Using a Novel Grey System Model to Forecast Natural Gas Consumption in China

Lifeng Wu, Sifeng Liu, Haijun Chen, and Na Zhang

1College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China
2Department of Mathematics and Physical, Handan College, Handan 056005, China

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

Received 14 April 2014; Revised 24 July 2014; Accepted 24 July 2014

1. Introduction

Energy has an influencing role in achieving economic and social progress. Forecasting energy constitutes a vital part of energy policy of a country, especially for a developing country like China whose economy is in a stage of energy transition: from low efficiency solid fuels to oil, gas, and electric power [1]. This has motivated many researchers to focus their research on energy forecasting. For example, Jian et al. used bayesian combination model to forecast energy demand of China [2]. Ji predicted petroleum consumption in China by comparing three models [3]. Meng and Niu analysed and predicted annual electricity consumption of China [4]. Zhang et al. forecasted Chinese transport energy demand based on partial least squares regression [5]. Unler proposed swarm intelligence to improve energy demand forecasting [6]. A large number of studies on energy consumption forecasting using grey model and improved grey model are reported, such as Meng et al. [7], Bianco et al. [8], Li et al. [9], Niu et al. [10], Pao and Tsai [11], and Wang et al. [12].

Although the first-order grey model with one variable (GM(1,1)) had been widely adopted, its predicting performance still could be improved. In this paper, based on the principle of new information priority, we consider the priority of the nth component and the n − 1th component, respectively, and the nth component is more prior than the n − 1th component. Similarly, the 2nd component is more prior than the 1st component [13]. Two case studies show the effectiveness of the proposed modeling method, which is particularly suitable for limited sample forecasting.

The rest of the paper proceeds as follows: an overview of the relevant grey models and a new GM(1,1) model with the principle of new information priority of grey system are present in Section 2. The advantage of the grey model proposed in this paper over the other grey models is proved by two real cases in China in Section 3. The prediction concerning natural gas consumption in China will be conducted in Section 4. Some conclusions are discussed in final section.

2. Methodology

2.1. Grey Model Based on Autoregressive Method. Assume $X = \{x(1), x(2), \ldots, x(n)\}$, where $X$ stands for a data sequence. The equation

$$x(k) = \alpha x(k-1) + \beta, \quad k = 2, 3, \ldots, n,$$

(1)

is called the autoregressive GM(1,1) model (ARGM(1,1)). Use the ordinary least squares method to estimate the parameters as follows:

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = (B^TB)^{-1}B^TY,$$

(2)
where
\[
Y = \begin{bmatrix}
x(2) \\
x(3) \\
\vdots \\
x(n - 1) \\
x(n)
\end{bmatrix}, \quad B = \begin{bmatrix}
x(1) & 1 \\
x(2) & 1 \\
\vdots & \vdots & \vdots \\
x(n - 2) & 1 \\
x(n - 1) & 1
\end{bmatrix}.
\]

(3)

**Theorem 1.** Assume that \(\hat{\alpha}, \hat{\beta}\) are the same as (1). Set \(\hat{x}(1) = x(1)\); then recursive function is given by
\[
\hat{x}(k + 1) = \hat{x}(1) + \hat{\alpha} \sum_{i=k}^{n-1} \hat{x}(i) + \frac{\hat{\beta}}{\hat{\alpha}}, \quad k = 1, 2, \ldots, n - 1.
\]

(4)

**Proof.** Substitute \(\hat{\alpha}, \hat{\beta}\) into (1); then
\[
\hat{x}(k + 1) = \hat{x}(1) + \sum_{i=k}^{n-1} \hat{x}(i) + \beta + \frac{\beta}{\hat{\alpha}} = \hat{x}(1) + \frac{1 - \hat{\alpha}^{-k}}{1 - \hat{\alpha}} x(1) + \frac{\beta}{1 - \hat{\alpha}}.
\]

(5)

Let \(\hat{x}(1) = x(1)\); then
\[
\hat{x}(k + 1) = \frac{1 - \hat{x}(k)}{1 - \hat{\alpha}} + \frac{\beta}{1 - \hat{\alpha}}.
\]

(6)

We also obtain a nonhomogenous exponential model as follows:
\[
\hat{x}(k) = \frac{1 - \hat{x}(k)}{1 - \hat{\alpha}} + \frac{\beta}{1 - \hat{\alpha}}, \quad k = 1, 2, \ldots, n - 1.
\]

(7)

Because grey autoregressive method of GM(1,1) is formulated by using the original data rather than the accumulated generation data, it does not need to inverse accumulated generating. We can see that (1) is an exponential model and (7) is a nonhomogenous exponential model. This method is called the ARGM(1,1).

2.2. The GM(1,1) Model with the Principle of New Information Priority via Weighted Least Squares Method. In this section, by weighted least squares method, the coefficients of (1) can be obtained as follows:

Minimize : \(\sum_{k=2}^{n} w_k [x(k) - \alpha x(k-1) - \beta]^2\),

(8)

where \(w_k\) can reflect the principle of new information priority, so (9) also pays more attention to the new error and embodies the principle of new information priority. We call this method WLSGM(1,1).

2.3. The GM(1,1) Model with the Principle of New Information Priority. Assume \(X = \{x(1), x(2), \ldots, x(n)\}\), where \(X\) stands for a data sequence. The original data of (1) is
\[
X = \begin{bmatrix}
x(2) & x(1) & 1 \\
x(3) & x(2) & 1 \\
\vdots & \vdots & \vdots \\
x(n - 1) & x(n - 2) & 1 \\
x(n) & x(n - 1) & 1
\end{bmatrix}.
\]

(10)

The data with the principle of new information priority is
\[
\begin{bmatrix}
1 & 1 & \cdots & 1 & 1 \\
0 & 1 & \cdots & 1 & 1 \\
\vdots & \vdots & \cdots & \vdots & \vdots \\
0 & 0 & \cdots & 1 & 1 \\
0 & 0 & \cdots & 0 & 1
\end{bmatrix} \begin{bmatrix}
x(2) & x(1) & 1 \\
x(3) & x(2) & 1 \\
\vdots & \vdots & \vdots \\
x(n - 1) & x(n - 2) & 1 \\
x(n) & x(n - 1) & 1
\end{bmatrix} = A \begin{bmatrix}
x(2) & x(1) & 1 \\
x(3) & x(2) & 1 \\
\vdots & \vdots & \vdots \\
x(n - 1) & x(n - 2) & 1 \\
x(n) & x(n - 1) & 1
\end{bmatrix}
\]

(11)

which is called the treatment of new information priority. The equation
\[
\sum_{i=k}^{n} x(i) = \alpha \sum_{i=k}^{n} x(i - 1) + (n - k + 1) \beta,
\]

(12)

is referred to as grey nonhomogenous exponential model with the principle of new information priority (abbreviated as NIGM(1,1)).

**Theorem 2.** Assume that \(\alpha, \beta\) are certain. Actually, if the error is \(\mu\), then
\[
x(k) = \alpha x(k - 1) + \beta + \mu_k, \quad k = 2, 3, \ldots, n.
\]

(13)

Via the treatment of new information priority, the mean of the error is
\[
E(A\mu) = \mu_0 + \frac{(n - 2) \mu_{n-1}}{n - 1} + \frac{2 \mu_3}{n - 1} + \frac{\mu_2}{n - 1}.
\]

(14)
where \( A \) is the same as (11). Use the ordinary least squares method to estimate the parameters as follows:

\[
\begin{bmatrix}
\hat{\alpha} \\
\hat{\beta}
\end{bmatrix} = (B^T B)^{-1} B^T Y,
\]

where

\[
Y = \begin{bmatrix}
\sum_{k=2}^{n} x(k) \\
\sum_{k=3}^{n} x(k) \\
\vdots \\
\sum_{k=n-1}^{n} x(k) \\
x(n)
\end{bmatrix},
\]

\[
B = \begin{bmatrix}
\sum_{k=2}^{n} x(k-1) & n-1 \\
\sum_{k=3}^{n} x(k-1) & n-2 \\
\vdots & \vdots \\
\sum_{k=n-1}^{n} x(k-1) & 2 \\
x(n-1) & 1
\end{bmatrix}.
\]

The variable of

\[
\begin{bmatrix}
\hat{\alpha} \\
\hat{\beta}
\end{bmatrix}
\]

is \( \text{Var} \left[ \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} \right] = (B^T A^T AB)^{-1} B^T A^T (A \mu \mu^T A^T) A B (B^T A^T AB)^{-1}, \)

where

\[
E(A \mu \mu^T A^T) = \begin{bmatrix}
\sum_{i=2}^{n} \mu_i^2 & \mu_3 & \ldots & \mu_n \\
\mu_3 & \sum_{i=2}^{n} \mu_i & \ldots & \mu_n \\
\vdots & \vdots & \ddots & \vdots \\
\mu_n & \mu_n & \ldots & \mu_n^2
\end{bmatrix}.
\]

Obviously,

\[
E(A \mu) = \mu_n + \frac{(n-2)\mu_{n-1}}{n-1} + \cdots + \frac{2\mu_3}{n-1} + \frac{\mu_2}{n-1}.
\]

Use the ordinary least squares method to estimate the parameters as follows:

\[
\begin{bmatrix}
\hat{\alpha} \\
\hat{\beta}
\end{bmatrix} = (B^T A^T AB)^{-1} B^T A^T A Y.
\]

Substituting

\[
AY = AB \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + A \mu
\]

into (21), we obtain

\[
\begin{bmatrix}
\hat{\alpha} \\
\hat{\beta}
\end{bmatrix} = (B^T A^T AB)^{-1} B^T A^T \left( AB \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + A \mu \right)
\]

\[
= \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + (B^T A^T AB)^{-1} B^T A^T A \mu.
\]

Then

\[
E\left[ \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} \right] = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + E\left[ (B^T A^T AB)^{-1} B^T A^T A \mu \right]
\]

\[
= \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + (B^T A^T AB)^{-1} B^T A^T E(A \mu).
\]

Actually, by

\[
E\left[ \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} \right] = \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + E\left( (B^T A^T AB)^{-1} B^T A^T A \mu \right)
\]

and \( E(A \mu) \), we obtain

\[
\lim_{n \to \infty} \frac{n-2}{n-1} - \frac{n-3}{n-1} = 0,
\]

\[
\ldots,
\]

\[
\lim_{n \to \infty} \frac{2}{n-1} - \frac{1}{n-1} = 0.
\]

The more the sample is, the less the attention we pay to the error from new information and vice versa. We can obtain

\[
\text{Var} \left[ \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} \right] = \begin{bmatrix} \text{Var}(\hat{\alpha}) & \text{Cov}(\hat{\beta}, \hat{\alpha}) \\ \text{Cov}(\hat{\beta}, \hat{\alpha}) & \text{Var}(\hat{\beta}) \end{bmatrix}.
\]

By

\[
\begin{bmatrix}
\hat{\alpha} \\
\hat{\beta}
\end{bmatrix} - \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = (B^T A^T AB)^{-1} B^T A^T A \mu,
\]

thus

\[
\text{Var} \left[ \begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} \right] = (B^T A^T AB)^{-1} B^T A^T E
\]

\[
\times (A \mu \mu^T A^T) A B \begin{bmatrix} B^T A^T AB \end{bmatrix}^{-1}.
\]
In most cases, we pay attention to the \( \text{Var}(\hat{\alpha}) \) and \( \text{Var}(\hat{\beta}) \) and overlook \( \text{Cov}(\hat{\beta}, \hat{\alpha}) \); thus we denote (27) as

\[
\text{Var} \left[ \begin{array}{c} \hat{\alpha} \\ \hat{\beta} \end{array} \right] = B^TA^T A B^{-1} E \left( A \hat{\mu}^T A^T \right). \tag{30}
\]

So

\[
E \left( A \hat{\mu}^T A^T \right) = \begin{bmatrix} \mu_3 & \cdots & \mu_n \\ \vdots & \ddots & \vdots \\ \mu_n & \cdots & \mu_3 \end{bmatrix}.
\]

Let \( E(A \hat{\mu}^T A^T) = D \); then the components from \( D_{11} \) to \( D_{nn} \) all contain \( \mu_n^2 \). The components from \( D_{11} \) to \( D_{n-1,n-1} \) all contain \( \mu_{n-1}^2 \). The components from \( D_{11} \) to \( D_{n-1,n-1} \) all contain \( \mu_{n-2}^2 \). Analogizing successively, the components \( D_{11} \) and \( D_{21} \), both contain \( \mu_2^2 \). Only \( D_{11} \) contains \( \mu_1^2 \). So

\[
\text{Var} \left[ \begin{array}{c} \hat{\alpha} \\ \hat{\beta} \end{array} \right] = \begin{bmatrix} \mu_3 & \cdots & \mu_n \\ \vdots & \ddots & \vdots \\ \mu_n & \cdots & \mu_3 \end{bmatrix} \tag{32}
\]

pay more attention to the variable of the new information in the modelling process. The modeling procedure of NIGM(1,1) can be concluded as follows.

**Step 1.** An original nonnegative series \( \{x(1), x(2), \ldots, x(n)\} \) is given.

**Step 2.** \( Y \) and \( B \) which are the same as Theorem 2 can be obtained via the treatment of new information priority.

**Step 3.** Least squares estimation of \( \alpha \) and \( \beta \) according to Theorem 2 is as follow.

\[
\left[ \begin{array}{c} \hat{\alpha} \\ \hat{\beta} \end{array} \right] = (B^TA^T A B)^{-1} B^TA^TY. \tag{33}
\]

**Step 4.** We obtain a nonhomogenous exponential model as follows:

\[
\bar{x}(k) = \hat{\alpha}^{k-1} x(1) - \frac{\hat{\beta}}{1 - \hat{\alpha}} + \frac{\hat{\beta}}{1 - \hat{\alpha}}, \tag{34}
\]

\[k = 1, 2, \ldots, n - 1. \]

**Step 5.** In order to solve the effect of the initial condition, we assume

\[
x(k) = c\hat{\alpha}^k + b. \tag{35}
\]

Use the ordinary least squares method to estimate the parameters

\[
\begin{bmatrix} \hat{c} \\ \hat{b} \end{bmatrix} = (B^TB)^{-1} B^TY, \tag{36}
\]

where

\[
Y = \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(n - 1) \\ x(n) \end{bmatrix}, \quad B = \begin{bmatrix} \hat{\alpha} \\ \hat{\alpha}^2 \\ \vdots \\ \hat{\alpha}^{n-1} \\ \hat{\alpha}^n \end{bmatrix} \tag{37}
\]

thus, we obtain

\[
\bar{x}(k) = \hat{c}\hat{\alpha}^k + \hat{b}, \quad k = 1, 2, \ldots, n, \tag{38}
\]

which is a nonhomogenous exponential model.

### 3. Results

In this section, the advantage of the NIGM(1,1) model over the other grey models is demonstrated by two real case studies of energy consumption. Mean absolute percentage error (MAPE = 100%(1/n) \( \sum_{k=1}^{n} |(x(k) - \bar{x}(k))/x(k)| \)) compares the real and forecasted values to evaluate the precision.

**Case 1** (natural gas consumption forecasting example in China [14]). This example is from the literature [14], in order to compare the result of the literature [14] with the result of the proposed model. The sample data is the same as the literature [14]. Using the data from 1995 to 2006 to construct NIGM(1,1),

\[
x(k) = 22.478 \times 1.276^k + 144.995. \tag{39}
\]

Then the value from 2007 to 2008 (out-of-sample) is predicted. Actual values and the forecasting values of five compared models are presented in Table 1. As can be seen from Table 1, NIGM(1,1) yielded the lowest MAPE compared with the other models during the period from 1995 to 2008.

**Case 2** (the transport energy demand forecasting in China [5]). The example from the literature [5] which provides the sample data is considered. The data from 1990 to 2004 are used to construct NIGM(1,1),

\[
x(k) = 11.315 \times 1.168^k + 32.49. \tag{40}
\]

Then the value from 2005 to 2006 (out-of-sample) is predicted. Actual values and the forecasting values of five compared models are presented in Table 2. As can be seen from Table 2, NIGM(1,1) reduces the MAPE of original GM(1,1) from 4.82% to 3.7%. This indicates that NIGM(1,1) can improve the forecasting error of GM(1,1) but not the partial least squares regression method in the in-sample model. From a short-term forecasting viewpoint, NIGM(1,1) yielded the lowest MAPE from 2005 to 2006 compared with the other five models, which means that the
NIGM(1,1) model reaches the objective of minimizing the forecast error and has highly accurate forecasting power. Therefore, we can conclude that NIGM(1,1) significantly enhances the precision of grey forecasting model.

4. Discussion and Policy Implications

Many experts have developed methods to forecast natural gas consumption [15–23] because, in recent years, the Chinese government has begun to recognize the disadvantages of the high reliance on coal. Current policies are different from previous policies. According to current policies, natural gas will be developed as a partial substitution for coal. From a short-term forecasting viewpoint, we must pay more attention to the recent data. So we choose data from 1996 to 2013 to construct NIGM(1,1):

\[ x(k) = 122.2219 \times 1.14736^k - 30.8291. \]  (41)

We use the latest (2012-2013) value to compare the estimated value with the actual data; the result is listed in Table 3.

Model (41) is used to predict the Chinese natural gas consumption based on the above comparison. The forecasted values and the trend from 2014 to 2018 are given in Table 4.
<table>
<thead>
<tr>
<th>Year</th>
<th>Actual value</th>
<th>Model (41)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2012</td>
<td>1463.0</td>
<td>1420.5</td>
</tr>
<tr>
<td>2013</td>
<td>1616.1</td>
<td>1634.3</td>
</tr>
<tr>
<td>MAPE</td>
<td></td>
<td>2.0</td>
</tr>
</tbody>
</table>

**Table 3:** Fitted values and MAPE of model (41).

<table>
<thead>
<tr>
<th>Year</th>
<th>Model (41)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2014</td>
<td>1879.7</td>
</tr>
<tr>
<td>2015</td>
<td>2161.2</td>
</tr>
<tr>
<td>2016</td>
<td>2484.3</td>
</tr>
<tr>
<td>2017</td>
<td>2854.9</td>
</tr>
<tr>
<td>2018</td>
<td>3280.2</td>
</tr>
</tbody>
</table>

**Table 4:** The forecasting values of model (41).

It can be seen that there is an increase in trend in the consumption data from Table 4. Continued fast growth makes it important to ensure the production and availability of the amount of natural gas to meet life's need. Further effort must be made to achieve a sustainable development of China's natural gas industry.

**5. Conclusion**

The incompleteness of information is the primary characteristic of grey system theory, so we put more emphasis on recent data. A heavier weight is assigned to the recent data with two cases which are used to examine the performance of new grey model. It may be used for other real cases for energy consumption forecasting.

**Conflict of Interests**

The authors declare that they have no competing interests in this paper.

**Acknowledgments**

The relevant researches done in this paper are supported by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme (Grant no. FP7-PIIF-GA-2013-629051), the National Natural Science Foundation of China (Grant nos. 71771113, 71271226, and 71401051), the Natural Science Foundation of Jiangsu Province (Grant nos. BK20130785 and BK20130786), Research Fund for the Doctoral Program of Higher Education (Grant no. 2013218120036), Project of Social Science Foundation of the China (Grant no. 12AZD102), Funding for Talent Introduction in NUAA (Grant no. 1009-YAH14003), and Funding of Nantong Science and Technology Program (Grant no. HS2013026).

**References**


Submit your manuscripts at
http://www.hindawi.com