

Article

A Novel Linear Time-Varying GM(1,N) Model for Forecasting Haze: A Case Study of Beijing, China

Pingping Xiong ^{1,*}, Jia Shi ¹, Lingling Pei ² and Song Ding ³¹ College of Mathematics and Statistics, Nanjing University of Information Science and Technology, Nanjing 210044, China² School of Business Administration, Zhejiang University of Finance and Economics, Hangzhou 310018, China³ School of Economics, Zhejiang University of Finance and Economics, Hangzhou 310018, China

* Correspondence: xiongpingsing@nuist.edu.cn

Received: 29 May 2019; Accepted: 10 July 2019; Published: 13 July 2019



Abstract: Haze is the greatest challenge facing China's sustainable development, and it seriously affects China's economy, society, ecology and human health. Based on the uncertainty and suddenness of haze, this paper proposes a novel linear time-varying grey model (GM)(1,N) based on interval grey number sequences. Because the original GM(1,N) model based on interval grey number sequences has constant parameters, it neglects the dynamic change characteristics of parameters over time. Therefore, this novel linear time-varying GM(1,N) model, based on interval grey number sequences, is established on the basis of the original GM(1,N) model by introducing a linear time polynomial. To verify the validity and practicability of this model, this paper selects the data of PM₁₀, SO₂ and NO₂ concentrations in Beijing, China, from 2008 to 2018, to establish a linear time-varying GM(1,3) model based on interval grey number sequences, and the prediction results are compared with the original GM(1,3) model. The result indicates that the prediction effect of the novel model is better than that of the original model. Finally, this model is applied to forecast PM₁₀ concentration for 2019 to 2021 in Beijing, and the forecast is made to provide a reference for the government to carry out haze control.

Keywords: haze; linear time-varying GM(1,N) model; interval grey number; Beijing; forecasting

1. Introduction

In recent years, the process of urbanization and industrialization has been accelerating in China, but air pollution is seriously increasing. At present, haze is one of the greatest atmospheric, environmental pollution problems in China. Haze not only has adverse effects on the ecological environment [1–3] but also poses a major threat to human health [4,5]. As one of the largest developing countries in the world, China is facing the biggest challenge of sustainable development. In 2018, China's government officially issued the 'Three-Year Plan of Action to Win the Blue Sky Defense War'. Therefore, an accurate study of haze is of great importance.

Beijing is one of key cities for haze control in China. 16 haze events occurred in Beijing from November, 2012 to January, 2013, and the minimum visibility was even 667 m during this period [6]. Besides, the haze event in January, 2013 may have been the cause of 690 deaths in Beijing, which could lead to 253.8 million dollars losses [7]. To reduce the haze event, China's government implemented the 'Air Pollution Prevention and Control Action Plan' in September, 2013. Since then, an increasing number of scholars started to research haze from different aspects. Current haze studies mainly focused on health impact [3,4,7], economic loss [7], chemical composition [8], statistical characteristics [9], trend prediction [10], formation mechanism [11], and the rest. Above researches mainly based on hourly or daily data, but the goal of Blue Sky Defense War is to reduce the average annual concentration

of pollutants. Therefore, using annual data to forecast haze is necessary. Due to the uncertainty and suddenness of haze, the data of its related indices are uncertain. In the system research, grey number refers to the uncertain number in an interval or a general number set [12]. In other words, due to the constraints of data acquisition tools, acquisition conditions and errors by the acquisition personnel, the value of haze may contain inevitable measurement errors in a certain range, thus, haze has obvious grey number features. Therefore, Wu et al. established a grey prediction model to forecast annual pollutant concentrations effecting haze by limited data for the first time [13]. Grey prediction model is an important part of grey system. As an emerging uncertain system theory, grey system theory is characterized by small data modeling to obtain accurate results, and widely applied in all professions and trades. The grey system theory was established by Professor Deng, and assists with solving an uncertain system of small data and poor information, mainly through the deep mining of the inherent law in the existing information research system. Besides, grey number is the basic unit of grey system [12]. To summarize, grey prediction model is an effective tool for studying haze.

At present, the original grey model (GM)(1,N) based on interval grey number sequences is used to estimate structural parameters by the least squares method, and the result of its structural parameters are fixed values that are not related to time. Although this method is simple, it ignores the dynamic change characteristics of the parameters over time, which may lead the model to have a high precision fit, but not an ideal predictive effect. Therefore, this paper selects the data of PM₁₀ concentration (PM₁₀ refers to particles with aerodynamic equivalent diameter less than or equal to 10 microns in ambient air, called inhalable particulate matter), SO₂ concentration and NO₂ concentration in Beijing, China, from 2008 to 2018, and establishes a novel linear time-varying GM(1,3) model based on interval grey number sequences to perform simulations and make predictions. Meanwhile, we establish an original GM(1,3) model based on interval grey number sequences as the contrast model, and then we compare the results of these two models.

This paper is organized as follows: the literature reviews of studying on haze and grey prediction model are presented in Section 2; the modeling algorithms and model testing methods of the linear time-varying GM(1,N) model based on interval grey number sequences are illustrated in Section 3; the linear time-varying GM(1,N) model based on interval grey number sequences is established to forecast PM₁₀ concentration of Beijing in Section 4; the main conclusions of this paper are summarized in Section 5.

2. Literature Reviews

2.1. Study of Haze

To control haze more scientifically and effectively, many scholars at home and abroad have carried out a series of extensive studies. PM_{2.5} (PM_{2.5} refers to particles with an aerodynamic equivalent diameter less than or equal to 2.5 microns in ambient air, called fine particulate matter) and PM₁₀ are main factors causing haze, which not only affect the quality of atmospheric environment and visibility, but also endanger human health. Voukantsis et al. used the neural network method of multilayer perceptive structure to predict the daily change of PM₁₀ concentration in Thessaloniki and Helsinki [14]. Kumar et al. used a principal component regression method to predict short-term air quality index days in Delhi [15]. Lang et al. generalized the season autoregressive integrated moving average (SARIMA) model to the small-scale time sequences and used this model to make a short-term prediction of PM_{2.5} concentration at 10 stations in Hangzhou [16]. Mishra et al. studied the meteorological factors and atmospheric pollutants affecting PM_{2.5} in Delingha, and predicted the PM_{2.5} concentration under haze conditions in Delingha using a combination of a neural network and fuzzy logic [17]. Konovalov et al. predicted future PM₁₀ values in Europe by combining the deterministic prediction method and the linear regression method [18]. Based on the data of PM_{2.5} concentration in 285 cities in China from 2001 to 2012, Cheng et al. used a dynamic space panel model to analyze the main driving factors affecting air pollution [19]. Wu et al. established an input-oriented zero sum gains-data envelopment analysis

(ZSG-DEA) model, and studied the allocation efficiency of PM_{2.5} emissions in China's provinces under fixed target amount conditions [20]. The novel air quality early warning system was proposed by Xu et al., which included an evaluation module, a prediction module and a feature estimation module, and this literature used a novel dynamic fuzzy comprehensive evaluation method to determine the air quality level and major pollutants in the study area [21]. Han et al. predicted the spatial distribution of lung cancer in males induced by PM_{2.5} in China from 2010 to 2015, and the spatial autocorrelation method was used to evaluate the spatial relationship between the incidence of lung cancer and the atmospheric level of satellite-derived PM_{2.5} from 2006 to 2009 [22].

Nowadays, there are some scholars studying haze with grey system theory. Xiong et al. established the multivariate grey model (MGM) (1,M) based on interval grey number sequences to predict the visibility and relative humidity during a haze period in Nanjing [23]. Gong et al. combined the GM(1,1) model with both the Markoff chain model and the residual error correction model to establish a modified grey Markoff chain model, and used this model to predict the PM_{2.5} concentration in Shanghai [24]. Wang studied the distribution characteristics of PM_{2.5} concentration in Huaian with the non-parametric hypothesis test, and predicted the PM_{2.5} concentration of Huaian in the next five years by using GM(1,1) model [25]. Chen et al. predicted the hourly PM_{2.5} and PM₁₀ in Taichung's Dali area by using several GM(1,1) models and back-propagation artificial neural network, and compared their predictive performances [26]. Wang et al. established a novel grey correlation degree model, and used this model to dynamically analyze the influencing factors of haze in southern China [27].

2.2. Study of Grey Prediction Model

The grey prediction model has been widely applied in many fields, and it occupies an important position in the grey system theory. The GM(1,1) model that has only one variable is the most widely used prediction model in the grey system, but it is most suitable for the time sequences of monotonic increasing or monotonic decreasing. Hence, many scholars have worked to improve it. Wu et al. proposed a new GM(1,1) model with the fractional order accumulation, and this model had a better predicted performance than the traditional model [28]. Focusing on why the discrete grey model simulated the constant value growth rate, Zhang et al. established a linear time-varying discrete grey model by introducing a linear time polynomial [29]. Wang introduced the time polynomial function into the GM(1,1) power model and optimized the power exponent of the model [30]. Zeng et al. built a self-adapting intelligent grey prediction model to predict the natural gas demand of China [31].

However, the GM(1,1) model ignores the effect of related factors on the system behavior data, and then, Professor Deng proposed the GM(1,N) model, which has one system behavior sequence and N-1 related factors sequences [12]. The GM(1,N) model, as an extension of the GM(1,1) model, fully considers the impact of related factors on the system behavior data. Therefore, the GM(1,N) model and its optimization models are increasing in researches and applications. Ding et al. established a novel GM(1,N) model combined with the changing trend of the driving term, and this model was applied to predict CO₂ emissions from fuel combustion in China [32]. Zeng et al. proposed the optimal background-value GM(1,N) model through optimizing the background-value coefficient with the particle swarm optimization algorithm [33]. Wang et al. constructed a nonlinear GM(1,N) model by introducing the power exponent of the related factors, and this model was used to predict the carbon emissions of fossil energy consumption in China [34]. The above studies on grey prediction models were mainly based on real number sequence. Recently, increasingly more scholars have begun to explore the modeling problem of grey prediction models based on interval grey number sequences. Ye et al. fully explored and expanded the axiom of generalized non-decreasing grey degrees and established a prediction model for interval grey number sequences [35]. Luo et al. established a discrete GM(1,1) model for kernel and measure sequences, and then restored the predicted value of the interval grey number [36]. Yang et al. established a prediction model for the normal distribution based on interval grey number sequences in the context of the normal distribution of uncertain information [37].

A new nonlinear GM(1,N) model was established based on interval grey number sequences by Xiong et al., which was applied to the prediction of air quality index in haze period [38].

3. Methodology

3.1. The Introduction of Interval Grey Number

In this section, we will mainly introduce the basic concepts of interval grey number, including the definitions of interval grey number, kernel and grey radius.

Definition 1. [12] grey number that has both lower bound a_k and upper bound b_k is called interval grey number, denoted as $\otimes_k \in [a_k, b_k]$.

Definition 2. [12] suppose that interval grey number $\otimes_k$ is a continuous function, $\tilde{\otimes} = (a_k + b_k)/2$ is called a kernel of interval grey number $\otimes_k$.

Definition 3. [12] when $\otimes_k$ is a continuous interval grey number, $r(k) = (b_k - a_k)/2$ is called a grey radius of interval grey number $\otimes_k$.

3.2. Linear Time-Varying GM(1,N) Model Based on Interval Grey Number Sequences

In this section, the modeling mechanism of the linear time-varying GM(1,N) model based on interval grey number sequences will be introduced. Besides, this model will be constructed from the kernel and grey radius sequences, respectively. Finally, the kernel and grey radius sequences will be restored to the interval grey number sequences. To illustrate the model more clearly, we will show the modeling steps of the linear time-varying GM(1,N) model based on interval grey number sequences in Figure 1.

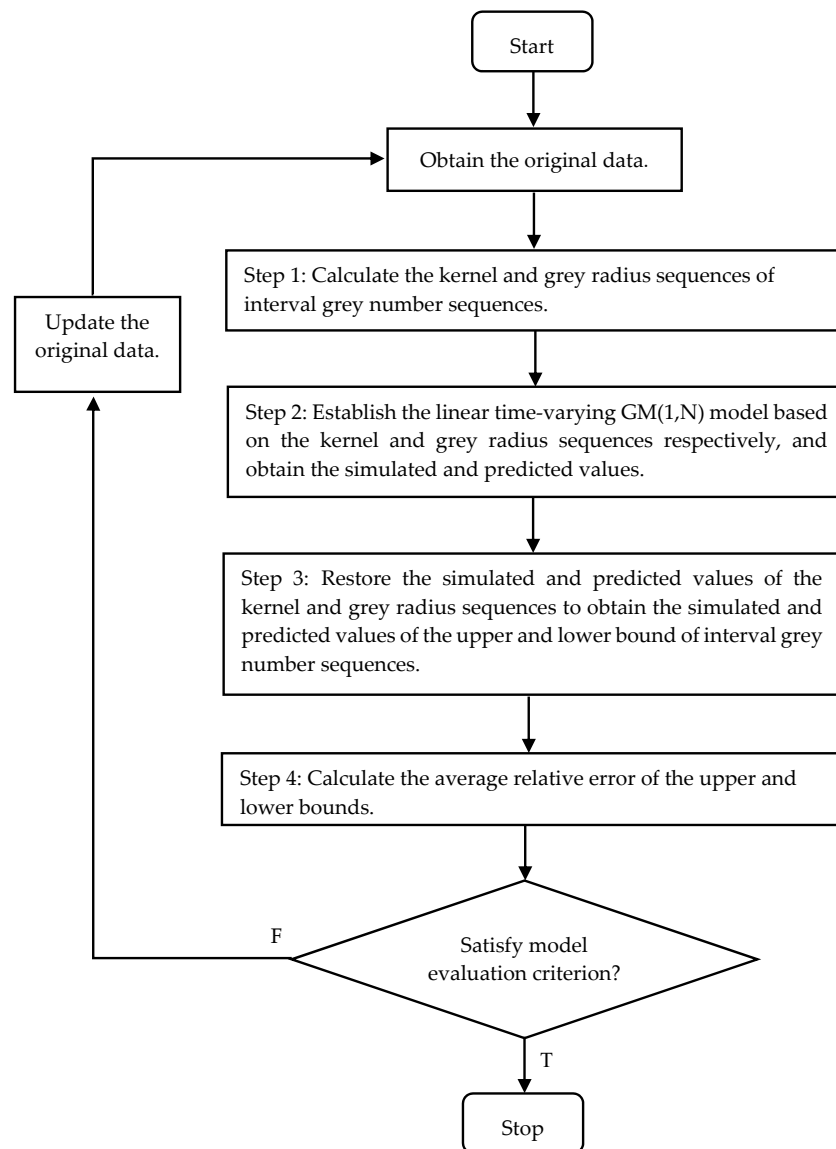


Figure 1. Modeling flow chart of the linear time-varying GM(1,N) model based on interval grey number sequences.

Then, this section will mainly introduce the linear time-varying GM(1,N) model based on kernel sequences as follows:

Definition 4. Suppose that the system behavior sequence is $\tilde{\otimes}_1^{(0)} = (\tilde{\otimes}_1^{(0)}(1), \tilde{\otimes}_1^{(0)}(2), \dots, \tilde{\otimes}_1^{(0)}(n))$, and the related factor sequences is $\tilde{\otimes}_i^{(0)} = (\tilde{\otimes}_i^{(0)}(1), \tilde{\otimes}_i^{(0)}(2), \dots, \tilde{\otimes}_i^{(0)}(n))$, $i = 2, 3, \dots, N$, where $\tilde{\otimes}_i^{(1)}$ is the first order accumulating generation sequence of $\tilde{\otimes}_i^{(0)}$, $i = 1, 2, \dots, N$; $z_1^{(1)}(k) = 0.5(\tilde{\otimes}_1^{(1)}(k) + \tilde{\otimes}_1^{(1)}(k-1))$ is the mean sequence generated by consecutive neighbors of $\tilde{\otimes}_1^{(1)}$; thus, the linear time-varying GM(1,N) model based on kernel sequences is shown as follows:

$$\tilde{\otimes}_1^{(0)}(k) + az_1^{(1)}(k) = \sum_{i=2}^N (b_{i1} + b_{i2}k)\tilde{\otimes}_i^{(1)}(k). \quad (1)$$

Also, the whitening equation of the linear time-varying model based on kernel sequences is shown as follows:

$$\frac{d\tilde{\otimes}_1^{(1)}}{dt} + a\tilde{\otimes}_1^{(1)} = \sum_{i=2}^N (b_{i1} + b_{i2}t)\tilde{\otimes}_i^{(1)}. \quad (2)$$

Theorem 1. The undetermined coefficient vector of the linear time-varying GM(1,N) model based on kernel sequences is $\hat{a} = [a, b_{21}, \dots, b_{N1}, b_{22}, \dots, b_{N2}]^T$, which can be estimated by utilizing the least squares method as follows:

- (1) When $n - 1 = 2N - 1$, that is $n = 2N$, and $|B| \neq 0$, $\hat{a} = B^{-1}Y$;
- (2) When $n - 1 > 2N - 1$, that is $n > 2N$, and $|B^T B| \neq 0$, $\hat{a} = (B^T B)^{-1} B^T Y$;
- (3) When $n - 1 < 2N - 1$, that is $n < 2N$, and $|B^T B| \neq 0$, $\hat{a} = B^T (BB^T)^{-1} Y$;

$$\text{where } B = \begin{bmatrix} -z_1^{(1)}(2) & \tilde{\otimes}_2^{(1)}(2) & \dots & \tilde{\otimes}_N^{(1)}(2) & 2\tilde{\otimes}_2^{(1)}(2) & \dots & 2\tilde{\otimes}_N^{(1)}(2) \\ -z_1^{(1)}(3) & \tilde{\otimes}_2^{(1)}(3) & \dots & \tilde{\otimes}_N^{(1)}(3) & 3\tilde{\otimes}_2^{(1)}(3) & \dots & 3\tilde{\otimes}_N^{(1)}(3) \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ -z_1^{(1)}(n) & \tilde{\otimes}_2^{(1)}(n) & \dots & \tilde{\otimes}_N^{(1)}(n) & n\tilde{\otimes}_2^{(1)}(n) & \dots & n\tilde{\otimes}_N^{(1)}(n) \end{bmatrix}, Y = \begin{bmatrix} \tilde{\otimes}_1^{(0)}(2) \\ \tilde{\otimes}_1^{(0)}(3) \\ \vdots \\ \tilde{\otimes}_1^{(0)}(n) \end{bmatrix}$$

The proof is similar to literature [39].

Theorem 2. After calculating the coefficient vector $\hat{a}$, the solution of the linear time-varying GM(1,N) model based on kernel sequences is shown as follows:

$$\hat{\otimes}_1^{(1)}(k+1) = \frac{1}{a} \sum_{i=2}^N (b_{i1} + b_{i2}k)\tilde{\otimes}_i^{(1)}(k+1) + e^{-ak} \left[\tilde{\otimes}_1^{(1)}(0) - \frac{1}{a} \sum_{i=2}^N (b_{i1} + b_{i2}k)\tilde{\otimes}_i^{(1)}(k+1) \right], \quad (3)$$

where the solution can be obtained using the initial condition $\tilde{\otimes}_1^{(1)}(0) = \tilde{\otimes}_1^{(0)}(1)$. Also, the inverse accumulating reduction equation is shown as follows:

$$\hat{\otimes}_1^{(0)}(k+1) = \hat{\otimes}_1^{(1)}(k+1) - \hat{\otimes}_1^{(1)}(k). \quad (4)$$

The modeling mechanism of the linear time-varying GM(1,N) model based on grey radius sequences are the same as the linear time-varying GM(1,N) model based on kernel sequences, thus we will not present the linear time-varying GM(1,N) model based on grey radius sequences repeatedly.

After obtaining the values of kernel and grey radius respectively, we will calculate the upper and lower bounds of the interval grey number sequences as follows [40]:

$$\begin{cases} \hat{a}_{k+1} = \hat{\otimes}_1^{(0)}(k+1) - \hat{r}_1^{(0)}(k+1) \\ \hat{b}_{k+1} = \hat{\otimes}_1^{(0)}(k+1) + \hat{r}_1^{(0)}(k+1) \end{cases}. \quad (5)$$

3.3. Model Evaluation Criterion

To analyze the reliability and credibility of the prediction model, we will show the model evaluation criterion for testing the model accuracy in this section. By comparing the relative error and average relative error of the upper and lower bounds of the interval grey number sequences to test the prediction model, the testing equations are shown as follows:

The relative error of the upper and lower bounds of the interval grey number sequences are shown as follows:

$$\begin{cases} \Delta_k = \frac{|\hat{a}_k - a_k|}{a_k} \times 100\% \\ \Delta_k = \frac{|\hat{b}_k - b_k|}{b_k} \times 100\% \end{cases}, k = 1, 2, \dots, n. \quad (6)$$

The average relative error of the upper and lower bounds of the interval grey number sequences are shown as follows:

$$\begin{cases} \bar{\hat{a}}_k = \frac{1}{n} \sum_{k=1}^n \frac{\Delta_k}{\bar{\Delta}_k} \\ \bar{\hat{b}}_k = \frac{1}{n} \sum_{k=1}^n \frac{\Delta_k}{\bar{\Delta}_k} \end{cases}, k = 1, 2, \dots, n. \quad (7)$$

Prediction accuracy is an important criterion for measuring the reliability of the prediction model. Therefore, this paper provides the prediction accuracy corresponding to the average relative error in Table 1.

Table 1. The average relative error criterion for testing model [41].

Average Relative Error	Prediction Accuracy
<10%	High prediction
10–20%	Good prediction
20–50%	Reasonable prediction
>50%	Weak prediction

4. Empirical Results and Discussion

In this section, the linear time-varying GM(1,N) model and the original GM (1,N) model, based on interval grey number sequences, will be established to simulate the development trend of haze in Beijing, and the model with high prediction accuracy will be selected to forecast the haze situation in Beijing from 2019 to 2021.

4.1. Data Selection and Processing

PM₁₀ can be formed by the interaction of sulfur oxides, nitrogen oxides and other compounds in the ambient air. Additionally, PM₁₀ is highly correlated with SO₂ and NO₂ in 31 cities of China [42]. Therefore, PM₁₀ concentration in Beijing from 2008 to 2018 is selected as the system behavior sequence, and SO₂ concentration and NO₂ concentration are selected as related factor sequences. In addition, the data are from the annual report of air quality in Beijing. In data processing, the maximum and minimum observed values over the previous three values are the upper and lower bounds of the third interval grey number. Moreover, the interval grey number sequences for 2008 to 2014 are used as the modeling data, and the interval grey number sequences for 2015 to 2018 are used as the prediction data. Besides, we denote the interval grey number sequences of PM₁₀ concentration, SO₂ concentration and NO₂ concentration as $X_1(\otimes)$, $X_2(\otimes)$, $X_3(\otimes)$ respectively. The original data is shown in Table 2.

Table 2. The interval grey number sequences of PM₁₀ concentration, SO₂ concentration and NO₂ concentration

k	Year	$X_1(\otimes)(\mu\text{g}/\text{m}^3)$	$X_2(\otimes)(\mu\text{g}/\text{m}^3)$	$X_3(\otimes)(\mu\text{g}/\text{m}^3)$
1	2008	[122,161]	[36,53]	[49,66]
2	2009	[121,148]	[34,47]	[49,66]
3	2010	[114,122]	[32,36]	[49,55]
4	2011	[114,121]	[28,34]	[53,55]
5	2012	[109,121]	[28,32]	[52,55]
6	2013	[108,114]	[27,28]	[52,56]
7	2014	[108,116]	[22,28]	[52,59]
8	2015	[102,116]	[14,27]	[50,59]
9	2016	[100,116]	[10,22]	[48,59]
10	2017	[95,108]	[8,14]	[46,50]
11	2018	[91,100]	[6,10]	[42,48]

4.2. Establishment and Comparison of Model

Step 1: on the basis of the data from 2008 to 2014 in Table 2, the kernel and grey radius sequences of PM₁₀ concentration, SO₂ concentration and NO₂ concentration are calculated according to Definition 2 and 3, and the results are shown in Table 3.

Table 3. The kernel and grey radius sequences of PM₁₀ concentration, SO₂ concentration and NO₂ concentration

k	$\tilde{\otimes}_1(k)$	$\tilde{\otimes}_2(k)$	$\tilde{\otimes}_3(k)$	$r_1(k)$	$r_2(k)$	$r_3(k)$
1	141.5	44.5	57.5	19.5	8.5	8.5
2	134.5	40.5	57.5	13.5	6.5	8.5
3	118	34	52	4	2	3
4	117.5	31	54	3.5	3	1
5	115	30	53.5	6	2	1.5
6	111	27.5	54	3	0.5	2
7	112	25	55.5	4	3	3.5

Step 2: after calculating model parameters by least squares method, a linear time-varying GM(1,3) model for the PM₁₀ concentration kernel sequence is established as follows:

$$\tilde{\otimes}_1^{(0)}(k) + 1.9972z_1^{(1)}(k) = (1.5419 - 0.0591k)\tilde{\otimes}_2^{(1)}(k) + (3.7449 - 0.0014k)\tilde{\otimes}_3^{(1)}(k) \quad (8)$$

Similarly, a linear time-varying GM(1,3) model for the PM₁₀ concentration grey radius sequence is established as follows:

$$r_1^{(0)}(k) + 1.6900z_1^{(1)}(k) = (-6.9833 + 2.2742k)r_2^{(1)}(k) + (9.5542 - 2.0577k)r_3^{(1)}(k). \quad (9)$$

According to the linear time-varying GM(1,3) model of PM₁₀ concentration kernel and grey radius sequences, the simulated values of PM₁₀ concentration kernel and grey radius sequences can be respectively obtained. The SO₂ concentration and NO₂ concentration sequences that are predicted by the GM(1,1) model are used as the related factor sequences for 2015 to 2018. Then, a linear time-varying GM(1,3) model is used to obtain the predicted values of PM₁₀ concentration kernel and grey radius sequences.

Step 3: according to Equation (5), the simulated and predicted values of the upper and lower bounds over the PM₁₀ concentration interval grey number sequence are obtained by restoring the simulated and predicted values of kernel and grey radius sequences. The results are shown in Table 4.

Step 4: the relative error and average relative error of the PM₁₀ concentration upper and lower bounds are calculated according to Equation (6) and Equation (7). The results are shown in Table 4.

At the same time, the original GM(1,N) model based on interval grey number sequences is selected as the contrast model. Establishing the GM(1,3) model based on interval grey number sequences to simulate and forecast PM₁₀ concentration in Beijing from 2008 to 2018, and the relative error and average relative error of upper and lower bounds are calculated. The results are shown in Table 4.

Table 4. The simulated and predicted values and the average relative error of the PM₁₀ concentration upper and lower bounds.

Year	Linear Time-Varying GM(1,3) Model				GM(1,3) Model		
	Actual Value ($\mu\text{g}/\text{m}^3$)	Simulated Value ($\mu\text{g}/\text{m}^3$)	Lower Bound Relative Error (%)	Upper Bound Relative Error (%)	Simulated Value ($\mu\text{g}/\text{m}^3$)	Lower Bound Relative Error (%)	Upper Bound Relative Error (%)
2008	[122,161]	[122.00,161.00]	0.00	0.00	[122.00,161.00]	0.00	0.00
2009	[121,148]	[107.05,130.05]	11.54	12.13	[105.05,126.89]	13.18	14.26
2010	[114,122]	[124.80,141.14]	9.47	15.69	[124.73,138.16]	9.41	13.25
2011	[114,121]	[121.66,123.27]	6.72	1.88	[115.38,125.57]	1.21	3.78
2012	[109,121]	[110.83,121.07]	1.68	0.06	[111.76,119.48]	2.53	1.26
2013	[108,114]	[110.18,115.51]	2.02	1.32	[110.10,114.59]	1.95	0.52
2014	[108,116]	[105.85,118.84]	1.99	2.45	[104.30,117.56]	3.43	1.35
Average simulated relative error (%)			4.78	4.79		4.53	4.92
Year	Actual Value ($\mu\text{g}/\text{m}^3$)	Predicted Value ($\mu\text{g}/\text{m}^3$)	Lower Bound Relative Error (%)	Upper Bound Relative Error (%)	Predicted Value ($\mu\text{g}/\text{m}^3$)	Lower Bound Relative Error (%)	Upper Bound Relative Error (%)
2015	[102,116]	[103.77,109.13]	1.74	5.92	[103.88,106.75]	1.85	7.98
2016	[100,116]	[100.99,106.51]	0.99	8.18	[101.46,103.48]	1.46	10.80
2017	[95,108]	[98.52,104.11]	3.70	3.60	[99.14,100.56]	4.36	6.89
2018	[91,100]	[96.31,101.91]	5.84	1.91	[96.95,97.95]	6.53	2.05
Average predicted relative error (%)			3.07	4.90		3.55	6.93

This paper will compare the simulated and predicted values, relative errors of upper and lower bounds of the PM₁₀ concentration interval grey number sequence. According to Table 4, the average simulation relative error of upper and lower bounds of linear time-varying GM(1,3) model are 4.79% and 4.78%, respectively, and the average simulation relative error of upper and lower bounds of GM(1,3) model are 4.92% and 4.53%, respectively. The average simulation relative errors of those two models are very close, and both belong to high simulation accuracy. However, the average predicted relative error of upper and lower bounds of linear time-varying GM(1,3) model are 4.90% and 3.07% respectively, which are less than those of GM(1,3) model.

For ease of comparison, we make comparison maps as shown in Figures 2–5. For simplicity, we briefly note the linear time-varying GM(1,N) model based on interval grey number sequences as LTVGM(1,N), and briefly note the GM(1,N) model based on interval grey number sequences as GM(1,N). From Figure 2 to Figure 5, the novel linear time-varying GM(1,3) model based on interval grey number sequences can more accurately describe the upper and lower bounds of PM₁₀ concentration in Beijing, which especially has the better prediction effect.



Figure 2. Comparison map of the PM₁₀ concentration lower bound values for LTVGM(1,N) model and GM(1,N) model.



Figure 3. Comparison map of the PM₁₀ concentration upper bound values for LTVGM(1,N) model and GM(1,N) model.

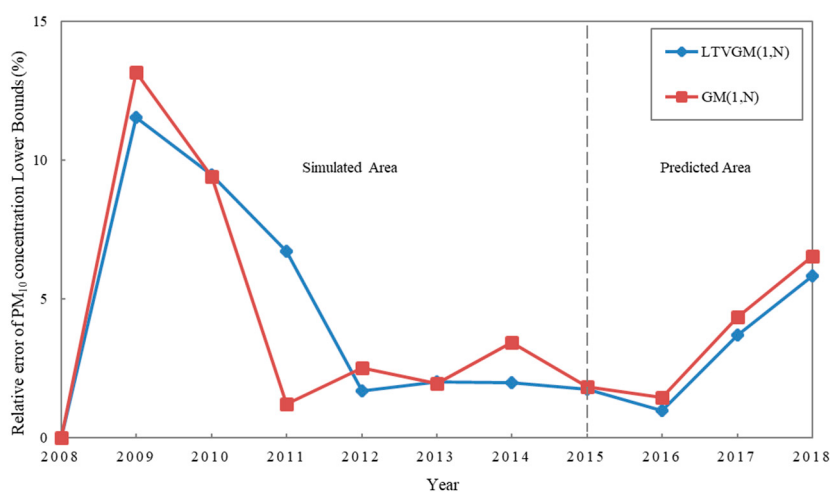


Figure 4. Comparison map of the relative errors of PM₁₀ concentration lower bound for LTVGM(1,N) model and GM(1,N) model.

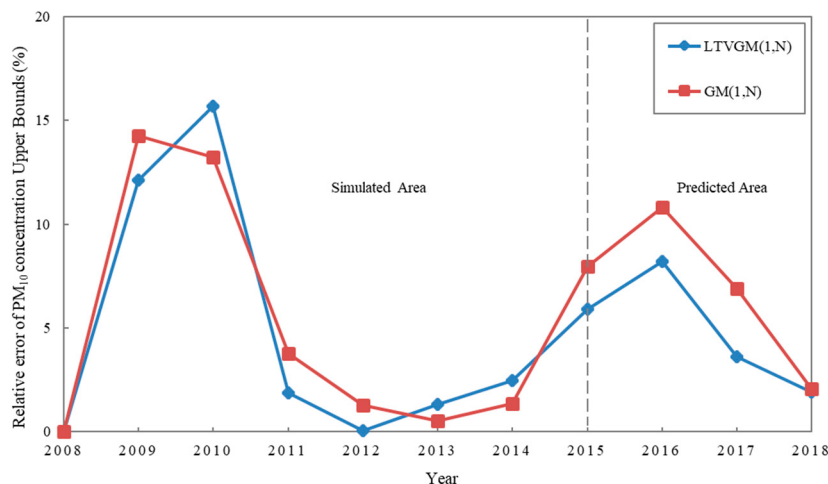


Figure 5. Comparison map of the relative errors of PM₁₀ concentration upper bound for LTVGM(1,N) model and GM(1,N) model.

4.3. Forecast Results and Discussion

To understand the trend of haze in Beijing in the future, the linear time-varying GM(1,3) model proposed in this paper is used to forecast PM₁₀ concentration in Beijing for 2019 to 2021, the forecast results are shown in Table 5. According to Table 5, PM₁₀ concentration in Beijing will decrease slowly from 2019 to 2021, but still exceed the China's environmental air quality standard [13].

Table 5. The predicted value of PM₁₀ concentration in Beijing for 2019 to 2021.

Year	2019	2020	2021
PM ₁₀ concentration (μg/m ³)	[94.36,99.88]	[92.62,98.07]	[91.06,96.29]

According to all the results considered, we will perform a discussion as follows: the average predicted relative errors of the linear time-varying GM(1,3) model are less than those of the original GM(1,3) model, which is attributed to the improved adaptability of the model to the dynamic change characteristics data. The model proposed in this paper is simple in the modeling method and convenient for calculation and applications. In addition, this novel model expands the range of predicted values from the real number to interval grey number, which can broaden the range of applications. To a certain extent, it can make up for the errors caused by data acquisition tools, acquisition conditions and acquisition personnel. However, this novel model still has certain limitations. It only considers that the related factor sequences are air pollutants, and it neglects the influence of meteorological factors on haze. Therefore, the future research will select both air pollutants and meteorological factors as the related factor sequences to study haze.

5. Conclusions

Aiming at forecasting haze of Beijing more accurately, this paper introduced a linear time polynomial into the GM(1,N) model based on interval grey number sequences, and established a novel linear time-varying GM(1,N) model based on interval grey number sequences. The data of PM₁₀ concentration, SO₂ concentration and NO₂ concentration in Beijing, China, were selected as modeling data to establish a linear time-varying GM(1,3) model from 2008 to 2018. The results indicate that the prediction model proposed in this paper has a higher prediction accuracy than the original model, and both the corresponding prediction errors are less than 5%, which proves the validity and practicability of this model. When using this novel model to make a forecast of PM₁₀ concentration in Beijing, the forecast shows that PM₁₀ concentration will be a downward trend for 2019 to 2021 in Beijing. This is because PM₁₀ concentration is not only determined by air pollutants, but also related to meteorological

factors, such as wind speed, precipitation, and so forth. Specifically, when wind speed or precipitation increases, haze will decrease in Beijing [43]. Meteorological factors vary from year to year, and the extreme weather event such as El Niño cannot be accurately forecasted a year in advance. When the extreme weather event occurs, the actual PM₁₀ concentration in Beijing may be different with our forecast for 2019 to 2021. If there is no extreme weather event in that year, the haze predicted value of this paper will be accurate. Therefore, this model can provide decision support for the government when working toward greater haze control.

The model proposed in this paper had a high accuracy of prediction, but there is much room for improvement. This paper only selected two pollutants as related factors, which were SO₂ and NO₂, in order to forecast PM₁₀ concentration in Beijing. In future research, we will consider a few meteorological factors as related factors to construct model, such as wind speed, relative humidity, air pressure, temperature, precipitation, etc. Besides, we will apply this novel model to several key cities in the Beijing-Tianjin-Hebei region of China. On this basis, we will use data mining technology to compare and analyze the haze of different cities in the Beijing-Tianjin-Hebei region.

Author Contributions: Conceptualization, J.S.; methodology, P.X.; software, J.S.; validation, J.S. and S.D.; formal analysis, P.X.; investigation, L.P.; resources, S.D.; data curation, L.P.; writing—original draft preparation, P.X. and J.S.; writing—review and editing, P.X. and J.S.; visualization, L.P. and S.D.; supervision, P.X.; project administration, P.X.; funding acquisition, P.X.

Funding: National Natural Science Foundation of China (71701105, 41505118); Major Program of the National Social Science Fund of China (17ZDA092); Ministry of Education Humanities and Social Sciences Research Youth Subsidy Project in China (17YJC630182, 17YJC630123); Key Research Project of Philosophy and Social Sciences in Universities of Jiangsu Province (2018SJZDI111).

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

References

1. Tao, M.; Chen, L.F.; Xiong, X.Z.; Zhang, M.G.; Ma, P.F.; Tao, J.H.; Wang, Z.F. Formation process of the widespread extreme haze pollution over northern China in January 2013: Implications for regional air quality and climate. *Atmos. Environ.* **2014**, *98*, 417–425. [\[CrossRef\]](#)
2. Lu, Y.L.; Wang, Y.; Zuo, J.; Jiang, H.Q.; Huang, D.C.; Rameezdeen, R. Characteristics of public concern on haze in China and its relationship with air quality in urban areas. *Sci. Total Environ.* **2018**, *637*, 1597–1606. [\[CrossRef\]](#) [\[PubMed\]](#)
3. Behera, S.N.; Cheng, J.P.; Huang, X.; Zhu, Q.Y.; Liu, P.; Balasubramanian, R. Chemical composition and acidity of size-fractionated inorganic aerosols of 2013–14 winter haze in Shanghai and associated health risk of toxic elements. *Atmos. Environ.* **2015**, *122*, 259–271. [\[CrossRef\]](#)
4. Yin, H.; Pizzol, M.; Jacobsen, J.B.; Xu, L.Y. Contingent valuation of health and mood impacts of PM_{2.5} in Beijing, China. *Sci. Total Environ.* **2018**, *630*, 1269–1282. [\[CrossRef\]](#) [\[PubMed\]](#)
5. Faridi, S.; Shamsipour, M.; Krzyzanowski, M.; Künzli, N.; Amini, H.; Azimi, F.; Malkawi, M.; Momeniha, F.; Gholampour, A.; Hassanvand, M.S.; et al. Long-term trends and health impact of PM_{2.5} and O₃ in Tehran, Iran, 2006–2015. *Environ. Int.* **2018**, *114*, 37–49. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Zhang, Q.; Quan, J.N.; Tie, X.X.; Li, X.; Liu, Q.; Gao, Y.; Zhao, D.L. Effects of meteorology and secondary particle formation on visibility during heavy haze events in Beijing, China. *Sci. Total Environ.* **2015**, *502*, 578–584. [\[CrossRef\]](#)
7. Gao, M.; Guttikunda, S.K.; Carmichael, G.R.; Wang, Y.S.; Liu, Z.R.; Stanier, C.O.; Saide, P.E.; Yu, M. Health impacts and economic losses assessment of the 2013 severe haze event in Beijing area. *Sci. Total Environ.* **2015**, *511*, 553–561. [\[CrossRef\]](#)
8. Tian, S.L.; Pan, Y.P.; Liu, Z.R.; Wen, T.X.; Wang, Y.S. Size-resolved aerosol chemical analysis of extreme haze pollution events during early 2013 in urban Beijing, China. *J. Hazard. Mater.* **2014**, *279*, 452–460. [\[CrossRef\]](#)
9. Zheng, Z.F.; Xu, G.R.; Yang, Y.J.; Wang, Y.T.; Li, Q.C. Statistical characteristics and the urban spillover effect of haze pollution in the circum-Beijing region. *Atmos. Pollut. Res.* **2018**, *9*, 1062–1071. [\[CrossRef\]](#)
10. Zhu, S.L.; Lian, X.Y.; Liu, H.X.; Hu, J.M.; Wang, Y.Y.; Che, J.X. Daily air quality index forecasting with hybrid models: A case in China. *Environ. Pollut.* **2017**, *231*, 1232–1244. [\[CrossRef\]](#)

11. Liu, X.G.; Li, J.; Qu, Y.; Han, T.; Hou, L.; Gu, J.; Chen, C.; Yang, Y.; Liu, X.; Yang, T.; et al. Formation and evolution mechanism of regional haze: A case study in the megacity Beijing, China. *Atmos. Chem. Phys.* **2013**, *13*, 4501–4514. [\[CrossRef\]](#)
12. Liu, S.F.; Dang, Y.G.; Fang, Z.G.; Xie, N.M. *Grey Systems Theory and Its Applications*, 5th ed.; Science Press: Beijing, China, 2010.
13. Wu, L.F.; Nu, L.; Yang, Y.J. Prediction of air quality indicators for the Beijing-Tianjin-Hebei region. *J. Clean. Prod.* **2018**, *196*, 682–687. [\[CrossRef\]](#)
14. 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.* **2011**, *409*, 1266–1276. [\[CrossRef\]](#) [\[PubMed\]](#)
15. Kumar, A.; Goyal, P. Forecasting of air quality in Delhi using principal component regression technique. *Atmos. Pollut. Res.* **2011**, *2*, 436–444. [\[CrossRef\]](#)
16. Lang, Y.C.; Xiao, L.; George, C. Forecasting concentrations of PM_{2.5} in main urban area of Hangzhou and mapping using SARIMA model and ordinary Kriging method. *Acta Sci. Circum.* **2018**, *1*, 62–70.
17. Mishra, D.; Goyal, P.; Upadhyay, A. Artificial intelligence based approach to forecast PM_{2.5}, during haze episodes: A case study of Delhi, India. *Atmos. Environ.* **2015**, *102*, 239–248. [\[CrossRef\]](#)
18. Konovalov, I.B.; Beekmann, M.; Meleux, F.; Dutot, A.; Foret, G. Combining deterministic and statistical approaches for PM₁₀, forecasting in Europe. *Atmos. Environ.* **2009**, *43*, 6425–6434. [\[CrossRef\]](#)
19. Cheng, Z.H.; Li, L.S.; Liu, J. Identifying the spatial effects and driving factors of urban PM_{2.5}, pollution in China. *Ecol. Indic.* **2017**, *82*, 61–75. [\[CrossRef\]](#)
20. Wu, X.H.; Tan, L.; Guo, J.; Wang, Y.Y.; Liu, H.; Zhu, W.W. A study of allocative efficiency of PM_{2.5}, emission rights based on a zero sum gains data envelopment model. *J. Clean. Prod.* **2016**, *113*, 1024–1031. [\[CrossRef\]](#)
21. Xu, Y.Z.; Du, P.; Wang, J.Z. Research and application of a hybrid model based on dynamic fuzzy synthetic evaluation for establishing air quality forecasting and early warning system: A case study in China. *Environ. Pollut.* **2017**, *223*, 435–448. [\[CrossRef\]](#)
22. Han, X.; Liu, Y.Q.; Gao, H.; Ma, J.M.; Mao, X.X.; Wang, Y.T.; Ma, X.D. Forecasting PM_{2.5}, induced male lung cancer morbidity in China using satellite retrieved PM_{2.5}, and spatial analysis. *Sci. Total Environ.* **2017**, *607*, 1009–1017. [\[CrossRef\]](#) [\[PubMed\]](#)
23. Xiong, P.P.; Zhang, Y.; Yao, T.X.; Zeng, B. Multivariate grey prediction model based on interval grey number sequences. *Math. Pract. Theor.* **2017**, *7*, 310–319.
24. Gong, M.; Ye, C.M. Prediction of the PM_{2.5} concentration in Shanghai Municipality based on Modified Grey Markov chain model. *J. Nat. Disasters* **2016**, *5*, 97–104.
25. Wang, Z.X. Distribution characteristics and grey prediction model of PM_{2.5} in Huai'an. *Math. Pract. Theor.* **2018**, *48*, 286–291.
26. Chen, L.; Pai, T.Y. Comparisons of GM(1,1) and BPNN for predicting hourly particulate matter in Dali area of Taichung City, Taiwan. *Atmos. Pollut. Res.* **2015**, *6*, 572–580. [\[CrossRef\]](#)
27. Wang, J.J.; Hipel, K.W.; Dang, Y.G. An improved grey dynamic trend incidence model with application to factors causing smog weather. *Expert Syst. Appl.* **2017**, *87*, 240–251. [\[CrossRef\]](#)
28. Wu, L.F.; Liu, S.F.; Yao, L.G.; Yan, S.L.; Liu, D.L. Grey system model with the fractional order accumulation. *Commun. Nonlinear Sci.* **2013**, *18*, 1775–1785. [\[CrossRef\]](#)
29. Zhang, K.; Liu, S.F. Linear time-varying parameters discrete grey forecasting model. *Syst. Eng. Theor. Pract.* **2010**, *30*, 1650–1657.
30. Wang, Z.X. GM(1,1) power model with time-varying parameters and its application. *Control Decis.* **2014**, *10*, 1828–1832.
31. Zeng, B.; Li, C. Forecasting the natural gas demand in China using a self-adapting intelligent grey model. *Energy* **2016**, *112*, 810–825. [\[CrossRef\]](#)
32. Ding, S.; Dang, Y.G.; Li, X.M.; Wang, J.J.; Zhao, K. Forecasting Chinese CO₂ emissions from fuel combustion using a novel grey multivariable model. *J. Clean. Prod.* **2017**, *162*, 1527–1538. [\[CrossRef\]](#)
33. Zeng, B.; Li, C. Improved multi-variable grey forecasting model with a dynamic background-value coefficient and its application. *Comput. Ind. Eng.* **2018**, *118*, 278–290. [\[CrossRef\]](#)
34. Wang, Z.X.; Ye, D.J. Forecasting Chinese carbon emissions from fossil energy consumption using non-linear grey multivariable models. *J. Clean. Prod.* **2017**, *142*, 600–612. [\[CrossRef\]](#)

35. Ye, J.; Dang, Y.G.; Ding, S. Grey prediction model of interval grey numbers based on axiom of generalized non-decrease grey degree. *Control Decis.* **2016**, *31*, 1831–1836.
36. Luo, D.; Li, L. Interval grey number prediction model based on kernel and measure. *Math. Pract. Theor.* **2014**, *44*, 96–100.
37. Yang, J.W.; Xiao, X.P.; Guo, J.H. Grey prediction model for interval grey number of normal distribution. *Control Decis.* **2015**, *9*, 1711–1716.
38. Xiong, P.P.; Yin, Y.; Shi, J.; Gao, H. Nonlinear multivariable GM(1,N) model based on interval grey number sequence. *J. Grey Syst.* **2018**, *30*, 33–47.
39. Zeng, B.; Duan, H.M.; Zhou, Y.F. A new multivariable grey prediction model with structure compatibility. *Appl. Math. Model.* **2019**, *75*, 385–397. [[CrossRef](#)]
40. Liu, J.F.; Liu, S.F.; Fang, Z.G. Continuous interval grey number prediction model based on kernel and grey radius. *Syst. Eng.* **2013**, *2*, 61–64.
41. Lewis, C. *Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting*; Butterworth Scientific: Oxford, UK, 1982.
42. Xie, Y.Y.; Zhao, B.; Zhang, L.; Luo, R. Spatiotemporal variations of PM_{2.5} and PM₁₀ concentrations between 31 Chinese cities and their relationships with SO₂, NO₂, CO and O₃. *Particuology* **2015**, *20*, 141–149. [[CrossRef](#)]
43. Chang, L.Y.; Xu, J.M.; Tie, X.X.; Wu, J.B. Impact of the 2015 El Nino event on winter air quality in China. *Sci. Rep.* **2016**, *6*, 34275. [[CrossRef](#)] [[PubMed](#)]



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).