Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 8914501, 13 pages
https://doi.org/10.1155/2020/8914501

Research Article

Short-Term Air Quality Prediction Based on Fractional Grey Linear Regression and Support Vector Machine

Meng Dun, Zhicun Xu, Yan Chen, and Lifeng Wu

School of Management Engineering and Business, Hebei University of Engineering, Handan 056038, China

Correspondence should be addressed to Lifeng Wu; wlf6666@126.com

Received 21 January 2020; Revised 7 April 2020; Accepted 2 May 2020; Published 18 May 2020

Academic Editor: Ana C. Teodoro

To predict the daily air pollutants, the fractional multivariable model is established. The hybrid model of the grey multivariable regression model with fractional order accumulation model (FGM(0, m)) and support vector regression model (SVR) is used to predict the air pollutants (PM_{10}, $PM_{2.5}$, and NO_2) from December 31, 2018, to January 3, 2019, in Shijiazhuang and Chongqing. The absolute percentage errors (APEs) are used to determine the weights of the FGM(0, m) and SVR. Meanwhile, the Holt–Winters model is used to predict the air quality pollutants for the same location and period. When the mean absolute percent error (MAPE) is 0%–20%, it indicates that the model has good accuracy of fitting and prediction. The MAPE of the hybrid model is less than 20%. It is shown that except for the $PM_{2.5}$ concentration prediction in Shijiazhuang (13.7%), the MAPE between the forecasting and actual values of the three air pollutants in Shijiazhuang and Chongqing was less than 10%.

1. Introduction

According to the statistical data in China [1], the 338 cities had an average of 79.3% of days with good air quality (meet the air quality standard), which increased to 1.3% compared with 2017. The number of days with heavy pollution was 2.2 percent, which fell to 0.3% compared with 2017. The $PM_{2.5}$ concentration was 39 microgram/m^3, which fell to 9.3% compared with 2017. The concentration of PM_{10} was 71 microgram/m^3, which fell to 5.3% compared with 2017. The air quality in China improved in 2018 on a whole, but only 121 of 338 cities meet air quality standards as shown in Table 1. When the concentration of the air pollutants (PM_{10}, $PM_{2.5}$, and NO_2) meets the standard, the air quality is regarded as good. Otherwise, the air quality will be regarded as poor. The 24-hour air pollutant standard implemented since 2016 in China is shown in Table 1 [2]. In addition, there were 822 days of severe pollution, 20 more than that of 2017. It indicates that the governance of air quality is still a problem that cannot be ignored.

In recent years, air quality has attracted more and more attention, and more and more research studies have been done on air quality. The impact of foreign direct investment and research as well as development on China's industrial CO_2 emission reduction has been studied and its trend has been predicted [3]. A seasonal stacked autoencoder model combining seasonal analysis and deep feature learning was proposed for forecasting the hourly $PM_{2.5}$ concentration in Beijing [4]. An integrated short-duration memory neural network was proposed for the prediction of hourly $PM_{2.5}$ concentration in Beijing [5]. The trend of the observational PM_{10} concentrations in Shimla city, India, was analyzed [6]. The predictive models can be divided into two categories (single model and the hybrid model). Some scholars used a single model to study air quality, and the multigene genetic programming was used to predict the concentrations of PM_{10} [7]. The grey Markov model was used to predict the concentration of air pollutants in Pingdingshan [8]. A single dependent variable partial least squares regression was used to predict $PM_{2.5}$ real-time concentration in Beijing [9]. The grey Holt–Winters Model was used to predict the air quality indexes of Shijiazhuang and Handan [10]. A microscale land use regression model was used to predict NO_2 concentrations at a heavy trafficked suburban area in Auckland, NZ [11]. The seasonal grey one variable model with fractional order accumulation was used to predict air quality indexes of

Table 1: 24-hour air pollutant standard (implemented since 2016 in China).

	Meet the standard	Fail to meet the standard
PM_{10}	0~150 $\mu g/m^3$	>150 $\mu g/m^3$
$PM_{2.5}$	0~75 $\mu g/m^3$	>75 $\mu g/m^3$
NO_2	0~80 $\mu g/m^3$	>80 $\mu g/m^3$

Xingtai and Handan [12]. The optimized particle swarm was used to predict the concentration of air pollutants in Aburrá Valley, Colombia [13]. The empirical mode decomposition based on the multifractal detrended fluctuation analysis method was used to study the daily $PM_{2.5}$ concentration in Hong Kong [14]. A land use regression model was used to estimate annual and seasonal PM_1, $PM_{2.5}$, and PM_{10} concentrations [15]. Many scholars combined two different models to study the air quality, and the hybrid model of the regression models and feedforward backpropagation models with principle component analysis was used to predict the daily PM_{10} concentrations [16]. The mixed model of information gain and least absolute shrinkage was used to predict the air quality index [17]. A linear and an artificial neural network statistical model have been developed and validated and established to forecast the short-term PM_{10} hourly concentrations in the city of Brescia (Italy) [18]. The mixed air quality assessment model was designed and applied to analyze the pollution sources of $PM_{2.5}$ [19]. A mixed forecasting model of daily air quality index considering air pollution factors in Beijing and Guilin was proposed [20]. A hybrid particle swarm optimization-support vector machine model based on clustering algorithm was used to forecast the short-term atmospheric pollutant concentration in Beijing [21]. A hybrid multiresolution multiobjective ensemble model was used to forecast the daily $PM_{2.5}$ concentrations [22]. The hybrid model of autoencoder with bidirectional long short-term memory neural networks was used to predict the $PM_{2.5}$ concentration [23].

In recent years, more and more hybrid models have been used by scholars to predict. However, few of them will use the hybrid model of the artificial intelligence algorithm and statistical algorithm and few scholars used the hybrid model to predict the air pollutants. It has been proved that the hybrid models with good prediction effect in M4 are the combination of artificial intelligence algorithm and statistical algorithm. In order to improve the prediction accuracy, a hybrid grey multivariable regression model with fractional order accumulation model [24] and support vector regression [25] model (FGM(0, m)-SVR) model is proposed to predict air pollutants ($PM_{2.5}$, PM_{10}, and NO_2) in this paper.

This paper is divided into five parts. In Section 2, the situation in Shijiazhuang and Chongqing is introduced. In Section 3, the hybrid model is introduced. In Sections 4 and 5, the process of calculation and the results of Shijiazhuang and Chongqing are shown, respectively. Through the analysis of the calculation results, some suggestions for the air quality in Shijiazhuang and Chongqing are given in Section 6. Meanwhile, the conclusions are summarized in Section 6.

2. Location and Air Quality in Shijiazhuang and Chongqing

2.1. Location in Shijiazhuang and Chongqing. Shijiazhuang is the capital city of Hebei Province. It is located in the north China plain, which is adjacent to Beijing and Tianjin in the north, Bohai in the east, Taihang Mountain in the west, and the economic zone in the south. Shijiazhuang is the gate of the capital city, 273 kilometers away from Beijing. It is located between latitude 37°27′~38°47′ and longitude 113°30′~115°20′ (as shown in Figure 1). Shijiazhuang is one of the most polluted cities in China. According to the statistics in 2018, Shijiazhuang ranked 168th among 169 cities with poor air quality in China.

Chongqing is an economic, financial, scientific, and technological innovation, shipping and commercial logistics center in the upper basins of the Yangtze River. It is located in the southwest of China's inland, Hubei and Hunan in the east, Guizhou in the south, Sichuan in the west, and Shanxi in the north. Chongqing is located longitude 105°17′~110°11′ and latitude 28°10′~32°13′ (as shown in Figure 1). Chongqing is also a heavily polluted city. Compared with previous years, the condition of air quality in Chongqing has improved significantly in 2019. But it is still a long way from the goal set by the Chongqing Ecology And Environment Bureau that ensures the number of days with good air quality in 2019 stays above 300. Chongqing has been known as "the city of fog" and the air quality ranks behind other cities in China.

2.2. Air Quality in Shijiazhuang and Chongqing. The number of days with good air quality (up to the air quality standard) as shown in Table 2 is obtained from the website of Shijiazhuang Environmental Protection Bureau (http://www.sjzhb.gov.cn) and Chongqing Environmental Protection Bureau (http://www.cepb.gov.cn/), respectively. The days with good air quality in Shijiazhuang and Chongqing from 2014 to 2018 can be seen from Figure 2 clearly.

As shown in Figure 2, despite the increasing efforts of government governance, the number of days with bad air quality had been increasing since 2015. According to the statistical data of the Hebei Province Environment Protection Hall in 2018, the air quality of Shijiazhuang is the worst in Hebei province. In addition, according to the statistics of China Environment Network, the air quality of Shijiazhuang was the worst among the 11 cities when it was ranked by air quality composite index. However, the target of air quality has been proposed in the "Three-year Action Plan for Shijiazhuang City to Win the Blue Sky Defense War (2018–2020)," and the days with good air quality in Shijiazhuang will exceed 176 days in 2019. It is mentioned in the plan that Shijiazhuang will complete the targets of "The 13th Five-year for Economic and Social Development of the People's obligatory" for air environmental quality until 2020, the main atmospheric pollutants emissions will be cut, and the rank of air quality strives to exit from the last 10 in 169 cities of China. Therefore, how to effectively predict the air quality is particularly important.

Figure 1: The geographical location of Shijiazhuang and Chongqing.

Table 2: Air quality standard.

Air quality index (unit: $\mu g/m^3$)	If up to standard
0–100	Up to standard
More than 100	Not up to standard

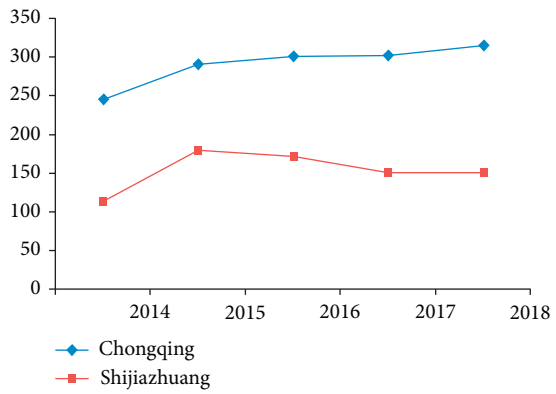

Figure 2: Days with good air quality in Shijiazhuang and Chongqing.

The number of days with heavy pollution in Chongqing is increasing year by year. The number of days with good air quality in Chongqing reached 316 in 2018. At the same time, the arrangements for the environmental protection work of 2019 had been made in the teleconference of Chongqing environmental protection on January 21, 2019: ensuring that the number of days with good air quality remains above 300 and that the average annual concentration of fine particulate matter is kept within 40 micrograms per cubic meter. From 2016 to 2018, the number of days with good air quality in Chongqing was 301, 303, and 316, which is inseparable from the government's governance measures. If the air pollutant can be predicted more accurately, the air pollution control will be more effective.

3. The Construction of Model

By accumulating generation operators, the FGM(0, m) model can transform the data from nonlinear to linear. At the same time, the data can be mapped from nonlinear to linear by using the SVR model. The historical parallelism was existed in the data of the same period. By processing the data through FGM(0, m) and SVR, better accuracy of prediction will be achieved, so the FGM(0, m) model and SVR model are used to forecast the air pollutants in this paper. In China, the number of people who choose to travel during the New Year's Day is larger, and people's travel is also affected by the environment, so it is particularly important to predict the air pollutants more effectively. Meanwhile, taking the air pollutant (PM_{10}) in Shijiazhuang as an example, the Holt–Winters model is used to calculate the fitting and predictive values. The basis of the Holt–Winters model is that the time series with linear trend, seasonal change, and random fluctuation can be decomposed and combines with exponential smoothing method to establish a forecasting model. In recent years, the Holt–Winters model is often used to predict the seasonal data, and it was used to predict the air pollutants in Shijiazhuang and Handan [10]. The Holt–Winters model was used to predict the air pollutant concentration (PM_{10}) in Shijiazhuang, and the results are contrasted with the hybrid model in this paper.

3.1. The Model of FGM(0, m). The GM(0, m) model has been widely used in recent years, and it was used to analyze the influence factor for the construction of the model in vocational high school [26]. The GM(0, m) model was used to analyze the consumer experience on small- and medium-sized enterprises in creative living industry [27]. In this paper, the hybrid model of FGM(0, m) and SVM is applied to predict air pollutants. The FGM(0, m) model is introduced as follows:

$X_1^{(0)} = (x_1^{(0)}(1), x_1^{(0)}(2), \cdots, x_1^{(0)}(n))$ is the system characteristics sequence.

$$\begin{aligned} X_2^{(0)} &= \left(x_2^{(0)}(1), x_2^{(0)}(2), \cdots, x_2^{(0)}(n)\right), \\ X_3^{(0)} &= \left(x_3^{(0)}(1), x_3^{(0)}(2), \cdots, x_3^{(0)}(n)\right), \\ &\vdots \\ X_m^{(0)} &= \left(x_m^{(0)}(1), x_m^{(0)}(2), \cdots, x_m^{(0)}(n)\right), \end{aligned} \tag{1}$$

are the sequences of related factors.

Then,

$$x_j^{(p/q)}(k) = \sum_{i=1}^{k} C_{k-i+((p/q)-1)}^{k-i} x_j^{(0)}(i), \tag{2}$$

is the p/q order accumulation generate operator. $C_{(p/q)-1}^{0} = 1$, $C_k^{k+1} = 0$, $k = 0, 1, \cdots, n-1$:

$$C_{k-i+((p/q)-1)}^{k-i} = \frac{(k-i+((p/q)-1)(k-i+((p/q)-2)))\cdots((p/q)+1)(p/q)}{(k-i)!}, \tag{3}$$

$X_j^{(p/q)} = (x_j^{(p/q)}(1), x_j^{(p/q)}(2), \cdots, x_j^{(p/q)}(n))$ is the p/q order accumulation sequence:

$$\begin{aligned} B &= \begin{bmatrix} 1 & x_2^{p/q}(1) & \cdots & x_m^{p/q}(1) \\ 1 & x_2^{p/q}(2) & \cdots & x_m^{p/q}(2) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_2^{p/q}(n) & \cdots & x_m^{p/q}(n) \end{bmatrix}, \\ Y &= \begin{bmatrix} x_1^{p/q}(1) \\ x_1^{p/q}(2) \\ \vdots \\ x_1^{p/q}(n) \end{bmatrix}, \\ [a, b_2, \cdots, b_m]^{\mathrm{T}} &= \left(B^{\mathrm{T}}B\right)^{-1}B^{\mathrm{T}}Y. \end{aligned} \tag{4}$$

Therefore, the equation of the FGM(0, m) model is given by

$$\begin{aligned} \hat{x}_1^{(p/q)}(k) &= a + b_2 x_2^{(p/q)}(k) + b_3 x_3^{(p/q)}(k) + L + b_n x_n^{(p/q)}(k) \\ &\quad \cdot (k = 1, 2, \cdots), \end{aligned} \tag{5}$$

where $a, b_2, b_3, \cdots, b_m$ are model parameters.

3.2. The Model of SVR. The SVR model is a linear separable model based on kernel functions that converts linearly indivisible data into high-dimensional space. In this paper, SVR model is used to learn the historical data, and according to the learning results, the model will make predictions. For the training sample set: $\left\{x_j^{(0)}(i), x_1^{(0)}(i)\right\}$, $j = 2, 3, \cdots, m$.

$x_j^{(0)}(i)$, $j = 1, 2, \cdots, m$ represents the input data of SVR, and $x_1^{(0)}(i)$ $(i = n+1, n+2, \cdots)$ represents the output data of SVR. The function model of SVR is given by

$$\hat{x}_1^{(0)} = w^{\mathrm{T}}\varphi\left(x_j^{(0)}\right) + b. \tag{6}$$

Among them, $\hat{x}_1^{(0)}$ represents predicted value, $\varphi(\bullet)$ represents a nonlinear mapping function, w represents the weight vector, and b is the bias.w and b give us the following formula:

$$\begin{aligned} &\min J = \frac{1}{2}|w|^2 + C\sum_{i=1}^{n}\left(\zeta_i^+ + \zeta_i^-\right), \\ &\text{s.t.} \begin{cases} y_i - w^{\mathrm{T}}\varphi\left(x_j^{(0)}(i)\right) - b \le \varepsilon + \zeta_i^+, \\ w^T\varphi\left(x_j^{(0)}(i)\right) + b - y_i \le \varepsilon + \zeta_i^-, \\ \zeta_i^+, \zeta_i^- \ge 0, \quad i = 1, 2, \cdots, n;\ j = 1, 2, \cdots, m, \end{cases} \end{aligned} \tag{7}$$

where C and ε represent the maximum error coefficients of the penalty coefficient and insensitive loss function, respectively. ζ_i^+ and ζ_i^- represent the relaxation coefficients. n represents the sample size of input data.

The weight vector w can be expressed as follows:

$$w = \sum_{i=1}^{n}(\beta_i^* - \beta_i)\varphi\left(x_j^{(0)}(i)\right), \quad j = 2, 3, \cdots, m, \tag{8}$$

where β_i^* and β_i represent Lagrange coefficient, respectively. The mathematical model equation of SVR is given by

$$\hat{y}_2 = \sum_{i=1}^{n}(\beta_i^* - \beta_i)K\left(x_1^{(0)}(i), x_j^{(0)}(i)\right) + b, \tag{9}$$

where $K(\cdot)$ represents the kernel function for calculating the inner product of two input vectors in high-dimensional eigenvectors. The function of sigmoid is used in this paper.

3.3. The Basis of Model Weights. Through the statistical interpretation of multiple prediction models, a robust short-term forecast of wind power generation under uncertainty is proposed [28]. A new method to determine the weight of the hybrid model is proposed. The root mean square errors (RMSEs) of model are used to determine the weight of the models. The RMSE is used as a performance index, which indicates the accuracy of the forecasting models due to awareness of RMSE over large errors in prediction. In this paper, the APEs are used to determine weights according to the same principle.

3.4. The Hybrid Model of FGM(0, m) and SVR. In this paper, the fitting values and predictive values obtained by FGM(0, m) model and SVR model are given different weights, respectively. Then, the results obtained by the method described above are summed up, which is the final result. The calculation process is as follows:

$$\begin{aligned} \mathrm{APE}_{\mathrm{FGM}(0,\mathrm{m})} &= \frac{\left|\widehat{x}^{(0)}_{\mathrm{FGM}(0,\mathrm{m})} - x^{(0)}_1\right|}{x^{(0)}_1}, \\ \mathrm{APE}_{\mathrm{SVR}} &= \frac{\left|\widehat{x}^{(0)}_{\mathrm{SVR}} - x^{(0)}_1\right|}{x^{(0)}_1}, \end{aligned} \tag{10}$$

where $\widehat{x}_{\mathrm{FGM}(0,\mathrm{m})}$ is the forecasting value of the FGM(0, m) model and $\widehat{x}_{\mathrm{SVR}}$ is the forecasting result of the SVR model. $x^{(0)}_1$ is the actual values.

The weights of the two calculation results are divided, and then, the results are summed up:

$$\omega_{\mathrm{FGM}(0,\mathrm{m})} = \frac{\mathrm{APE}^{-1}_{\mathrm{FGM}(0,\mathrm{m})}}{\mathrm{APE}^{-1}_{\mathrm{FGM}(0,\mathrm{m})} + \mathrm{APE}^{-1}_{\mathrm{SVR}}}, \tag{11}$$

$$\omega_{\mathrm{SVR}} = \frac{\mathrm{APE}^{-1}_{\mathrm{SVR}}}{\mathrm{APE}^{-1}_{\mathrm{FGM}(0,\mathrm{m})} + \mathrm{APE}^{-1}_{\mathrm{SVR}}}, \tag{12}$$

$$\widehat{x}^{(0)}_1 = \omega_{\mathrm{FGM}(0,\mathrm{m})} \times \widehat{x}^{(0)}_{\mathrm{FGM}(0,\mathrm{m})} + \omega_{\mathrm{SVR}} \times \widehat{x}^{(0)}_{\mathrm{SVR}}. \tag{13}$$

The APE is calculated between the final results and the actual values:

$$\mathrm{APE} = \frac{\left|\widehat{x}^{(0)}_1 - x^{(0)}_1\right|}{x^{(0)}_1}. \tag{14}$$

The MAPE from December 20 to 30 in 2018 is taken as the fitting error. The MAPE from December 31, 2018, to January 3, 2019, is taken as the predictive error. The process of modeling is shown in Figure 3.

4. The Calculation Process and Results in Shijiazhuang

In order to verify the accuracy of the hybrid model, three air pollutants in Shijiazhuang are used (PM_{10}, $PM_{2.5}$, and NO_2) from December 20 to January 3 between 2014 and 2017 and December 20 to December 30, 2018, as the samples for forecasting. The FGM(0, m) model and the SVM model are used to predict the values from December 31 in 2018 to January 3 in 2019. According to equations (11)-(12), the weights of the FGM(0, m) model and the SVM model are determined. Then, the calculated results of the FGM(0, m) model and SVR model are multiplied by the above weights, respectively. According to equation (13), the fitting and forecasting values of the hybrid model can be obtained. The accuracy of the hybrid model is proved by calculating the MAPE between the final results and the actual values of the three models. The original data of the air pollutants are obtained from the website (http://www.airitilibrary.com/Publication/alDetailedMesh?docid=10289488-201903-201903060025-201903060025-s19-28). There are eight monitoring stations, and seventeen monitoring stations have been set up in Shijiazhuang and Chongqing, respectively. According to the mean of the pollutant concentration in each monitoring station for a natural day (24 hours), the 24-hour air pollutant concentration can be obtained.

4.1. Calculation Results of PM_{10}

4.1.1. Calculation Result of PM_{10} by Using FGM(0, m) Model. Air quality has attracted more and more attention of the government in recent years. Since the meteorological indicators in the same historical period are similar, the meteorological conditions (favourable or unfavourable to pollutant dispersion) affect air quality, and similar meteorological condition generally affects air quality in a similar way. The governance measures and efforts of governments in the same region are not different from each other, and it accounts for that the air quality in the same region is also similar. The same period data (December 20 to January 3) from 2014 to 2017 are used as the independent variables ($X^{(0)}_j$), and the data from December 20 to January 3, 2018, are used as the dependent variable ($X^{(0)}_1$). Data of December 20 to December 30, 2018, are fitted and data of December 31, 2018, to January 3, 2019, are forecasted. The calculation process of PM_{10} in Shijiazhuang is taken as an example, and the original data of PM_{10} in Shijiazhuang is shown in Table 3.

FGM(0, m) model is used to calculate the fitting values from December 20 to December 30, 2018, and the values from December 31, 2018, to January 3, 2019, are predicted. The calculation process is shown as follows.

A FGM(0, 4) model is established:

$$X^{(0)}_1 = (215, 153, 228, 132, 193, 248, 114, 78, 68, 79, 122), \tag{15}$$

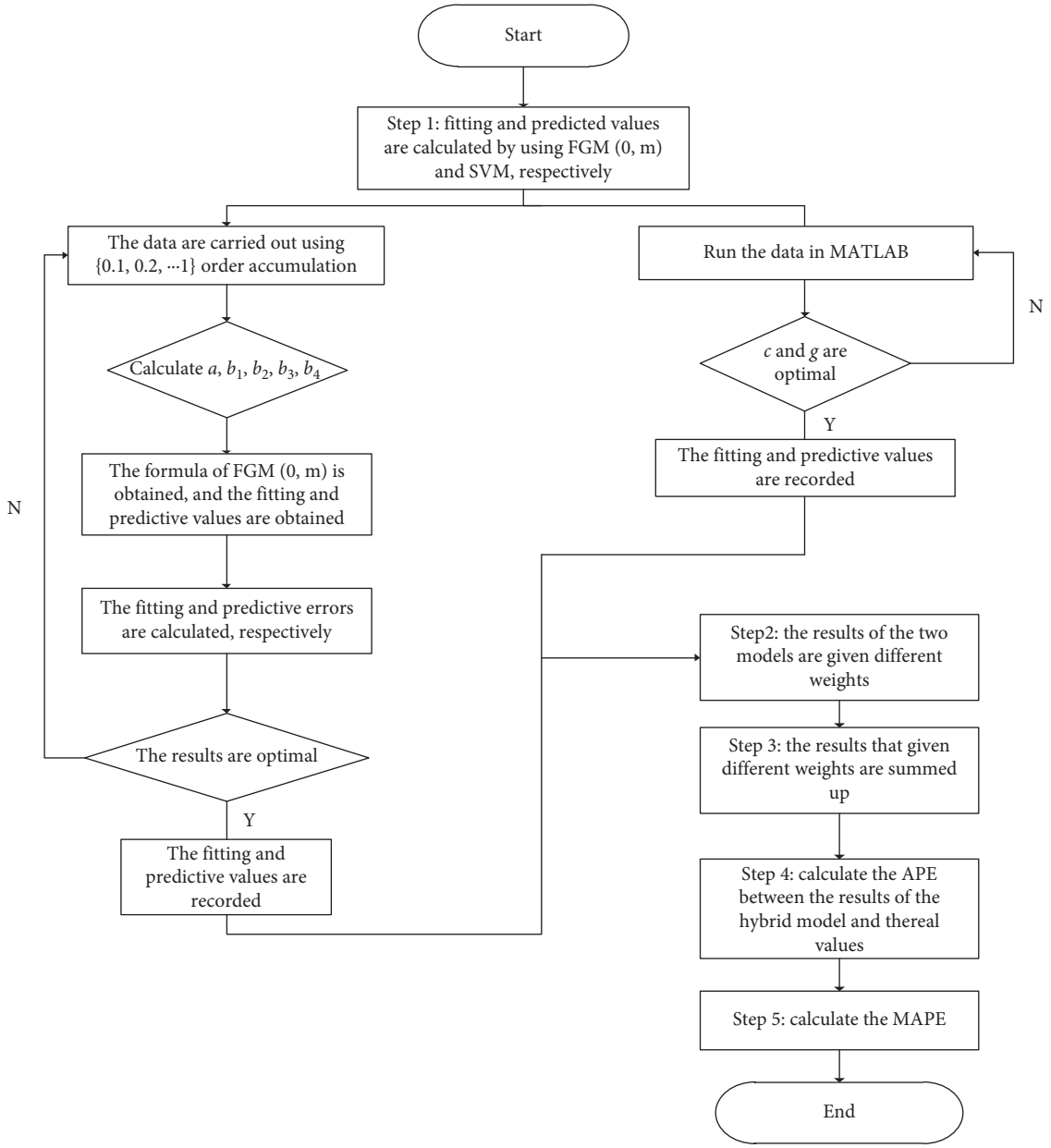

Figure 3: The process of modeling.

is the system characteristics sequence, and

$$\begin{aligned} X_2^{(0)} &= (87, 97, 108, 333, 237, 191, 353, 478, 237, 218, 192, 122, 156, 220, 358), \\ X_3^{(0)} &= (238, 490, 383, 412, 376, 295, 278, 51, 174, 378, 155, 207, 414, 617, 485), \\ X_4^{(0)} &= (866, 765, 181, 142, 258, 282, 298, 170, 342, 308, 413, 710, 617, 493, 471), \\ X_5^{(0)} &= (85, 188, 268, 304, 117, 116, 146, 198, 222, 428, 262, 95, 154, 148, 146), \end{aligned} \tag{16}$$

Table 3: Original data of PM_{10} (unit: $\mu g/m^3$) in Shijiazhuang.

	$X_2^{(0)}$ 2014	$X_3^{(0)}$ 2015	$X_4^{(0)}$ 2016	$X_5^{(0)}$ 2017	$X_1^{(0)}$ 2018
12–20	87	238	866	85	215
12–21	97	490	765	188	153
12–22	108	383	181	268	228
12–23	333	412	142	304	132
12–24	237	376	258	117	193
12–25	191	295	282	116	248
12–26	353	278	298	146	114
12–27	478	51	170	198	78
12–28	237	174	342	222	68
12–29	218	378	308	428	79
12–30	192	155	413	262	122
12–31	122	207	710	95	200
01–01	156	414	617	154	198
01–02	220	617	493	148	273
01–03	358	485	471	146	299

are the sequences of related factors.

When $(p/q) = 0.2$, the MAPE of FGM(0, m) model is the smallest after the repeated experiments. Then, the system characteristic sequence and relevant factor sequence are carried out using 0.2-order accumulation, respectively; then

$$\begin{aligned}
x_1^{(0.2)}(1) &= 215.00,\\
x_1^{(0.2)}(2) &= 196.00,\\
x_1^{(0.2)}(3) &= 284.40,\\
x_1^{(0.2)}(4) &= 214.88,\\
x_1^{(0.2)}(5) &= 275.36,\\
x_1^{(0.2)}(6) &= 345.99,\\
x_1^{(0.2)}(7) &= 234.49,\\
x_1^{(0.2)}(8) &= 187.92,\\
x_1^{(0.2)}(9) &= 167.91,\\
x_1^{(0.2)}(10) &= 172.23,\\
x_1^{(0.2)}(11) &= 213.77.
\end{aligned} \tag{17}$$

Then,

$$\begin{aligned}
Y &= (215, 196, 284.4, 214.88, 275.36, 345.99, 234.49, 187.92, 167.91, 172.23, 213.77)^{\mathrm{T}},\\
B &= \begin{bmatrix}
1 & 87 & 238 & 866 & 85\\
1 & 114.4 & 537.6 & 938.2 & 25\\
1 & 137.84 & 509.56 & 437.92 & 315.8\\
\vdots & \vdots & \vdots & \vdots & \vdots\\
1 & 540.11 & 800.31 & 864.02 & 305.90
\end{bmatrix}.
\end{aligned} \tag{18}$$

Thus,

$$\begin{aligned}
[a, b_2, b_3, b_4, b_5]^{\mathrm{T}} &= \left(B^{\mathrm{T}}B\right)^{-1}B^{\mathrm{T}}Y\\
&= [8.789, 0.156, 0.478, 0.003, -0.200]^{\mathrm{T}}.
\end{aligned} \tag{19}$$

Therefore, the estimation equation of FGM(0, 4) is given by

$$\begin{aligned}
\widehat{x}_i^{(0.2)}(k) &= 8.789 + 0.156x_2^{(0.2)}(k) + 0.478x_3^{(0.2)}(k)\\
&\quad + 0.003x_4^{(0.2)}(k) - 0.200x_5^{(0.2)}(k), \quad k = 1, 2, \cdots.
\end{aligned} \tag{20}$$

Therefore, the fitting values for PM_{10} from December 20 to December 30, 2018, and the forecasting values from December 31, 2018, to January 3, 2019, are shown in Table 4.

4.1.2. Calculation Result of PM_{10} by Using SVR Model. In the research process of this paper, the firefly optimization algorithm and support vector regression model are established by using the toolkit in Matlab. The two parameters C and g of SVR method are optimized. In the process of selection, the optimal result can be obtained when the MAPE between the actual values and the fitting values is the smallest. In this paper, the optimal results of C and g are $C = 3.2103$ and $g = 0.023268$, respectively. The kernel function sigmoid is used for learning. The optimal calculation results are obtained after the data are operated 50 times.

SVR is used to calculate the fitting values from December 20, 2018, to December 30, 2018, and the values from December 31, 2018, to January 3, 2019, are forecasted. In this paper, the data from December 20 to December 30 in 2014–2017 were used as the training set. Meanwhile, the data from December 31 to January 3 of the following year in 2014–2017 were used as the testing set. The MAPE was used as the standard to measure the effect of fitting. Taking the concentration of PM_{10} in Shijiazhuang as an example, the MAPE between the fitting values and the original data is 0.48%. It can be seen that SVR has a good fitting effect and can be used for the forecasting of air pollutants. The calculation results are shown in Table 5.

4.1.3. Calculation Result of PM_{10} by Using the Hybrid Model. According to equations (11)-(12), the weights are given to the calculation results obtained by the FGM(0, m) model and SVR model, respectively:

Table 4: The results of FGM(0, m) for PM_{10}.

	Actual value (unit: $\mu g/m^3$)	Predictive value (unit: $\mu g/m^3$)	APE (%)
12–20	215	122	43.3
12–21	153	223	45.8
12–22	228	156	31.6
12–23	132	197	49.2
12–24	193	203	5.2
12–25	248	157	36.7
12–26	114	168	47.4
12–27	78	69	11.5
12–28	68	86	26.5
12–29	79	139	75.9
12–30	122	62	49.2
MAPE			38.4
12–31	200	110	45.0
01–01	198	202	2
01–02	273	310	13.6
01–03	299	269	10
MAPE			17.7

Table 5: The results of SVM for PM_{10}.

	Actual value (unit: $\mu g/m^3$)	Predictive value (unit: $\mu g/m^3$)	APE (%)
12–20	215	206	4.2
12–21	153	162	5.9
12–22	228	219	3.9
12–23	132	123	6.8
12–24	193	202	4.7
12–25	248	215	13.3
12–26	114	123	7.9
12–27	78	85	9
12–28	68	116	70.6
12–29	79	88	11.4
12–30	122	113	7.4
MAPE			13.2
12–31	200	206	3
01–01	198	184	7.1
01–02	273	161	41
01–03	299	134	55.2
MAPE			26.6

$$\begin{aligned} \widehat{x}^{(0)}_{\mathrm{FGM}(0,\mathrm{m})} &= 122, \\ \widehat{x}^{(0)}_{\mathrm{SVR}} &= 206. \end{aligned} \tag{21}$$

Thus,

$$\begin{aligned} \mathrm{APE}_{\mathrm{FGM}(0,\mathrm{m})} &= \frac{\left|\widehat{x}^{(0)}_{\mathrm{FGM}(0,\mathrm{m})} - x^{(0)}_1\right|}{x^{(0)}_1} = \frac{|122 - 215|}{215} = 43.256\%, \\ \mathrm{APE}_{\mathrm{SVR}} &= \frac{\left|\widehat{x}^{(0)}_{\mathrm{SVR}} - x^{(0)}_1\right|}{x^{(0)}_1} = \frac{|206 - 215|}{215} = 4.186\%, \\ \omega_{\mathrm{FGM}(0,\mathrm{N})} &= \frac{\mathrm{APE}^{-1}_{\mathrm{FGM}(0,\mathrm{m})}}{\mathrm{APE}^{-1}_{\mathrm{FGM}(0,\mathrm{m})} + \mathrm{APE}^{-1}_{\mathrm{SVR}}} = \frac{43.256\%^{-1}}{43.256\%^{-1} + 4.186\%^{-1}} = 0.09, \\ \omega_{\mathrm{SVR}(0,\mathrm{m})} &= \frac{\mathrm{APE}^{-1}_{\mathrm{SVR}}}{\mathrm{APE}^{-1}_{\mathrm{FGM}(0,\mathrm{m})} + \mathrm{APE}^{-1}_{\mathrm{SVR}}} = \frac{4.186\%^{-1}}{43.256\%^{-1} + 4.186\%^{-1}} = 0.91. \end{aligned} \tag{22}$$

Then, the predicted value of the hybrid model is given by

$$\widehat{x}_1^{(0)}(12) = \omega_{\mathrm{FGM}(0,\mathrm{m})} \times \widehat{x}_{\mathrm{FGM}(0,\mathrm{m})}^{(0)} + \omega_{\mathrm{SVR}} \times \widehat{x}_{\mathrm{SVR}}^{(0)} = 0.09 \times 122 + 0.91 \times 206 = 198. \tag{23}$$

The APE of the hybrid model is given by

$$\mathrm{APE} = \frac{\left|\widehat{x}_1(0) - x_1^{(0)}\right|}{x_1^{(0)}} = \frac{|198 - 215|}{215} = 7.9\%. \tag{24}$$

Similarly, the fitting values of the hybrid model from December 20 to December 30, 2018, and the predicted values from December 31, 2018, to January 3, 2019, are shown in Table 5. Meanwhile, the comparison results between the hybrid model and the Holt–Winters model are shown in Table 6.

The MAPE of FGM(0, m) is 38.4% and 17.7%, respectively. The MAPE of SVR is 13.2% and 26.6%, respectively. However, the MAPE of the hybrid model is 12.3% and 4.4%, respectively. Although the fitting accuracy is improved slightly, the predictive accuracy is improved clearly. However, the MAPE of the hybrid model is relatively small, so the prediction effect of the hybrid model is better. The fitting and prediction accuracies of the hybrid model are 22.1% and 37.1% higher than those of the Holt–Winters model, respectively. It indicates that the fitting and prediction accuracies of the hybrid model are significantly higher than those of the Holt–Winters model.

4.2. Calculation Result of Hybrid Model for $PM_{2.5}$. It can be seen from Table 7 that the MAPE is 14.1% and 13.7%, respectively. It can be seen that the MAPEs of the fitting values and prediction values of the hybrid model are all smaller, which indicates that the fitting accuracy and prediction accuracy are higher.

4.3. Calculation Results of NO_2 by Using the Hybrid Model. According to the calculation results of Table 8, the MAPE for NO_2 of the hybrid model is 12.4% and 2.3%, respectively. The MAPE of the hybrid model is very small, and it indicates that both the fitting and prediction accuracies of the hybrid model are high.

5. Calculation Results of the Air Pollutants in Chongqing

Similarly, three pollutants of Chongqing (PM_{10}, $PM_{2.5}$, and NO_2) from December 20 to January 1 between 2014 and 2017, and December 20 to December 30, 2018, are used as samples for forecasting.

5.1. Calculation Result of PM_{10} by Using Hybrid Model. In this paper, three pollutants (PM_{10}, $PM_{2.5}$, and NO_2) are predicted by using the hybrid model. The same method is used to verify that the hybrid model has higher accuracy than the FGM(0, m) model and SVR model. The FGM(0, m) model and SVR model are used to calculate the fitting values from December 20 to December 30, 2018, and the predictive values from December 31, 2018, to January 1, 2019, respectively. The calculation results of the two models are multiplied by their weights, respectively. The results are shown in Table 9.

Table 6: The comparison results of the two models (unit: $\mu g/m^3$).

	Actual value	Predictive value	
		Holt–Winters	The hybrid model
12–20	215	120	198
12–21	153	159	169
12–22	228	97	212
12–23	132	123	132
12–24	193	102	202
12–25	248	91	199
12–26	114	111	129
12–27	78	92	78
12–28	68	100	94
12–29	79	137	95
12–30	122	105	106
MAPE		34.4%	12.3%
12–31	200	117	200
01–01	198	138	198
01–02	273	152	273
01–03	299	150	247
MAPE		41.5%	4.4%

The MAPE of fitting and forecasting is 22.6% and 1.9%, respectively. The MAPE of the forecasting values is small, which indicates that the prediction accuracy of the hybrid model is high.

5.2. Calculation Result of $PM_{2.5}$ by Using Hybrid Model. It can be seen from Table 10 that, for the hybrid model of FGM(0, m) model and SVR model, the MAPE of fitting and forecasting is 9.5% and 3.4%, respectively. The MAPE of the fitting and forecasting values is small, and it can be seen that the precision of the hybrid model has been significantly improved.

5.3. Calculation Result of NO_2 by Using Hybrid Model. The hybrid model of the FGM(0, m) model and SVR model is used to calculate the fitting values from December 20 to December 30, 2018, and the forecasting values from December 31, 2018, to January 3, 2019, respectively. As shown in Table 11, the fitting and prediction accuracies are 4.7% and 8.4%, respectively. The fitting and prediction errors of the hybrid model are smaller. It indicates that the fitting and prediction accuracies of the hybrid model are higher. The error level of fitting and prediction is shown in Table 12.

According to the criteria mentioned above, the prediction accuracy of two air pollutants (PM_{10} and NO_2) in Shijiazhuang is 4.4% and 2.3%, respectively. The prediction accuracy of three air pollutants (PM_{10}, $PM_{2.5}$, and NO_2) in Chongqing is 1.9%, 3.4%, and 8.4%, respectively. Although the prediction accuracy of $PM_{2.5}$ is 13.7%, the prediction accuracy appears to be in the relatively superior range. It indicates that the prediction accuracy of the hybrid model is relatively high.

Table 7: The results of the hybrid model for $PM_{2.5}$

	Actual value (unit: $\mu g/m^3$)	Predictive value (unit: $\mu g/m^3$)	APE (%)
12–20	134	124	7.5
12–21	79	90	13.9
12–22	138	128	7.2
12–23	51	61	19.6
12–24	90	99	10.0
12–25	148	127	14.2
12–26	66	66	0
12–27	40	40	0
12–28	33	48	45.5
12–29	39	48	23.1
12–30	72	62	13.9
MAPE			14.1
12–31	142	131	7.7
01–01	127	119	6.3
01–02	182	182	0
01–03	203	120	40.9
MAPE			13.7

Table 8: The results of the hybrid model for NO_2.

	Actual value (unit: $\mu g/m^3$)	Predictive value (unit: $\mu g/m^3$)	APE (%)
12–20	83	83	0
12–21	74	79	6.8
12–22	91	71	22.0
12–23	37	50	35.1
12–24	75	70	6.7
12–25	89	61	31.5
12–26	47	47	0
12–27	35	38	8.6
12–28	40	48	20.0
12–29	50	53	6
12–30	60	60	0
MAPE			12.4
12–31	76	70	7.9
01–01	77	78	1.3
01–02	93	93	0
01–03	100	100	0
MAPE			2.3

Table 9: The results of the hybrid model for PM_{10}.

	Actual value (unit: $\mu g/m^3$)	Predictive value (unit: $\mu g/m^3$)	APE (%)
12–20	130	123	5.4
12–21	78	85	9.0
12–22	90	69	23.3
12–23	79	69	12.7
12–24	56	56	0
12–25	60	55	8.3
12–26	56	49	12.5
12–27	52	63	21.2
12–28	34	62	82.4
12–29	25	40	60.0
12–30	49	42	14.3
MAPE			22.6
12–31	57	57	0
01–01	77	81	5.2
01–02	85	87	2.4
01–03	85	85	0
MAPE			1.9

Table 10: The results of the hybrid model for $PM_{2.5}$

	Actual value (unit: $\mu g/m^3$)	Predictive value (unit: $\mu g/m^3$)	APE (%)
12–20	101	94	6.9
12–21	61	67	9.8
12–22	68	62	8.8
12–23	58	54	6.9
12–24	41	41	0
12–25	45	41	8.9
12–26	42	42	0
12–27	41	42	2.4
12–28	26	33	26.9
12–29	21	27	28.6
12–30	36	34	5.6
MAPE			9.5
12–31	45	45	0
01–01	57	61	7.0
01–02	60	64	6.7
01–03	63	63	0
MAPE			3.4

Table 11: The results of the hybrid model for NO_2.

	Actual value (unit: $\mu g/m^3$)	Predictive value (unit: $\mu g/m^3$)	APE (%)
12–20	55	53	3.6
12–21	51	49	3.9
12–22	42	37	11.9
12–23	42	42	0
12–24	42	42	0
12–25	41	39	4.9
12–26	36	38	5.6
12–27	34	35	2.9
12–28	32	37	15.6
12–29	27	28	3.7
12–30	32	32	0
MAPE			4.7
12–31	33	31	6.1
01–01	38	38	0
01–02	43	40	7.0
01–03	39	31	20.5
MAPE			8.4

Table 12: The error level of fitting and prediction.

MAPE	Accuracy
0%–10%	High
10%–20%	Good
20%–30%	General
Exceed 30%	Bad

6. Conclusions and Suggestions

6.1. Suggestion for Shijiazhuang. In 2018, Shijiazhuang ranked 168th among 169 key cities with poor air quality in China, but this year, it plans to move out from the last 10. The plan represents an effort to improve the atmosphere. Combined with the specific situation of Shijiazhuang, the suggestions are given from the following aspects:

(1) Firstly, as a city with a relatively dense population, Shijiazhuang should actively reduce the burden of the city and guide the transfer of population to surrounding counties and cities. This will not only reduce the pressure on the city, but also reduce the living pollution, traffic, and environmental pollution of Shijiazhuang.

(2) The transformation of the industrial structure should be accelerated, and the transformation of urban development should be promoted from heavy to light. In addition, "reduce the weight" of the city should be taken seriously. The transformation of industries from high-emission industries to emerging industries and environmental protection industries should be guided by the government actively.

6.2. Suggestion for Chongqing. As a city with developed tourism and economy, the annual pollution brought by tourism cannot be underestimated. Therefore, it is necessary

to reduce the content of the pollutants in the air. Suggestions are given from the following aspects:

(1) Firstly, the main pollution should be cleaned up, such as dust pollution, coal pollution, and industrial pollution, which will lead to the increase of the pollutant content in the air and eventually lead to the decline of the air quality.

(2) Secondly, because of the developed tourism in Chongqing, the traffic pollution caused by the huge population flow cannot be ignored. In order to reduce air pollution, citizens and the public should be encouraged to take public transportation instead of private cars.

7. Conclusions

It has been proved that the hybrid model is of great research significance and practicability in the M4 competition in 2019. However, the hybrid models with good prediction effect in M4 are the combination of artificial intelligence algorithm and statistical algorithm. It is shown that the forecasting effect of the hybrid model (FGM(0, m)-SVR) is better. The hybrid model is used to predict three air pollutants (PM_{10}, $PM_{2.5}$, and NO_2) in Shijiazhuang and Chongqing, and it is shown that the prediction accuracy of the hybrid model is significantly higher than that of the single model. The hybrid model can also be used to predict other air pollutants in other cities.

Data Availability

The data used in this study can be accessed via the website of Shijiazhuang Environmental Protection Bureau (http://www.sjzhb.gov.cn) and Chongqing Environmental Protection Bureau (http://www.cepb.gov.cn/).

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The relevant researches in this paper are supported by the National Natural Science Foundation of China (71871084), the Excellent Young Scientist Foundation of Hebei Education Department (SLRC2019001), and the Project of High-Level Talent in Hebei Province.

References

[1] Asian Clean Air Centre, *The Atmospheric Report in China of 2019*, Asian Clean Air Centre, 2019.

[2] Ministry of Environmental Protection of the People's Republic of China, *Ambient Air Quality Standards*, Ministry of Environmental Protection of the People's Republic of China, Beijing, China, 2012.

[3] Y. Yang and W. Xu, "Impact of FDI and R&D on China's industrial CO_2 emissions reduction and trend prediction," *Atmospheric Pollution Research*, vol. 10, no. 5, pp. 1627–1635, 2019.

[4] Y. Bai, B. Zeng, C. Li, and J. Zhang, "An ensemble long short-term memory neural network for hourly $PM_{2.5}$ concentration forecasting," *Chemosphere*, vol. 222, pp. 286–294, 2019.

[5] Y. Bai, Y. Li, B. Zeng, C. Li, and J. Zhang, "Hourly $PM_{2.5}$ concentration forecast using stacked autoencoder model with emphasis on seasonality," *Journal of Cleaner Production*, vol. 224, pp. 739–750, 2019.

[6] R. Ganguly, D. Sharma, and P. Kumar, "Trend analysis of observational PM_{10} concentrations in Shimla city, India," *Sustainable Cities and Society*, vol. 51, p. 101719, 2019.

[7] J. C. M. Pires, M. C. M. Alvim-Ferraz, M. C. Pereira, and F. G. Martins, "Prediction of PM_{10} concentrations through multi-gene genetic programming," *Atmospheric Pollution Research*, vol. 1, pp. 305–310, 2010.

[8] J. Yang and B. Sun, "Prediction of air pollutant concentration in Pingdingshan based on gray markov model," *Journal of Mathematics in Practice and Theory*, vol. 44, pp. 64–70, 2014.

[9] Z. Wang, X. Chen, and Q. Zhang, "Application of univariate PLS model in $PM_{2.5}$ real-time concentration prediction," *Computer Simulation*, vol. 34, pp. 387–391, 2017.

[10] L. Wu, X. Gao, Y. Xiao, S. Liu, and Y. Yang, "Using grey Holt-Winters model to predict the air quality index for cities in China," *Natural Hazards*, vol. 88, pp. 1003–1012, 2017.

[11] L. F. Weissert, J. A. Salmond, G. Miskell, M. Alavi-Shoshtari, and D. E. Williams, "Development of a microscale land use regression model for predicting NO_2 concentrations at a heavy trafficked suburban area in Auckland, New Zealand," *Science of The Total Environment*, vol. 619-620, pp. 112–119, 2018.

[12] Q. Wu and H. Lin, "A novel optimal-hybrid model for daily air quality index prediction considering air pollutant factors," *Science of The Total Environment*, vol. 683, pp. 808–821, 2019.

[13] J. Murillo-Escobar, J. P. Sepulveda-Suescun, M. A. Correa, and D. Orrego-Metaute, "Forecasting concentrations of air pollutants using support vector regression improved with particle swarm optimization: case study in Aburrá Valley, Colombia," *Urban Climate*, vol. 29, Article ID 100473, 2019.

[14] C. Zhang, X. Wang, S. Chen, Le Zou, X. Zhang, and C. Tang, "A study on daily $PM_{2.5}$ concentration in Hong Kong using the EMD-based MFDFA method," *Physica A: Statistical Mechanics and Its Applications*, vol. 530, Article ID 121182, 2019.

[15] M. Miri, Y. Ghassoun, A. Dovlatabadi, E. Ali, and M.-O. Löwner, "Estimate annual and seasonal PM_1, $PM_{2.5}$ and PM_{10} concentrations using land use regression model," *Ecotoxicology and Environmental Safety*, vol. 174, pp. 137–145, 2019.

[16] A. Z. Ul-Saufie, A. S. Yahaya, N. A. Ramli, N. Rosaida, and H. Abdul Hamid, "Future daily PM_{10} concentrations prediction by combining regression models and feedforward backpropagation models with principle component analysis (PCA)," *Atmospheric Environment*, vol. 77, pp. 621–630, 2013.

[17] B. Liu, H. Zheng, and B. Zhang, "Study on mixed prediction of air quality index based on IG-LASSO model," *Environmental Science & Technology*, vol. 40, pp. 144–148, 2017.

[18] G. Gualtieri, F. Carotenuto, S. Finardi, M. Tartaglia, P. Toscano, and B. Gioli, "Forecasting PM_{10} hourly concentrations in northern Italy: insights on models performance and PM_{10} drivers through self-organizing maps," *Atmospheric Pollution Research*, vol. 9, pp. 1204–1213, 2018.

[19] H.-C. Lai, H.-W. Ma, C.-R. Chen, M.-C. Hsiao, and B.-H. Pan, "Design and application of a hybrid assessment of air quality models for the source apportionment of $PM_{2.5}$," *Atmospheric Environment*, vol. 212, pp. 116–127, 2019.

[20] K.-T. Wu, C.-S. Lai, M.-L. Shyu, and Chung-Jen, "Chin. Apply GM(0, N) to consumer experience on small and medium-sized enterprises in creative living industry," *Journal of Grey System*, vol. 3, pp. 55–60, 2019.

[21] S. Chen, J.-Q. Wang, and H.-Y. Zhang, "A hybrid PSO-SVM model based on clustering algorithm for short-term atmospheric pollutant concentration forecasting," *Technological Forecasting & Social Change*, vol. 146, pp. 41–54, 2019.

[22] H. Liu, Z. Duan, and C. Chao, "A hybrid multi-resolution multi-objective ensemble model and its application for forecasting of daily $PM_{2.5}$ concentrations," *Information Sciences*, vol. 516, pp. 266–292, 2020.

[23] B. Zhang, H. Zhang, G. Zhao, and J. Lian, "Constructing a $PM_{2.5}$ concentration prediction model by combining auto-encoder with Bi-LSTM neural networks," *Environmental Modelling & Software*, vol. 124, Article ID 104600, 2020.

[24] L. Wu, S. Liu, L. Yao, S. Yan, and D. Liu, "Grey system model with the fractional order accumulation," *Communications in Nonlinear Science and Numerical Simulation*, vol. 18, pp. 1775–1785, 2013.

[25] C. Cortes and V. Vapnik, "Support-vector networks," *Springer Journal*, vol. 195, no. 20, pp. 273–297, 1995.

[26] D.-C. Wang and J.-R. Wang, "Appling GM(0, N) model on the influence factor analysis for the construction of the model in vocational high school," *Journal of Grey System*, vol. 3, pp. 19–28, 2019.

[27] L. F. Wu, Nu Li, and T. Zhao, "Using the seasonal FGM(1, 1) model to predict the air quality indicators in Xingtai and Handan," *Environmental Science and Pollution Research*, vol. 26, pp. 14683–14688, 2019.

[28] N. Korprasertsak and T. Leephakpreeda, "Robust short-term prediction of wind power generation under uncertainty via statistical interpretation of multiple forecasting models," *Energy*, vol. 180, pp. 387–397, 2019.