

## Research Article

# Nonlinear Grey Prediction Model with Convolution Integral NGMC (1, $n$ ) and Its Application to the Forecasting of China's Industrial SO<sub>2</sub> Emissions

Zheng-Xin Wang<sup>1,2</sup>

<sup>1</sup> School of Economics & International Trade, Zhejiang University of Finance & Economics, Hangzhou, Zhejiang 310018, China

<sup>2</sup> College of Economics & Management, Nanjing University of Aeronautics & Astronautics, Nanjing, Jiangsu 210016, China

Correspondence should be addressed to Zheng-Xin Wang; [jenkin107@hotmail.com](mailto:jenkin107@hotmail.com)

Received 11 December 2013; Accepted 14 February 2014; Published 24 March 2014

Academic Editor: Hui-Shen Shen

Copyright © 2014 Zheng-Xin Wang. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The grey prediction model with convolution integral GMC (1,  $n$ ) is a multiple grey model with exact solutions. To further improve prediction accuracy and describe better the relationship between cause and effect, we introduce nonlinear parameters into GMC (1,  $n$ ) model and additionally apply a convolution integral to produce an improved forecasting model here designated as NGMC (1,  $n$ ). The model solving process applied the least-squares method to evaluate the structure parameters of the model: convolution was used to obtain an exact solution with this improved grey model. The nonlinear optimisation took the parameters as the decision variables with the objective of minimising forecasting errors. The GMC (1, 2) and NGMC (1, 2) models were used to predict China's industrial SO<sub>2</sub> emissions from the basis of the economic output level as the influencing factor. Results indicated that NGMC (1, 2) can effectively describe the nonlinear relationship between China's economic output and SO<sub>2</sub> emissions with an improved accuracy over current GMC (1, 2) models.

## 1. Introduction

Due to the complexity of both the internal and external environmental factors in any such system, the behavioural data are often sparse. Statistical analyses can effectively address the problem where datasets are large; for small sample sizes ( $n < 15$ ), solution becomes difficult. Grey system theory, pioneered by Deng in 1982 [1], is an uncertainty theory dealing specifically with the analysis, modelling, prediction, and control of information-poor systems. Grey theory considers that although the objective system appears complex, with sparse data, it always has an innate overall governing relationship: key to accurate forecast is the choice of appropriate methods of data-mining and utilisation. In grey theory [2, 3], every stochastic process has grey variables changing at a certain amplitude and periodicity. An accumulated generating operation (AGO) is a basic method to render a grey process white. With accumulation, the changing trend becomes more apparent and the innate governing relationship is revealed. Based on this, a grey differential equation is built to describe

the dynamic law of grey accumulation. Then, an inverse accumulated generating operation (IAGO) is used to reduce the whitening of this grey process. Of course, it is not merely the case that greater accumulation leads to better forecasts. Tien demonstrated that the  $k$ -AGO ( $k > 1$ ) data of the original series cannot be used as intermediate information in grey prediction model building [4]. In practical applications, the system generated by real data is considered grey: on that basis, grey models can be both effective past simulators and future predictors.

In general, a grey model can be denoted by GM ( $m, n$ ), where  $m$  is the order and  $n$  is the number of variables of the grey equation. GM (1, 1) is the most widely used grey prediction model with the simplest structure [2]. Its core principle is to regard real systems as generalised energy systems with exponentially changing trends under no external agency's influence. Therefore, only when 1-AGO original data series are consistent with this exponential change, can accurate prediction by GM (1, 1) be achieved. Many scholars have made useful improvements to basic GM (1, 1) models [5–8]. In

spite of this, it is not appropriate to use GM (1, 1) predictions based upon data mining from series which are themselves subject to external agents. The GM (1,  $n$ ) model with  $n - 1$  relative factors acting as the associated series is a multiple, grey, prediction model [2, 3]. This model can fully exploit information provided the associated series unlike a GM (1, 1) model which only includes information about the predicted series in its modelling process. So, from the point of view of provision of supplementary information, the prediction accuracy of GM (1,  $n$ ), should, under these conditions, be higher than that of GM (1, 1). However, the solution of the whitening differential equation of GM (1,  $n$ ) is inexact and on occasion wrong thereby producing significant practical forecasting errors [4]. The model therefore has found few real-world applications [9–11] to date.

A grey prediction model with convolution integral GMC (1,  $n$ ) proposed by Tien [12] is a new model to improve upon traditional GM (1,  $n$ ) versions. The values modelled by GMC (1,  $n$ ) are theoretically the exact solution of the traditional GM (1,  $n$ ) model, and the grey control parameter like that of GM (1, 1) is introduced into the model. GMC (1,  $n$ ) reduces to GM (1, 1) for  $n = 1$ . The GMC (1,  $n$ ) model improved the prediction accuracy of multiple grey models and has successfully been applied to different areas [12–14]. Tien proposed three improved models over the basic GMC (1,  $n$ ) model to meet various application requirements. The three models are deterministic GMC (1,  $n$ ) (DGDMC (1,  $n$ )) [15]; interval GDMC (1,  $n$ ) (IGDMC (1,  $n$ )) [16]; and first pair-of-data GMC (1,  $n$ ) (FGMC (1,  $n$ )) [17]. The first derivative of the 1-AGO data of each associated series is introduced into the DGDMC (1,  $n$ ) model [15] to strengthen the indicative significance while evaluating its 1-AGO predicted data by a convolution integral. The IGDMC (1,  $n$ ) model [16] draws lessons from the prediction method of its linear regression interval, extending the prediction method of DGDMC (1,  $n$ ) from point to interval prediction. Certain components of the model may be removed to better satisfy hypothesis tests based upon system parameters. The building of the improved model FGMC (1,  $n$ ) [17] only needs  $n + 2$  pairs of historical data: for the first pair of entries, the message was shown to be independent from the modelling results. Therefore, FGMC (1,  $n$ ) is usually more satisfactory and stable than GMC (1,  $n$ ) because it extracts the first pair of entries and bases its predictions on the message enshrined in the original series.

Nowadays, GMC (1,  $n$ ) and variants thereon are linear: the high-precision predictions are predicated upon the similarity in trend between the predicted variables and their influencing factors. At present, most applications meet this premise. For improved predictions, Wu and Chen [13] used grey relational analysis to analyse the similarities between predicted, and associated, series curves of differing periodicities. They then built GMC (1,  $n$ ) using data from the largest period of grey relational analysis and obtained satisfactory predictions. However, the nonlinear interaction between the internal and external system factors is ubiquitous in reality: the linear structure of these GMC (1,  $n$ ) models is not conducive to high-precision prediction of dynamic systems. Therefore, nonlinear processing to the forms of the relative factors on the right-hand side of the GMC (1,  $n$ ) equation

was applied here. A power exponent was added to the 1-AGO series of every associated series to reflect the nonlinear interaction of the associated series with that predicted. Although these newly added parameters were initially unknown, they were estimated and subjected to refinement by a least-squares method. After obtaining the time response function of the predicted series, the optimum values of parameters reflecting their nonlinear interactions were derived with regard to the aforementioned minimum error criterion. This improved GMC (1,  $n$ ) model, a nonlinear grey prediction model with the addition of a convolution integral, is hereafter designated NGMC (1,  $n$ ).

In recent years, industrial SO<sub>2</sub> emissions have become a major source of harmful air pollution in China as their industrial growth continues. To cater for these increased emissions, an enhanced understanding of likely future industrial SO<sub>2</sub> emissions is needed. Though China's industrial output values can be used as a basis to interpret growing industrial emissions, there is no a simultaneous increase in industrial emissions and industrial development [18]. Data also show that China's SO<sub>2</sub> emissions and industrial output values are not governed by a simple linear relationship but an uncertain nonlinear one. We present GMC (1,  $n$ ) and NGMC (1,  $n$ ) models where the influencing, causative, factors are industrial output; predicted series is China's SO<sub>2</sub> emissions. Through comparison of their modelling and prediction accuracy, the paper demonstrated the effectiveness of the proposed NGMC (1,  $n$ ).

The remainder of this paper is organized as follows. An introduction to the existing GMC (1,  $n$ ) model and the modelling method of nonlinear grey model with convolution integral NGMC (1,  $n$ ) proposed in this paper are given in Section 2. Section 3 demonstrates the effectiveness of the proposed NGMC (1,  $n$ ) model by forecasting China's SO<sub>2</sub> emissions. Finally, the paper concludes with some comments in Section 4.

## 2. Modelling Method

*2.1. The Existing GMC (1,  $n$ ) Model.* Assume that pairs of observations  $X^{(0)} = (X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)})$  are available at time intervals of  $n - 1$  inputs  $X_2^{(0)}, X_3^{(0)}, \dots, X_n^{(0)}$  and an output  $X_1^{(0)}$  from some dynamic system. The existing GMC (1,  $n$ ) modelling process [12] is carried out as follows.

Consider the following original predicted series:

$$X_1^{(0)} = \{X_1^{(0)}(rp + 1), X_1^{(0)}(rp + 2), \dots, X_1^{(0)}(rp + r)\}; \quad (1)$$

and the original associated series:

$$X_i^{(0)} = \{X_i^{(0)}(1), X_i^{(0)}(2), \dots, X_i^{(0)}(r)\}, \quad i = 2, 3, \dots, n \quad (2)$$

then the first order accumulated generation (1-AGO) data for  $X_1^{(0)}, X_2^{(0)}, \dots, X_n^{(0)}$  are given by (3) and (4), respectively:

$$X_1^{(1)}(rp+k) = \sum_{j=1}^k X_1^{(0)}(rp+j), \quad (3)$$

$$X_i^{(1)}(k) = \sum_{j=1}^k X_i^{(0)}(j), \quad i = 2, 3, \dots, n. \quad (4)$$

The grey prediction model based on the predicted 1-AGO series:

$$X_1^{(1)} = \{X_1^{(1)}(rp+1), X_1^{(1)}(rp+2), \dots, X_1^{(1)}(rp+r)\} \quad (5)$$

and the associated 1-AGO series:

$$X_i^{(1)} = \{X_i^{(1)}(1), X_i^{(1)}(2), \dots, X_i^{(1)}(r)\}, \quad i = 2, 3, \dots, n \quad (6)$$

is given by the differential equation:

$$\begin{aligned} \frac{dX_1^{(1)}(rp+t)}{dt} + b_1 X_1^{(1)}(rp+t) \\ = b_2 X_2^{(1)}(t) + b_3 X_3^{(1)}(t) + \dots + b_n X_n^{(1)}(t) + u, \quad (7) \\ t = 1, 2, \dots, rf, \end{aligned}$$

where  $b_1, b_2, \dots, b_n, u$  are parameters to be estimated and  $r$  is the data number used in model building;  $rp$  is a delay period, and  $rf$  is the number of entries to be forecast. Equation (7)

is called the  $n$ -factor grey prediction model with convolution integral and is denoted by GMC (1,  $n$ ) [12]; the 1 represents the first-order derivative of the 1-AGO series of  $X_1^{(1)}$ ; the  $n$  represents the total of  $n$  relative series introduced into the grey differential equation.

The grey derivative for the first-order grey differential equation with 1-AGO is represented as follows:

$$\frac{dX_1^{(1)}(rp+t)}{dt} = \lim_{\Delta t \rightarrow 0} \frac{X_1^{(1)}(rp+t) - X_1^{(1)}(rp+t-\Delta t)}{\Delta t}, \quad (8)$$

and when  $\Delta t \rightarrow 1$

$$\begin{aligned} \frac{dX_1^{(1)}(rp+t)}{dt} &= \frac{\Delta X_1^{(1)}(rp+t)}{\Delta t} \\ &= X_1^{(1)}(rp+t) - X_1^{(1)}(rp+t-1) \quad (9) \\ &= X_1^{(0)}(rp+t). \end{aligned}$$

The background value of the grey derivative  $dX_1^{(1)}(rp+t)/dt$  is taken as the mean of  $X_1^{(1)}(rp+t)$  and  $X_1^{(1)}(rp+t-1)$ , and those of the associated series  $X_i^{(1)}(t)$  are also taken as the mean of  $X_i^{(1)}(t)$  and  $X_i^{(1)}(t-1)$  for  $i = 2, 3, \dots, n$ , respectively, in the determination of model parameters by GMC (1,  $n$ ).

The least-squares solution to the model parameters of GMC (1,  $n$ ) [12] in (7) by  $t$  from 1 to  $r$  is

$$(b_1, b_2, \dots, b_n, u)^T = (B^T B)^{-1} B^T Y_R, \quad (10)$$

where

$$B = \begin{bmatrix} \frac{X_1^{(1)}(rp+1) + X_1^{(1)}(rp+2)}{2} & \frac{X_2^{(1)}(1) + X_2^{(1)}(2)}{2} & \frac{X_3^{(1)}(1) + X_3^{(1)}(2)}{2} & \dots & \frac{X_n^{(1)}(1) + X_n^{(1)}(2)}{2} & 1 \\ \frac{X_1^{(1)}(rp+2) + X_1^{(1)}(rp+3)}{2} & \frac{X_2^{(1)}(2) + X_2^{(1)}(3)}{2} & \frac{X_3^{(1)}(2) + X_3^{(1)}(3)}{2} & \dots & \frac{X_n^{(1)}(2) + X_n^{(1)}(3)}{2} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \frac{X_1^{(1)}(rp+r-1) + X_1^{(1)}(rp+r)}{2} & \frac{X_2^{(1)}(r-1) + X_2^{(1)}(r)}{2} & \frac{X_3^{(1)}(r-1) + X_3^{(1)}(r)}{2} & \dots & \frac{X_n^{(1)}(r-1) + X_n^{(1)}(r)}{2} & 1 \end{bmatrix}, \quad (11)$$

$$Y_R = (X_1^{(0)}(rp+2), X_1^{(0)}(rp+3), \dots, X_1^{(0)}(rp+r))^T. \quad (12)$$

In summary, the right-hand side of (7), the discrete function  $f(t)$  [12] can be obtained as

$$\begin{aligned} f(t) &= b_2 X_2^{(1)}(t) + b_3 X_3^{(1)}(t) + \dots + b_n X_n^{(1)}(t) + u, \\ &t = 1, 2, \dots, rf. \end{aligned} \quad (13)$$

The 1-AGO modelling values of the predicted series [12] can be derived with the initial condition  $\widehat{X}_1^{(1)}(rp+1) = X_1^{(1)}(rp+1)$  as

$$\begin{aligned} \widehat{X}_1^{(1)}(rp+t) &= X_1^{(1)}(rp+1) e^{-b_1(t-1)} + \int_1^t e^{-b_1(t-\tau)} f(\tau) d\tau \\ &t = 1, 2, \dots, rf. \end{aligned} \quad (14)$$

The modelling values,  $\widehat{X}_1^{(1)}(rp + t)$ ,  $t = 1, 2, \dots, rf$  in (14) of the 1-AGO data of the predicted series [12] can be evaluated approximately by

$$\widehat{X}_1^{(1)}(rp + 1) = X_1^{(1)}(rp + 1) = X_1^{(0)}(rp + 1), \quad (15)$$

$$\begin{aligned} \widehat{X}_1^{(1)}(rp + t) &\cong X_1^{(0)}(rp + 1)e^{-b_1(t-1)} + u(t-2) \\ &\times \left\{ \sum_{k=2}^t e^{-b_1(t-k+(1/2))} \cdot \frac{1}{2} [f(t) + f(t-1)] \right\}, \end{aligned} \quad (16)$$

where  $u(t-2)$  is the unit step function.

Applying 1-IAGO to (16) yields the following modelled values together with the forecasts:

$$\widehat{X}_1^{(1)}(rp + 1) = X_1^{(1)}(rp + 1) = X_1^{(0)}(rp + 1), \quad (17)$$

$$\begin{aligned} \widehat{X}_1^{(0)}(rp + t) &= \widehat{X}_1^{(1)}(rp + t) - \widehat{X}_1^{(1)}(rp + t - 1), \\ &t = 2, 3, \dots, rf. \end{aligned} \quad (18)$$

Assume the system parameters in (7) to be constants in the postsampling period and then by using the postsampling data, combined with the given data for the corresponding associated series, as a new input series, the corresponding forecasts or values of indirect measurement for the predicted series can be derived.

**2.2. Nonlinear Grey Modelling with Convolution Integral: NGMC (1, n).** To adapt GMC (1, n) to reality's ubiquitous nonlinearity, the improved model—NGMC (1, n)—with its convolution integral was developed and used.

**2.2.1. Representation of the New Model: NGMC (1, n).** Suppose that the grey prediction model based on the predicted 1-AGO series:

$$X_1^{(1)} = \{X_1^{(1)}(rp + 1), X_1^{(1)}(rp + 2), \dots, X_1^{(1)}(rp + r)\} \quad (19)$$

and the associated 1-AGO series:

$$X_i^{(1)} = \{X_i^{(1)}(1), X_i^{(1)}(2), \dots, X_i^{(1)}(r)\}, \quad i = 2, 3, \dots, n \quad (20)$$

is given by the differential equation:

$$\begin{aligned} \frac{dX_1^{(1)}(rp + t)}{dt} + b_1 X_1^{(1)}(rp + t) \\ = b_2 [X_2^{(1)}(t)]^{\beta_2} + b_3 [X_3^{(1)}(t)]^{\beta_3} + \dots + b_n [X_n^{(1)}(t)]^{\beta_n} + u, \\ t = 1, 2, \dots, rf, \end{aligned} \quad (21)$$

where  $b_1, b_2, \dots, b_n, u, \beta_2, \beta_3, \dots, \beta_n$  are parameters to be estimated,  $r$  is the data number used in model building,  $rp$  is a delay period, and  $rf$  the number of entries to be forecast. Equation (21) denotes the  $n$ -factor nonlinear grey prediction model with the addition of a convolution integral (NGMC (1, n)), where 1 represents the first-order derivative of the 1-AGO series of  $X_1^{(1)}$  and  $n$  represents the total number relative series introduced into the grey differential equation.

Compared with GMC (1, n),  $n - 1$  unknown parameters  $\beta_2, \beta_3, \dots, \beta_n$  are introduced into this NGMC (1, n) model. They are taken as the  $n - 1$  power exponents of the predicted variables' relevant factors to reflect the effect of these upon nonlinear system behaviours and interactions.

When  $\beta_2 = \beta_3 = \dots = \beta_n = 1$ , (20) is reduced to a traditional GMC (1, n) model [12].

When  $\beta_2 = \beta_3 = \dots = \beta_n = 1$  and  $u = 0$ , (20) reduces to a traditional GM (1, n) model [2, 4].

When  $n = 1$  or  $\beta_2 = \beta_3 = \dots = \beta_n = 0$ , (20) reduces to a traditional GM (1, 1) model [2].

**2.2.2. The Evaluation of System Parameters  $b_1, b_2, \dots, b_n, u$ , and  $f_N(t)$ .** To evaluate all parameters in the model, the power exponents  $\beta_2, \beta_3, \dots, \beta_n$  of  $X_2^{(1)}, X_3^{(1)}, \dots, X_n^{(1)}$  are assumed to be known; by applying a method similar to that of GMC (1, n), estimates of the corresponding parameters  $b_1, b_2, \dots, b_n, u$  are made. After obtaining the time response function of  $\widehat{X}_1^{(0)}$ , the corresponding optimisation algorithm is applied to solve for  $\beta_2, \beta_3, \dots, \beta_n$ .

The grey derivative for the first-order grey differential equation with 1-AGO is conventionally represented by

$$\begin{aligned} \frac{dX_1^{(1)}(rp + t)}{dt} &= X_1^{(1)}(rp + t) - X_1^{(1)}(rp + t - 1) \\ &= X_1^{(0)}(rp + t). \end{aligned} \quad (22)$$

The background value of the grey derivative  $dX_1^{(1)}(rp + t)/dt$  is taken as the means of  $X_1^{(1)}(rp + t)$  and  $X_1^{(1)}(rp + t - 1)$ , and those of the associated series  $X_i^{(1)}(t)$  are also taken as the means of  $X_i^{(1)}(t)$  and  $X_i^{(1)}(t - 1)$  for  $i = 2, 3, \dots, n$ , respectively, when determining the model parameters by NGMC (1, n).

The least-squares solution to the model parameters of NGMC (1, n) in (20) by  $t$  from 1 to  $r$  is

$$(b_1, b_2, \dots, b_n, u)^T = (B^T B)^{-1} B^T Y_R, \quad (23)$$

where

$$B = \begin{bmatrix} -\frac{X_1^{(1)}(rp+1) + X_1^{(1)}(rp+2)}{2} & \left[\frac{X_2^{(1)}(1) + X_2^{(1)}(2)}{2}\right]^{\beta_2} & \left[\frac{X_3^{(1)}(1) + X_3^{(1)}(2)}{2}\right]^{\beta_3} & \dots & \left[\frac{X_n^{(1)}(1) + X_n^{(1)}(2)}{2}\right]^{\beta_n} & 1 \\ -\frac{X_1^{(1)}(rp+2) + X_1^{(1)}(rp+3)}{2} & \left[\frac{X_2^{(1)}(2) + X_2^{(1)}(3)}{2}\right]^{\beta_2} & \left[\frac{X_3^{(1)}(2) + X_3^{(1)}(3)}{2}\right]^{\beta_3} & \dots & \left[\frac{X_n^{(1)}(2) + X_n^{(1)}(3)}{2}\right]^{\beta_n} & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ -\frac{X_1^{(1)}(rp+r-1) + X_1^{(1)}(rp+r)}{2} & \left[\frac{X_2^{(1)}(r-1) + X_2^{(1)}(r)}{2}\right]^{\beta_2} & \left[\frac{X_3^{(1)}(r-1) + X_3^{(1)}(r)}{2}\right]^{\beta_3} & \dots & \left[\frac{X_n^{(1)}(r-1) + X_n^{(1)}(r)}{2}\right]^{\beta_n} & 1 \end{bmatrix}, \tag{24}$$

$$Y_R = (X_1^{(0)}(rp+2), X_1^{(0)}(rp+3), \dots, X_1^{(0)}(rp+r))^T. \tag{25}$$

In summary, the right-hand side of (21), the discrete function  $f(t)$ , can be obtained as

$$f_N(t) = b_2[X_2^{(1)}(t)]^{\beta_2} + b_3[X_3^{(1)}(t)]^{\beta_3} + \dots + b_n[X_n^{(1)}(t)]^{\beta_n} + u, \tag{26}$$

$$t = 1, 2, \dots, rf.$$

As seen in (26),  $f_N(t)$  is a nonlinear function on  $X_i^{(1)}$ ,  $i = 2, 3, \dots, n$ , but it, in essence, is still linear. If the new variable  $Y_i^{(1)}(t) = [X_i^{(1)}(t)]^{\beta_i}$ ,  $i = 2, 3, \dots, n$ ,  $f_N(t)$  can be written as

$$f_N(t) = b_2Y_2^{(1)}(t) + b_3Y_3^{(1)}(t) + \dots + b_nY_n^{(1)}(t) + u, \tag{27}$$

$$t = 1, 2, \dots, rf;$$

therefore, (27) remains a linear function, essentially the same as (13).

2.2.3. *The Determination of Unit Impulse Response Function  $h(t)$ .* The unit impulse response function  $h(t)$  of the system characterised by (21) can be derived by Laplace transform. From (21):

$$\frac{dX_1^{(1)}(t)}{dt} + b_1X_1^{(1)}(t) = \delta(t), \tag{28}$$

where  $\delta(t)$  is the unit impulse function to determine the corresponding unit impulse response function. Applying a Laplace transform to (28) with the initial condition  $X_1^{(1)}(1) = 0$  gives

$$s\bar{X}_1^{(1)}(s) + b_1\bar{X}_1^{(1)}(s) = 1 \tag{29}$$

or

$$\bar{X}_1^{(1)}(s) = \frac{1}{s + b_1}. \tag{30}$$

The inverse transform of  $\bar{X}_1^{(1)}(s)$  is

$$X_1^{(1)}(t) = e^{-b_1t}. \tag{31}$$

That is, the unit impulse response function  $h(t)$  of the system is

$$h(t) = e^{-b_1t}. \tag{32}$$

The process continues by outlining the evaluation of  $\widehat{X}_1^{(0)}$ .

2.2.4. *The Evaluation of  $\widehat{X}_1^{(0)}$ .* The 1-AGO modelling values of the predicted series can be derived with the initial condition  $\widehat{X}_1^{(1)}(rp+1) = X_1^{(1)}(rp+1)$  as

$$\widehat{X}_1^{(1)}(rp+t) = X_1^{(1)}(rp+1)e^{-b_1(t-1)} + \int_1^t e^{-b_1(t-\tau)} f_N(\tau) d\tau, \tag{33}$$

$$t = 1, 2, \dots, rf.$$

The second term on the right-hand side of (33) can be evaluated approximately by two-point Gaussian numerical integration. Thus, the modelling values in (33) of the 1-AGO data of the predicted series can be evaluated approximately by

$$\widehat{X}_1^{(1)}(rp+1) = X_1^{(1)}(rp+1) = X_1^{(0)}(rp+1), \tag{34}$$

$$\begin{aligned} &\widehat{X}_1^{(1)}(rp+t) \\ &\cong X_1^{(0)}(rp+1)e^{-b_1(t-1)} + u(t-2) \\ &\times \left\{ \sum_{k=2}^t e^{-b_1(t-k+(1/2))} \cdot \frac{1}{2} [f_N(t) + f_N(t-1)] \right\}, \end{aligned} \tag{35}$$

where  $u(t-2)$  is the unit step function.

Applying 1-IAGO to (35) produced the following modelling values and their forecasts:

$$\widehat{X}_1^{(1)}(rp + 1) = X_1^{(1)}(rp + 1) = X_1^{(0)}(rp + 1), \quad (36)$$

$$\widehat{X}_1^{(0)}(rp + t) = \widehat{X}_1^{(1)}(rp + t) - \widehat{X}_1^{(1)}(rp + t - 1), \quad (37)$$

$$t = 2, 3, \dots, rf.$$

2.2.5. *An Overall Measure of Accuracy for Forecasts.* To evaluate forecast performance, Tien's standard [12] test using the root mean squared percentage error (RMSPE) for the priori-sample period (RMSPEPR) and postsample periods (RMSPEPO), respectively, is used. Generally, the RMSPEPR and RMSPEPO are defined as

$$\text{RMSPEPR} = \sqrt{\frac{1}{r} \sum_{t=rp+1}^{rp+r} \frac{[\widehat{X}_1^{(0)}(t) - X_1^{(0)}(t)]^2}{[X_1^{(0)}(t)]^2}} \times 100\%, \quad (38)$$

$$\text{RMSPEPO} = \sqrt{\frac{1}{rf} \sum_{t=rp+r+1}^{rp+r+rf} \frac{[\widehat{X}_1^{(0)}(t) - X_1^{(0)}(t)]^2}{[X_1^{(0)}(t)]^2}} \times 100\%. \quad (39)$$

2.2.6. *The Determination of Power Exponents.* The discussion above is proposed under the condition that  $\beta_i, i = 2, 3, \dots, n$  is known, whereas, in reality, we cannot really know the values of them that are unknown. This work suggests that the minimisation of the root mean squared percentage error for the presample period should be the objective, and the unknown parameters  $\beta_i = 1, i = 2, 3, \dots, n$  are solved by building the optimisation model below.

Consider the following:

$$\begin{aligned} \text{Min}_{\beta_i} \quad & \text{RMSPEPR} = \sqrt{\frac{1}{r} \sum_{t=rp+1}^{rp+r} \frac{[\widehat{X}_1^{(0)}(t) - X_1^{(0)}(t)]^2}{[X_1^{(0)}(t)]^2}} \times 100\%, \quad i = 2, 3, \dots, n \\ \text{s.t.} \quad & \left\{ \begin{aligned} & \widehat{X}_1^{(1)}(rp + 1) = X_1^{(1)}(rp + 1) = X_1^{(0)}(rp + 1) \\ & \widehat{X}_1^{(1)}(rp + t) = X_1^{(0)}(rp + 1) e^{-b_1(t-1)} + u(t-2) \\ & \quad \times \left\{ \sum_{k=2}^t e^{-b_1(t-k+(1/2))} \cdot \frac{1}{2} [f_N(t) + f_N(t-1)] \right\}, \quad t = 2, 3, \dots, r \\ & f_N(t) = b_2 [X_2^{(1)}(t)]^{\beta_2} + b_3 [X_3^{(1)}(t)]^{\beta_3} + \dots + b_n [X_n^{(1)}(t)]^{\beta_n} + u, \quad t = 1, 2, \dots, r \\ & (b_1, b_2, \dots, b_n, u)^T = (B^T B)^{-1} B^T Y_R \end{aligned} \right. \end{aligned} \quad (40)$$

The optimisation problem above can be solved by proprietary software. Once  $\beta_i = 1, i = 2, 3, \dots, n$  is confirmed, the structure parameters  $b_1, b_2, \dots, b_n, u$  of the model are also determined. Then, the discrete function  $f_N(t)$  can be derived by back-substituting the derived results into (26): final predictions can then be obtained by applying  $f_N(t)$  to (35) to (37).

### 3. Forecasting China's Industrial SO<sub>2</sub> Emissions

Industry is the dominant factor in promoting China's economic development. However, SO<sub>2</sub> emissions discharged during increasing industrialisation pose a significant threat to the eco-system and biosphere. The effective management of industrial SO<sub>2</sub> emissions demands accurate forecasts thereof. The paper argues that industrial output activity levels are considered to be the main factor affecting industrial SO<sub>2</sub> emissions. Local governments and enterprises, to increase industrial output, continue to expand the scale of production scale and thereby increase SO<sub>2</sub> emissions. China's industrial SO<sub>2</sub> emissions and gross industrial output levels from 2003 to 2010 are given in Table 1: Figure 1 shows the trends in China's industrial SO<sub>2</sub> emissions and gross industrial output.

As seen in Figure 1, China's industrial SO<sub>2</sub> emissions and industrial output values are not simply linearly related. With the increasing industrial output, SO<sub>2</sub> emissions initially increased and then decreased. Generally, the reasons for this phenomenon are complex; multifarious and interrelated but basically may be summarised as combinations of the following: structural economic change, technical progress, different modalities of demand, and more effective regulatory regimes, and so forth.

GMC (1, 2) and NGMC (1, 2) models were used to forecast—from known industrial output values—China's SO<sub>2</sub> emissions as a time series. Through comparing the modelling and prediction accuracy of the two models, the effectiveness of the proposed NGMC (1, 2) was demonstrated.

3.1. *Forecasting by GMC (1, 2).* Applying to the GMC (1, 2) model of (7) to (18), the values of parameters  $n, r, rf,$  and  $rp$  in (7), estimates of model parameters in (10) can be obtained and are listed in Table 2. The GMC (1, 2) model from (7) becomes

$$\begin{aligned} \frac{dX_1^{(1)}(t)}{dt} - 0.557995X_1^{(1)}(t) \\ = -0.015596X_2^{(1)}(t) + 1806.203, \quad t = 1, 2, \dots, 8. \end{aligned} \quad (41)$$

TABLE 1: China's industrial SO<sub>2</sub> emissions and gross industrial output values of 2003–2010.

Time (year)	1 (2003)	2 (2004)	3 (2005)	4 (2006)	5 (2007)	6 (2008)	7 (2009)	8 (2010)
Industrial SO <sub>2</sub> emissions (10 <sup>4</sup> Tons)	2158.50	2254.90	2549.40	2588.80	2468.09	2321.23	2214.40	2185.15
Industrial output (RMB¥100 Million)	54945.53	65210.03	77230.78	91310.94	110534.88	130260.24	135239.95	160867.01

N.B. Data sourced from the Chinese National Bureau of Statistics Website (<http://www.stats.gov.cn/>).

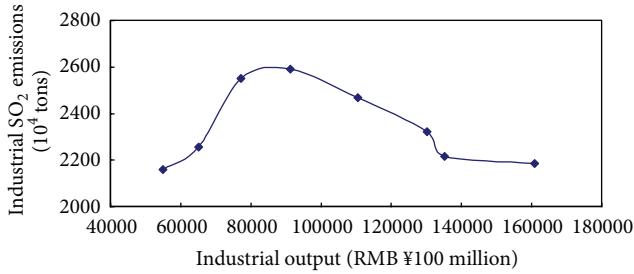


FIGURE 1: China's industrial SO<sub>2</sub> emissions correlated with gross industrial output values.

TABLE 2: The values of parameters  $n, r, rf$ , and  $rp$  in (7), the estimates of model parameters  $b_1, b_2$ , and  $u$  in (10), and the values of RMSPEPR and RMSPEPO in (38) and (39).

Parameters	Values
$n$	2
$r$	5
$rf$	3
$rp$	0
$b_1$	-0.557995
$b_2$	-0.015596
$u$	1806.203
RMSPEPR (%)	8.02
RMSPEPO (%)	185.25

In summary, the right-hand side of (7), the discrete function  $f(t)$  in (13) for the GMC (1, 2) model is obtained and listed in Table 3; the values of RMSPEPR and RMSPEPO in (38) and (39), respectively, are also listed in Table 2. The modelling values and forecasts for China's industrial SO<sub>2</sub> emissions by GMC (1, 2) are listed in Table 4.

As seen in Table 2, the root mean squared percentage error for the presample period (RMSPEPR) was 8.02%, but the root mean squared percentage error for the postsample period (RMSPEPO) was as high as 185%. This result was unacceptable. Table 4 shows that the relative errors of modelling and prediction gradually increased over time, especially the relative errors for the period 2008 to 2010, being up to -48%, -128%, and -290%, respectively. This indicated that the traditional GMC (1, 2) model cannot be used to describe the relationship between China's industrial SO<sub>2</sub> emissions and industrial output values nor could it be used to predict future industrial SO<sub>2</sub> emissions.

TABLE 3: The discrete function  $f(t)$  in (13).

$t$	$f(t)$
1	949.250
2	-67.792
3	-1272.31
4	-2696.44
5	-4420.38
6	-6451.98
7	-8561.23
8	-11070.18

3.2. Forecasting by NGMC (1, 2). Applying the NGMC (1, 2) model of (19) to (37), the values of parameters  $n, r, rf$ , and  $rp$  in (21), the estimates of model parameters  $b_1, b_2$  and  $u$  in (23), the optimised parameter  $\beta_2$  in (40) can be obtained and values are listed in Table 5. The NGMC (1, 2) model from (21) becomes

$$\frac{dX_1^{(1)}(t)}{dt} - 0.222770X_1^{(1)}(t) = -46938.3(X_2^{(1)}(t))^{-0.06305} + 25887.98, \quad t = 1, 2, \dots, 8. \tag{42}$$

In summary, the right-hand side of (21), the discrete function  $f_N(t)$  in (26) for the NGMC (1, 2) model, is obtained and values are listed in Table 6; the values of RMSPEPR and RMSPEPO in (38) and (39), respectively, are also listed in Table 5. The modelling values and forecasts of China's industrial SO<sub>2</sub> emissions by NGMC (1, 2) are listed in Table 7. Among them, the values for 2003 to 2007 are derived by modelling and those for 2008 to 2010 are forecasts.

As seen in Table 5, the RMSPEPR and RMSPEPO of NGMC (1, 2) were 2.44% and 5.48%, respectively, significantly less than the corresponding RMSPEPR and RMSPEPO values for the GMC (1, 2) model as shown in Table 2. Table 7 also shows the modelling and prediction values for the period 2003 to 2010 by NGMC (1, 2) to be close to the actual values. This indicated that the traditional NGMC (1, 2) model could effectively describe the relationship between China's industrial SO<sub>2</sub> emissions and industrial output values and accurately forecast industrial SO<sub>2</sub> emissions.

#### 4. Conclusions

Grey prediction modelling with the addition of a convolution integral to GMC (1,  $n$ ) system greatly improved the

TABLE 4: The modelling values and forecasts of China’s industrial SO<sub>2</sub> emissions by GMC (1, 2).

Time	Modelling values or forecast	Observations	Percent errors (%)
1 (2003)	2158.50	2158.50	0.00
2 (2004)	2195.32	2254.90	-2.64
3 (2005)	2367.35	2549.40	-7.14
4 (2006)	2398.87	2588.80	-7.34
5 (2007)	2110.67	2468.09	-14.48
6 (2008)	1205.64	2321.23	-48.06
7 (2009)	-630.24	2214.40	-128.46
8 (2010)	-4153.31	2185.15	-290.07

TABLE 5: The values of parameters  $n$ ,  $r$ ,  $rf$ , and  $rp$  in (21), the estimates of model parameters  $b_1$ ,  $b_2$ ,  $u$  in (23) and the optimized parameter  $\beta_2$  in (40), and the values of RMSPEPR and RMSPEPO in (38) and (39).

Parameters	Values
$n$	2
$r$	5
$rf$	3
$rp$	0
$b_1$	-0.222770
$b_2$	-46938.3
$u$	25887.98
$\beta_2$	-0.06305
RMSPEPR (%)	2.44
RMSPEPO (%)	5.48

TABLE 6: The discrete function  $f_N(t)$  in (26).

$t$	$f_N(t)$
1	2301.635
2	3437.039
3	4128.824
4	4644.269
5	5074.072
6	5441.376
7	5732.536
8	6006.089

prediction accuracy of the traditional multiply-grey model GM (1,  $n$ ) and was successfully applied in practice. The current GMC (1,  $n$ ) model is linear; therefore, the premise of obtaining high-precision modelling results by GMC (1,  $n$ ) is the existence of similar trends between the predicted variables and their influencing factors. The research introduced a power exponent into a 1-AGO series of relative factors of GMC (1,  $n$ ) to reflect the nonlinear interaction of the associated series with that predicted. The nonlinear optimisation model is built to solve location parameters. The prediction examples related to China’s industrial SO<sub>2</sub> emissions showed that the GMC (1, 2) model cannot effectively describe the nonlinear relationship between China’s industrial SO<sub>2</sub> emissions and their industrial output: the ensuing

TABLE 7: The modelling values and forecasts of China’s industrial SO<sub>2</sub> emissions by NGMC (1,2).

Time	Modelling values or forecast	Observations	Percent errors (%)
1 (2003)	2158.50	2158.50	0.00
2 (2004)	2135.84	2254.90	-5.28
3 (2005)	2526.60	2549.40	-0.89
4 (2006)	2562.03	2588.80	-1.03
5 (2007)	2473.20	2468.09	0.21
6 (2008)	2335.84	2321.23	0.63
7 (2009)	2163.90	2214.40	-2.28
8 (2010)	1984.36	2185.15	-9.19

prediction errors were insurmountable. However, the NGMC (1, 2) model proposed in this paper effectively described the relationship and achieved satisfactory prediction accuracy.

Compared with the traditional GMC (1,  $n$ ),  $n - 1$  power exponents of the predicted variables’ relevant factors  $\beta_2, \beta_3, \dots, \beta_n$  are introduced into the NGMC (1,  $n$ ) model to reflect the effect of these upon nonlinear system behaviours and interactions. The unknown parameters are determined by a computer program, which calculates the minimum average relative percentage error of the forecasting model. This strengthens the adaptability of the NGMC (1,  $n$ ) model towards the original data and eventually improves the forecast accuracy.

### Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

### Acknowledgments

The author is grateful to the editors and the anonymous reviewers for their insightful comments and suggestions. The author also thanks the National Natural Science Foundation of China (Grant no. 71101132), the Philosophy and Social Science Foundation of Zhejiang Province, China (Grant no. 13ZJQN029YB), the Academic Climbing Project for Young and Middle-aged Leading Academic in the Universities of Zhejiang Province, China (Grant no. PD2013275),



the Postdoctoral Science Foundation of China (Grant no. 2013M540448), and the Postdoctoral Science Foundation of Jiangsu Province, China (Grant no. 1302139C) for financially supporting this study.

## References

- [1] J. L. Deng, "Control problem of grey systems," *Systems & Control Letters*, vol. 1, no. 5, pp. 288–294, 1982.
- [2] J. L. Deng, *The Elements on Grey Theory*, HUST Press, Wuhan, China, 2002.
- [3] S. Liu, Y. Dang, and Z. Fang, *Grey Systems Theory and Its Applications*, Science Press, Beijing, China, 3rd edition, 2004.
- [4] T.-L. Tien, "A research on the grey prediction model  $GM(1, n)$ ," *Applied Mathematics and Computation*, vol. 218, no. 9, pp. 4903–4916, 2012.
- [5] C.-H. Wang and L.-C. Hsu, "Using genetic algorithms grey theory to forecast high technology industrial output," *Applied Mathematics and Computation*, vol. 195, no. 1, pp. 256–263, 2008.
- [6] N.-m. Xie and S.-f. Liu, "Discrete grey forecasting model and its optimization," *Applied Mathematical Modelling*, vol. 33, no. 2, pp. 1173–1186, 2009.
- [7] Z. Zhao, J. Wang, J. Zhao, and Z. Su, "Using a Grey model optimized by differential evolution algorithm to forecast the per capita annual net income of rural households in China," *Omega*, vol. 40, no. 5, pp. 525–532, 2012.
- [8] D.-C. Li, C.-J. Chang, C.-C. Chen, and W.-C. Chen, "Forecasting short-term electricity consumption using the adaptive grey-based approach—an Asian case," *Omega*, vol. 40, no. 6, pp. 767–773, 2012.
- [9] L.-C. Hsu, "Forecasting the output of integrated circuit industry using genetic algorithm based multivariable grey optimization models," *Expert Systems with Applications*, vol. 36, no. 4, pp. 7898–7903, 2009.
- [10] L.-C. Hsu and C.-H. Wang, "Forecasting integrated circuit output using multivariate grey model and grey relational analysis," *Expert Systems with Applications*, vol. 36, no. 2, pp. 1403–1409, 2009.
- [11] Z. X. Wang, Y. G. Dang, and J. J. Zhao, "Integration of regional science and technology resources based on GM, (1, 3)," *The Journal of Grey System*, vol. 22, pp. 129–136, 2010.
- [12] T.-L. Tien, "The indirect measurement of tensile strength of material by the grey prediction model  $GMC(1, n)$ ," *Measurement Science and Technology*, vol. 16, no. 6, pp. 1322–1328, 2005.
- [13] W.-Y. Wu and S.-P. Chen, "A prediction method using the grey model  $GMC(1, n)$  combined with the grey relational analysis: a case study on internet access population forecast," *Applied Mathematics and Computation*, vol. 169, no. 1, pp. 198–217, 2005.
- [14] C.-R. Su, C.-K. Chen, W.-L. Liu, and H.-Y. Lai, "Estimation for inner surface geometry of furnace wall using inverse process combined with grey prediction model," *International Journal of Heat and Mass Transfer*, vol. 52, no. 15-16, pp. 3595–3605, 2009.
- [15] T.-L. Tien, "The deterministic grey dynamic model with convolution integral  $DGDMC(1, n)$ ," *Applied Mathematical Modelling*, vol. 33, no. 8, pp. 3498–3510, 2009.
- [16] T.-L. Tien, "The indirect measurement of tensile strength for a higher temperature by the new model  $IGDMC(1, n)$ ," *Measurement*, vol. 41, no. 6, pp. 662–675, 2008.
- [17] T.-L. Tien, "The indirect measurement of tensile strength by the new model  $FGMC(1, n)$ ," *Measurement*, vol. 44, no. 10, pp. 1884–1897, 2011.
- [18] L. Li, "China's industrial emissions control performance and emission reduction potential," *Industrial Technology Economy*, vol. 29, pp. 134–137, 2010.