Forecasting smog in Beijing using a novel time-lag GM(1,N) model based on interval grey number sequences

Abstract

Purpose - Smog seriously affects the ecological environment and poses a threat to public health. Therefore, smog control has become a key task in China, which requires reliable prediction.

Design/methodology/approach - Based on interval grey number sequences, the traditional GM(1,N) model neglects the time-lag effect of driving terms, hence this paper introduces the time-lag parameters into driving terms of the traditional GM(1,N) model and proposes a novel time-lag GM(1,N) model. Firstly, calculating kernel and degree of greyness of the interval grey number sequence respectively. Then, establishing the time-lag GM(1,N) model of kernel and degree of greyness sequences respectively to obtain values of them after determining the time-lag parameters of two models. Finally, the upper and lower bounds of interval grey number sequences are obtained by restoring the values of kernel and degree of greyness.

Findings - In order to verify the validity and practicability of the model, the monthly concentrations of $PM_{2.5}$, SO_2 and NO_2 in Beijing during August 2017 to September 2018 are selected to establish the time-lag GM(1,3) model for kernel and degree of greyness sequences respectively. Compared with three existing models, the proposed model in this paper has better simulation accuracy. Therefore, the new model is applied to forecast monthly $PM_{2.5}$ concentration for October to December 2018 in Beijing, and provides a reference basis for the government to formulate smog control policies.

Originality/value - The proposed model can simulate and forecast system characteristic data with the time-lag effect more accurately, which shows that the time-lag GM(1,N) model proposed in this paper is practical and effective.

Key words Smog, Time-lag GM(1,N) model, Interval grey number, Kernel and degree of greyness, Forecasting

1 Introduction

In recent years, the smog problem in China has attracted wide public attention. Smog not only affects urban air quality (Lu et al., 2018), damages social economy (Hao et al., 2018), but also seriously endangers human health (Asraf et al., 2019). The Chinese government has been committed to smog control, and has put forward clear requirements for improving air quality in both “the 13th Five-Year Plan of Environmental Planning” and “the Three-Year Action Plan for Winning the Blue Sky Defending War”, and has designated Beijing as one of the key control areas. Beijing is the capital and central city of China, as well as the national political center, cultural center and international exchange center. Therefore, accurately forecasting the smog situation of Beijing is of great significance, which can provide a scientific basis for the government’s smog control work.

For the sake of forecasting the trend of smog more accurately, many scholars have established the statistical model (Zhang et al., 2018), the neural network model (Biancofiore et al., 2017), the support vector machine (García Nieto et al., 2013), the weather research and forecasting model with chemistry (Saide et al., 2011), and the combined model (Zhu et al., 2017) to conduct a series of smog studies. By comparing the above methods, it is found that these methods mainly forecast smog data with high temporal resolution. However, it is not always possible to have data with high temporal resolution, and a number of scholars have applied grey prediction models to forecast smog based on data with low temporal resolution (Wu et al., 2018). Grey prediction model is an important branch of grey system theory. As a new uncertainty system theory, grey system theory is characterized by small data modeling to obtain accurate results. Since Professor Deng founded grey system theory in 1982, the research object of grey systems has been the uncertain systems with small data and poor information, mainly through in-depth mining of existing information and extracting valuable information, so as to study the system intrinsic law (Liu et al., 2010). Considering the unexpected suddenness of smog, the smog related indicator data is usually uncertain. In the system research, due to the limitation of human cognitive ability, it is difficult to fully understand the information reflecting the system operation behavior, resulting in people only obtain the value range of system elements or parameters. Usually, the number that only knows the value range but does not know its exact value is called grey number. In other words, the value of smog has obvious characteristics of grey numbers which is uncertain within a certain range.

GM(1,1) model and GM(1,N) model are important components of grey prediction models. GM(1,1) model is the basic model of grey prediction models, which mainly simulates and forecasts the behavior sequence of a single system. Scholars not only optimized GM(1,1) model from the aspects of background value optimization (Wang et al., 2018; Zeng et al., 2020; Wu et al., 2013), coefficient improvement(Zeng et al., 2016; Wang et al., 2019), but also combined it with the statistical model to solve practical problems (Yuan et al., 2016). GM(1,1) model does not count the influence of relative factors on the change of system characteristic data, and GM(1,N) model, as an extension of GM(1,1) model, fully considers the interference of relative factors on system characteristic data. Therefore, scholars established GM(1,N) model to conduct overall and global dynamic analysis of the system, and improved the system via background value (Wang et al., 2016; Ma et al., 2018), driving terms optimization (Wu et al., 2018; Ding et al., 2018; Zeng et al, 2019), discretization (Ding, 2019; Ma et al., 2019), coefficient improvement (Wu, 2018; Wang, 2014) etc. These researches focus mainly on real number sequences, and optimize the GM(1,N) model on both system characteristic data and relative factors changing at the same time. However, these models are not suitable for systems with lag. Time-lag effect is to describe the delay relationship in the system, which is widely used in transportation (Zhou et al., 2017), economy (Lee et al., 2016), meteorology (Zheng et al., 2020) and other fields, so many scholars have also improved the GM(1,N) model from the angle of system lag, and constructed time-lag GM(1,N) model and the derived model(Zhai et al., 1996; Wang et al.,2015), discrete time-lag GM(1,N) model(Zhang et al., 2015; Dang et al., 2017; Ding et al., 2017), time-lag GM(1,N) model with fractional order accumulation (Mao et al., 2015). Through comparison, it is found that the literatures of time-lag mainly are focusing on the discrete grey prediction model, and the research results are relatively similar. Therefore, systematic studying on the non-discrete GM(1,N) model with time-lag dynamic change characteristics of is necessary.

Due to limited cognitive ability, observation errors and measurement errors, a lot of real number variables cannot be accurately described by real numbers, so many scholars have begun to explore the modeling mechanism of grey prediction model based on interval grey number sequences (Dang et al., 2018; Xie et al., 2018). At present, the existing GM(1,N) model based on interval grey number sequences (Xiong et al., 2018) can only model simultaneous changing variables, and cannot consider the relationship of time-lag cumulative effect in the real social system, which may lead to larger prediction error. Therefore, this paper will analyze the time-lag effect of previous relative factors on the current system characteristic data, construct a time-lag GM(1,N) model based on interval grey number sequences, and determine time-lag parameters of the model respectively. Meanwhile, this paper will consider the time-lag effect of

SO_2 and NO_2 on $PM_{2.5}$, establish the time-lag GM(1,3) model to simulate and forecast $PM_{2.5}$ concentration in Beijing, and compare the result with the time-lag discrete GM(1,3) model, traditional GM(1,3) model and multiple linear regression model. Besides, the time-lag GM(1,3) model is also applied to forecast the smog situation of Beijing.

The other research arrangements in this paper are as follows: the basic concepts of interval grey numbers and traditional GM(1,N) models are introduced in Section 2; the modeling mechanism of the time-lag GM(1,N) model based on kernel and degree of greyness sequences are presented respectively, determination of time-lag parameters and model checking method are also introduced in Section 3; the time-lag GM(1,3) model based on interval grey number sequences is applied to simulate and forecast monthly $PM_{2.5}$ concentration of Beijing in Section 4; conclusions are summarized in Section 5.

2 Basic concepts

2.1 Interval grey number

Definition 1 (Liu et al., 2010) The grey number with both lower bound a_k and upper bound b_k is called interval grey number, and interval grey number is denoted as $\otimes_k \in [a_k, b_k]$.

Definition 2 (Liu et al., 2010) Assume that the interval grey number is $\otimes_k \in [a_k, b_k](a_k < b_k)$, in the absence of interval grey number value distribution information, we can obtain:

(1) When $\otimes_k$ is a continuous function, so $\widetilde{\otimes} = (a_k + b_k)/2$ is the kernel of interval grey number $\otimes_k$;

(2) When $\otimes_k$ is a discrete function, and $a_i \in [a_k, b_k](k = 1,2,\cdots,n)$ is all possible values of interval grey number $\otimes_k$, so $\widetilde{\otimes} = \frac{1}{n}\sum_{i=1}^{n} a_i$ is the kernel of interval grey number $\otimes_k$.

Definition 3 (Liu et al., 2010) When the background or domain of interval grey number $\otimes_k$ is Ω, and $\mu(\otimes_k)$ is a measure on Ω, so $g^{\circ}(\otimes_k) = \mu(\otimes_k)/\mu(\Omega)$ is called the degree of greyness in interval grey number $\otimes_k$.

Definition 4 (Liu et al., 2010) The sequence consisting of interval grey number $\otimes_k \in [a_k, b_k], k = 1,2,\cdots,n$ is called the interval grey number sequence $X(\otimes)$. The sequence of all upper bounds in $X(\otimes)$ is called the upper bound sequence of $X(\otimes)$, denoted as $X_b = (b_1, b_2, \cdots, b_n)$. The sequence of all lower bounds in $X(\otimes)$ is called the lower bound sequence of $X(\otimes)$, denoted as $X_a = (a_1, a_2, \cdots, a_n)$. The sequence of all kernels in $X(\otimes)$ is called the kernel sequence of $X(\otimes)$, denoted as $X_{\widetilde{\otimes}} = (\widetilde{\otimes}_1, \widetilde{\otimes}_2, \cdots, \widetilde{\otimes}_n)$. The sequence of all degree of greyness in $X(\otimes)$ is called the degree of greyness sequence in $X(\otimes)$, denoted as $X_{g^{\circ}} = (g_1^{\circ}, g_2^{\circ}, \cdots, g_n^{\circ})$.

Lemma 1 (Zeng, 2011) Assume that $\otimes_k \in [a_k, b_k]$ is the continuous interval grey number, $\widetilde{\otimes} = (a_k + b_k)/2$ is the kernel of interval grey number $\otimes_k$, and $g^{\circ}(\otimes_k) = \mu(\otimes_k)/\mu(\Omega)$ is the degree of greyness in interval grey number $\otimes_k$, so we have $a_k = \widetilde{\otimes}_k - 0.5 g^{\circ}(\otimes_k)\mu(\otimes_k)$, $b_k = \widetilde{\otimes}_k + 0.5 g^{\circ}(\otimes_k)\mu(\otimes_k)$.

2.2 GM(1,N) model

Definition 5 (Liu et al., 2010) Assume that the system characteristic data sequence is $X_1^{(0)} = \left(x_1^{(0)}(1), x_1^{(0)}(2), \cdots, x_1^{(0)}(n)\right)$, and relative factor sequences are shown as follows:

$$X_2^{(0)} = \left(x_2^{(0)}(1), x_2^{(0)}(2), \cdots, x_2^{(0)}(n)\right),$$

$$X_3^{(0)} = \left(x_3^{(0)}(1), x_3^{(0)}(2), \cdots, x_3^{(0)}(n)\right),$$

$$\cdots\cdots$$

$$X_N^{(0)} = \left(x_N^{(0)}(1), x_N^{(0)}(2), \cdots, x_N^{(0)}(n)\right),$$

when $X_i^{(1)}$ is the first order accumulating generation sequence of $X_i^{(0)}, i = 1,2,\cdots,N$, and $Z_1^{(1)} = (z_1^{(1)}(2), z_1^{(1)}(3)\cdots, z_1^{(1)}(n))$ is the mean sequence generated by the consecutive neighbor of $X_1^{(1)}$,and $z_1^{(1)}(k) = 0.5(x_1^{(1)}(k) + x_1^{(1)}(k-1))$, the GM(1,N) model is shown as follows:

$$x_1^{(0)}(k) + a z_1^{(1)}(k) = \sum_{i=2}^{N} b_i x_i^{(1)}(k).$$

Definition 6 (Liu et al., 2010) In the GM(1,N) model, a represents the system development coefficient, $b_i x_i^{(1)}(k)$ represents the driving term, b_i represents the driving coefficient, and $\hat{a} = [a, b_2,\cdots, b_N]^T$ represents the undetermined coefficient vector.

Theorem 1 (Liu et al., 2010) When $X_1^{(0)}$ is the system characteristic data sequence, $X_i^{(0)}(i = 2,3,\dots,N)$ is the relative factor sequence, $X_i^{(1)}$ is the first order accumulating generation sequence of $X_i^{(0)}$, and $Z_1^{(1)}$ is the mean sequence generated by the consecutive neighbor of $X_1^{(1)}$, where

$$B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & \cdots & x_N^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & \cdots & x_N^{(1)}(3) \\ \vdots & \vdots & & \vdots \\ -z_1^{(1)}(n) & x_2^{(1)}(n) & \cdots & x_N^{(1)}(n) \end{bmatrix}, Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{bmatrix}.$$

Therefore, the undetermined coefficient vector of the GM(1,N) model is $\hat{a} = [a, b_2,\cdots, b_N]^T$, which can be estimated by the least squares method as follows:

$$\hat{a} = (B^T B)^{-1} B^T Y.$$

Theorem 2 (Liu et al., 2010) The coefficient vector $\hat{a}$ of the GM(1,N) model satisfies above conditions, thus we obtain as follows:

(1) The whitening equation is shown as follows:

$$\frac{dx_1^{(1)}}{dt} + a x_1^{(1)} = \sum_{i=2}^{N} b_i x_i^{(1)},$$

and the solution of the whitening equation can be obtained as follows:

$$x^{(1)}(t) = e^{-at}\left[x_1^{(1)}(0) - t\sum_{i=2}^{N} b_i x_i^{(1)}(0) + \sum_{i=2}^{N}\int b_i x_i^{(1)}(t) e^{at}\, dt\right].$$

(2) When the change of $X_i^{(1)}$ is very small, we suppose that $\sum_{i=2}^{N} b_i x_i^{(1)}(k)$ is a grey constant. Therefore, the approximate time response of the GM(1,N) model is shown as follows:

$$\hat{x}_1^{(1)}(k+1) = \frac{1}{a}\sum_{i=2}^{N} b_i x_i^{(1)}(k+1) + e^{-ak}\left[x_1^{(1)}(0) - \frac{1}{a}\sum_{i=2}^{N} b_i x_i^{(1)}(k+1)\right],$$

where the solution can be obtained with the initial condition $x_1^{(1)}(0) = x_1^{(0)}(1)$.

(3) The inverse accumulating reduction equation is shown as follows:

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

3 Methodology

3.1 Time-lag GM(1,N) model based on kernel sequences

Definition 7 Assume that the kernel sequence in the interval grey number sequence is $\widetilde{\otimes}_i^{(0)} = \left(\widetilde{\otimes}_i^{(0)}(1), \widetilde{\otimes}_i^{(0)}(2), \cdots, \widetilde{\otimes}_i^{(0)}(n)\right), i = 1,2,\cdots,N,$ and the kernel is shown as follows:

$$\widetilde{\otimes}_i^{(0)}(j) = \frac{(a_{ij} + b_{ij})}{2}, i = 1,2,\cdots,N, j = 1,2,\cdots,n. \tag{1}$$

Definition 8 Assume that the system characteristic data of kernel sequence is shown as follows:

$$\widetilde{\otimes}_1^{(0)} = \left(\widetilde{\otimes}_1^{(0)}(1), \widetilde{\otimes}_1^{(0)}(2), \cdots, \widetilde{\otimes}_1^{(0)}(n)\right),$$

and relative factor sequences of kernel sequence are shown as follows:

$$\widetilde{\otimes}_2^{(0)} = \left(\widetilde{\otimes}_2^{(0)}(-l), \cdots \widetilde{\otimes}_2^{(0)}(-2), \widetilde{\otimes}_2^{(0)}(-1), \widetilde{\otimes}_2^{(0)}(0), \widetilde{\otimes}_2^{(0)}(1), \cdots, \widetilde{\otimes}_2^{(0)}(n)\right),$$

$$\widetilde{\otimes}_3^{(0)} = \left(\widetilde{\otimes}_3^{(0)}(-l), \cdots \widetilde{\otimes}_3^{(0)}(-2), \widetilde{\otimes}_3^{(0)}(-1), \widetilde{\otimes}_3^{(0)}(0), \widetilde{\otimes}_3^{(0)}(1), \cdots, \widetilde{\otimes}_3^{(0)}(n)\right),$$

$$\cdots\cdots$$

$$\widetilde{\otimes}_N^{(0)} = \left(\widetilde{\otimes}_N^{(0)}(-l), \cdots \widetilde{\otimes}_N^{(0)}(-2), \widetilde{\otimes}_N^{(0)}(-1), \widetilde{\otimes}_N^{(0)}(0), \widetilde{\otimes}_N^{(0)}(1), \cdots, \widetilde{\otimes}_N^{(0)}(n)\right),$$

where n, l are positive integer, $\widetilde{\otimes}_N^{(0)}(1), \cdots, \widetilde{\otimes}_N^{(0)}(n)$ means current data and $\widetilde{\otimes}_N^{(0)}(0), \cdots, \widetilde{\otimes}_N^{(0)}(-l)$ means previous data. Thus, the time-lag relative factors of kernel sequence are shown as follows:

$$\widetilde{\otimes}_2^{(0)} = \left(\widetilde{\otimes}_2^{(0)}(1-\tau_2), \widetilde{\otimes}_2^{(0)}(2-\tau_2), \cdots, \widetilde{\otimes}_2^{(0)}(n-\tau_2)\right),$$

$$\widetilde{\otimes}_3^{(0)} = \left(\widetilde{\otimes}_3^{(0)}(1-\tau_3), \widetilde{\otimes}_3^{(0)}(2-\tau_3), \cdots, \widetilde{\otimes}_3^{(0)}(n-\tau_3)\right),$$

$$\cdots\cdots$$

$$\widetilde{\otimes}_N^{(0)} = \left(\widetilde{\otimes}_N^{(0)}(1-\tau_N), \widetilde{\otimes}_N^{(0)}(2-\tau_N), \cdots, \widetilde{\otimes}_N^{(0)}(n-\tau_N)\right),$$

where τ_i, n are positive integer, and $\tau_i \in [0, l+1]$ $(i = 2,3,\cdots,N)$ is the time-lag parameter for relative factors. When $\widetilde{\otimes}_i^{(1)}$ is the first order accumulating generation sequence of $\widetilde{\otimes}_i^{(0)}$, $i = 1,2,\cdots,N$, and

$Z_1^{(1)}$ is the mean sequence generated by the consecutive neighbor of $\tilde{\otimes}_1^{(1)}$, the time-lag GM(1,N) model based on kernel sequences is shown as follows:

$$\tilde{\otimes}_1^{(0)}(k) + a z_1^{(1)}(k) = \sum_{i=2}^{N} b_i \tilde{\otimes}_i^{(1)}(k - \tau_i) . \tag{2}$$

Especially, when $\tau_i = 0$ indicates that all relative factors and system characteristic data sequences are synchronous variables, the time-lag GM(1,N) model based on kernel sequences degenerates to the GM(1,N) model based on kernel sequences.

Definition 9 In the time-lag GM(1,N) model based on kernel sequences, a represents the system development coefficient, $b_i \tilde{\otimes}_i^{(1)}(k - \tau_i)$ represents the driving term, b_i represents the driving coefficient, and $\hat{a} = [a, b_2, \cdots, b_N]^T$ represents the undetermined coefficient vector.

Theorem 3 When $\tilde{\otimes}_1^{(0)}$ is the system characteristic data sequence, $\tilde{\otimes}_i^{(0)}$ $(i = 2,3,\ldots,N)$ is the relative factor sequence, $\tilde{\otimes}_i^{(1)}$ is the first order accumulating generation sequence of $\tilde{\otimes}_i^{(0)}$, $Z_1^{(1)}$ is the mean sequence generated by the consecutive neighbor of $\tilde{\otimes}_1^{(1)}$, and τ_i is the time-lag parameters of relative factors, where

$$\boldsymbol{B} = = \begin{bmatrix} -z_1^{(1)}(2) & \tilde{\otimes}_2^{(1)}(2 - \tau_2) & \cdots & \tilde{\otimes}_N^{(1)}(2 - \tau_n) \\ -z_1^{(1)}(3) & \tilde{\otimes}_2^{(1)}(3 - \tau_2) & \cdots & \tilde{\otimes}_N^{(1)}(3 - \tau_n) \\ \vdots & \vdots & & \vdots \\ -z_1^{(1)}(n) & \tilde{\otimes}_2^{(1)}(n - \tau_2) & \cdots & \tilde{\otimes}_N^{(1)}(n - \tau_n) \end{bmatrix}, \boldsymbol{Y} = \begin{bmatrix} \tilde{\otimes}_1^{(0)}(2) \\ \tilde{\otimes}_1^{(0)}(3) \\ \vdots \\ \tilde{\otimes}_1^{(0)}(n) \end{bmatrix}.$$

Therefore, the undetermined coefficient vector of the time-lag GM(1,N) model based on kernel sequences is $\hat{a} = [a, b_2, \cdots, b_N]^T$, which can be estimated by the least squares method as follows:

(1) When $n = \mathrm{N} + 1$, and $|\boldsymbol{B}| \neq 0$, so $\hat{\boldsymbol{a}} = \boldsymbol{B}^{-1}\boldsymbol{Y}$;

(2) When $n > \mathrm{N} + 1$, and $|\boldsymbol{B}^T\boldsymbol{B}| \neq 0$, so $\hat{\boldsymbol{a}} = (\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Y}$;

(3) When $n < \mathrm{N} + 1$, and $|\boldsymbol{B}^T\boldsymbol{B}| \neq 0$, so $\hat{\boldsymbol{a}} = \boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{B}^T)^{-1}\boldsymbol{Y}$.

Proof: By substituting $k = 2,3,\cdots,n$ into formula (2), equations are obtained as follows:

$$\begin{gathered} \tilde{\otimes}_1^{(1)}(2) = -a z_1^{(1)}(2) + b_2 \tilde{\otimes}_2^{(1)}(2 - \tau_2) + \cdots + b_N \tilde{\otimes}_N^{(1)}(2 - \tau_N), \\ \tilde{\otimes}_1^{(1)}(3) = -a z_1^{(1)}(3) + b_2 \tilde{\otimes}_2^{(1)}(3 - \tau_2) + \cdots + b_N \tilde{\otimes}_N^{(1)}(3 - \tau_N), \\ \vdots \\ \tilde{\otimes}_1^{(1)}(n) = -a z_1^{(1)}(n) + b_2 \tilde{\otimes}_2^{(1)}(n - \tau_2) + \cdots + b_N \tilde{\otimes}_N^{(1)}(n - \tau_N). \end{gathered}$$

Therefore, we can obtain by the least square method as follows:

$$\boldsymbol{Y} = \boldsymbol{B}\boldsymbol{a}.$$

(1) When $n = \mathrm{N} + 1$ and $|\boldsymbol{B}| \neq 0$, and $\boldsymbol{B}$ exists an inverse matrix, the equation has unique solutions, that is $\hat{\boldsymbol{a}} = \boldsymbol{B}^{-1}\boldsymbol{Y}$.

(2) When $n > N + 1$, and $\boldsymbol{B}$ is a column full rank matrix, we have the full rank decomposition of $\boldsymbol{B}$ is $\boldsymbol{B} = \boldsymbol{D}\boldsymbol{C}$, then the generalized inverse matrix of $\boldsymbol{B}$ can be obtained as follows:

$$\boldsymbol{B}^{+} = \boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{C}^T)^{-1}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T,$$

$$\hat{\boldsymbol{a}} = \boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{C}^T)^{-1}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{Y}.$$

Due to $\boldsymbol{B}$ is a column full rank matrix, when $\boldsymbol{C}$ is an identity matrix, we have $\boldsymbol{B} = \boldsymbol{D}\boldsymbol{I_n}$, $\boldsymbol{B} = \boldsymbol{D}$, that is

$$\hat{\boldsymbol{a}} = \boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{C}^T)^{-1}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{Y} = (\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{Y} = (\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Y}.$$

(3) When $n < N+1$, $\boldsymbol{B}$ is a row full rank matrix, when $\boldsymbol{D}$ is an identity matrix, we have $\boldsymbol{B} = \boldsymbol{I_n}\boldsymbol{C}$, $\boldsymbol{B} = \boldsymbol{C}$, that is

$$\hat{\boldsymbol{a}} = \boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{C}^T)^{-1}(\boldsymbol{D}^T\boldsymbol{D})^{-1}\boldsymbol{D}^T\boldsymbol{Y} = \boldsymbol{C}^T(\boldsymbol{C}\boldsymbol{C}^T)^{-1}\boldsymbol{Y} = \boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{B}^T)^{-1}\boldsymbol{Y}.$$

Definition 10 Let $\hat{\boldsymbol{a}} = [a, b_2, \cdots, b_N]^T$, the whitening equation for a time-lag GM(1,N) model based on kernel sequences is shown as follows:

$$\frac{d\,\widetilde{\otimes}_1^{(1)}}{dt} + a\,\widetilde{\otimes}_1^{(1)} = \sum_{i=2}^{N} b_i\,\widetilde{\otimes}_i^{(1)}\,(t-\tau_i)\,. \tag{3}$$

Theorem 4 The coefficient vector $\hat{a}$ of the time-lag GM(1,N) model based on kernel sequences satisfies above conditions, thus we obtain the solutions as follows:

(1) The solution of the whitening equation can be obtained as follows:

$$\widetilde{\otimes}^{(1)}\,(t) = e^{-at}\left[\widetilde{\otimes}_1^{(1)}\,(0) - t\sum_{i=2}^{N} b_i\,\widetilde{\otimes}_i^{(1)}\,(0) + \sum_{i=2}^{N}\int b_i\,\widetilde{\otimes}_i^{(1)}\,(t-\tau_i)e^{at}\,dt\right]. \tag{4}$$

(2) When the change of $\widetilde{\otimes}_i^{(1)}$ is very small, we suppose that $\sum_{i=2}^{N} b_i\,\widetilde{\otimes}_i^{(1)}\,(k-\tau_i)$ is a grey constant. Therefore, the approximate time response of the time-lag GM(1,N) model based on kernel sequences is shown as follows:

$$\begin{aligned}\hat{\widetilde{\otimes}}_1^{(1)}\,(k+1) &= \frac{1}{a}\sum_{i=2}^{N} b_i\,\widetilde{\otimes}_i^{(1)}\,(k+1-\tau_i) \\ &\quad + e^{-ak}\left[\widetilde{\otimes}_1^{(1)}\,(0) - \frac{1}{a}\sum_{i=2}^{N} b_i\,\widetilde{\otimes}_i^{(1)}\,(k+1-\tau_i)\right].\end{aligned} \tag{5}$$

where the solution can be obtained with the initial condition $\widetilde{\otimes}_1^{(1)}\,(0) = \widetilde{\otimes}_1^{(0)}\,(1)$.

(3) The inverse accumulating reduction equation is shown as follows:

$$\hat{\widetilde{\otimes}}_1^{(0)}\,(k+1) = \hat{\widetilde{\otimes}}_1^{(1)}\,(k+1) - \hat{\widetilde{\otimes}}_1^{(1)}\,(k). \tag{6}$$

3.2 Time-lag GM(1,N) model based on degree of greyness sequences

Definition 11 Assume that the degree of greyness sequence in the interval grey number sequence is $g_i^{\circ(0)} = \left(g_i^{\circ(0)}(1), g_i^{\circ(0)}(2), \cdots, g_i^{\circ(0)}(n)\right), i = 1,2,\cdots,N,$ and the degree of greyness is shown as follows:

$$g_i^{\circ(0)}(j) = \frac{(b_{ij} - a_{ij})}{\mu(\Omega)}, i = 1,2,\cdots,N, j = 1,2,\cdots,n. \tag{7}$$

Definition 12 Assume that the system characteristic data sequence of the degree of greyness is shown as follows:

$$g_1^{\circ(0)} = \left(g_1^{\circ(0)}(1), g_1^{\circ(0)}(2), \cdots, g_1^{\circ(0)}(n)\right),$$

and the relative factor sequences of the degree of greyness are shown as follows:

$$g_2^{\circ(0)} = \left(g_2^{\circ(0)}(-l), \cdots g_2^{\circ(0)}(-2), g_2^{\circ(0)}(-1), g_2^{\circ(0)}(0), g_2^{\circ(0)}(1), \cdots, g_2^{\circ(0)}(n)\right),$$

$$g_3^{\circ(0)} = \left(g_3^{\circ(0)}(-l), \cdots g_3^{\circ(0)}(-2), g_3^{\circ(0)}(-1), g_3^{\circ(0)}(0), g_3^{\circ(0)}(1), \cdots, g_3^{\circ(0)}(n)\right),$$

$$\cdots\cdots$$

$$g_N^{\circ(0)} = \left(g_N^{\circ(0)}(-l), \cdots g_N^{\circ(0)}(-2), g_N^{\circ(0)}(-1), g_N^{\circ(0)}(0), g_N^{\circ(0)}(1), \cdots, g_N^{\circ(0)}(n)\right),$$

where n, l are positive integer, $g_N^{\circ(0)}(1), \cdots, g_N^{\circ(0)}(n)$ means current data and $g_N^{\circ(0)}(0), \cdots, g_N^{\circ(0)}(-l)$ means previous data. Thus, the time-lag relative factors of the degree of greyness sequence are shown as follows:

$$g_2^{\circ(0)} = \left(g_2^{\circ(0)}(1-\tau_2), g_2^{\circ(0)}(2-\tau_2), \cdots, g_2^{\circ(0)}(n-\tau_2)\right),$$

$$g_3^{\circ(0)} = \left(g_3^{\circ(0)}(1-\tau_3), g_3^{\circ(0)}(2-\tau_3), \cdots, g_3^{\circ(0)}(n-\tau_3)\right),$$

$$\cdots\cdots$$

$$g_N^{\circ(0)} = \left(g_N^{\circ(0)}(1-\tau_N), g_N^{\circ(0)}(2-\tau_N), \cdots, g_N^{\circ(0)}(n-\tau_N)\right),$$

where τ_i, n are positive integer, and $\tau_i \in [0, l+1]$ $(i = 2,3,\cdots,N)$ is the time-lag parameter for relative factors. When $g_i^{\circ(1)}$ is the first order accumulating generation sequence of $g_i^{\circ(0)}$, $i = 1,2,\cdots,N$, and $Z_1^{(1)}$ is the mean sequence generated by the consecutive neighbor of $g_1^{\circ(1)}$, the time-lag GM(1,N) model based on degree of greyness sequences is shown as follows:

$$g_1^{\circ(0)}(k) + az_1^{(1)}(k) = \sum_{i=2}^{N} b_i g_i^{\circ(1)}(k-\tau_i)\ . \tag{8}$$

Especially, when $\tau_i = 0$ indicates that all relative factors and system characteristic data sequences are synchronous variables, the time-lag GM(1,N) model based on degree of greyness sequences degenerates to the GM(1,N) model based on degree of greyness sequences.

Definition 13 In the time-lag GM(1,N) model based on degree of greyness sequences, a represents the system development coefficient, $b_i g_i^{\circ(1)}(k-\tau_i)$ represents the driving term, b_i represents the driving coefficient, and $\hat{a} = [a, b_2, \cdots, b_N]^T$ represents the undetermined coefficient vector.

Theorem 5 $g_1^{\circ(0)}$ is the system characteristic data sequence, $g_i^{\circ(0)}(i = 2,3,\dots,N)$ is the relative factor sequence, $g_i^{\circ(1)}$ is the first order accumulating generation sequence of $g_i^{\circ(0)}$, $Z_1^{(1)}$ is the mean sequence generated by the consecutive neighbor of $g_1^{\circ(1)}$, and τ_i is the time-lag parameter of relative factors, where

$$\boldsymbol{B}=\begin{bmatrix}-z_1^{(1)}(2) & g_2^{\circ(1)}(2-\tau_2) & \cdots & g_N^{\circ(1)}(2-\tau_N)\\ -z_1^{(1)}(3) & g_2^{\circ(1)}(3-\tau_2) & \cdots & g_N^{\circ(1)}(3-\tau_N)\\ \vdots & \vdots & & \vdots\\ -z_1^{(1)}(n) & g_2^{\circ(1)}(n-\tau_2) & \cdots & g_N^{\circ(1)}(n-\tau_N)\end{bmatrix},\boldsymbol{Y}=\begin{bmatrix}g_1^{\circ(0)}(2)\\ g_1^{\circ(0)}(3)\\ \vdots\\ g_1^{\circ(0)}(n)\end{bmatrix},$$

The undetermined coefficient vector of the time-lag GM(1,N) model based on degree of greyness sequences is $\hat{a}=[a,b_2,\cdots,b_N]^T$, which can be estimated by the least squares method as follows:

(1) When $n=\mathrm{N}+1$, and $|\boldsymbol{B}|\neq 0$, so $\widehat{\boldsymbol{a}}=\boldsymbol{B}^{-1}\boldsymbol{Y}$;

(2) When $n>\mathrm{N}+1$, and $|\boldsymbol{B}^T\boldsymbol{B}|\neq 0$, so $\widehat{\boldsymbol{a}}=(\boldsymbol{B}^T\boldsymbol{B})^{-1}\boldsymbol{B}^T\boldsymbol{Y}$;

(3) When $n<\mathrm{N}+1$, and $|\boldsymbol{B}^T\boldsymbol{B}|\neq 0$, so $\widehat{\boldsymbol{a}}=\boldsymbol{B}^T(\boldsymbol{B}\boldsymbol{B}^T)^{-1}\boldsymbol{Y}$.

The proof is same as Theorem 3.

Definition 14 Let $\widehat{\boldsymbol{a}}=[a,b_2,\cdots,b_N]^T$, the whitening equation for a time-lag GM(1,N) model based on degree of greyness sequences is shown as follows:

$$\frac{dg_1^{\circ(1)}}{dt}+ag_1^{\circ(1)}=\sum_{i=2}^{N}b_i g_i^{\circ(1)}(t-\tau_i)\ . \tag{9}$$

Theorem 6 The coefficient vector $\hat{a}$ of the time-lag GM(1,N) model based on degree of greyness sequences satisfies above conditions, thus we obtain the solutions as follows:

(1) The solution of the whitening equation can be obtained as follows:

$$g^{\circ(1)}(t)=e^{-at}\left[g_1^{\circ(1)}(0)-t\sum_{i=2}^{N}b_i g_i^{\circ(1)}(0)+\sum_{i=2}^{N}\int b_i g_i^{\circ(1)}(t-\tau_i)e^{at}\,dt\right]. \tag{10}$$

(2) When the change of $g{}_i^{\circ(1)}$ is very small, we suppose that $\sum_{i=2}^{N}b_i g{}_i^{\circ(1)}(k-\tau_i)$ is a grey constant. Therefore, the approximate time response of the time-lag GM(1,N) model based on degree of greyness sequences is shown as follows:

$$\begin{aligned}\widehat{g^{\circ}}{}_1^{(1)}(k+1)=&\frac{1}{a}\sum_{i=2}^{N}b_i g_i^{\circ(1)}(k+1-\tau_i)\\ &+e^{-ak}\left[g_1^{\circ(1)}(0)-\frac{1}{a}\sum_{i=2}^{N}b_i g_i^{\circ(1)}(k+1-\tau_i)\right].\end{aligned} \tag{11}$$

where the solution can be obtained with the initial condition $g_1^{\circ(1)}(0)=g_1^{\circ(0)}(1)$.

(3) The inverse accumulating reduction equation is shown as follows:

$$\widehat{g_1^{\circ}}^{(0)}(k+1)=\widehat{g_1^{\circ}}^{(1)}(k+1)-\widehat{g_1^{\circ}}^{(1)}(k). \tag{12}$$

3.3 Identification method of time-lag parameter τ_i

In the modeling process of the time-lag GM(1,N) model based on interval grey number sequences, identifying the time-lag parameter τ_i is the most critical step, which directly affects the modeling and prediction accuracy of the model.

This paper considers from the angle of improving the model prediction accuracy, by minimizing the average relative error of modeling as the optimal objective function and taking the relationship between model parameters as the constraint condition, the objective function of the time-lag GM(1,N) model based on kernel sequences is constructed as follows:

$$\min_{\tau_i} avg(e(k)) = \frac{1}{n-1}\sum_{k=2}^{n}\left|\frac{\widehat{\tilde{\otimes}}_k - \tilde{\otimes}_k}{\tilde{\otimes}_k}\right|, \tag{13}$$

$$s.t.\begin{cases} \hat{\boldsymbol{a}} = [a, b_2, \cdots, b_N]^T; \\ \widehat{\tilde{\otimes}}_1^{(0)}(k) = \widehat{\tilde{\otimes}}_1^{(1)}(k) - \widehat{\tilde{\otimes}}_1^{(1)}(k-1), k = 2,3,\cdots,n; \\ \widehat{\tilde{\otimes}}_1^{(1)}(k) = \frac{1}{a}\sum_{i=2}^{N} b_i\, \tilde{\otimes}_i^{(1)}(k-\tau_i) + e^{-a(k-1)}\left[\tilde{\otimes}_1^{(1)}(0) - \frac{1}{a}\sum_{i=2}^{N} b_i\, \tilde{\otimes}_i^{(1)}(k-\tau_i)\right]; \end{cases}$$

Similarly, the objective function of the time-lag GM(1,N) model based on degree of greyness sequences is constructed as follows:

$$\min_{\tau_i} avg(e(k)) = \frac{1}{n-1}\sum_{k=2}^{n}\left|\frac{\widehat{g_k^{\circ}} - g_k^{\circ}}{g_k^{\circ}}\right|, \tag{14}$$

$$s.t.\begin{cases} \hat{\boldsymbol{a}} = [a, b_2, \cdots, b_N]^T; \\ \widehat{g_1^{\circ}}^{(0)}(k) = \widehat{g_1^{\circ}}^{(1)}(k) - \widehat{g_1^{\circ}}^{(1)}(k-1), k = 2,3,\cdots,n; \\ \widehat{g^{\circ}}_1^{(1)}(k) = \frac{1}{a}\sum_{i=2}^{N} b_i g_i^{\circ(1)}(k-\tau_i) + e^{-ak}\left[g_1^{\circ(1)}(0) - \frac{1}{a}\sum_{i=2}^{N} b_i g_i^{\circ(1)}(k-\tau_i)\right]; \end{cases}$$

Formula (13) and formula (14) can be solved by LINGO, MATLAB or some intelligent optimization algorithms (such as particle swarm optimization, genetic algorithm, etc.). After the optimal solution of the time-lag parameter τ_i is determined, the structural parameters of the model can be estimated, and then the simulation and prediction results can be obtained through the optimal time response function.

3.4 Model error testing

In order to study the reliability of the prediction results, it is necessary to test the error of the model. In this paper, the prediction model is tested by comparing the relative error and average relative error of upper and lower bounds in interval grey number sequences, the average relative error of grey elements in interval grey number sequences, and the model average relative error. The calculation formulas are shown as follows:

The relative error of grey elements upper and lower bounds in interval grey number sequences is shown as follows:

$$\begin{cases} \overline{\Delta_k} = \dfrac{|\hat{a}_k - a_k|}{a_k} \times 100\% \\ \underline{\Delta_k} = \dfrac{|\hat{b}_k - b_k|}{b_k} \times 100\% \end{cases}, k = 1,2,\cdots,n. \tag{15}$$

The average relative error of grey elements upper and lower bounds in interval grey number sequences is shown as follows:

$$\begin{cases} \overline{\hat{a}_k} = \frac{1}{n}\sum_{k=1}^{n}\overline{\Delta_k} \\ \overline{\hat{b}_k} = \frac{1}{n}\sum_{k=1}^{n}\underline{\Delta_k} \end{cases}, k = 1,2,\cdots,n. \tag{16}$$

The average relative error of grey elements in interval grey number sequences is shown as follows:

$$\Delta_k = \frac{1}{2}\sum_{k=1}^{n}\left(\overline{\Delta_k} + \underline{\Delta_k}\right), k = 1,2,\cdots,n. \tag{17}$$

The model average relative error is shown as follows:

$$\Delta = \frac{1}{n}\sum_{k=1}^{n}\Delta_k = \frac{1}{2}\sum_{k=1}^{n}\left(\overline{\hat{a}_k} + \overline{\hat{b}_k}\right). \tag{18}$$

The prediction accuracy is an important criterion to measure the quality of prediction model. This paper specifies the prediction accuracy corresponding to the average relative error, that is the prediction model passing the model error testing when the model average relative error is less than 10%, as shown in Table 1.

Table1 Prediction accuracy corresponding to average relative error (Lewis, 1982)

Average relative error (%)	Prediction accuracy
<10	High
10-20	Good
20-50	Reasonable
>50	Weak

3.5 Modeling steps

The time-lag GM(1,N) model based on interval grey sequences can be established as illustrated as follows:

Step 1: After obtaining the raw data, both kernel and degree of greyness in the interval grey sequence of the system characteristic data and relative factors are calculated.

Step 2: By minimizing the average relative error of modeling as the optimal objective function, the time-lag parameter τ_i of the time-lag GM(1,N) model are determined in both kernel and degree of greyness sequences respectively.

Step 3: After calculating the coefficient vector $\hat{a} = [a, b_2,\cdots, b_N]^T$ of the time-lag GM(1,N) model based on kernel and degree of greyness sequences, the time-lag GM(1,N) model can be built based on kernel and degree of greyness sequences respectively, so values of both kernel and degree of greyness can be respectively obtained.

Step 4: The upper and lower bounds of the interval grey number sequence are obtained by restoring the simulated and predicted values of kernel and degree of greyness sequences.

Step5: Calculating the model errors, if the model average relative error passes the model error testing, the modeling process will end. If not, the raw data will be updated, and start again from Step 1.

In order to present the modeling process more clearly, modeling steps are as shown in Fig.1.

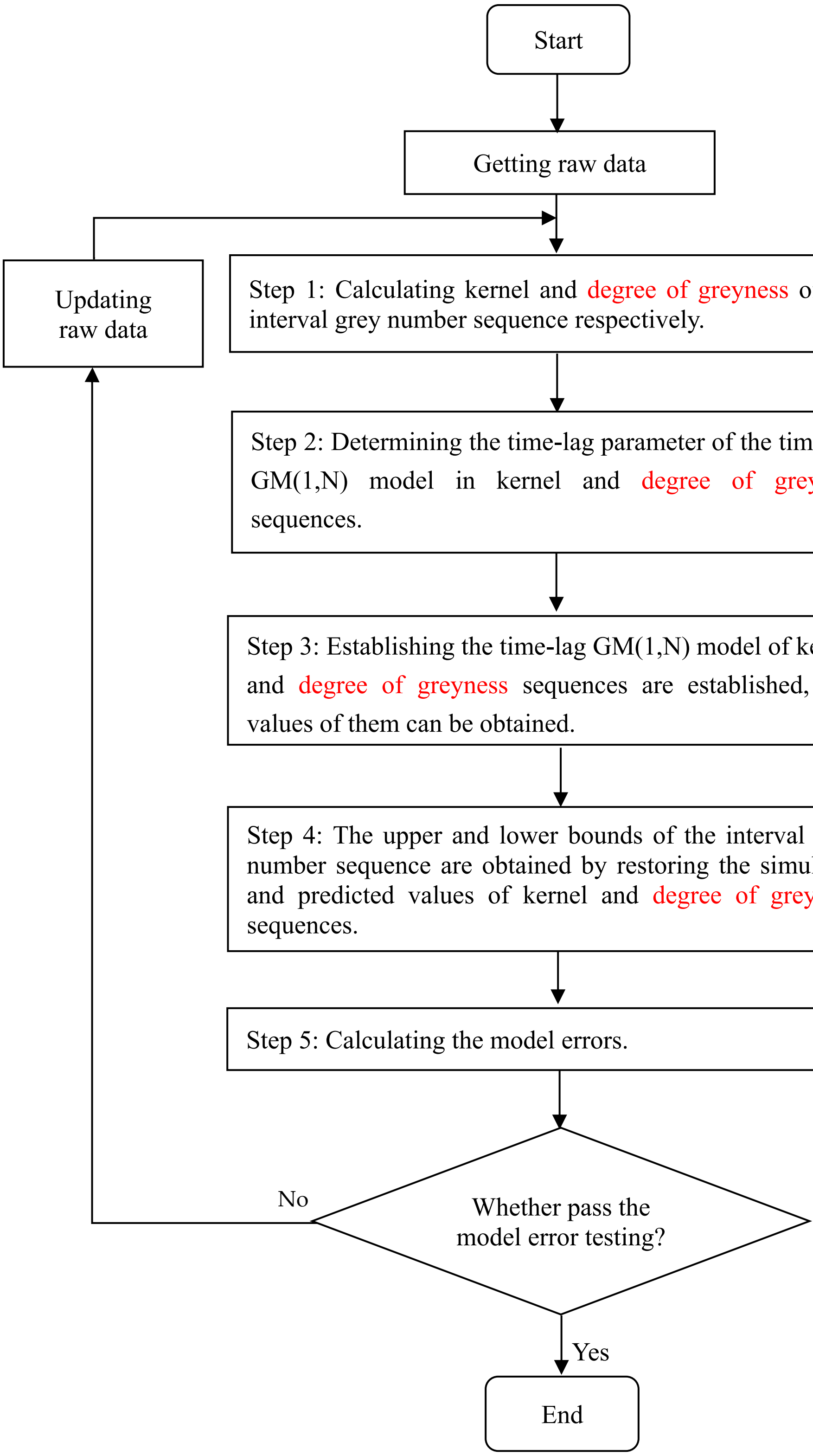

Fig.1 Modeling flow chart of the time-lag GM(1,N) model based on interval grey number sequences

4 Case Study

Smog is mainly composed of inhalable particulate matter, sulfur dioxide and nitrogen oxides, among which the former is particulate matter, and the latter two are atmospheric pollutants. When the three are combined with fog, the atmosphere becomes turbid so that visibility decreases. Under the meteorological condition that is difficult to diffusion, the accumulation of atmospheric pollutants is conducive to the formation of smog, in other words, the accumulation of atmospheric pollutants and smog production has the time-lag effect (Shao et al., 2018; Zhang et al., 2018), and the previous atmospheric pollutants will have a certain impact on the current smog. Therefore, the time-lag GM(1,N) model based on interval grey number sequences is suitable to conduct smog prediction.

4.1 Selection and processing of data

$PM_{2.5}$ is the main indicator of smog in Beijing (Gao et al., 2018; Qian et al., 2019), so this paper takes $PM_{2.5}$ concentration as the system characteristic data of the model for studying the change trend of smog in Beijing. Furthermore, $PM_{2.5}$ can be transformed from SO_2 and NO_2 in air pollutants (Zhang et al., 2013), thus SO_2 and NO_2 concentrations are considered as the relative factors of the model. Due to the limitation of data acquisition tools, acquisition conditions and artificial errors, there may be unavoidable measurement errors in smog related data, which makes the data of $PM_{2.5}$, SO_2 and NO_2 concentrations uncertain, so their values have certain grey number characteristics in a certain range. This paper mainly considers the situation of interval grey numbers.

The monthly data of $PM_{2.5}$, SO_2 and NO_2 concentrations at 35 monitoring stations in Beijing from August 2017 to September 2018 were collected from http://beijingair.sinaapp.com. In data processing, we averaged the hourly data of 35 monitoring stations into daily data, and then the daily data into monthly data. Moreover, the maximum and minimum values of the observed values in a moving window of three months are taken as the upper and lower bounds of the interval grey number for the third month in the corresponding moving window. We denote that the interval grey number corresponding to $PM_{2.5}$, SO_2 and NO_2 concentrations as $X_1(\otimes)$, $X_2(\otimes)$, $X_3(\otimes)$ respectively. Using January 2018 to September 2018 as current data and August 2017 to December 2017 as previous data, the raw data is shown in Table 2. Based on the data of $PM_{2.5}$, SO_2 and NO_2 concentrations in Beijing in the past two years, their domains are determined as $\mu(\Omega_1) = \mu(\Omega_2) = \mu(\Omega_3) = 500$.

Table 2 Interval grey number sequence of $PM_{2.5}$, SO_2 and NO_2 concentrations

k	Time	$X_1(\otimes)$(μg/m^3)	$X_2(\otimes)$(μg/m^3)	$X_3(\otimes)$(μg/m^3)
-4	Aug. 2017	[38,52]	[3,6]	[34,38]
-3	Sep. 2017	[38,58]	[3,4]	[34,49]
-2	Oct. 2017	[38,58]	[3,4]	[35,49]
-1	Nov. 2017	[46,58]	[3,5]	[44,48]
0	Dec. 2017	[44,57]	[4,9]	[44,47]
1	Jan. 2018	[34,46]	[5,8]	[34,49]
2	Feb. 2018	[34,50]	[8,10]	[34,49]
3	Mar. 2018	[34,85]	[8,10]	[34,58]
4	Apr. 2018	[50,85]	[7,10]	[34,58]

5	May 2018	[48,85]	[6,8]	[40,58]
6	Jun. 2018	[48,74]	[5,7]	[36,50]
7	Jul. 2018	[48,74]	[4,7]	[36,50]
8	Aug. 2018	[36,74]	[3,5]	[28,36]
9	Sep. 2018	[34,59]	[3,4]	[26,34]

4.2 Establishment and comparison of the model

Step 1: According to Table 2, the data for January 2018 to July 2018 are used as the training set, and the data for August 2018 to September 2018 are used as the test set. According to formula (1) and formula (7), the kernel and degree of greyness in $PM_{2.5}$, SO_2 and NO_2 concentrations are calculated, and the results are shown in Table 3.

Table 3 Kernel and degree of greyness sequences in $PM_{2.5}$, SO_2 and NO_2 concentrations

k	$\tilde{\otimes}_1(k)$	$\tilde{\otimes}_2(k)$	$\tilde{\otimes}_3(k)$	$r_1(k)$	$r_2(k)$	$r_3(k)$
-4	45	4.5	36	0.028	0.006	0.008
-3	48	3.5	41.5	0.04	0.002	0.03
-2	48	3.5	42	0.04	0.002	0.028
-1	52	4	46	0.024	0.004	0.008
0	50.5	6.5	45.5	0.026	0.01	0.006
1	40	6.5	41.5	0.024	0.006	0.03
2	42	9	41.5	0.032	0.004	0.03
3	59.5	9	46	0.102	0.004	0.048
4	67.5	8.5	46	0.07	0.006	0.048
5	66.5	7	49	0.074	0.004	0.036
6	61	6	43	0.052	0.004	0.028
7	61	5.5	43	0.052	0.006	0.028

Step 2: According to formula (2) and formula (13), the time-lag parameters of the time-lag GM(1,3) model based on kernel sequences are determined by MATLAB, which are $\tau_2 = 2,\ \tau_3 = 2$.
According to formula (8) and formula (14), the time-lag parameters of the time-lag GM(1,3) model based on degree of greyness sequences are determined by MATLAB, which are $\tau_2 = 3,\ \tau_3 = 1$.

Step 3: The model parameters are determined according to Theorem 3 and Theorem 5, and according to Formula (2), the time-lag GM(1,3) model based on kernel sequences with $PM_{2.5}$ concentration is established as follows:

$$\tilde{\otimes}_1^{(0)}(k) + 1.53 z_1^{(1)}(k) = 8.73\, \tilde{\otimes}_2^{(1)}(k-2) + 0.59\, \tilde{\otimes}_3^{(1)}(k-2).$$

According to Formula (8), the time-lag GM(1,3) model based on degree of greyness sequences with $PM_{2.5}$ concentration is established as follows:

$$g_1^{\circ(0)}(k) + 1.35 z_1^{(1)}(k) = 12.56 g_2^{\circ(1)}(k-3) + 0.55 g_3^{\circ(1)}(k-1).$$

According to the time-lag GM(1,3) model based on kernel and degree of greyness sequences with $PM_{2.5}$ concentration, the simulated and predicted values of kernel and degree of greyness sequences with $PM_{2.5}$ concentration are respectively obtained.

Step 4: According to Lemma 1, the upper and lower bounds of the interval grey number sequence are calculated by restoring the simulated and predicted values of kernel and degree of greyness sequences with $PM_{2.5}$ concentration, and the results are shown in tables 4 and 5.

Step 5: According to formula (15) and formula (16), relative errors and average relative errors of $PM_{2.5}$ concentration upper and lower bounds are calculated, and the results are shown in Table 4 and Table 5. According to formula (17) and formula (18), average relative errors of $PM_{2.5}$ concentration interval grey number and model are calculated, and the results are shown in table 6.

In order to test the superiority of the model proposed in this paper, we select the time-lag discrete GM(1,3) model and the traditional GM(1,3) model based on interval grey number sequences as the comparison model. To verify the superiority of the new model in dealing with small data, the multivariate linear regression model is also selected as the non-grey comparison model, because the multiple regression model and GM(1,N) model both study the influence of multiple independent variables on one dependent variable. The time-lag discrete GM(1,3) model, the traditional GM(1,3) model and the multivariate linear regression model are established as follows:

The time-lag discrete GM(1,3) model based on kernel and degree of greyness sequences with $PM_{2.5}$ concentration are shown as follows:

$$\tilde{\otimes}_1^{(1)}(k) - 0.04\,\tilde{\otimes}_1^{(1)}(k-1) = 2.86\,\tilde{\otimes}_2^{(1)}(k-3) - 2.53\,\tilde{\otimes}_3^{(1)}(k-2) + 144.53,$$
$$g_1^{\circ(1)}(k) + 0.97g_1^{\circ(1)}(k-1) = -7.02g_2^{\circ(1)}(k-1) + 3.38g_3^{\circ(1)}(k) - 0.19.$$

The GM(1,3) model based on kernel and degree of greyness sequences with $PM_{2.5}$ concentration are shown as follows:

$$\tilde{\otimes}_1^{(0)}(k) + 0.53z_1^{(1)}(k) = 4.10\,\tilde{\otimes}_2^{(1)}(k) + 0.14\,\tilde{\otimes}_3^{(1)}(k),$$
$$g_1^{\circ(0)}(k) + 0.94z_1^{(1)}(k) = -7.61g_2^{\circ(1)}(k) + 2.65g_3^{\circ(1)}(k).$$

The multivariate linear regression model based on kernel and degree of greyness sequences with $PM_{2.5}$ concentration are shown as follows:

$$\tilde{\otimes}_1^{(0)}(k) = 48.89 - 2.77\,\tilde{\otimes}_2^{(0)}(k) + 0.68\,\tilde{\otimes}_3^{(0)}(k),$$
$$g_1^{\circ(0)}(k) = 0.01 - 7.97g_2^{\circ(0)}(k) + 2.31g_3^{\circ(0)}(k).$$

According to the time-lag discrete GM(1,3) model and the GM(1,3) model and multiple linear regression model based on kernel and degree of greyness sequences, we can obtain the simulated and predicted values of $PM_{2.5}$ concentration kernel and degree of greyness sequences. The simulated and predicted results and average relative errors are shown in Table 4 to Table 6. For convenience, the time-lag GM(1,3) model, time-lag discrete GM(1,3) model, GM(1,3) model and multiple linear regression model are recorded as TLGM(1,3), TLDGM(1,3), GM(1,3) and MLR respectively.

Table 4 $PM_{2.5}$ concentration simulated and predicted values of the lower bound in the interval grey number and their relative error

Time	Actual Value	TLGM(1,3)		TLDGM(1,3)		GM(1,3)		MLR	
		Simulated Value	Relative Error(%)	Simulated Value	Relative Error(%)	Simulated Value	Relative Error(%)	Simulated Value	Relative Error(%)
Jan. 2018	34	34.00	0.00	34.00	0.00	34.00	0.00	50.12	47.41

Feb. 2018	34	34.21	0.62	36.10	6.18	31.96	6.00	39.21	15.32
Mar. 2018	34	32.73	3.74	35.59	4.69	52.29	53.79	31.87	6.26
Apr. 2018	50	49.79	0.42	56.76	13.51	63.75	27.50	37.24	25.52
May 2018	48	55.05	14.69	53.80	12.08	61.65	28.44	46.39	3.35
Jun. 2018	48	53.05	10.52	52.31	8.98	57.12	19.00	49.69	3.52
Jul. 2018	48	41.69	13.15	50.44	5.09	54.69	13.94	55.06	14.71
Average simulated error (%)			6.16		7.22		21.24		16.59
		Predicted Value	Relative Error(%)	Predicted Value	Relative Error(%)	Predicted Value	Relative Error(%)	Predicted Value	Relative Error(%)
Aug. 2018	36	38.35	6.53	32.81	8.86	48.56	34.89	56.19	56.08
Sep. 2018	34	35.29	3.79	29.37	13.62	44.17	29.91	58.77	72.86
Average predicted error (%)			5.16		11.24		32.40		64.47

Table 5 $PM_{2.5}$ concentration simulated and predicted values of the upper bound in the interval grey number and their relative error

Time	Actual Value	TLGM(1,3)		TLDGM(1,3)		GM(1,3)		MLR	
		Simulated Value	Relative Error(%)	Simulated Value	Relative Error(%)	Simulated Value	Relative Error(%)	Simulated Value	Relative Error(%)
Jan. 2018	46	46.00	0.00	46.00	0.00	46.00	0.00	68.29	48.46
Feb. 2018	50	51.50	3.00	57.09	14.17	51.55	3.10	65.36	30.72
Mar. 2018	85	86.47	1.73	87.85	3.35	103.56	21.84	78.84	7.25
Apr. 2018	85	90.78	6.80	85.94	1.11	112.31	32.13	76.24	10.31
May 2018	85	84.90	0.12	82.39	3.07	100.00	17.65	79.48	6.49
Jun. 2018	74	79.54	7.49	81.76	10.48	82.47	11.45	73.53	0.64
Jul. 2018	74	75.57	2.12	80.97	9.42	70.85	4.26	70.92	4.16
Average simulated error (%)			3.03		5.94		12.92		15.43
		Predicted Value	Relative Error(%)	Predicted Value	Relative Error(%)	Predicted Value	Relative Error(%)	Predicted Value	Relative Error(%)
Aug. 2018	74	62.79	15.15	62.74	15.22	66.98	9.49	74.36	0.49
Sep. 2018	59	60.12	1.90	54.35	7.89	59.54	0.92	74.63	26.49
Average predicted error (%)			8.53		11.55		5.21		13.49

Table 6 $PM_{2.5}$ concentration interval grey number and average relative error

Time	TLGM(1,3)	TLDGM(1,3)	GM(1,3)	MLR
	Relative Error(%)	Relative Error(%)	Relative Error(%)	Relative Error(%)
Jan. 2018	0.00	0.00	0.00	47.94
Feb. 2018	1.81	10.18	4.55	23.02
Mar. 2018	2.74	4.02	37.82	6.76
Apr. 2018	3.61	7.31	29.82	17.92
May 2018	7.41	7.57	23.05	4.92
Jun. 2018	9.01	9.73	15.23	2.08
Jul. 2018	7.64	7.25	9.10	9.44
Average simulated error (%)	4.60	6.58	17.08	16.01
Aug. 2018	10.84	12.04	22.19	28.29
Sep. 2018	2.85	10.76	15.42	49.68

Average predicted error (%)	6.85	11.40	18.81	38.99

Table 4 shows that the average relative errors of lower bounds in training sets for TLGM(1,3), TLDGM(1,3), GM(1,3) and MLR models are 6.16%, 7.22%, 21.24%, 16.59%, and those of test sets are 5.16%, 11.24%, 32.40% and 64.47% respectively. It shows that the proposed TLGM(1,3) model has a good ability to simulate and forecast lower bounds in interval grey number sequences, and this novel model is superior to the existing time-lag discrete model. Besides, Table 5 shows that the average relative errors of upper bounds in training sets for TLGM(1,3), TLDGM(1,3), GM(1,3) and MLR models are 3.03%, 5.94%, 12.92%, 15.43%, and those of test sets are 8.53%, 11.55%, 5.21% and 13.49% respectively. Although the upper bound prediction accuracy of GM(1,3) model is higher than that of TLGM(1,3), the upper bound simulation accuracy of GM(1,3) model is worse than that of TLGM(1,3) model, which is higher than 10%. It shows that GM(1,3) model does not consider the time-lag effect among SO_2, NO_2 and $PM_{2.5}$, and the effect of previous relative factors on the current smog lag accumulation. Moreover, Table 6 shows that the average relative errors of training sets for TLGM(1,3), TLDGM(1,3), GM(1,3) and MLR models are 4.60%, 6.58%, 17.08%, 16.01%, and those of test sets are 6.85%, 11.40%, 18.81% and 38.99% respectively. It shows that TLGM(1,3) model can describe the system characteristic data more reasonably, and the reason is that TLGM(1,3) model takes full account of the latency and time-lag effects of previous relative factor sequences on the current system characteristic data sequences.

According to Table 4 to Table 6, we compare the simulated and predicted values, the relative errors and the average relative errors of $PM_{2.5}$ concentration upper and lower bounds in interval grey numbers in Beijing from January 2018 to September 2018, and the comparison charts are as shown in Fig. 2 to Fig. 6.

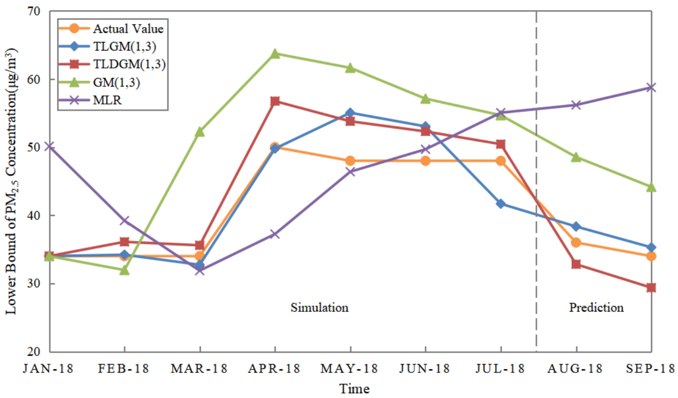

Fig.2 Simulated and predicted values of $PM_{2.5}$ concentration lower bound

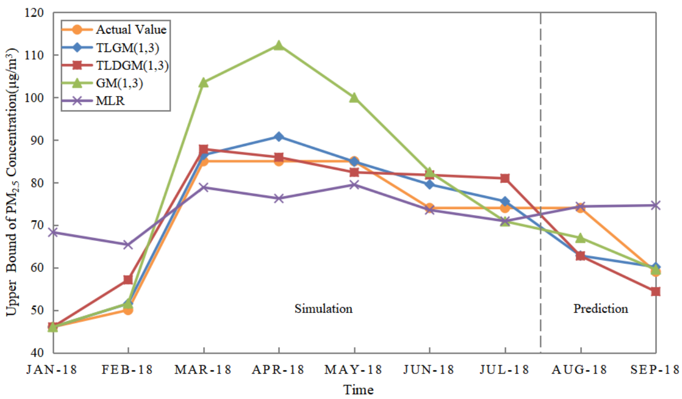

Fig.3 Simulated and predicted values of $PM_{2.5}$ concentration upper bound

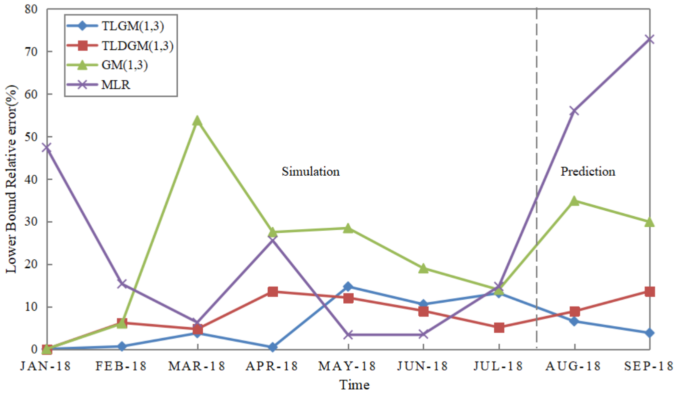

Fig.4 Relative error of $PM_{2.5}$ concentration lower bound

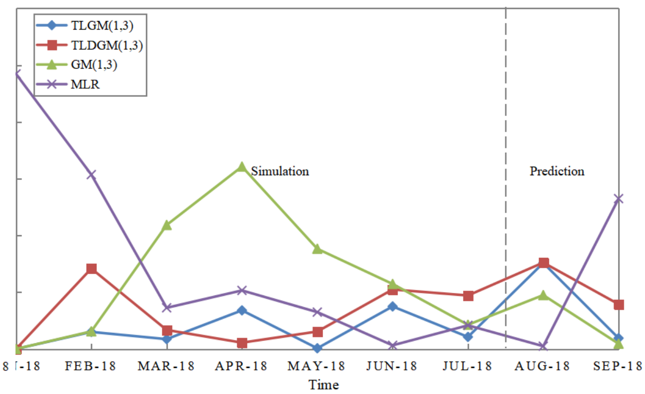

Fig.5 Relative error of $PM_{2.5}$ concentration upper bound

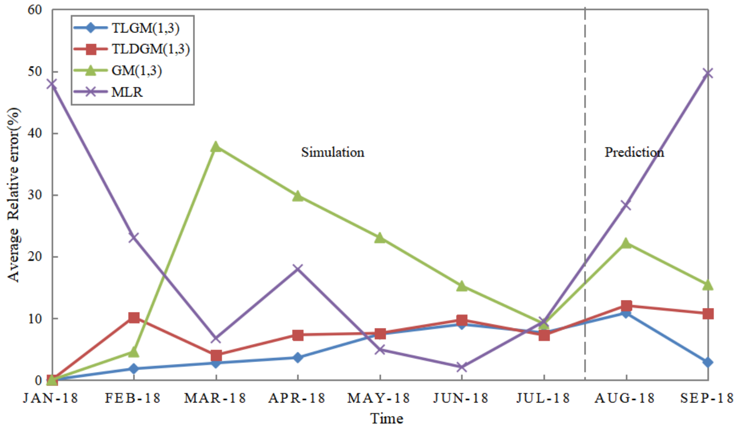

Fig.6 Average relative error of $PM_{2.5}$ concentration interval grey number

4.3 Forecasting smog in Beijing

For understanding the smog situation of Beijing in the future, the time-lag GM(1,3) model proposed in this paper is used to forecast $PM_{2.5}$ concentration in Beijing in the next three months of 2018, and the prediction results are shown in Table 7.

Table7 Forecasting $PM_{2.5}$ monthly average concentration in Beijing

Time	October 2018	November 2018	December 2018
$PM_{2.5}$ concentration ($\mu g/m^3$)	[29.51,63.25]	[26.75,60.29]	[24.38,57.50]
Maximum value ($\mu g/m^3$)	63.25	60.29	57.50
Minimum value ($\mu g/m^3$)	29.51	26.75	24.38
Average value ($\mu g/m^3$)	46.38	43.52	40.94

Table 7 shows that $PM_{2.5}$ concentration of Beijing will continue to decrease in the next three months of 2018, indicating that the smog situation of Beijing will ease in autumn and winter. If the government strictly controls coal combustion, industry, motor vehicles and other pollution sources in Beijing, the average $PM_{2.5}$ concentration of Beijing may reach 24.38 $\mu g/m^3$ in December 2018. Besides, the predicted results for 2018 is compared with $PM_{2.5}$ concentration of Beijing in the past four years from October to December, the result shows that the predicted average $PM_{2.5}$ concentration of Beijing in 2018 may be the lowest value in the five years, and the comparison chart is shown in Fig. 7. We also compare the predicted average $PM_{2.5}$ concentration of Beijing in 2018 against the observed average values in the previous 4 years for other months, and the comparison chart is shown in Fig. 8. Fig. 7 and Fig. 8 show that the monthly average $PM_{2.5}$ concentration of Beijing has a fluctuating downward trend in the five years, and overall characteristics are high in winter and low in summer. Particularly, the $PM_{2.5}$ concentration of Beijing decreased sharply in autumn and winter, and it was the lowest monthly average concentration level of the past five years from August 2018 to December 2018.

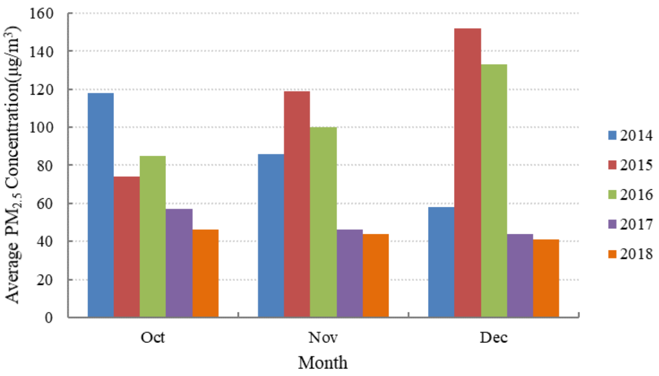

Fig.7 Monthly average $PM_{2.5}$ concentration of Beijing (October to December)

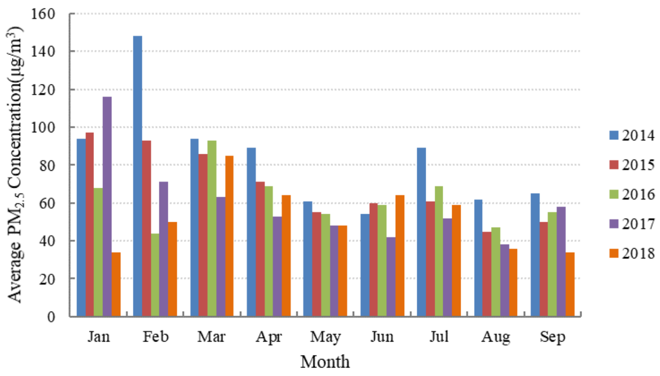

Fig.8 Annual average $PM_{2.5}$ concentration of Beijing (January to September)

After averaging $PM_{2.5}$ concentration monthly data, the $PM_{2.5}$ concentration annual estimated value of Beijing in 2018 can be obtained. Moreover, the resulted predicted annual $PM_{2.5}$ concentration of Beijing in 2018 is compared with the observed values in the past four years, and the comparison chart is shown in Fig. 9. Fig. 9 shows that since the State Council promulgated “the Action Plan for the Prevention and Control of Air Pollution” in 2013, $PM_{2.5}$ concentration of Beijing has decreased year by year in the past five years, and the annual average $PM_{2.5}$ concentration in Beijing was 58 μg/m^3 in 2017, which had fulfilled the target task on schedule. By averaging the original monthly data and predicted monthly value of $PM_{2.5}$ concentration, we can obtain the annual average $PM_{2.5}$ concentration of Beijing may be 51 μg/m^3 in 2018, which will be probably the lowest value in the past five years. It shows that the smog control effect in Beijing has been better in the past five years, but it still does not meet the national standards. Therefore, it is necessary to further improve the control policy and strengthen the smog control efforts.

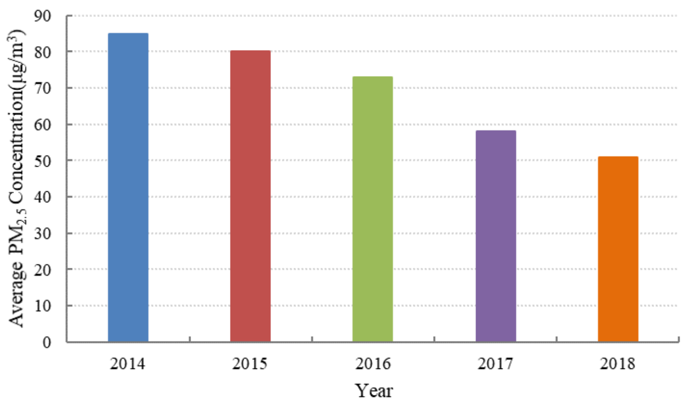

Fig.9 Annual average $PM_{2.5}$ concentration of Beijing

5 Conclusions

The GM(1,N) model based on interval grey number sequences does not consider the time-lag effect mechanism of driving terms, thus this paper introduces the time-lag parameter into driving terms, and proposes a time-lag GM(1,N) model based on interval grey number sequences. The following conclusions are drawn:

The monthly data of $PM_{2.5}$, SO_2 and NO_2 concentrations of Beijing for August 2017 to September 2018 are selected, which are used to establish the time-lag GM(1,3) model based on interval grey number sequences, and the associated time-lag parameters are also determined. Meanwhile, the time-lag discrete GM(1,3) model, traditional GM(1,3) model and multiple linear regression model are selected as the models for comparison, and the comparing result shows that the prediction model proposed in this paper has better simulation and prediction accuracy, while the corresponding simulation and prediction errors are controlled within 10%.

$PM_{2.5}$ concentration in Beijing is forecasted for October to December 2018 by the time-lag GM(1,3) model proposed in this paper. Besides, we compare the $PM_{2.5}$ concentration of Beijing in the past five years, and the result illustrates that $PM_{2.5}$ concentration of Beijing had a fluctuating downward trend in the past five years, especially those of autumn and winter declined sharply in the past five years, which indicates that the model proposed in this paper can provide a basis for the government's smog control and decision-making.

Acknowledgements

This research was supported by 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).

References

Lu, Y.L., Wang, Y., Zuo, J., et al. (2018), "Characteristics of public concern on haze in China and its relationship with air quality in urban areas", *Science of the Total Environment*, Vol.s637-638, pp.1597-1606.

Hao, Y., Peng, H., Temulun, T., et al. (2018), "How harmful is air pollution to economic development? New evidence from $PM_{2.5}$ concentrations of Chinese cities", *Journal of Cleaner Production*, Vol.172, pp.743-757.

Ashraf, A., Butt, A., Khalid, I., et al. (2019), "Smog analysis and its effect on reported ocular surface diseases: A case study of 2016 smog event of Lahore", *Atmospheric Environment*, Vol.198, pp.257-264.

Zhang, L.Y., Lin, J., Qiu, R.Z., et al. (2018), "Trend analysis and forecast of $PM_{2.5}$ in Fuzhou, China using the ARIMA model", *Ecological Indicators*, Vol.95, pp.702-710.

Biancofiore, F., Busilacchio, M., Verdecchia, M., et al. (2017), "Recursive neural network model for analysis and forecast of PM_{10} and $PM_{2.5}$", *Atmospheric Pollution Research*, Vol.8(4), pp.652-659.

García Nieto, P.J., Combarro, E.F., del Coz Díaz, J.J., et al. (2013), "A SVM-based regression model to study the air quality at local scale in Oviedo urban area (Northern Spain): A case study", *Applied Mathematics and Computation*, Vol.219(17), pp.8923-8937.

Saide, P.E., Carmichael, G.R., Spak, S.N., et al. (2011), "Forecasting urban PM_{10} and $PM_{2.5}$ pollution episodes in very stable nocturnal conditions and complex terrain using WRF–Chem CO tracer model", *Atmospheric Environment*, Vol.45(16), pp. 2769-2780.

Zhu, S.L., Lian, X.Y., Liu, H.X., et al. (2017), "Daily air quality index forecasting with hybrid models: a case in China", *Environmental Pollution*, Vol.231, pp.1232-1244.

Wu, L.F., Nu, L., Yang, Y.J. (2018), "Prediction of air quality indicators for the Beijing-Tianjin-Hebei region", *Journal of Cleaner Production*, Vol.196, pp.682-687.

Liu, S.F., Dang, Y.G., Fang, Z.G., et al. (2011), "Grey system theory and its application (Fifth Edition)", *Science Press*.

Wang, Z.X., Li, Q., Pei, L.L. (2018), "A seasonal GM(1,1) model for forecasting the electricity consumption of the primary economic sectors", *Energy*, Vol.154, pp.522-534.

Zeng, B., Tong, M.Y., Ma X. (2020), "A new-structure grey Verhulst model: development and performance comparison", *Applied Mathematical Modelling*, Vol.81, pp.522-537.

Wu, L.F., Liu, S.F., Yao, L.G. et al. (2013), "Grey system model with the fractional order accumulation", *Communications in Nonlinear Science and Numerical Simulation*, Vol.18, pp.1775-1785.

Zeng, B., Li, C. (2016), "Forecasting the natural gas demand in China using a self-adapting intelligent grey model", *Energy*, Vol.112, pp. 810-825.

Wang, Z.X., Li, Q. (2019), "Modelling the nonlinear relationship between CO_2 emissions and economic growth a PSO algorithm-based grey Verhulst model", *Journal of Cleaner Production*, Vol.207, pp.214-224.

Yuan, C.Q. Liu, S.F., Fang, Z.G. (2016), "Comparison of China's primary energy consumption forecasting by using ARIMA (the autoregressive integrated moving average) model and GM(1,1) model", *Energy*, Vol.100, pp.384-390.

Wang, Z.X., Hao, P. (2016), "An improved grey multivariable model for predicting industrial energy consumption in China", *Applied Mathematical Modelling*, Vol.40, pp. 5745-5758.

Ma, X., Liu, Z.B. (2018), "The kernel-based nonlinear multivariate grey model", *Applied Mathematical Modelling,* Vol.56(C), pp.217-238.

Wu, L.F. Zhang, Z.Y. (2018), "Grey multivariable convolution model with new information priority accumulation", *Applied Mathematical Modelling*, Vol.62, pp.595-604.

Wang, Z.X. (2014), “A GM (1,N) - based economic cybernetics model for the high - tech industries in China”, *Kybernetes*, Vol.43No.5, pp.672-685.

Ding, S. Dang, Y.G. Xu, N. et al. (2018), “A novel grey model based on the trends of driving factors and its application”, *The Journal of Grey System*, Vol.30, pp.105-126.

Zeng, B., Duan, H.M., Zhou, Y.F. (2019), “A new multivariable grey prediction model with structure compatibility", *Applied Mathematical Modelling*, Vol.75, pp.385–397.

Ding, S. (2019), “A novel discrete grey multivariable model and its application in forecasting the output value of China’s high-tech industries”, *Computers & Industrial Engineering*, Vol.127, pp.749-760.

Ma, X., Xie, M., Wu, W.Q. et al. (2019), “The novel fractional discrete multivariate grey system model and its applications”, *Applied Mathematical Modelling*, Vol.70, pp.402-424.

Wu, L.F., Gao, X.H., Xiao, Y.L., et al. (2018), “Using a novel multi-variable grey model to forecast the electricity consumption of Shandong Province in China”, Energy, Vol.157, pp.327-335.

Zhou, X.B., Hong, H.K., Xing, X.X., et al. (2017), “Discovering spatio-temporal dependencies based on time-lag in intelligent transportation data”, *Neurocomputing,* Vol.259, pp.76-84.

Lee, T.H., Zhang, Y.S., Jeong, B.H. (2016), “A multi-period output DEA model with consistent time lag effects”, *Computers & Industrial Engineering*, Vol.93, pp.267-274.

Zheng, L.F., Gatti, C.M., Gamarro, E.G., et al. (2020), “Modeling the time-lag effect of sea surface temperatures on ciguatera poisoning in the South Pacific: Implications for surveillance and response”, *Toxicon*, Vol.182, pp.21-29.

Zhai, J., Feng, Y.J., Sheng, J.M. (1996), “GM(1,2) model with time-lag and its application”, *System Engineering*, Vol.14No.6, pp.66-68+51.

Wang, Z.X. (2015), “Multivariable time-lag GM(1,N) model and its application”, *Control and Decision*, Vol.30No.12, pp. 2298-2304.

Zhang, K., Qu, P.P., Zhang, Y.T. (2015), “Time-lag multi-variables discrete grey model and its application”, *Systems Engineering – Theory & Practice*, Vol.35No.8, pp.2092-2103.

Dang, Y.G., Wei, L., Ding, S. (2017), “Grey multivariable discrete time-lag model based on driving information control term and its application”, *Control and Decision*, Vol.32No.9, pp.1672-1680.

Ding, S. Dang, Y.G. Xu, N. et al. (2017), “Research on multivariate discrete grey prediction model based on time-lag effect”, *Control and Decision*, Vol.32No.11, pp.1997-2004.

Mao, S.H., Gao, M.Y., Xiao, X.P. (2015), “Fractional order accumulation time-lag GM(1,N,τ) model and its application”, *Systems Engineering – Theory & Practice*. Vol.35No.2, pp.430-436.

Dang, Y.G., Ye, J. (2018), “An optimized grey prediction model of interval grey numbers based on residual corrections”, *Control and Decision*, Vol.33No.06, pp.1147-1152.

Xie, N.M., Chen, N. (2018), “Flexible job shop scheduling problem with interval grey processing time”, *Applied Soft Computing*, Vol.70, pp.513-524.

Xiong, P.P., Yin, Y., Shi, J. et al. (2018), “Nonlinear multivariable GM(1,N) model based on interval gray number sequence", *The Journal of Grey System*, Vol.30No.3, pp.33-47.

Zeng, B. (2011), Prediction model of interval grey number based on kernel and grey degree, *Journal of Systems Engineering and Electronics*, Vol.33No.4, pp.821-824.

Lewis, C.D. (1982), "Industrial and business forecasting methods: a practical guide to exponential smoothing and curve fitting", *Butterworth Scientific*.

Shao, P.Y., Tian, H.Z., Sun, Y.J., et al. (2018), "Characterizing remarkable changes of severe haze events and chemical compositions in multi-size airborne particles (PM_1, $PM_{2.5}$ and PM_{10}) from January 2013 to 2016-2017 winter in Beijing, China", *Atmospheric Environment*, Vol.189, pp.133-144.

Zhang, X.L., Shi, M.J., Li, Y.J. et al. (2018), "Correlating $PM_{2.5}$ concentrations with air pollutant emissions: A longitudinal study of the Beijing-Tianjin-Hebei region", *Journal of Cleaner Production*, Vol.179, pp.103-113.

Gao, Y., Ji, H.B. (2018), "Microscopic morphology and seasonal variation of health effect arising from heavy metals in $PM_{2.5}$ and PM_{10}: One-year measurement in a densely populated area of urban Beijing", *Atmospheric Research*, Vol.212, pp.213-226.

Qian, W.H., Huang, J. (2019), "Applying the anomaly-based weather analysis on Beijing severe haze episodes", *Journal of Cleaner Production*, Vol.647, pp.878-887.

Zhang, R.S., Jing, J., Tao, J., et al. (2013), "Chemical characterization and source apportionment of $PM_{2.5}$ in Beijing: seasonal perspective", *Atmospheric Chemistry and Physics*, Vol.13No.14, pp.7053-7074.