybrid Kalman filter for gas turbine engines," *IEEE Transactions on Control Systems Technology*, vol. 24, no. 4, pp. 1184–1200, 2016.

[32] S. Vrazic, "Robust driver heartbeat estimation: a q-Hurst exponent based automatic sensor change with interactive multi-model EKF," in *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, Italy, August 2015.

[33] X. R. Li and V. P. Jilkov, "Survey of maneuvering target tracking. Part V: multiple-model methods," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 41, no. 4, pp. 1255–1321, 2005.