

Article

Atmospheric PM_{2.5} Prediction Based on Multiple Model Adaptive Unscented Kalman Filter

Jihan Li ¹ , Xiaoli Li ^{1,2,*}, Kang Wang ¹ and Guimei Cui ³

¹ Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; ljhbgd2017@emails.bjut.edu.cn (J.L.); wangkang@bjut.edu.cn (K.W.)

² Beijing Key Laboratory of Computational Intelligence and Intelligent System, Engineering Research Center of Digital Community, Ministry of Education, Beijing 100124, China

³ School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou 014010, China; cuiguimei@imust.edu.cn

* Correspondence: lixiaolibjut@bjut.edu.cn

Abstract: The PM_{2.5} concentration model is the key to predict PM_{2.5} concentration. During the prediction of atmospheric PM_{2.5} concentration based on prediction model, the prediction model of PM_{2.5} concentration cannot be usually accurately described. For the PM_{2.5} concentration model in the same period, the dynamic characteristics of the model will change under the influence of many factors. Similarly, for different time periods, the corresponding models of PM_{2.5} concentration may be different, and the single model cannot play the corresponding ability to predict PM_{2.5} concentration. The single model leads to the decline of prediction accuracy. To improve the accuracy of PM_{2.5} concentration prediction in this solution, a multiple model adaptive unscented Kalman filter (MMAUKF) method is proposed in this paper. Firstly, the PM_{2.5} concentration data in three time periods of the day are taken as the research object, the nonlinear state space model frame of a support vector regression (SVR) method is established. Secondly, the frame of the SVR model in three time periods is combined with an adaptive unscented Kalman filter (AUKF) to predict PM_{2.5} concentration in the next hour, respectively. Then, the predicted value of three time periods is fused into the final predicted PM_{2.5} concentration by Bayesian weighting method. Finally, the proposed method is compared with the single support vector regression-adaptive unscented Kalman filter (SVR-AUKF), autoregressive model-Kalman (AR-Kalman), autoregressive model (AR) and back propagation neural network (BP). The prediction results show that the accuracy of PM_{2.5} concentration prediction is improved in whole time period.

Keywords: support vector regression; adaptive unscented Kalman filter; Bayesian; multiple model



Citation: Li, J.; Li, X.; Wang, K.; Cui, G. Atmospheric PM_{2.5} Prediction Based on Multiple Model Adaptive Unscented Kalman Filter. *Atmosphere* **2021**, *12*, 607. <https://doi.org/10.3390/atmos12050607>

Academic Editors: Yongming Han, Nanjun Chen and Merja Tölle

Received: 4 March 2021

Accepted: 29 April 2021

Published: 7 May 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In recent years, the global economic integration has developed rapidly. The Beijing–Tianjin–Hebei region economic belt (Beijing–Tianjin–Hebei) as the development area has gradually been formed. The severe haze event happened in Beijing in 2013 [1]. The fine particle matter (PM_{2.5}) is a particle with a diameter of 2.5 μm or less in the atmosphere. PM_{2.5} has a long atmospheric residence time, it has an important impact on environmental quality, atmospheric visibility, human health and climate change. PM_{2.5} has become the primary air pollutant in China. It mainly contains polycyclic aromatic hydrocarbons and heavy metals [2]. The harmful substances mainly including heavy metals, microorganisms and organic volatile compounds [3]. The human respiratory system is damaged by these harmful substances, which leads to damage to human health and death [4]. The government has also formulated a large number of policies, and a large number of software and automatic monitoring equipment are installed throughout every corner of the city. Once PM_{2.5} concentration monitored exceeds the specified range, the environmental supervision

department will send effective information to inform people to reduce social activities to avoid long-term exposure to the environment with high $PM_{2.5}$ concentration [5].

The $PM_{2.5}$ prediction method can be usually divided into physical and chemical methods and statistical methods [6]. Physical and chemical reaction methods do not need a large number of historical meteorological data. The chemical reactions of exhaust gas from pollution sources, emissions and fuel energy consumption need to have a deterministic description source. For example, Sun et al. proposed a cuckoo search algorithm based on principal component analysis and least squares support vector machine (PCA-LSSVM) to predict $PM_{2.5}$ concentration [7]. Zhang et al. observed the pollution characteristics of $PM_{2.5}$ and PM_{10} in autumn in Beijing and identified the source regions of $PM_{2.5}$ and PM_{10} by the positive matrix decomposition (PMF), the backward trajectory and potential source contribution function (PSCF) model [8].

On the other hand, the statistical methods are gradually applied in the field of artificial intelligence. Some statistical method-based AIs have been adopted. For example, for regression model, neural network, machine learning algorithm, time series method and Kalman filter, the relationship of pollutants and prediction variables are established by different functions or network models. Coburn W. G. et al. proposed an enhanced $PM_{2.5}$ air quality prediction model based on nonlinear regression (NLR) and reverse trajectory concentration to predict $PM_{2.5}$ in Louisville metropolitan area, Kentucky [9]. Elangasinghe, M.A. et al. analyzed the PM_{10} and $PM_{2.5}$ concentration time series by artificial neural network and K-means clustering [10,11]. For this method, the artificial neural network (ANN) model is used, and it has better performance in picking up high concentrations. The artificial neural network sometimes falls into the problem of local over fitting; but, the neural network model needs a lot of historical data for training, which increases the complexity and time of calculation. Zhou et al. proposed an ensemble empirical model decomposition generalized regression neural network (EEMD-GRNN) model to predict $PM_{2.5}$ concentration in a day [12]. The EEMD method is used to determine different information scales in the original time series for this method. In addition, each decomposed IMF has similar frequency characteristics. Thus, the efficiency and accuracy of prediction is improved. There are many other experts who are beginning to focus on application of artificial intelligence algorithm to propose the corresponding algorithm to predict $PM_{2.5}$. For example, Zhan et al. studied the spatial and temporal daily $PM_{2.5}$ concentration prediction based on spatial explicit machine learning algorithm [13]. In this study, geographically weighted gradient boosting machine (GW-GBM) method is proposed. The deviation of $PM_{2.5}$ concentration estimation is overcome, so that the prediction is more accurate. Some other scholars use time series method to predict the $PM_{2.5}$ concentration [14–18]. For example, Akhil Kadiyala used a vector time series and back propagation neural network (VTS-BPNN) method to review articles of hybrid indoor air quality (IAQ) models. In dynamic atmosphere, it is sometimes impossible to measure every variable. The Kalman filter can estimate the missing information by the limited, indirect and noisy measurement information, which has the characteristics of small data storage and handling multidimensional and non-stationary random processes. The Kalman filtering is also used to predict the future trend of dynamic system [19]. For example, Federico Cassola used a Kalman filter to forecast weather by 2-year-long data sets. The traditional Kalman filtering method is only suitable for linear system, and the observation equation must also be linear. The nonlinear Kalman filters are applied to deal with the estimation of nonlinear systems [20–25]. However, the extended Kalman filter (EKF) method needs to linearize the nonlinear function with the first or second order term of Taylor formula, and the Jacobian matrix is calculated. The $PM_{2.5}$ concentration prediction is not very accurate by the EKF method. To solve this problem, the unscented Kalman filter (UKF) method is applied to predict the future trend.

The appeal method is used to predict $PM_{2.5}$ concentration by a single model. In practice, the $PM_{2.5}$ concentration is always changing in different environments or time periods. A single model can usually predict the $PM_{2.5}$ concentration in the current situation. Once the original environment changes, the original model will not match the new environment,

and the prediction error will be generated. Thus, it is the main problem to establish the corresponding prediction model for different states. Blom, H.A.P. et al. have proposed a multiple model theory [26]. For example, the problem of outlet temperature of flow pressure controlled tubular heat exchanger is studied [27]. This method is based on the linear generalized predictive control (LGPC). The main idea of the strategy uses multiple linear models to represent the operating environment of multiple systems. Each system is accurately identified by the intelligent decision mechanism (IDM) at every moment, the target can be controlled more accurately. In addition, Xie et al. have proposed a comprehensive vehicle trajectory prediction method based on the combination of physics and mobility [28]. An interactive multiple model prediction method based on the two prediction models is proposed, which realizes recursive adjustment and switching by the probability of each model occurrence, and the lane change is analyzed to realize vehicle trajectory. Li et al. used multiple model theory to control variables under different working conditions or different states to improve control accuracy [29–32]. Although the idea of multiple models is widely applied in the field of control, it has not been applied to the prediction field of atmospheric $PM_{2.5}$ concentration.

To predict $PM_{2.5}$ concentration more accurately, the idea of multiple model is introduced into the atmosphere prediction, and a multiple model adaptive unscented Kalman filter method is proposed to predict $PM_{2.5}$ concentration in the next hour. The contributions of this paper are given as follows:

- (1) The $PM_{2.5}$ concentration prediction values of different time periods are fused to predict the change trend and fluctuation of $PM_{2.5}$ concentration by the multiple model adaptive method.
- (2) The proposed MMAUKF method also considers the estimation of noise in the prediction process. When the $PM_{2.5}$ concentration is predicted, the noise is also estimated adaptively by estimator.
- (3) The proposed MMAUKF method improves the accuracy and stability of $PM_{2.5}$ concentration prediction.

The remaining sections of the paper are described as follows. Section 2 introduces the establishment of $PM_{2.5}$ concentration prediction system. Section 3 mainly describes $PM_{2.5}$ concentration prediction based on SVR-AUKF method. Section 4 is a multiple model adaptive unscented Kalman filter for $PM_{2.5}$ concentration prediction. Section 5 is the simulation results. Section 6 is the conclusion of this paper.

2. The Establishment of $PM_{2.5}$ Concentration Prediction System

Most scholars pay attention to the prediction of $PM_{2.5}$ concentration. A large number of atmospheric $PM_{2.5}$ monitoring points are established in the city. In this paper, a monitoring point of a campus in the Chaoyang District of Beijing is taken as the research area, and the monitoring of $PM_{2.5}$ concentration on campus is provided 24 h a day. The $PM_{2.5}$ concentration is predicted by monitoring the historical atmospheric meteorological data monitoring platform. The data monitoring platform is composed of an outdoor air sensor probe, upper computer and solar power supply system. The measurement error of the monitoring probe is less than 0.05%, which has a stable operating environment and avoids the interference of the external environment. The monitor can directly access the monitored data by the network cloud. The power supply system of the equipment is a stable power supply system, which is provided by the renewable energy solar panel for the real-time and stable operation. The data monitoring platform is shown in Figure 1.

The simulation experimental data are obtained by this platform in this section. The main monitoring data of the atmospheric monitoring platform include $PM_{2.5}$, PM_{10} , SO_2 , NO_2 , CO , O_3 , humidity (RH), temperature (T) and air quality index (AQI). The historical data of $PM_{2.5}$ concentration in July 2018 are used as the study object in this paper. The relationship between $PM_{2.5}$ concentration and other potential variables was illustrated by the scatter plot of $PM_{2.5}$ concentration in Figure 2 (the blue points represent $PM_{2.5}$ and other air pollutants, the green points represent $PM_{2.5}$ and meteorological conditions).

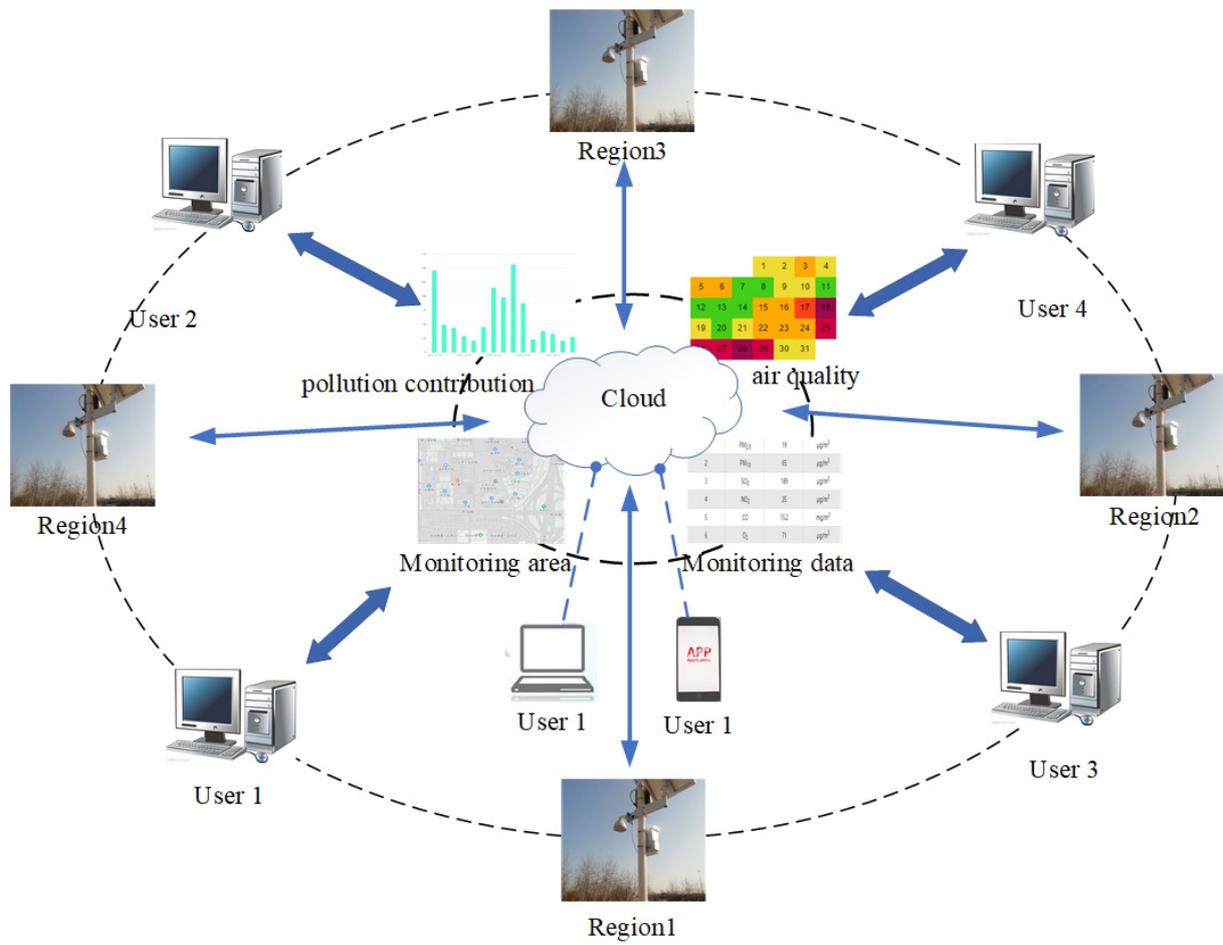


Figure 1. Schematic of atmospheric data monitoring platform.

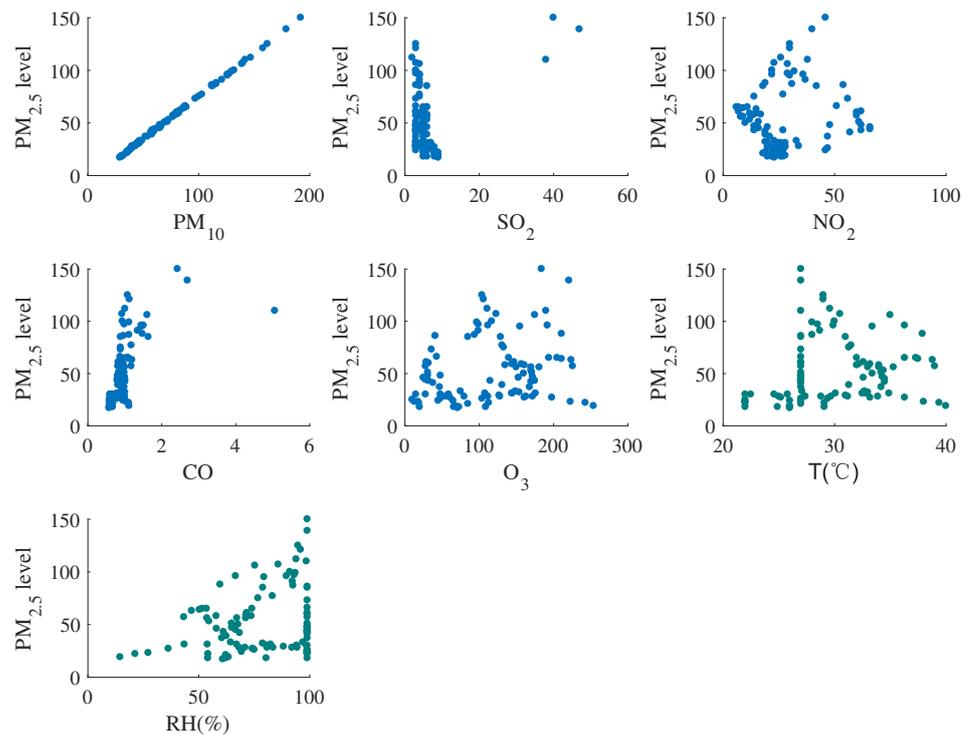


Figure 2. The scatterplot of PM_{2.5} concentration with other monitored variables.

The mean and variance of pollutants and meteorological conditions are given by statistical methods in Table 1. The mean represents the mean of time series, and standard deviation represent the mean change of the time series of the monitored data.

Table 1. Descriptive statistics of pollutants and meteorological conditions.

	Unit	Mean	Standard Deviation
PM _{2.5}	µg/m ³	51.37	24.38
PM ₁₀	µg/m ³	71.20	29.99
SO ₂	µg/kg	5.57	4.38
NO ₂	mg/m ³	28.77	11.84
CO	mmol/L	0.98	0.41
O ₃	mg/m	127.04	65.70
RH	%	31.50	4.28
T	°C	71.21	19.42

According to the analysis of literature [12], the PM_{2.5} concentration is closely related to other monitoring variables, the monitored variables (PM_{2.5}, PM₁₀, SO₂, NO₂, CO, O₃, humidity (RH), temperature (T)) are used as the input variables of the SVR model, and PM_{2.5} concentration is used as the output variables of the model. To improve the accuracy of the proposed method in predicting PM_{2.5} concentration, the prediction accuracy of the proposed method is compare with the other single model method by three error indexes. The three error statistics methods are given as follows:

$$MAE = \frac{1}{n} \sum_{k=1}^n |y_{k+1} - \hat{y}_{k+1}| \quad (1)$$

$$MAPE = \frac{1}{n} \sum_{k=1}^n \left| \frac{y_{k+1} - \hat{y}_{k+1}}{y_{k+1}} \right| \times 100\% \quad (2)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{k=1}^n (y_{k+1} - \hat{y}_{k+1})^2} \quad (3)$$

where y_{k+1} represents the actual PM_{2.5} concentration value, \hat{y}_{k+1} is PM_{2.5} concentration prediction value at time $k + 1$, ($k = 0, 1, \dots, n$), n is the number points in the test dataset. The MAE is mean absolute error. The MAPE is mean absolute percentage error. The RMSE denotes root mean square error. These performance indexes analyze the error between the predicted value and the actual value, and compare the prediction of PM_{2.5} concentration by different methods.

3. PM_{2.5} Concentration Prediction Based on the SVR-AUKF Method

3.1. Support Vector Regression

In this section, the state space equation framework based on SVR is introduced. The state equation framework is combined with adaptive Kalman filter to achieve state regression and adaptive noise estimation. The support vector machine (SVM) is a theoretical method of pattern recognition, which is usually used for classification and regression [33–35]. The artificial intelligence is developed rapidly, the regression problem is also widely used. The SVR is one of the emerging methods [36]. In the regression problem, the correlation function is establish between input variables and output variables. The SVR method mainly constructs linear decision function in high-dimensional space to realize

model prediction. In SVR, a set of training samples $\{x_i, y_i\}$, ($i = 1, 2, \dots, n$) is usually given. The sample regression function is described as follows:

$$f(x) = \langle w, x \rangle + b \quad (4)$$

where w is the regression coefficient, and b is bias. $\langle \cdot, \cdot \rangle$ represents the dot product, which can be solved by convex optimization problems [37]. When the f exists and approximates (x_i, y_i) , the errors are allowed. The slack variables ξ, ξ_i^* is introduced to cope with infeasible constraints of the optimization problem. The formula is shown as follows:

$$\begin{aligned} \min & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) \\ \text{s. t.} & \begin{cases} y_i - \langle w, x \rangle - b \leq \varepsilon + \xi_i \\ -y_i + \langle w, x \rangle + b \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned} \quad (5)$$

where C is trade off parameter between f and larger than ε , ε is the tolerance coefficient. The dual optimization problem is obtained by using Lagrange multipliers, which is described as follows:

$$\text{Max} - \frac{1}{2} \sum_{i=1, j=1}^N (\alpha_j - \alpha_i^*)(\alpha_j - \alpha_i^*)K(x_i, x_j) - \sum_{i=1}^N (\alpha_i + \alpha_i^*)\varepsilon + \sum_{i=1}^N (\alpha_i - \alpha_i^*)y_i \quad (6)$$

where α_i, α_i^* is Lagrange multipliers, $K(x_i, x_j)$ represents kernel function. $\sum_{i=1}^N y_i(\alpha_i - \alpha_i^*) = 0$ is constraint condition. $0 < \alpha_i, \alpha_i^* < C$. The regression function equation can be obtained as follows:

$$f(x) = \sum_{i=1}^N (\alpha_i - \alpha_i^*)K(x, x_i^*) + b \quad (7)$$

where $\langle w, x \rangle$ can be obtained as follows:

$$\langle w, x \rangle = \sum_i^N (\alpha_i - \alpha_i^*)x_i \quad (8)$$

$$K(x, x_i) = \exp\left(-\frac{\|x_i - x_j\|^2}{\sigma^2}\right) \quad (9)$$

where σ is the radial basis function (RBF) width. The b is calculated by Karush–Kuhn–Tucker (KKT) conditions [37]. In this paper, the optimal choice of C and σ are respectively 2^5 and 0.0825 by multi-fold cross-validation.

$$b = y_i - \langle w, x_i \rangle - \varepsilon, 0 < \alpha_i < C \quad (10)$$

$$b = y_i - \langle w, x_i \rangle + \varepsilon, 0 < \alpha_i^* < C \quad (11)$$

When the PM_{2.5} concentration is predicted, the environment is always complex and changeable. There are always various external noises, uncertainties and randomness, which results in the prediction accuracy deviation [38]. In this paper, the unscented Kalman filter (UKF) method is introduced to deal with the uncertainty and randomness by constantly updating the status.

3.2. Unscented Kalman Filter Method

The traditional KF method often loses its optimality [39]. The UKF is introduced in this section. The UKF method is based on UT transform and adopts Kalman linear filter framework. The SVR-UKF is used in this paper. The nonlinear state equation frame of SVR method is given as follows:

$$x_{k+1} = f(x_k) + w_k \tag{12}$$

$$y_k = h(x_k) + v_k \tag{13}$$

where f and h are the nonlinear equation and observation function, respectively. w_k and v_k are the Gaussian white noise of x_k and y_k , respectively. q and r are the mean values of w_k and v_k , respectively. Q and R are covariance matrices of w_k and v_k , respectively. Assuming nonlinear change $y = f(x)$, the state vector is an n dimensional random variable. The \bar{x} is mean value, and the P is variance. x_k denotes the state vector of PM_{2.5} concentration. y_k represents the observation vector of PM_{2.5} concentration. The n is state dimension. The initial value of the filter is calculated as follows:

$$\bar{x}_0 = E(x_0) \tag{14}$$

$$P_0 = E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] \tag{15}$$

where E is expected average, \bar{x}_0 is mean value, P_0 is covariance. The derivation process of UKF is given in [38]. The final derivation result of UKF is given directly in this paper. The $\hat{x}_{k+1|k}$ is given as follows:

$$\hat{x}_{k+1|k} = \sum_{i=0}^{2n} \omega^{(i)} x_{k+1|k}^{(i)} + q_k \tag{16}$$

where ω is a non-negative weight coefficient. The covariance matrix is calculated as follows:

$$P_{k+1|k} = \sum_{i=0}^{2n} \omega^{(i)} [\hat{x}_{k+1|k} - x_{k+1|k}^{(i)}][\hat{x}_{k+1|k} - x_{k+1|k}^{(i)}]^T + Q_k \tag{17}$$

The predicted observation value is obtained as follows:

$$y_{k+1|k}^{(i)} = h[x_{k+1|k}^{(i)}] \tag{18}$$

The predicted mean and covariance of the system are described by weighted sum as follows:

$$\bar{y}_{k+1|k} = \sum_{i=0}^{2n} \omega^{(i)} y_{k+1|k}^{(i)} + r_k \tag{19}$$

The Kalman gain is calculated as follows:

$$K_{k+1} = P_{x_k y_k} P_{y_k y_k}^{-1} \tag{20}$$

where $P_{y_k y_k}$, $P_{x_k y_k}$ is calculated separately as follows:

$$P_{y_k y_k} = \sum_{i=0}^{2n} \omega^{(i)} [y_{k+1|k}^{(i)} - \bar{y}_{k+1|k}][y_{k+1|k}^{(i)} - \bar{y}_{k+1|k}]^T + R_k \tag{21}$$

$$P_{x_k y_k} = \sum_{i=0}^{2n} \omega^{(i)} [x_{k+1|k}^{(i)} - \bar{y}_{k+1|k}][x_{k+1|k}^{(i)} - \bar{y}_{k+1|k}]^T \tag{22}$$

The state update is calculated as follows:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1}[y_{k+1} - \hat{y}_{k+1|k}] \tag{23}$$

The covariance update is calculated as follows:

$$P_{k+1|k+1} = P_{k+1|k} - K_{k+1}P_{y_k y_k}K_{k+1}^T \tag{24}$$

For PM_{2.5} concentration prediction problems, statistical characteristics of interference signals are always not known. To solve the influence of noise on PM_{2.5} concentration prediction, the AUKF method is used to estimate the noise in the next section.

3.3. Adaptive Unscented Kalman Filter Noise Estimation for PM_{2.5} Concentration Prediction

When the statistical characteristics of the prior noise are given, the SVR-UKF prediction accuracy is usually accurate [37]. In complex atmospheric environments, the PM_{2.5} concentration prediction is accompanied by various kinds of noise, and the process noise (q_k, Q_k) and the measurement noise covariance (r_k, R_k) are usually unknown or incorrect, which affects accuracy of prediction. The AUKF method is used to estimate noise adaptively and effectively avoid the deviation caused by noise in this paper. The Figure 3 shows the schematic diagram of the SVR-AUKF PM_{2.5} concentration prediction method.

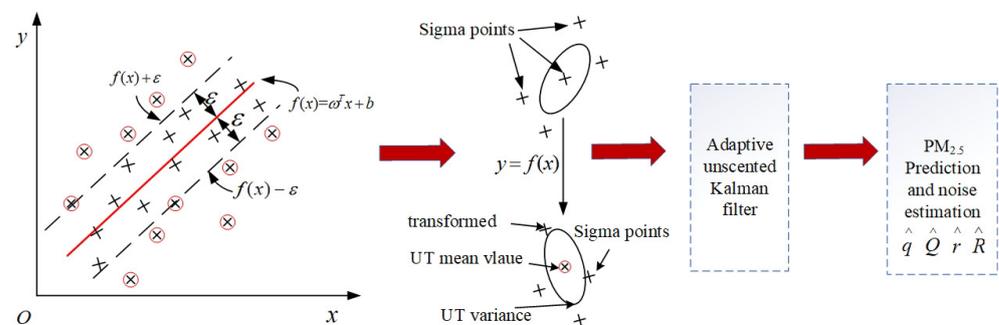


Figure 3. The schematic diagram of the SVR-AUKF algorithm for PM_{2.5} concentration prediction.

The noise estimation formula is given as follows. The mean value of process noise is given as follows:

$$\hat{q}_{k+1} = (1 - d_{k+1})\hat{q}_k + d_{k+1}[\hat{x}_{k|k} - \sum_{i=0}^{2n} \omega_i^{(m)} f(x^{(i)}(k))] \tag{25}$$

where d_{k+1} is obtained as follows:

$$d_{k+1} = \frac{1 - b}{1 - b^{k+1}} \tag{26}$$

where b is the forgetting factor, which is set as 0.98 in this paper. The process noise covariance is estimated as follows:

$$\hat{Q}_{k+1} = (1 - d_{k+1})\hat{Q}_k + d_{k+1}[K_{k+1}F_{k+1}F_k^T K_{k+1} + P_{k+1|k+1} - \sum_{i=0}^{2n} \omega_i^{(m)} (x_{k+1|k}^{(i)} - \hat{x}_{k+1|k})(x_{k+1|k}^{(i)} - \hat{x}_{k+1|k})^T] \tag{27}$$

where F_{k+1} is obtained as follows:

$$F_{k+1} = y_{k+1} - \hat{y}_{k+1} \tag{28}$$

The mean value of measured noise is estimated as follows:

$$\hat{r}_{k+1} = (1 - d_{k+1})\hat{r}_k + d_{k+1}[y_{k+1} - \sum_{i=0}^{2n} \omega_i^{(m)} h(x_{k+1|k}^{(i)})] \tag{29}$$

The measurement noise covariance is estimated as follows:

$$\hat{R}_{k+1} = (1 - d_{k+1})\hat{R}_k + d_{k+1}[F_{k+1}F_{k+1}^T - \sum_{i=0}^{2n} \omega_i^{(c)} (y_{k+1|k}^{(i)} - \bar{y}_{k+1|k})(y_{k+1|k}^{(i)} - \bar{y}_{k+1|k})^T] \quad (30)$$

The SVR-AUKF method has been given in this section. The PM_{2.5} concentration is predicted based on the SVR-AUKF method, and the prediction result of this method is shown as follows:

It can be seen from Figure 4 that the curve of PM_{2.5} concentration predicted by single model SVR-AUKF method is basically consistent with the curve of actual measured value. There will be large error at some individual time points. The concentration of PM_{2.5} also changed in different time periods. The single model SVR-AUKF method can not fully reflect the change of PM_{2.5} concentration in the whole process and then produce errors. In order to fully reflect the change of PM_{2.5} concentration, the multiple model prediction method is introduced to improve the prediction accuracy of PM_{2.5} concentration and reduce error. The PM_{2.5} concentration prediction of multiple model adaptive unscented Kalman filter is described in the next section.

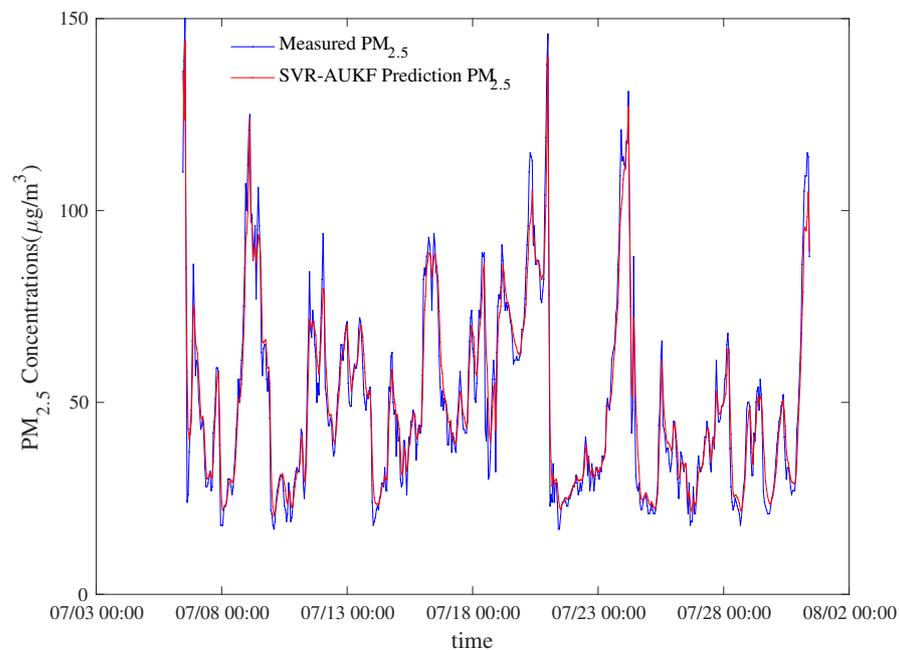


Figure 4. The PM_{2.5} concentration prediction of single SVR-AUKF method.

4. Multiple Model Adaptive Unscented Kalman Filter for PM_{2.5} Concentration Prediction

In the process of predicting PM_{2.5} concentration, the single model is often not accurate enough, because the dynamic characteristics of the model will be affected by a variety of variables in different time periods. According to the reference [40], even in the same day, the pollutants will fluctuate greatly in different time periods. The model will also be affected. To predict PM_{2.5} concentration more accurately in different time periods, the MMAUKF method is proposed. Due to the change of temperature and humidity in the morning, middle and evening, the concentration of PM_{2.5} also changes, the typical time of the day is mainly in the morning, noon and evening. The three representative time periods are selected as the research object. The predicted PM_{2.5} concentration data are divided into three time periods, which are 23:00~6:00, 7:00~14:00, 15:00~22:00, respectively. The PM_{2.5} concentration model is established based on the data of different time periods. The PM_{2.5} concentration data in different time periods are shown in Figure 5.

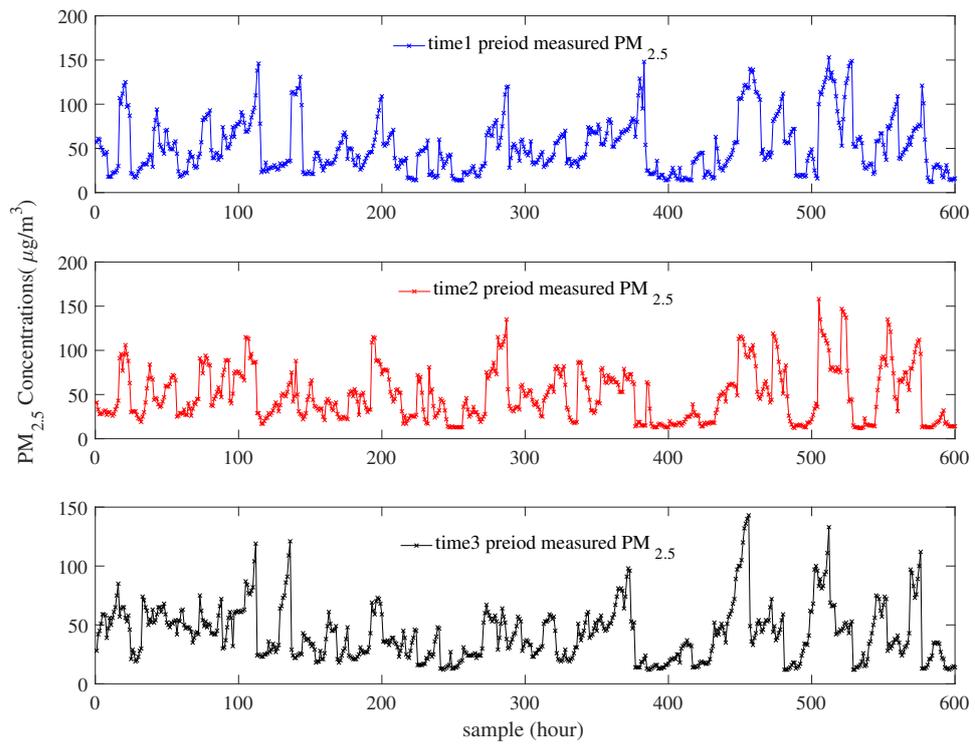


Figure 5. The modeling data of PM_{2.5} concentration in three time periods.

The obtained PM_{2.5} concentration prediction values from different time periods are fused by Bayesian method. The proposed MMAUKF method is implemented to predict PM_{2.5} concentration. The structure chart of MMAUKF method is shown as follows:

As shown in Figure 6, the structure chart of MMAUKF method for PM_{2.5} prediction concentration based on SVR-AUKF method is given. The final predicted value is obtained by weighting. The feedback of each adaptive Kalman filter total output is used as the initial value for predicting the next moment. When the MMAUKF method is used to predict PM_{2.5} concentration, the weight value calculation is the key to solving the PM_{2.5} concentration prediction. The weight value is determined by the probability of selected model in the current period. The prediction error reflects the accuracy of the method. The error is small, which indicates the prediction accuracy of the selected model is high, and the probability of the selected model is high. The final fusion PM_{2.5} concentration value is obtained by Bayesian method.

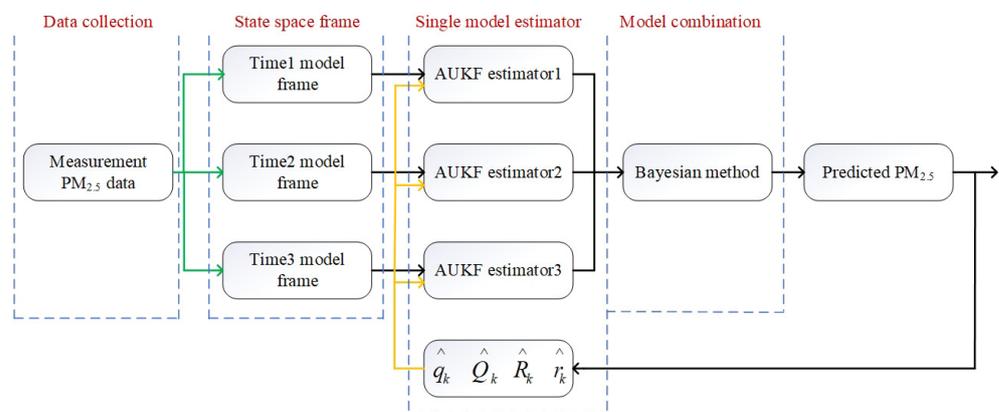


Figure 6. The structure chart of the MMAUKF method for PM_{2.5} concentration prediction.

In this section, the Bayesian theory is adopted to predict PM_{2.5}. The \hat{y}_1 , \hat{y}_2 and \hat{y}_3 are obtained for PM_{2.5} concentration prediction value at the corresponding time period. When the predicted value is close to the measured value in the current time period, the weight value is larger. The predicted values in three time periods are calculated as follows:

$$\hat{y}_{k+1} = W_1 \cdot \hat{y}_{1,k+1} + W_2 \cdot \hat{y}_{2,k+1} + W_3 \cdot \hat{y}_{3,k+1} \quad (31)$$

where W is the weight values. The weighted sum is given as follows:

$$W_1 + W_2 + W_3 = 1 \quad (32)$$

The prediction error ε is given as follows:

$$\varepsilon_{i,k+1} = y_{i,k+1} - \hat{y}_{i,k+1} \quad (33)$$

where $y_{i,k+1}$ is the measured PM_{2.5} concentration value, ($i = 0, 1, 2, \dots, n$). $\hat{y}_{i,k+1}$ is the estimated PM_{2.5} concentration value. k is the sampling period. Thus, the weight or the probability of selected model are given as follows: the initial values probability of the i th model are equal for each time period. The initial values are calculated as follows:

$$P(M_i|y_{(0)}) = 1/M_i \quad (34)$$

The conditional probability density function of the measured value is obtained as follows:

$$f(y_{i,k+1}|M_i) = \frac{1}{(2\pi)^{m/2} Q_{i,k+1}^{1/2}} \exp\left(-\frac{1}{2} \varepsilon_{i,k+1}^T Q_{i,k+1}^{-1}(k) \varepsilon_{i,k+1}\right) \quad (35)$$

where M_i represents the parameter set of the i th ($i = 1, 2, 3$) model. m is the number of measurement value, $m = 1$. $Q_{i,k}$ represents the process noise variance corresponding to the i th filter estimator.

$$W_j = P(M_j|y_{i,k+1}) = \frac{f(y_{i,k+1}|M_j)P(M_j)}{\sum_{i=1}^3 f(y_{i,k+1}|M_i)P(M_i)} \quad (36)$$

The conditional probability density of the measured values of each time period model according to Bayesian theory is calculated as follows:

$$P(M_i|y_{i,k}) = \frac{P(y_{i,k}|M_i)P(M_i)}{P(y_{i,k})} \quad (37)$$

where $P(y_{i,k}|M_i) = P(y_{i,k})$. The formula (36) is deduced as follows:

$$W_{j,k+1} = P(M_j|y_{i,k+1}) = \frac{f(y_{i,k+1}|M_j)P(M_j|y_{i,k})}{\sum_{i=1}^3 f(y_{i,k+1}|M_i)P(M_i|y_{i,k})} \quad (38)$$

where j is observer number. According to the above formula, the weight is calculated in three different time periods for PM_{2.5} concentration prediction, and the fusion prediction result is obtained by formula (31).

5. The Simulation Results

In this section, the prediction results of different methods are described, and the performance indexes of the error are given. In order to verify the accuracy of the proposed method, the three time periods of PM_{2.5} concentration are predicted, and the PM_{2.5} concentration of the whole time period is predicted by the proposed method. Through the comparison of different methods, a more accurate prediction method is shown. The predicted PM_{2.5} concentration value by SVR-AUKF1 is given in Figure 7.

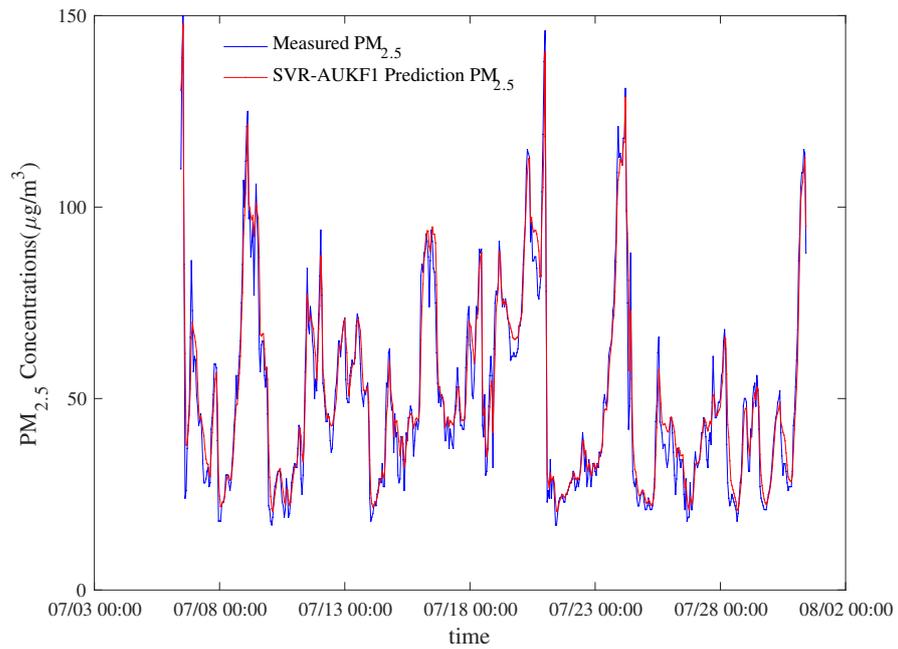


Figure 7. The $PM_{2.5}$ concentration prediction of SVR-AUKF1.

The prediction residuals of SVR-AUKF1 are shown in Figure 8.

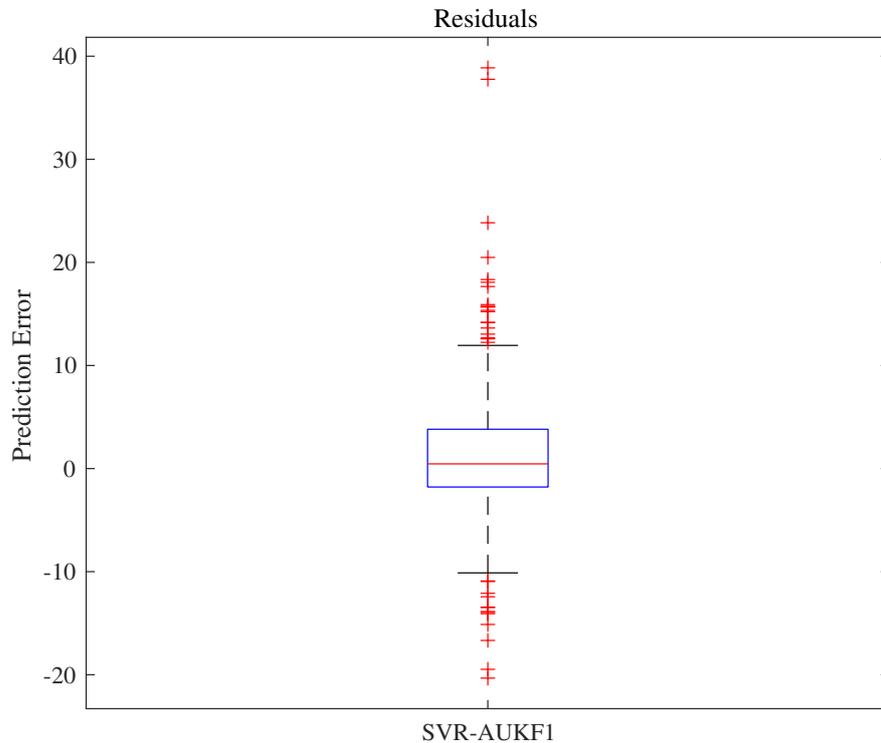


Figure 8. Boxplot of SVR-AUKF1 prediction residuals.

For further simulation, this paper assumes that the unknown noise is Gaussian white noise. According to the actual $PM_{2.5}$ concentration in three time periods, the initial states $x_{1,(0)} = 120$, $x_{2,(0)} = 58$ and $x_{3,(0)} = 78$ are given, respectively. Figure 7 shows the $PM_{2.5}$ concentration prediction of SVR-AUKF1. It can be seen that there are still some errors for the whole sampling period in Figure 7. The MAE is 4.1994, the MAPE is 9.6523 and the RMSE is 6.2302. When the prediction is relatively accurate at the beginning of the time

period. The errors will be produced in some point of time period. The SVR-AUKF1 method cannot predict $PM_{2.5}$ concentration accurately in the whole time period. The predicted $PM_{2.5}$ concentration value by SVR-AUKF2 is given in Figure 9.

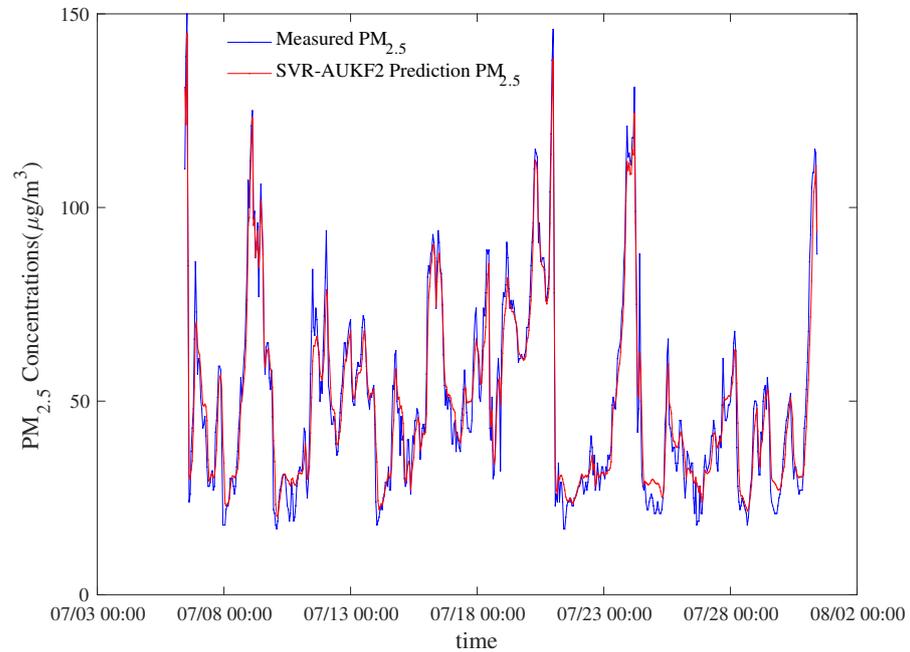


Figure 9. The $PM_{2.5}$ concentration prediction of SVR-AUKF2.

The prediction residuals of SVR-AUKF2 are shown in Figure 10.

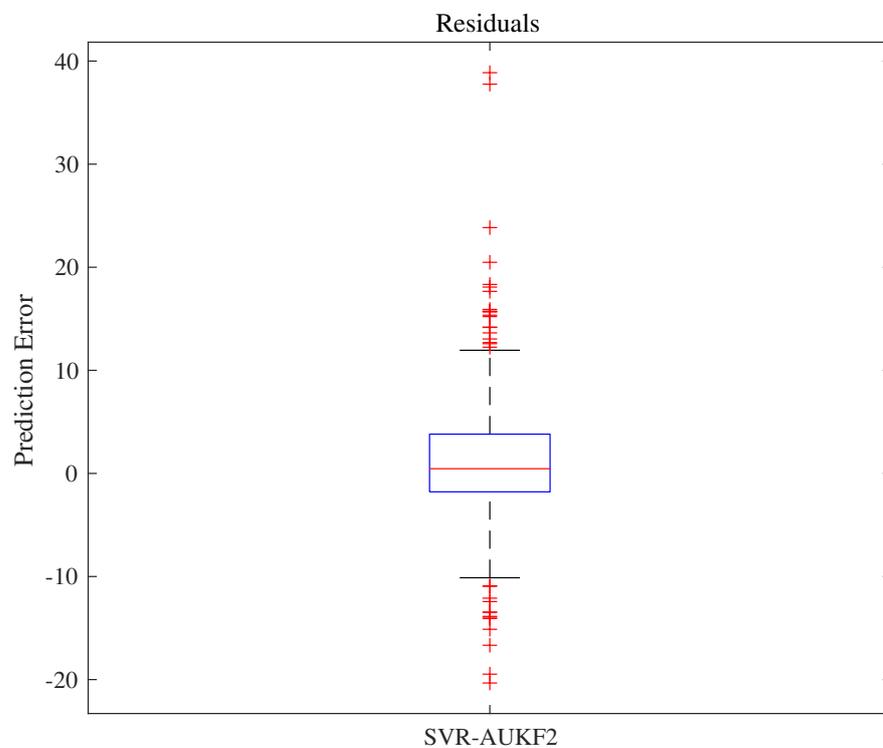


Figure 10. Boxplot of SVR-AUKF2 prediction residuals.

Figure 9 shows the predicted $PM_{2.5}$ concentration based on SVR-AUKF2 method. The prediction value is closer to the measured $PM_{2.5}$ concentration in the time range of (07/08

00:00~07/13 00:00). The MAE is 4.3706, the MAPE is 10.7784 and the RMSE is 5.9424. Although the single model can accurately predict the PM_{2.5} concentration in the current time period, the model will be affected by the environmental changes in different periods, which will affect the accuracy of PM_{2.5} concentration prediction. The predicted PM_{2.5} concentration value by SVR-AUKF3 is given in Figure 11.

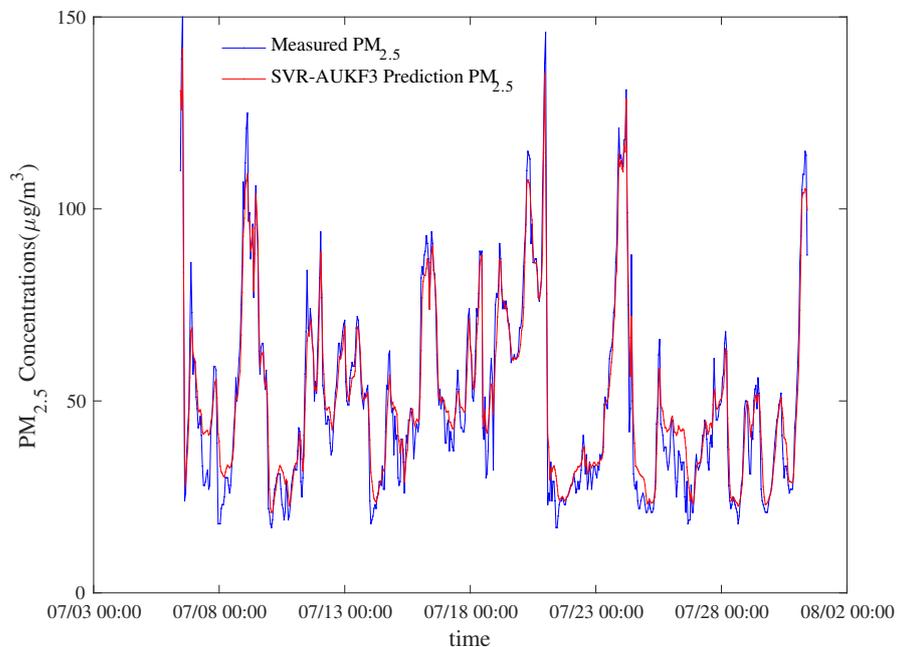


Figure 11. The PM_{2.5} concentration prediction of SVR-AUKF3.

The prediction residuals of SVR-AUKF3 are shown in Figure 12.

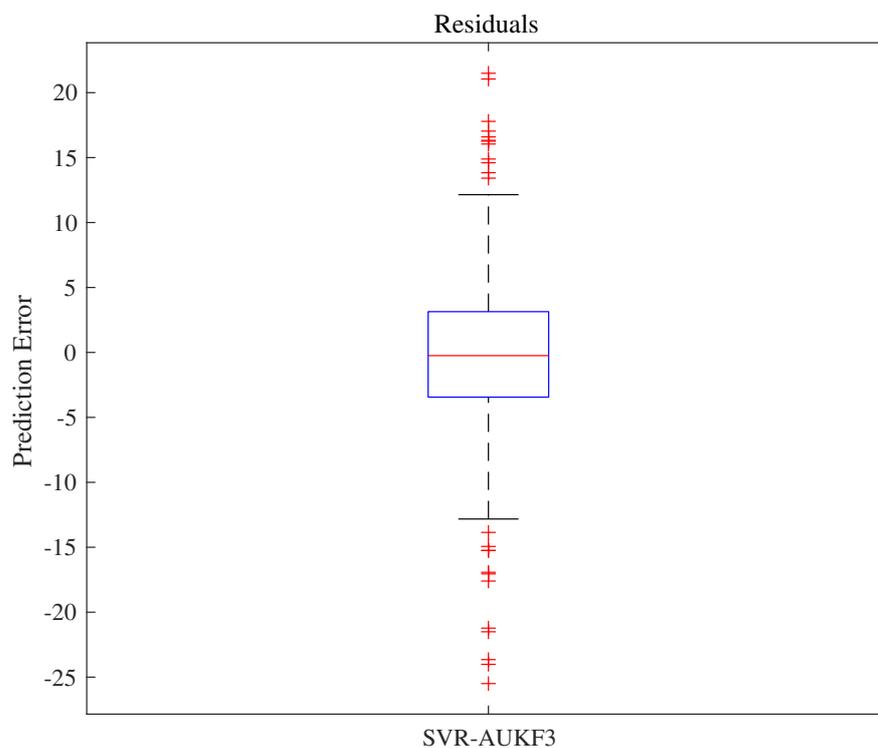


Figure 12. Boxplot of SVR-AUKF3 prediction residuals.

The Figure 11 shows the predicted $PM_{2.5}$ concentration value based on the SVR-AUKF3 method. The prediction value is close to the measured value in the time range of (07/13 00:00~07/18 00:00). The MAE is 4.4815 the MAPE is 11.8585 and the RMSE is 6.2018. In other time periods, the prediction accuracy of individual points can still be improved. In order to further improve the prediction accuracy, the multiple model method should be used to predict $PM_{2.5}$ concentration in different time periods.

The proposed MMAUKF method is compared with SVR-AUKF1, SVR-AUKF2 and SVR-AUKF3 methods in Figure 13.

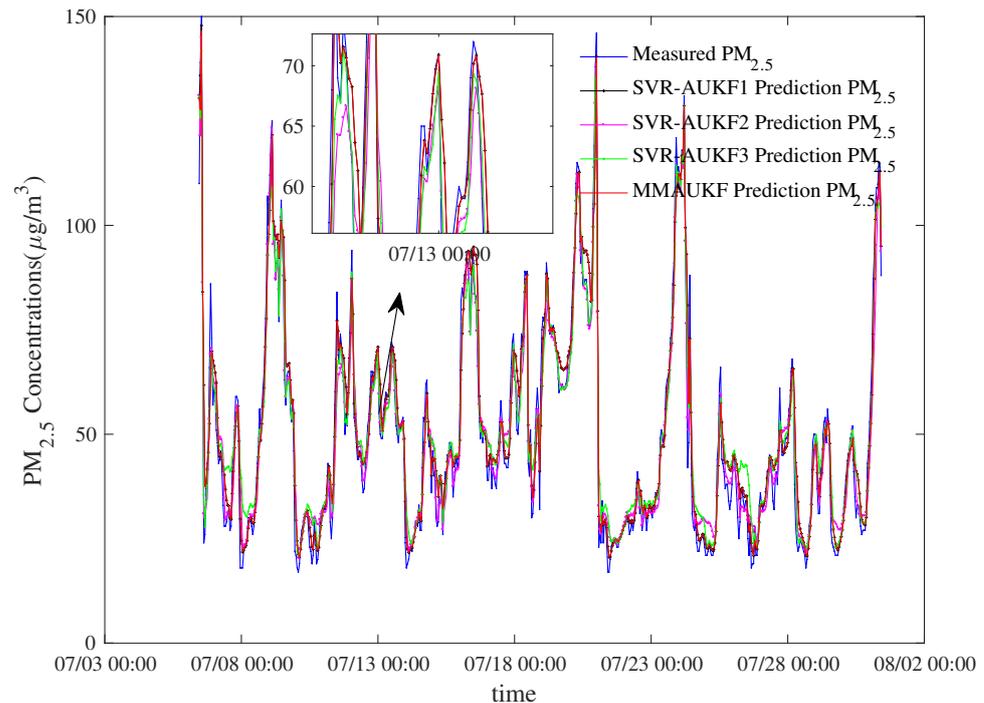


Figure 13. The comparison results of the proposed MMAUKF method and three single SVR-AUKF methods.

The statistical errors based on SVR-AUKF methods for $PM_{2.5}$ concentration in three time periods are shown in Table 2.

Table 2. The error performance index of three SVR-AUKF and MMAUKF methods.

Method	MAE	MAPE(%)	RMSE
MMAUKF	4.0492	9.5337	5.9130
SVR-AUKF1	4.1994	9.6523	6.2302
SVR-AUKF2	4.3706	10.7784	5.9424
SVR-AUKF3	4.4815	11.8585	6.2018

In Figure 13, the precision of MMAUKF is compared with the three single model of SVR-AUKF to predict $PM_{2.5}$ concentration. The MAE of MMAUKF method is 4.0492, the MAPE is 9.5337, the RMSE is 5.9130. The MAE of SVR-AUKF1 is 4.1994, the MAPE is 9.6523, and the RMSE is 6.2302. The MAE of SVR-AUKF2 is 4.3706, the MAPE is 10.7784, and the RMSE is 5.9424. The MAE of SVR-AUKF3 is 4.4815, the MAPE is 11.8585 and the RMSE is 6.2018. By the comparison of statistical errors, the error of MMAUKF method is smaller than the other method, and the proposed method overcomes that a single model cannot predict $PM_{2.5}$ concentration accurately under different time periods. The Table 2 shows the statistical errors based on SVR-AUKF methods for $PM_{2.5}$ concentration in three time periods. The boxplot of prediction residuals is shown in Figure 14,

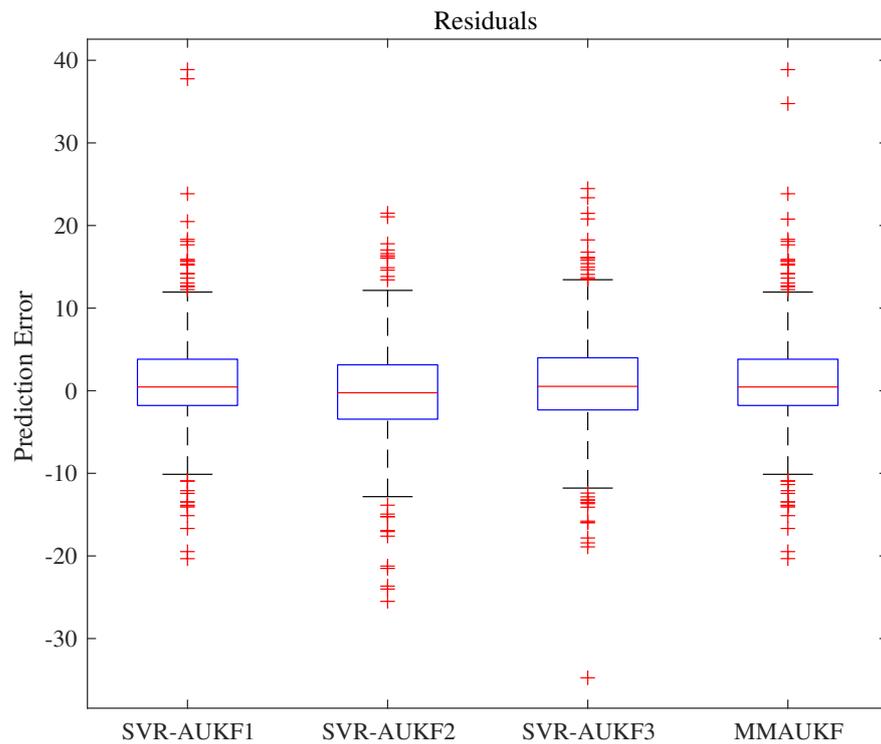


Figure 14. Boxplot of SVR-AUKF1, SVR-AUKF2, SVR-AUKF3 and MMAUKF prediction residual.

Figure 14 shows that 75% of the residual error predicted by the proposed MMAUKF method is approximately up to 4, which is smaller than other methods. Thus, the proposed MMAUKF method is more accurate for predicting $PM_{2.5}$ concentration. The proposed MMAUKF method and SVR-AUKF, AR-Kalman, AR and BP methods is used to predicted $PM_{2.5}$ concentration in Figure 15.

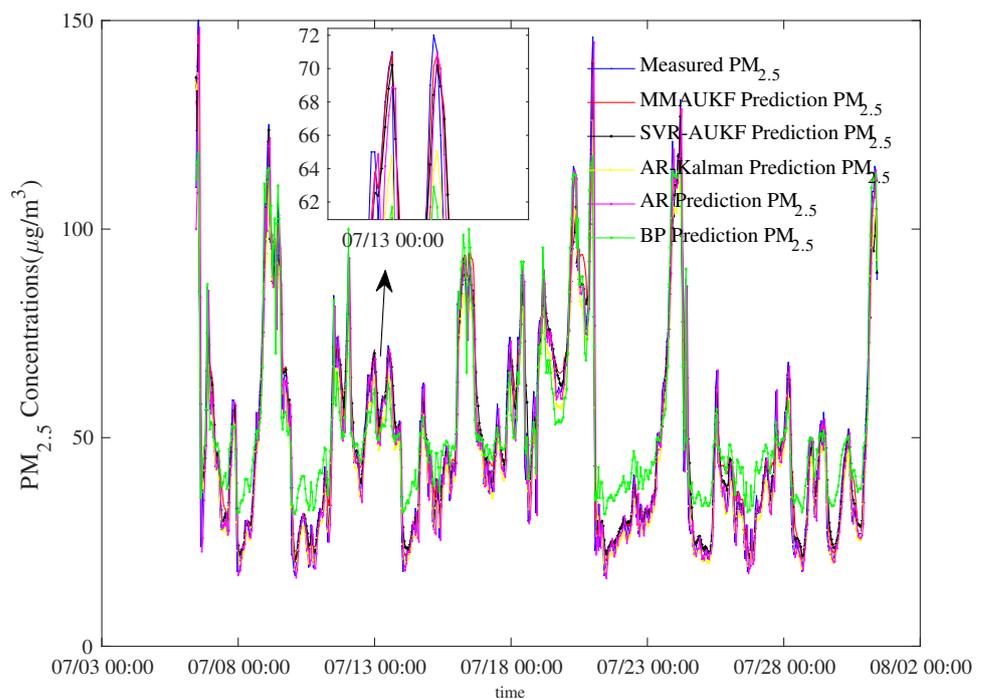


Figure 15. The comparison results of the proposed MMAUKF method and SVR-AUKF, AR-Kalman, AR and BP methods.

The prediction results of the proposed method is displayed, the fitting results between the predicted values and the actual values for different methods are shown in Figure 16.

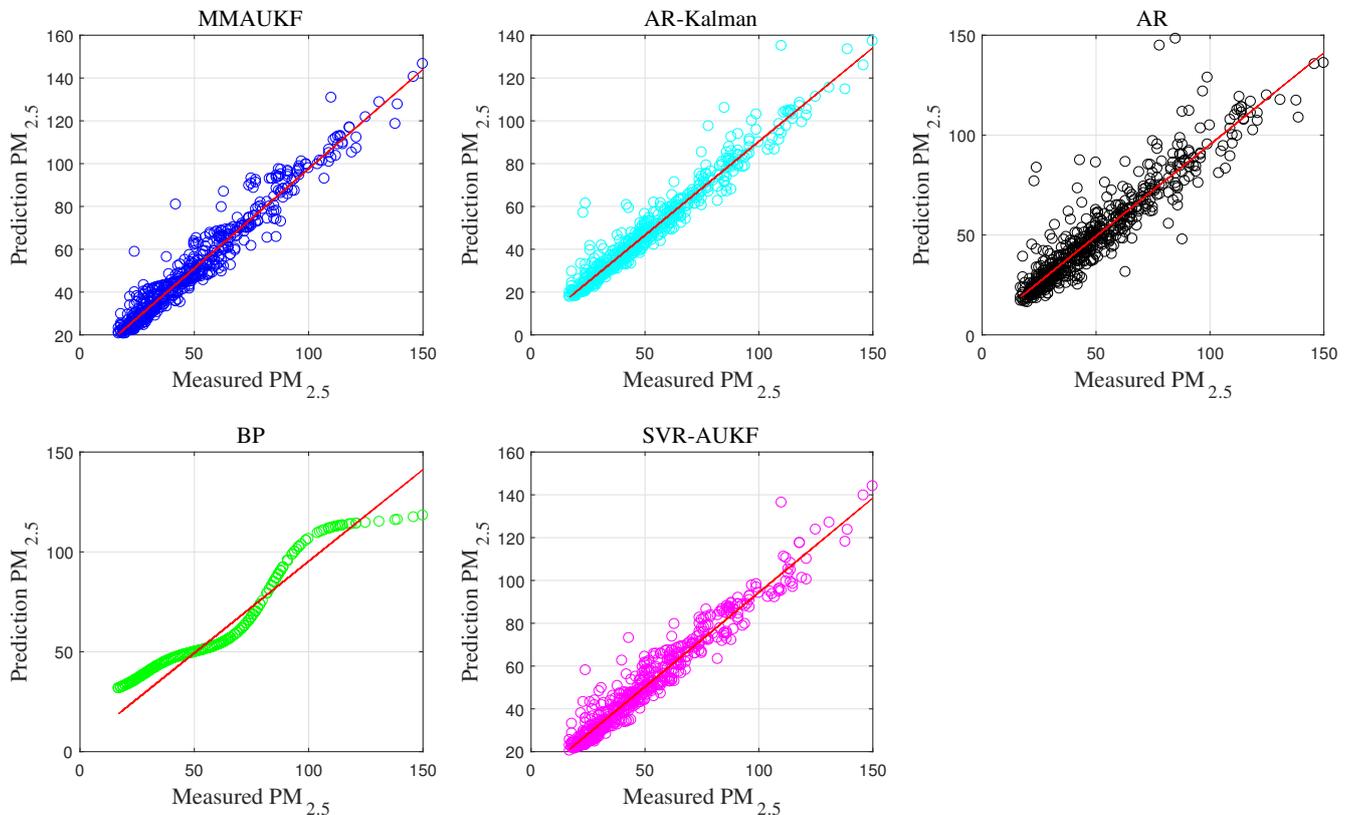


Figure 16. Boxplot of different methods’ prediction residual.

It can be seen from Figure 16 that the fit result of MMAUKF method is better than the other prediction methods. Figure 15 shows the comparison between the proposed method and other methods. The statistical error is shown in Table 3. The MAE is 4.0492, the MAPE is 95,337 and the RMSE is 5.9130. The prediction accuracy of other methods is not as good as the proposed method. The proposed MMAUKF method provides better prediction results than other single model methods. The proposed MMAUKF method can globally predict the PM_{2.5} concentration in the next hour, and which has good stability.

Table 3. The error performance index of MMAUKF and other single model methods.

Method	MAE	MAPE(%)	RMSE
MMAUKF	4.0492	9.5337	5.9130
SVR-AUKF	4.2304	9.7704	6.1347
AR-Kalman	5.07	10.2621	6.8705
AR	5.8735	12.7705	9.5259
BP	7.2173	21.2940	8.3226

In order to better reflect the prediction accuracy of the proposed method, the prediction residuals for different methods are given by the method of statistical analysis. The prediction errors of proposed MMAUKF method and SVR-AUKF, ARKalman, AR and BP methods are shown Figure 17.

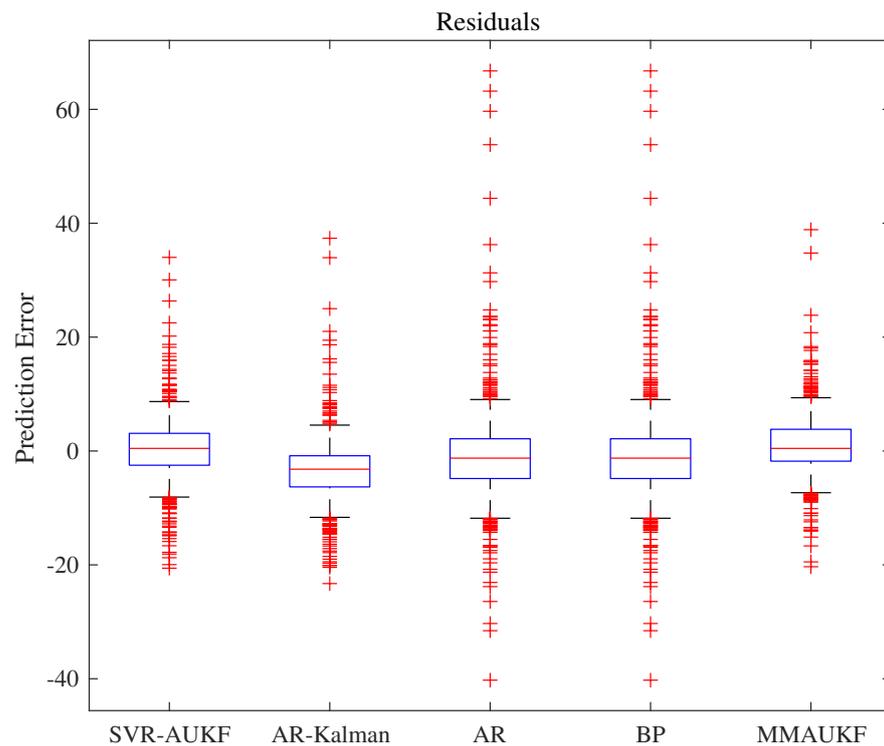


Figure 17. The comparison results of the proposed MMAUKF method and SVR-AUKF, AR-Kalman, AR and BP methods.

It can be seen from Figure 17 that the residual error of 75% of the proposed MMAUKF method is very small. The results show the prediction accuracy has been improved. In this section, the proposed MMAUKF method is used to predict $PM_{2.5}$ concentration, which has important practical significance to the public. The prediction results have reference significance for government departments to establish early warning system, which reduces the impact of public travel and avoids people exposed to high $PM_{2.5}$ concentration for a long time.

6. Conclusions

Previous studies found that $PM_{2.5}$ will affect people's life and travel. In order to protect people from the impact of $PM_{2.5}$ concentration on air quality, an accurate $PM_{2.5}$ concentration prediction model that can deal with different environments is an urgent requirement. In this paper, the multiple model adaptive unscented Kalman filter (MMAUKF) is proposed to predict the atmospheric $PM_{2.5}$ concentration in the next hour. In addition, to evaluate the prediction applicability effect of the proposed method on $PM_{2.5}$ concentration in a campus in the previous hour. The proposed MMAUKF method is compared with SVR-AUKF, AR-Kalman, AR and BP methods. The error performance index of several different methods is analyzed, which includes MAE, MAPE and RMSE. The box chart of statistical analysis is given. The experimental results show that the proposed method has the smallest error and can accurately predict $PM_{2.5}$ concentration in different time periods. The MAE of the proposed MMAUKF method is 4.0492, MAPE is 9.5337 and RMSE is 5.9130. The prediction results fully show that the multiple model prediction method overcomes the shortcomings of the single model which is easily affected by the environment. The proposed method not only guarantees the accuracy of predicting $PM_{2.5}$ concentration when the noise is inaccurate, but also overcomes the limitations of the single model method. The validity and effectiveness of the proposed method is also proved. The $PM_{2.5}$ concentration is also closely related to spatial distribution. In the future, $PM_{2.5}$ concentration spatial dynamic model and tracing to the source may be considered and studied, which is combined with advanced prediction algorithm to predict $PM_{2.5}$ concentration more accurately.

Author Contributions: Conceptualization, J.L. and X.L.; methodology, J.L.; software, J.L.; validation, K.W., X.L. and J.L.; formal analysis, J.L.; investigation, J.L.; resources, X.L.; data curation, J.L.; writing—original draft preparation, J.L.; writing—review and editing, J.L.; visualization, K.W., G.C. and J.L.; supervision, X.L.; funding acquisition, X.L. All authors have read and agreed to the published version of the manuscript.

Funding: This study is support by Beijing Natural Science Foundation (4204087), National Key Research and Development Project (2018YFB1702704, 2018YFC1602704) and the National Natural Science Foundation of China (61873006, 61673053).

Data Availability Statement: The pollutant and meteorological data used in this study was obtained from the ongoing monitoring system of Beijing University of Technology. The monitoring system includes data from regional monitoring stations.

Acknowledgments: The author would like to thank other members of this paper for their contributions.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gao, J.; Tian, H.; Cheng, K.; Lu, L.; Zheng, M.; Wang, S.; Hao, J.; Wang, K.; Hua, S.; Zhu, C.; et al. The variation of chemical characteristics of PM_{2.5} and PM₁₀ and formation causes during two haze pollution events in urban Beijing, China. *Atmos. Environ.* **2015**, *107*, 1–8. [[CrossRef](#)]
2. Wang, Y.X.; Jiang, H.; Zhang, S.Q.; Xu, J.H.; Lu, X.H.; Jin, J.X.; Wang, C.L. Estimating and source analysis of surface PM_{2.5} concentration in the Beijing–Tianjin–Hebei region based on MODIS data and air trajectories. *Int. J. Remote Sens.* **2016**, *37*, 4799–4817. [[CrossRef](#)]
3. Peng, W.; Ge, S.; Liu, Z.; Furuta, Y. Adsorption characteristics of sulfur powder by bamboo charcoal to restrain sulfur allergies. *Saudi J. Biol. Sci.* **2016**, *24*, 103–107. [[CrossRef](#)]
4. You, S.; Tong, Y.W.; Neoh, K.G.; Dai, Y.; Wang, C.H. On the association between outdoor PM_{2.5} concentration and the seasonality of tuberculosis for Beijing and Hong Kong. *Saudi J. Biol. Sci.* **2016**, *218*, 1170–1179. [[CrossRef](#)] [[PubMed](#)]
5. Wang, Y.; Kloog, I.; Coull, B.A.; Kosheleva, A.; Zanobetti, A.; Schwartz, J.D. Estimating Causal Effects of Long-Term PM_{2.5} Exposure on Mortality in New Jersey. *Environ. Health Perspect.* **2016**, *124*, 1182–1188. [[CrossRef](#)] [[PubMed](#)]
6. Kavousi-Fard, A.; Samet, H.; Marzbani, F. A new hybrid modified firefly algorithm and support vector regression model for accurate short term load forecasting. *Expert Syst. Appl.* **2014**, *41*, 6047–6056. [[CrossRef](#)]
7. Sun, W.; Sun, J. Daily PM_{2.5} concentration prediction based on principal component analysis and LSSVM optimized by cuckoo search algorithm. *J. Environ. Manag.* **2017**, *188*, 144–152. [[CrossRef](#)]
8. Zhang, Y.; Lang, J.; Cheng, S.; Li, S.; Zhou, Y.; Chen, D.; Zhang, H.; Wang, H. Chemical composition and sources of PM₁ and PM_{2.5} in Beijing in autumn. *Sci. Total Environ.* **2018**, *630*, 72–82. [[CrossRef](#)]
9. Cobourn, W. An enhanced PM_{2.5} air quality forecast model based on nonlinear regression and back-trajectory concentrations. *Atmos. Environ.* **2010**, *44*, 3015–3023. [[CrossRef](#)]
10. Elangasinghe, M.A.; Singhal, N.; Dirks, K.N.; Salmond, J.A.; Samarasinghe, S. Complex time series analysis of PM₁₀ and PM_{2.5} for a coastal site using artificial neural network modelling and k-means clustering. *Atmos. Environ.* **2014**, *94*, 106–116. [[CrossRef](#)]
11. Voukantsis, D.; Karatzas, K.; Kukkonen, J.; Räsänen, T.; Karppinen, A.; Kolehmainen, M. Intercomparison of air quality data using principal component analysis, and forecasting of PM₁₀ and PM_{2.5} concentrations using artificial neural networks, in Thessaloniki and Helsinki. *Sci. Total Environ.* **2014**, *140*, 220–232. [[CrossRef](#)] [[PubMed](#)]
12. Zhou, Q.; Jiang, H.; Wang, J.; Zhou, J. A hybrid model for PM_{2.5} forecasting based on ensemble empirical mode decomposition and a general regression neural network. *Sci. Total Environ.* **2014**, *496*, 264–274. [[CrossRef](#)]
13. Zhan, Y.; Luo, Y.; Deng, X.; Chen, H.; Grieneisen, M.L.; Shen, X.; Zhu, L.; Zhang, M. Spatiotemporal prediction of continuous daily PM_{2.5} concentrations across China using a spatially explicit machine learning algorithm. *Atmos. Environ.* **2017**, *155*, 129–139. [[CrossRef](#)]
14. Kadlyala, A.; Kumar, A. Vector—time—series—based back propagation neural network modeling of air quality inside a public transportation bus using available software. *Environ. Prog. Sustain. Energy.* **2017**, *35*, 7–13. [[CrossRef](#)]
15. das Chagas Moura, M.; Zio, E.; Lins, I.D.; Droguett, E. Failure and reliability prediction by support vector machines regression of time series data. *Reliab. Eng. Syst. Saf.* **2011**, *96*, 1527–1534. [[CrossRef](#)]
16. Sivakumar, S. Marginally Stable Triangular Recurrent Neural Network Architecture for Time Series Prediction. *IEEE Trans. Cybern.* **2018**, *48*, 2836–2850. [[CrossRef](#)]
17. Dai, X.L.; Liu, J.J.; Li, Y.L. A recurrent neural network using historical data to predict time series indoor PM_{2.5} concentrations for residential buildings. *Indoor Air.* **2021**, 1–10. [[CrossRef](#)]
18. Qiu, X.; Ren, Y.; Suganthan, P.N.; Amaratunga, G.A. Empirical Mode Decomposition based ensemble deep learning for load demand time series forecasting. *Appl. Soft. Comput.* **2017**, *54*, 246–255. [[CrossRef](#)]
19. Cassola, F.; Burlando, M. Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output. *Appl. Energy* **2012**, *99*, 154–166. [[CrossRef](#)]

20. Cassola, F.; Burlando, M. Short-term load forecasting using SVR (support vector regression)-based radial basis function neural network with dual extended Kalman filter. *Energy*. **2013**, *49*, 413–422.
21. Nasserri, M.; Moeini, A.; Tabesh, M. Forecasting monthly urban water demand using Extended Kalman Filter and Genetic Programming. *Expert Syst. Appl.* **2011**, *38*, 7387–7395. [[CrossRef](#)]
22. Santamaría-Bonfil, G.; Reyes-Ballesteros, A.; Gershenson, C. Wind speed forecasting for wind farms: A method based on support vector regression. *Renew. Energy* **2016**, *85*, 790–809. [[CrossRef](#)]
23. Bisoi, R.; Dash, P.K. A hybrid evolutionary dynamic neural network for stock market trend analysis and prediction using unscented Kalman filter. *Appl. Softw. Comput.* **2014**, *19*, 41–56. [[CrossRef](#)]
24. Meng, J.; Luo, G.; Gao, F. Lithium Polymer Battery State-of-Charge Estimation Based on Adaptive Unscented Kalman Filter and Support Vector Machine. *IEEE Trans. Power Electron.* **2016**, *31*, 2226–2238. [[CrossRef](#)]
25. Sun, F.; Hu, X.; Zou, Y.; Li, S.G. Adaptive unscented Kalman filtering for state of charge estimation of a lithium-ion battery for electric vehicles. *Energy* **2011**, *36*, 3531–3540. [[CrossRef](#)]
26. Blom, H.A.P.; Bar-Shalom, Y. The interacting multiple model algorithm for systems with Markovian switching coefficients. *IEEE Trans. Autom. Control.* **1988**, *33*, 780–783. [[CrossRef](#)]
27. Mazinan, A.H.; Sadati, N. Fuzzy predictive control based multiple models strategy for a tubular heat exchanger system. *Appl. Intell.* **2010**, *33*, 247–263. [[CrossRef](#)]
28. Xie, G.; Gao, H.; Qian, L.; Huang, B.; Li, K.; Wang, J. Vehicle Trajectory Prediction by Integrating Physics- and Maneuver-Based Approaches Using Interactive Multiple Models. *IEEE Trans. Ind. Electron.* **2018**, *33*, 5999–6008. [[CrossRef](#)]
29. Li, X.L.; Jia, C.; Wang, K.; Wang, J. Trajectory tracking of nonlinear system using multiple series-parallel dynamic neural networks. *Neurocomputing.* **2015**, *168*, 1–12. [[CrossRef](#)]
30. Li, X.L.; Jia, C.; Liu, D.X.; Ding, D.W. Nonlinear adaptive control using multiple models and dynamic neural networks. *Neurocomputing.* **2014**, *136*, 190–200. [[CrossRef](#)]
31. Li, X.L.; Zhang, X.F.; Jia, C.; Liu, D.X. Multi-model adaptive control based on fuzzy neural networks. *J. Intell. Fuzzy Syst.* **2014**, *27*, 965–975. [[CrossRef](#)]
32. Li, X.L.; Liu, D.X.; Jia, C.; Chen, X.Z. Multi-model control of blast furnace burden surface based on fuzzy SVM. *Neurocomputing.* **2014**, *148*, 209–215. [[CrossRef](#)]
33. Breton, R.G.; Lloyd, R.G. Support Vector Machines for classification and regression. *Analyst* **2009**, *23*, 230–267.
34. Yang, W.; Deng, M.; Xu, F.; Wang, H. Prediction of hourly PM_{2.5} using a space-time support vector regression model. *Atmos. Environ.* **2018**, *181*, 12–19. [[CrossRef](#)]
35. Li, X.L.; Luo, A.R.; Li, J.G.; Li, Y. Air Pollutant Concentration Forecast Based on Support Vector Regression and Quantum-Behaved Particle Swarm Optimization. *Atmos. Environ.* **2019**, *24*, 205–222. [[CrossRef](#)]
36. Smola, A.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
37. Gass, S.I.; Fu, M.C. Karush-Kuhn-Tucker (KKT) Conditions. In *Encyclopedia of Operations Research and Management Science*; Springer: New York, NY, USA, 2013; pp. 833–834.
38. Chen, K.; Yu, J. Short-term wind speed prediction using an unscented Kalman filter based state-space support vector regression approach. *Appl. Energy* **2014**, *113*, 690–705. [[CrossRef](#)]
39. Liu, H.; Tian, H.; Li, Y. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. *Appl. Energy* **2012**, *98*, 415–424. [[CrossRef](#)]
40. Wang, Z.; Chen, L.; Ding, Z.; Chen, H. An enhanced interval PM_{2.5} concentration forecasting model based on BEMD and MLP~1 with influencing factors. *Atmos. Environ.* **2020**, *223*, 117200.1–117200.16. [[CrossRef](#)]