Atmospheric PM$_{2.5}$ concentration prediction and noise estimation based on adaptive unscented Kalman filtering

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

Abstract
Due to the randomness and uncertainty in the atmospheric environment, and accompanied by a variety of unknown noise. Accurate prediction of PM$_{2.5}$ concentration is very important for people to prevent injury effectively. In order to predict PM$_{2.5}$ concentration more accurately in this environment, a hybrid modelling method of support vector regression and adaptive unscented Kalman filter (SVR-AUKF) is proposed to predict atmospheric PM$_{2.5}$ concentration in the case of incorrect or unknown noise. Firstly, the PM$_{2.5}$ concentration prediction model was established by support vector regression. Secondly, the state space framework of the model is combined with the adaptive unscented Kalman filter method to estimate the uncertain PM$_{2.5}$ concentration state and noise through continuous updating when the model noise is incorrect or unknown. Finally, the proposed method is compared with SVR-UKF method, the simulation results show that the proposed method is more accurate and robust. The proposed method is compared with SVR-UKF, AR-Kalman, AR and BP methods. The simulation results show that the proposed method has higher prediction accuracy of PM$_{2.5}$ concentration.

Keywords
Support vector regression, adaptive unscented Kalman filtering, noise estimation, PM$_{2.5}$ prediction

Date received: 4 January 2021; accepted: 28 January 2021

Introduction
The rapid development of our industrialization process and economy in recent years, air pollution has already affected in developed regions of China such as Beijing, Tianjin and Hebei, the Pearl River Delta and the Yangtze River Delta. PM$_{2.5}$ is one of particulate matter causing pollution. The sources of PM$_{2.5}$ are mainly divided into primary emission and secondary transformation sources, and the secondary transformation is the main source. The people’s health, vegetation, agriculture and building materials are harmed in the Los Angeles Smog in 1930s, so people must use protective equipment to avoid being affected. To make people avoid air pollution in advance, China’s environment department has also formulated a large number of measures and policies.

The prediction of PM$_{2.5}$ concentration is the first research topic at present. Many methods are proposed to predict PM$_{2.5}$ concentration by a large number of experts and scholars. For example, Zhou et al.$^9,10$ proposes a generalized regression neural network prediction by data processing and analysis integrated empirical mode decomposition to predict PM$_{2.5}$. This method can improve the prediction accuracy of PM$_{2.5}$ concentration. Elangasinghe et al.$^{11}$ studies the PM$_{2.5}$ concentration in the region of New Zealand by artificial neural network (ANN). Feng et al.$^{12}$ uses different stations to establish a hybrid model to improve the prediction accuracy of daily average PM$_{2.5}$ concentration by artificial neural network (ANN). The influencing factors between different stations is used as the inputs of the hybrid model by the data of 13 pollution monitoring stations. The data is decomposed by wavelet changes and applied to the corresponding models. The method can establish powerful nonlinear network and improve prediction accuracy. Some scholars use time series method to study PM$_{2.5}$ concentration. Erdem

---

$^1$Faculty of Information Technology, Beijing University of Technology, Beijing, China
$^2$Beijing Key Laboratory of Computational Intelligence and Intelligent System, Engineering Research Center of Digital Community, Ministry of Education, Beijing, China
$^3$School of Information Engineering, Inner Mongolia University of Science and Technology, Baotou, China

Corresponding author:
Xiaoli Li, Faculty of Information Technology, Beijing University of Technology, 100 Ping Le Yuan, Chaoyang District, Beijing 100124, China.
Email: lixiaolibjut@bjut.edu.cn

Creative Commons CC BY: This article is distributed under the terms of the Creative Commons Attribution 4.0 License (https://creativecommons.org/licenses/by/4.0/) which permits any use, reproduction and distribution of the work without further permission provided the original work is attributed as specified on the SAGE and Open Access pages (https://us.sagepub.com/en-us/nam/open-access-at-sage).
et al.\textsuperscript{13} studies that the hourly average wind speed is predicted by the autoregressive moving average (ARMA) method. This method is used to predict the single sequence PM$_{2.5}$ concentration. This method improves the prediction accuracy of non-stationary wind speed time series. Lee and Tong\textsuperscript{14} uses a hybrid forecasting method to solve nonlinear time series prediction by improved autoregressive integral moving average method. The support vector machine and other statistical methods are also widely applied. Kavousi-Fard et al.\textsuperscript{19} studies that a hybrid forecasting algorithm to forecast power load. The support vector regression is combined with improved firefly method to obtain optimal parameters and improve the prediction accuracy.\textsuperscript{15} Lin et al.\textsuperscript{16} gives an improved combination algorithm of support vector regression (SVR) to forecast seasonal revenues. However, the above-mentioned advanced methods need a large number of sample data, and the algorithm is complex and the amount of calculation is large. In complex environment, the most required variables are usually difficult to obtain. For example, the monitoring of SO$_2$ emissions from coal-fired power plant, the environment of monitoring flue gas is bad. The cost of installing monitoring equipment is high, so the SO$_2$ emissions are difficult to obtain. Thus, it is difficult to apply online in atmospheric environment. For this problem, the Kalman filter (KF) method doesn't need a lot of calculation, and the calculation is simple. The Kalman model can estimate the missing information by limited measurement information. The KF method is widely applied.\textsuperscript{17,18} The traditional KF method can only be used in linear system, and the application scope is limited. The extended Kalman filter (EKF) and unscented Kalman filter (UKF)\textsuperscript{19,20} in nonlinear methods are used to apply in prediction area. Nasser et al.\textsuperscript{19} studies the EKF is combined with Genetic Programming (GP) method to predict the water demand in Tehran. Zhao et al.\textsuperscript{21} proposes a robust iterated extended Kalman filter method. The EKF method needs to linearize the nonlinear function and calculate the Jacobian matrix by the Taylor's formula. The prediction accuracy of this method is not ideal. Qi et al.\textsuperscript{22} proposes an improved UKF method (UKF-GPS) to improve numerical stability for power system. The shortcomings of the EKF method are solved.\textsuperscript{23}

However, most Kalman filter applications need to assume that the noise statistics are accurate or known. Such as noise variance of system and measurement noise, which is correct.

Once the noise statistics are not accurate, the PM$_{2.5}$ concentration prediction based on Kalman method will be inaccurate or even unstable. For this problem, Dai et al.\textsuperscript{24} uses adaptive Kalman filter to estimate battery temperature. Liu and Partovibakhsh\textsuperscript{25} studies adaptive unscented Kalman filter to estimate the state of lithium battery. This method uses covariance matching technology to adjust the noise covariance adaptively, which makes state of charge (SoC) estimation with high accuracy. Xiong et al.\textsuperscript{26} uses the adaptive extended Kalman filter to realize the accurate SOC estimation of battery pack and ensure the accuracy of each battery estimation.

In addition, in the process of PM$_{2.5}$ concentration prediction. The statistical noise is usually assumed to be zero mean Gaussian white noise. In actual situation, the assumption is difficult to realize. Because the actual environment is complex and changeable, it is easy to be affected by external interference, which leads to the deviation of noise distribution, and affects the prediction accuracy. For this problem, an adaptive unscented Kalman filter based on support vector regression is proposed and applied in the area of PM$_{2.5}$ concentration prediction and on-line noise estimation. The prediction accuracy of atmospheric PM$_{2.5}$ concentration is improved from two aspects in this paper. Firstly, a framework of state equation based on support vector regression (SVR) is constructed. The framework of the model takes advantage of the SVR which can be used to do regression under the condition of small samples, and the kernel function can be used to solve the non-linear problems. The SVR method avoids the shortcomings of model over fitting. Secondly, for the randomness and uncertainty of PM$_{2.5}$ concentration, an adaptive unscented Kalman filter (AUKF) is used to update the status continuously, and this method can predict the atmospheric PM$_{2.5}$ concentration when the noise statistics are incorrect or incorrect Gaussian. Then, the proposed method is compared with SVR-UKF. The simulation results show that the noise is estimated adaptively when the noise statistics of atmospheric PM$_{2.5}$ model is incorrect or exactly Gaussian. The accuracy and stability of the proposed method are proved. Finally, the proposed method is compared with SVR-UKF, AR-Kalman, AR and BP method. The simulation results prove the proposed method is more accurate in predicting PM$_{2.5}$ concentration.

The structure of the remainder of this paper is described as follows: the second section describes the development of PM$_{2.5}$ monitoring system. The third section is the PM$_{2.5}$ concentration prediction method. The simulation analysis is shown in fourth section. The last part draws the conclusions.

The development of PM$_{2.5}$ monitoring system

The monitoring system is introduced in this section. Beijing is a city with distinct seasons. The historical data of PM$_{2.5}$ concentration in 2018 is collected in a campus monitoring station area. The data is selected as research object from May to August. About 70% of the monitoring data is used as the training set, and the remaining 30% is the test data. The air monitoring system is shown in Figure 1, and the data is monitored by outdoor monitoring sensor probe. The system of platform equipment is composed of a solar panel power supply system and a power distribution box, which uses solar energy to supply power for the whole equipment.
And the monitoring data is transferred to network cloud. The monitoring data concentration and historical trend are directly accessed the network by the upper computer. The trend of air pollution data can be monitored by trend graph. The operation status and geographical location of equipment at different monitoring areas can be seen in the monitoring area map. The users can monitor normal operation of equipment at different areas by different computer. The monitoring device can also monitor air quality at different times and display the PM$_{2.5}$ mean concentration. The monitored PM$_{2.5}$ concentration is used to verify the proposed method by this system.

Through this monitoring platform, the proposed method in this paper is applied to predict the atmospheric PM$_{2.5}$ concentration in the next hour. Thus, the proposed is mainly introduced in next section.

**Method**

**Support vector regression**

The support vector machine (SVM) has usually classification and regression. The SVM method is one of the SVM methods. The SVR is widely used in the field of regression. In this paper, the SVR is used to predict PM$_{2.5}$ concentration. The SVR method mainy constructs linear decision function in high-dimensional space to realize model prediction. The linear decision function is usually found by solving convex quadratic programming problems. This solution process embodies the principle of structural risk minimization, which avoids the over-fitting phenomenon of traditional methods and improves the prediction ability. The SVR gives an $n$-dimensional input variable and a target variable. The samples $(x_1, y_1), (x_2, y_2), \ldots, (x_i, y_i)$ is usually given with $x_i \in \mathbb{R}^d, y_i \in \mathbb{R}$, $i = 1, 2, \ldots, n$. The regression function of sample is described as follows:

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

where $w$ is regression coefficient, $b$ is the bias, $\langle \cdot, \cdot \rangle$ represents denotes the dot product, and it can be solved by convex optimization problems. When the $f$ exists and approximates $(x_i, y_i)$ with $\varepsilon$ precision. Some errors are allowed. Thus, the slack variables $\xi, \xi^*$ is introduced to cope with infeasible constraints of the optimization problem. The formula is shown as follows:
\[
\min \frac{1}{2} ||w||^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) \quad (2)
\]
\[
s.t. \begin{cases}
    y_i - \langle w, x_i \rangle - b \leq \xi_i \\
    -y_i - \langle w, x_i \rangle + b \leq \xi_i^* \\
    \xi_i, \xi_i^* \geq 0, \quad i = 1, \ldots, N
\end{cases} \quad (3)
\]

where \( C \) is trade off parameter between \( f \) and larger than \( \varepsilon \), \( \varepsilon \) is tolerance coefficient. The dual optimization problem is obtained by using Lagrange method, which is described as follows:

\[
L = \frac{1}{2} ||w||^2 + C \sum_{i=1}^{N} (\xi_i + \xi_i^*) - C \sum_{i=1}^{N} (\eta_i \xi_i + \eta_i^* \xi_i^*) - \sum_{i=1}^{N} \alpha_i (\xi_i + \xi_i^* - y_i + \langle w, x_i \rangle + b) - \sum_{i=1}^{N} \alpha_i^* (\xi_i + \xi_i^* - y_i - \langle w, x_i \rangle - b)
\]

(4)

where \( L \) is Lagrangian, \( \eta_i, \eta_i^*, \alpha_i, \alpha_i^* \) are Lagrangian multipliers. The partial derivatives of \( L \) to \( b, \omega \) and \( \xi_i^* \) are calculated as follows:

\[
\frac{\partial L}{\partial b} = \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) = 0 \quad (5)
\]

\[
\frac{\partial L}{\partial \omega} = \omega - \sum_{i=1}^{N} (\alpha_i^* - \alpha_i) x_i = 0 \quad (6)
\]

\[
\frac{\partial L}{\partial \xi_i^*} = (C - \alpha_i^* - \eta_i^*) = 0 \quad (7)
\]

By substituting formula (5) to (7) into formula (4), the optimization problem is given as follows:

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

(8)

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

\[
f(x) = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) K(x, x_i^*) + b \quad (9)
\]

where \( \langle w, x \rangle \) is calculated as follows:

\[
\langle w, x \rangle = \sum_{i=1}^{N} (\alpha_i - \alpha_i^*) x_i \quad (10)
\]

The Radial Basis Function (RBF) is selected as kernel function in this paper, which is described as follows:

\[
K(x, x_i) = \exp \left( -\frac{||x_i - x_j||^2}{\sigma^2} \right) \quad (11)
\]

where \( \sigma \) is the RBF width. The \( b \) is calculated by Karush-Kuhn-Tucker (KKT) conditions.\(^{29}\)

\[
b = y_i - \langle w, x_i \rangle - \varepsilon, \quad 0 < \alpha_i < C \quad (12)
\]

\[
b = y_i - \langle w, x_i \rangle + \varepsilon, \quad 0 < \alpha_i^* < C \quad (13)
\]

The optimal choice of \( C \) and \( \sigma \) are respectively \( 2^4 \) and 0.0625 by multi-fold cross-validation in this paper. There are always randomness and uncertainty in the prediction process, the UKF based on state space frame of SVR can solve this problem by recursive state estimation.

**The adaptive unscented Kalman filter method**

The sampling points in the original state distribution are selected by a certain rule for this method, so that the mean and covariance of original state distribution are replaced by the sampling points. The state space frame of SVR is given as follows:

\[
x_{k+1} = f(x_k) + w_k \quad (14)
\]

\[
y_k = h(x_k) + v_k \quad (15)
\]

where \( f \) is a nonlinear equation function, \( h \) is the nonlinear observation equation function. \( w_k \) and \( v_k \) are respectively the Gaussian white noise of \( x_k \), \( y_k \). \( Q_k \) is covariance matrix of \( w_k \). The covariance matrix of \( v_k \) is \( R_k \). The mean is \( \bar{x} \), and the variance is \( P \). \( x_k \) denotes the state vector PM2.5 concentration data. \( y_k \) represents the observation PM2.5 concentration data. The \( w \) represents weights. \( n \) represents state dimension. The mean and covariance of the initialization state is given as follows:

\[
\begin{cases}
    x_0 &= E(x_0) \\
    P_0 &= E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T]
\end{cases}
\]

(16)

where \( E \) is expected average, \( \bar{x}_0 \) is mean value, \( P_0 \) is covariance. The steps of calculating the sigma point are as follows:

\[
x^{(0)} = \bar{x}, \quad i = 0 \quad (17)
\]

\[
x^{(i)} = \bar{x} + (\sqrt{n + \lambda} P), \quad i = 1, \ldots, n \quad (18)
\]

\[
x^{(i)} = \bar{x} - (\sqrt{n + \lambda} P), \quad i = 1, \ldots, n \quad (19)
\]

where \((\sqrt{P})^{(i)}(\sqrt{P}) = P, \quad (\sqrt{P})_{ij}\) represents the \( i \) column of matrix square root.

\[
\lambda = \alpha^2(n + \kappa) - n \quad (20)
\]

where \( \lambda \) and \( \alpha \) are respectively scaling parameter. The distribution state of the sigma sampling points is determined by its selection. \( \alpha \) is set to 0.01 in this paper. \( \kappa \) is set to 0. The corresponding weight of the sampling point is calculated as follows:
The update covariance and measurement covariance are obtained as follows:

\[
\Omega_{m}^{(i)} = \frac{\lambda}{n+\lambda}, \quad i = 1\sim 2n
\]

\[
\omega_{c}^{(i)} = \frac{\lambda}{2(n+\lambda)}, \quad i = 1\sim 2n
\]

where \( m \) and \( c \) are mean value and covariance respectively. \( i \) is the number of sampling points. \( \omega \) is nonnegative weight coefficient. In this paper, \( \beta \) is set to 2. The mean and covariance updates of system state is calculated as follows:

\[
x_{k+1/k}^{(i)} = [\hat{x}_{k/k} \hat{x}_{k/k} + (n+\lambda)P_{k/k}] [\hat{x}_{k/k} - (n+\lambda)P_{k/k}]
\]

The one step prediction of calculating sigma point set is obtained as follows:

\[
x_{k+1/k}^{(i)} = f[k, x_{k/k}^{(i)}]
\]

The one step prediction of system state is obtained by the weighted average value as follows:

\[
\hat{x}_{k+1/k} = \sum_{i=0}^{2n} \omega_{c}^{(i)} x_{k+1/k}^{(i)} + q_{k}
\]

The one step prediction of covariance matrix is calculated as follows:

\[
P_{k+1/k} = \sum_{i=0}^{2n} \omega_{c}^{(i)} [\hat{x}_{k+1/k} - x_{k+1/k}^{(i)}] + Q_{k}
\]

The new sigma point is generated by the UT transform according to the predicted value as follows:

\[
x_{k+1/k}^{(i)} = [\hat{x}_{k+1/k} \hat{x}_{k+1/k} + (n+\lambda)P_{k+1/k}] [\hat{x}_{k+1/k} - (n+\lambda)P_{k+1/k}]
\]

The predicted observations are obtained as follows:

\[
y_{k+1/k}^{(i)} = h[x_{k+1/k}]
\]

The predicted mean of the system is respectively described by weighted sum as follows:

\[
\bar{x}_{k+1/k} = \sum_{i=0}^{2n} \omega_{c}^{(i)} y_{k+1/k}^{(i)} + r_{k}
\]

The update covariance and measurement covariance are obtained as follows:

\[
\left\{ \begin{array}{l}
P_{m_{k+1}} = \sum_{i=0}^{2n} \omega_{m}^{(i)} [y_{k+1/k} - \hat{x}_{k+1/k}] [y_{k+1/k} - \hat{x}_{k+1/k}]^{T} + R_{k} \\
P_{m_{k+1}} = \sum_{i=0}^{2n} \omega_{m}^{(i)} [y_{k+1/k} - \hat{x}_{k+1/k}] [y_{k+1/k} - \hat{x}_{k+1/k}]^{T}
\end{array} \right.
\]

The Kalman gain is calculated by formulate (25):

\[
K_{k+1} = P_{m_{k+1}}^{-1} P_{m_{k+1}}^{T}
\]

The state update and covariance update of the system are given as follows:

\[
\hat{x}_{k+1/k+1} = \hat{x}_{k+1/k} + K_{k+1} [y_{k+1/k} - \hat{y}_{k+1/k}]
\]

\[
P_{k+1/k+1} = P_{k+1/k} - K_{k+1} P_{m_{k+1}} K_{k+1}^{T}
\]

where \( \hat{x}_{k+1/k+1} \) represents the predicted PM2.5 concentration at \( (k + 1) \)th time.

When the statistic noise is assumed to set, the SVR-UKF method can predict PM2.5 concentration accurately. However, the noise is usually unknown or inaccurate. For example, assuming that the noise is zero mean Gaussian noise, \( q_{k} / C_{0} \), \( r_{k} / C_{1} \). The actual problem is the opposite. The AUKF method is used to solve this problem. The SVR-AUKF method is applied to estimate noise adaptively when the noise is unknown or inaccurate. The SVR-AUKF method is given as follows:

The mean value of process noise is given as follows:

\[
\dot{q}_{k+1} = (1 - d_{k+1}) \dot{q}_{k} + d_{k+1}
\]

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

\[
d_{k+1} = \frac{(1 - b)}{(1 - b^{k+1})}, \quad k = 0, 1, 2, 3..i
\]

where \( b \) is the forgetting factor. The value is set to 0.98 in this paper. The process noise covariance is estimated as follows:

\[
Q_{k+1} = (1 - d_{k+1}) \dot{Q}_{k} + d_{k+1}
\]

\[
[K_{k+1} P_{k+1/k} K_{k+1}^{T} + P_{k+1/k+1}]
\]

\[
- \sum_{i=0}^{2n} w_{m}^{(i)} (x_{k+1/k} - \hat{x}_{k+1/k}) (x_{k+1/k} - \hat{x}_{k+1/k})^{T}
\]

where \( F_{k+1} \) is given as follows:

\[
F_{k+1} = y_{k+1/k} - \hat{y}_{k+1/k}
\]

The mean value of measured noise is estimated as follows:

\[
\dot{r}_{k+1} = (1 - d_{k+1}) \dot{r}_{k} + d_{k+1} y_{k+1/k} - \sum_{i=0}^{2n} w_{m}^{(i)} h(x_{k+1/k})
\]

The measurement noise covariance is estimated as follows:
The different noise statistics are used for simulation research and comparison in this section. When the process noises are incorrect or unknown, the SVR-AUKF method is compared with SVR-UKF method. The prediction accuracy and validity of SVR-AUKF is verified by comparison. Firstly, the measurement noise ($r_0 = 0.3$, $R_0 = 0.08$) is assumed to be known, and the incorrect process noise ($q_0$, $Q_0$) is selected randomly. Two groups of incorrect process noise values are given. The Tables 1 and 2 give the inaccurate initial values of process noise.

The SVR-AUKF method is improved, the noise is estimated adaptively. The state space framework of SVR-AUKF method is shown in Figure 2.

The flowchart shows the operation process of SVR-AUKF method. Firstly, we select the RBF as kernel function. The state space framework of SVR is established by the monitored PM$_{2.5}$ concentration data. Secondly, the appropriate set of sigma points is selected by a new set of PM$_{2.5}$ concentration data. Then, the recursive calculation of UKF method is carried out by the established SVR framework. Finally, the SVR-AUKF method is realized by this operation process.

**Simulation analysis**

In this section, the advantages and effectiveness of the proposed method are verified by simulation analysis. The PM$_{2.5}$ concentration is collected every hour as sampling period by monitor monitoring data system. The first 700 sets of sample data are used for training data, the other 300 sets are used for test data. The different performance indexes are used to evaluate the prediction precision, which includes mean absolute error (MAE), mean absolute percentage error (MAPE), root mean squared error (RMSE), which is respectively given as follows:

\[
MAE = \frac{1}{n} \sum_{k=1}^{n} |y_{k+1} - \hat{y}_{k+1}|
\]  

\[
MAPE = \frac{1}{n} \sum_{k=1}^{n} \frac{|y_{k+1} - \hat{y}_{k+1}|}{y_{k+1}} \times 100\%
\]  

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

where $y_{k+1}$ represents the actual PM$_{2.5}$ concentration value, $\hat{y}_{k+1}$ is the predicted PM$_{2.5}$ concentration value.

### PM$_{2.5}$ concentration prediction with inaccurate process noises

The different noise statistics are used for simulation research and comparison in this section. When the process noises are incorrect or unknown, the SVR-AUKF method is compared with SVR-UKF method. The prediction accuracy and validity of SVR-AUKF is verified by comparison. Firstly, the measurement noise ($r_0 = 0.3$, $R_0 = 0.08$) is assumed to be known, and the incorrect process noise ($q_0$, $Q_0$) is selected randomly. Two groups of incorrect process noise values are given. The Tables 1 and 2 give the inaccurate initial values of process noise ($q_0$, $Q_0$). In addition, the initial $x_0$ is 157.

The prediction result of SVR-AUKF method is compared with SVR-UKF method. The Figure 3 shows the prediction result of SVR-AUKF and SVR-UKF with the inaccurate process noise. The comparison results are given as follows:

Figure 3 shows the prediction results based on SVR-AUKF method and SVR-UKF method with incorrect noise statistics. It can be seen that when the inaccurate noise statistics are given, the SVR-AUKF can predict PM$_{2.5}$ concentration. The SVR-UKF can also predict PM$_{2.5}$ concentration. SVR-UKF has no noise
Once the noise is changed, the prediction accuracy will be affected. In order to further prove this situation, the Figure 4 gives boxplot of the prediction error based on the two methods.

The Figure 4 shows that the SVR-AUKF method is more accurate in predicting PM$_{2.5}$ concentration. More uniform error distribution. The upper and lower limits of error are smaller. However, SVR-UKF has no noise estimator. It is more susceptible to the given incorrect noise than SVR-AUKF. The prediction accuracy is also affected. For example, the Table 2 gives another set of incorrect noises is given.

The Figure 5 shows the prediction result of SVR-AUKF and SVR-UKF with another set of inaccurate process noise. The comparison results are given as follows:

It is obvious that incorrect noise statistics has a great impact on SVR-UKF from Figures 5 and 6. The prediction error is easy to be affected and changed, and the prediction accuracy will also be affected when noise change. The SVR-AUKF method is not affected greatly, and the noise is estimated adaptively. When the noise of SVR-AUKF method is 0.3, 0.05, 0.3, 0.08. The MAE is 2.7104, the MAPE is 0.9035, the RMSE is 4.2086. The MAE of SVR-UKF is 5.1892, the MAPE is 1.7297, the RMSE is 8.0205. Therefore, the SVR-AUKF method can adaptively realize noise statistics and improve PM$_{2.5}$ concentration prediction accuracy when the process noise is not correct. The SVR-AUKF has better robustness.

In order to further prove that the SVR-AUKF method has higher accuracy, the five different methods are compared. When the noise ($q_0 = 0.2$, $Q_0 = 0.02$, $r_0 = 0.3$, $R_0 = 0.08$) is given randomly, the Figure 7 gives the comparison results. The prediction residuals of five method are described in Figure 8.

In order to further prove the prediction accuracy of the proposed method for PM$_{2.5}$ concentration, the proposed method is compared with other methods. The statistical error of different methods is given as follows:
The Table 3 shows that the performance error of SVR-AUKF method is smallest. The SVR-AUKF method is superior to other method. The proposed SVR-AUKF method not only has PM2.5 concentration prediction accuracy, but it also has higher robustness.

**PM2.5 concentration prediction with inaccurate measurement noise**

In this section, the initial value of process noise ($q_0 = 0.15$, $Q_0 = 0.06$) is assumed to be known. As shown in Tables 4 and 5. Two groups of incorrect measurement noise values are given, which proves SVR-UKF with no noise estimator will be affected, the prediction accuracy is not as high as SVR-AUKF. The Table 4 shows the inaccurate initial values of measurement noise ($r_0$, $R_0$). The accuracy of SVR-AUKF and SVR-UKF method is compared. The initial values of measurement noise are given as follows:

![Figure 7](image1.png)  
**Figure 7.** The comparison results of PM2.5 concentration prediction based on five methods.

![Figure 8](image2.png)  
**Figure 8.** Boxplot of the prediction residuals based on five method.

The Table 3 shows that the performance error of SVR-AUKF method is smallest. The SVR-AUKF method is superior to other method. The proposed SVR-AUKF method not only has PM2.5 concentration prediction accuracy, but it also has higher robustness.

**Table 3.** The statistical error of different method.

| Method      | MAE   | MAPE (%) | RMSE  |
|-------------|-------|----------|-------|
| SVR-AUKF    | 2.9736| 0.9912   | 4.6270|
| SVR-UKF     | 5.1892| 1.7297   | 8.0205|
| AR-Kalman   | 6.0400| 2.0133   | 8.6136|
| AR          | 6.4632| 2.1544   | 9.5624|
| BP          | 9.9571| 3.3190   | 12.3944|

**Table 4.** The given process statistics and the inaccurate measurement noise.

| Method      | $q_0$ | $Q_0$ | $r_0$ | $R_0$ |
|-------------|-------|-------|-------|-------|
| SVR-AUKF    | 0.15  | 0.06  | 0.2   | 0.04  |
| SVR-UKF     | 0.15  | 0.06  | 0.2   | 0.04  |

**Table 5.** The given process statistics and the inaccurate measurement noise.

| Method      | $q_0$ | $Q_0$ | $r_0$ | $R_0$ |
|-------------|-------|-------|-------|-------|
| SVR-AUKF    | 0.15  | 0.06  | 0.25  | 0.02  |
| SVR-UKF     | 0.15  | 0.06  | 0.25  | 0.02  |

![Figure 9](image3.png)  
**Figure 9.** PM2.5 concentration prediction based on SVR-AUKF and SVR-UKF with the inaccurate measurement noise.

The Figure 9 shows the prediction result of SVR-AUKF and SVR-UKF with the inaccurate measurement noises. The comparison results are given as follows:

The Figures 9 and 10 shows that the prediction accuracy of the SVR-AUKF method with the inaccurate measurement noise. The prediction accuracy of the SVR-AUKF method is higher than the SVR-UKF method. The inaccurate noise is easy to produce large error. When the inaccurate measurement noise is given, the SVR-AUKF method still has higher prediction accuracy, and it is not affected by the noise, which can
be estimate adaptively. The proposed method is superior to SVR-UKF, and it has good robustness, and the effectiveness of SVR-AUKF method is also proved. In order to prove that SVR-UKF method is easily affected without noise estimator, another set of incorrect measurement noise is given as follows:

The Figure 11 shows the prediction result of SVR-AUKF and SVR-UKF with the inaccurate measurement noises. The comparison results are given as follows:

It can be seen that the prediction accuracy of SVR-AUKF method is still more accurate from Figures 11 and 12. The Figure 12 shows the prediction residuals of SVR-AUKF and SVR-UKF. The error upper limit distribution of SVR-AUKF is smaller and more uniform. When another inaccurate measurement noise is given, the SVR-AUKF method still has higher prediction accuracy, it is not affected and estimate adaptively. In order to better prove the prediction accuracy of the proposed method, when the noise \( q_0 = 0.16, Q_0 = 0.09, r_0 = 0.2, R_0 = 0.40 \) is given randomly, the SVR-AUKF method is compared with other methods. The Figure 13 gives the comparison results of the different methods.

From the Figures 13 and 14, it can be concluded that the SVR-AUKF method with noise estimator is more accurate and robust in predicting PM\(_{2.5}\) concentration. The error statistical analysis of Table 6 is given. The MAE of SVR-AUKF is 3.0418, the MAPE is 1.0139, the RMSE is 5.0469. The MAE of AR-Kalman method is 5.7451, the MAPE is 1.9150, the RMSE is 8.2389. The MAE of SVR-UKF method is 4.8615, the MAPE is 1.6205, the RMSE is 7.6176. The error of SVR-AUKF method is the smallest in five methods. Due to the SVR-UKF method lacks noise estimator,
the SVR-UKF method is easy to be affected with noise. The proposed SVR-AUKF method has higher prediction accuracy than other methods, the statistical error is smaller and the robustness is better. The statistical error is shown as follows:

### Conclusions

In this paper, a hybrid modeling method based on support vector regression state space and adaptive unscented Kalman filter is proposed. Firstly, the support vector regression model is established and the state space framework of PM$_{2.5}$ concentration series is described. Then, the adaptive unscented Kalman filter is used to dynamically update the state when the process and measurement noise are unknown or incorrect, and the noise is correctly estimated in the iterative process, so as to improve the prediction accuracy of PM$_{2.5}$ concentration. Finally, the proposed method is compared with the SVR-UKF method which lacks the ability of noise estimation, the results show that the proposed method can not only accurately predict the PM$_{2.5}$ concentration, but also has higher robustness when the process and measurement noises are incorrect. The proposed is compared with SVR-UKF, AR-Kalman, AR and BP methods, the simulation results prove that the proposed method has higher prediction accuracy in predicting PM$_{2.5}$ concentration.

However, the PM$_{2.5}$ concentration always changes with complex environment in atmospheric environment. The single model method is difficult to satisfy the current changing environment. Thus, the PM$_{2.5}$ concentration prediction of multiple models may be considered and studied in the future.

### Acknowledgement

The author would like to thank other members for their contributions to this work.

### Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

### Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This study is support by National Key Research and Development Project (2018YFB1702704, 2018YFC1602704) and the National Natural Science Foundation of China (61873006, 61673053)

### ORCID iD

Jihan Li https://orcid.org/0000-0003-3490-5820

### Data availability

The code for this study can be obtained from the corresponding author.

### References

1. Wang Z, Fang C and Xu G. Spatial-temporal characteristics of the PM$_{2.5}$ in China in 2014. Acta Geogr Sin 2015; 70(11): 1720–1734.
2. Haywood JM, Roberts DL, Slingo A, et al. General circulation model calculations of the direct radiative forcing by anthropogenic sulfate and fossil-fuel soot aerosol. J Clim 1997; 10(7): 1562–1577.
3. Hunt A, Abraham JL, Judson B, et al. Toxicologic and epidemiologic clues from the characterization of the 1952 London smog fine particulate matter in archival autopsy lung tissues. Environ Health Perspect 2003; 111(9): 1209–1214.
4. Schiller R. The Los Angeles smog. Nat Munic Rev 1955; 44(11): 558–564.
5. Yan S, Cao H, Chen Y, et al. Spatial and temporal characteristics of air quality and air pollutants in 2013 in Beijing. Environ Sci Pollut Res Int 2016; 23(14): 13996–14007.
6. Gao M, Guttiukonda SK, Carmichael GR, et al. Health impacts and economic losses assessment of the 2013 severe haze event in Beijing area. Sci Total Environ 2015; 511(1): 553–561.
7. Liang F, Xiao Q, Gu D, et al. Satellite-based short- and long-term exposure to PM$_{2.5}$ and adult mortality in urban Beijing, China. Environ Pollut 2018; 242: 492–499.
8. Zheng S, Yi H and Li H. The impacts of provincial energy and environmental policies on air pollution...
control in China. *Renew Sustain Energy Rev* 2015; 49: 386–394.
9. Tai APK, Mickley LJ and Jacob DJ. Correlations between fine particulate matter (PM$_{2.5}$) and meteorological variables in the United States: implications for the sensitivity of PM$_{2.5}$ to climate change. *Atmos Environ* 2010; 44(32): 3976–3984.
10. Zhou Q, Jiang H, Wang J, et al. 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(15): 264–274.
11. Elangasinghe MA, Singhal N, Dirks KN, et al. Complex time series analysis of PM$_{10}$ and PM$_{2.5}$ for a coastal site using artificial neural network modelling and k-means clustering. *Atmos Environ* 2014; 94: 106–116.
12. Feng X, Li Q, Zhu Y, et al. Artificial neural networks forecasting of PM$_{2.5}$ pollution using air mass trajectory based geographic model and wavelet transformation. *Atmos Environ* 2015; 107: 118–128.
13. Erdem E and Shi J. ARMA based approaches for forecasting the tuple of wind speed and direction. *Appl Energy* 2011; 88(4): 1405–1414.
14. Lee YS and Tong LI. Forecasting time series using a methodology based on autoregressive integrated moving average and genetic programming. *Knowl Based Syst* 2011; 24(1): 66–72.
15. Kavousi-Fard A, Samet H and Marzbani F. A new hybrid Modified Firefly Algorithm and Support Vector Regression model for accurate short-term load forecasting. *Expert Syst Appl* 2014; 41(13): 6047–6056.
16. Lin KP, Pai PF, Lu YM, et al. Revenue forecasting using a least-squares support vector regression model in a fuzzy environment. *Inf Sci* 2011; 220(1): 196–209.
17. Cassola F and Burlando M. Wind speed and wind energy forecast through Kalman filtering of numerical weather prediction model output. *Appl Energy* 2012; 99: 154–166.
18. Liu H, Tian HQ and Li YF. Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction. *Appl Energy* 2012; 98: 415–424.
19. Nasseri M, Moeini A and Tabesh M. Forecasting monthly urban water demand using Extended Kalman Filter and genetic programming. *Renew Energy* 2016; 85: 790–809.
20. Bisio R and Dash PK. A hybrid evolutionary dynamic neural network for stock market trend analysis and prediction using unscented Kalman filter. *Appl Soft Comput* 2014; 19: 41–56.
21. Zhao J, Netto M and Mill L. A robust iterated extended Kalman filter for power system dynamic state estimation. *IEEE Trans Power Syst* 2017; 32(4): 3205–3216.
22. Qi J, Sun K, Wang J, et al. Dynamic state estimation for multi-machine power system by unscented Kalman filter with enhanced numerical stability. *IEEE Trans Smart Grid* 2016; 9(2): 1184–1196.
23. Chen K and 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.
24. Dai H, Zhu L, Zhu J, et al. Adaptive Kalman filtering based internal temperature estimation with an equivalent electrical network thermal model for hard-cased batteries. *J Power Sources* 2015; 293: 351–365.
25. Liu G and Partovibakhsh M. An adaptive unscented Kalman filtering approach for online estimation of model parameters and state-of-charge of lithium-ion batteries for autonomous mobile robots. *IEEE Trans Control Syst Technol* 2014; 293(2): 357–363.
26. Xiong R, Sun F, Gong X, et al. Adaptive state of charge estimator for lithium-ion cells series battery pack in electric vehicles. *J Power Source* 2013; 242(15): 699–713.
27. Brereton RG and Lloyd GR. Support vector machines for classification and regression. *Analyst* 2009; 23: 230–267.
28. Smola AJ and Scholköpf B. A tutorial on support vector regression. *Stat Comput* 2004; 14(3): 199–222.
29. Gass SI and Fu MC. Karush-Kuhn-Tucker (KKT) conditions. In: Gass SI (ed.) Encyclopedia of operations research & management science. Boston, MA: Springer, 2013, pp.833–834.