A Grey NGM(1, 1, k) Self-Memory Coupling Prediction Model for Energy Consumption Prediction

Xiaojun Guo, 1,2 Sifeng Liu, 1 Lifeng Wu, 1 and Lingling Tang 3

1 College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
2 School of Science, Nantong University, Nantong 226019, China
3 School of Electrical and Computer Engineering, Cornell University, Ithaca, NY 14853, USA

Correspondence should be addressed to Xiaojun Guo; guoxj159@163.com

Received 6 January 2014; Revised 6 May 2014; Accepted 21 May 2014; Published 18 June 2014

Academic Editor: Constantin Papaodysseus

Copyright © 2014 Xiaojun Guo et al. 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.

Energy consumption prediction is an important issue for governments, energy sector investors, and other related corporations. Although there are several prediction techniques, selection of the most appropriate technique is of vital importance. As for the approximate nonhomogeneous exponential data sequence often emerging in the energy system, a novel grey NGM(1, 1, k) self-memory coupling prediction model is put forward in order to promote the predictive performance. It achieves organic integration of the self-memory principle of dynamic system and grey NGM(1, 1, k) model. The traditional grey model's weakness as being sensitive to initial value can be overcome by the self-memory principle. In this study, total energy, coal, and electricity consumption of China is adopted for demonstration by using the proposed coupling prediction technique. The results show the superiority of NGM(1, 1, k) self-memory coupling prediction model when compared with the results from the literature. Its excellent prediction performance lies in that the proposed coupling model can take full advantage of the systematic multitime historical data and catch the stochastic fluctuation tendency. This work also makes a significant contribution to the enrichment of grey prediction theory and the extension of its application span.

1. Introduction

As one of the most significant national strategic resources, the energy issue is an important factor which restricts the state economy and social development. Along with the ongoing economic growth and the acceleration of industrialization, energy consumption and production will increase even more rapidly. Therefore, it is meaningful to identify and analyze the energy issue legitimately, especially for predicting the future energy consumption correctly and scientifically. Energy consumption is featured by its information uncertainty and few useful analyzing samples. There exist many influential factors (economy condition, industry framework, climatic variation, government policy and so on), which are difficult to determine how exactly they affect energy consumption [1, 2]. Owing to the apparent uncertain characters embodied in the complicated energy system, the energy consumption prediction can be regarded as a grey system exactly.

Most traditional prediction techniques for time series concentrate mainly on the statistical analysis methods, such as simple regression, multivariate regression, exponential smoothing, and so forth, and possess the advantage of accurately approximating the evolutionary trend [3–5]. Nevertheless, the methods must be accomplished with the assumption of realizing the system structure and the limitation of requiring a large amount of historical data. And owing to the increasing complexity, uncertainty, and chaos of the system's structure, it is very difficult to accurately predict random fields using the traditional statistical methods. To overcome this drawback, the grey systems theory was initially proposed by Deng to study the uncertainty of systems [6]. As an important theoretical component, the grey prediction approach represented by GM(1, 1) model can weaken the randomness of original statistical data by means of accumulated generating operation [7]. The superiority of grey models is that they only require a limited amount of statistical data without
knowing their statistical distribution. GM(1, 1) model has ideal predictive effect for approximate homogenous exponential sequence and has already been effectively utilized in numerous fields, such as social economy, geographical environment, engineering science, public transit, and so on [8–12]. Meanwhile, several improved GM(1, 1) prediction models were developed for the prediction of electricity load, energy consumption, and so forth [13–16]. Nevertheless, certain time series of energy system often show large stochastic fluctuations due to some uncertain influence factors and present the characteristics of approximate nonhomogeneous exponential. It is inevitable to generate the apparent modeling errors by using GM(1, 1) model, so the grey NGM(1, 1, k) model appropriate for the approximate nonhomogeneous exponential law with stochastic fluctuation was put forward [17].

On the basis of the retrieved modeling, the self-memory principle of dynamic system was developed firstly by Cao [18]. As a mathematic realization of integrating the deterministic and random theories, the principle is a statistically-dynamic method to solve problems of nonlinear dynamic systems [19, 20]. The self-memory principle can retrieve ideal nonlinear dynamic models by means of the practical observational data. It can overcome not only the weakness as being sensitive to initial value of the initial value problem for differential equations, but also the limitation as irrelevant to mechanism equation due to the utilization of the historical materials. It has been utilized increasingly into time series forecasting in numerous fields from meteorology to engineering to economics [21–23]. In recent years, some scholars preliminarily attempted to introduce the self-memory principle into certain basic grey prediction models. Fan derived the self-memory numerical method for solving GM(1, 1) model and established a novel grey model recollecting the last several data [24]. Chen et al. established a coupled equation by combining DHGM(2, 2) grey differential equation with the self-memory principle to forecast the flood [25]. Guo et al. established an interval grey number self-memory coupling prediction model based on the grey degree of compound grey number [26]. Accordingly, for the purpose of extending the applicable range of grey prediction model and promoting its predictive performance, the self-memory principle is firstly introduced into grey NGM(1, 1, k) model.

The purpose of this paper is to construct a novel grey NGM(1, 1, k) self-memory coupling prediction model appropriate for the approximate nonhomogeneous exponential data sequence with stochastic fluctuation emerging in the energy consumption prediction. The novel prediction model synthesizes the advantages of the self-memory principle and grey NGM(1, 1, k) model through organically coupling the above two prediction methods. Its excellent predictive performance lies in that the grey model’s weakness as being sensitive to initial value can be overcome by using multi-time-point initial field instead of only single-time-point initial field.

The remaining content is organized as follows. Section 2 provides an overview of the relevant literature on generalized GM(r, h) model, GM(1, 1) model, and NGM(1, 1, k) model. Section 3 presents the detailed algorithm of a novel NGM(1, 1, k) self-memory coupling prediction model and a step-by-step procedure. In Section 4, the illustrative examples of total energy, coal, and electricity consumption prediction in China are adopted to demonstrate the adaptability and effectiveness of the proposed NGM(1, 1, k) self-memory coupling prediction model. Finally, some conclusions are drawn in Section 5.

2. Reviewing the Generalized Grey System Model and NGM(1, 1, k) Model

2.1. Generalized Grey System Model and Symbols Description

Assume that the sequence

\[ F_1^0 = \{ f_1^0 | t \in 1, 2, \ldots, n \} \] (1)

is an original time series, where \( f_1^0 \) denotes the observational data at time \( t \); then the first-order accumulated generation (abbreviated as 1-AGO) value \( f_1^1 \) of the original time series data \( f_1^0 \) is obtained as

\[ f_1^1 = \sum_{k=1}^{t} f_k^0, \quad t = 1, 2, \ldots, n. \] (2)

And the sequence

\[ F_1^1 = \{ f_1^1 | t \in 1, 2, \ldots, n \} \] (3)

is called the 1-AGO time series of original time series \( f_1^0 \).

The \( r \)-order differential equation

\[ \frac{d^r F_1^1}{dt^r} + a_1 \frac{d^{r-1} F_1^1}{dt^{r-1}} + \cdots + a_{r-1} \frac{d F_1^1}{dt} + a_r F_1^1 = b \] (4)

is called the whitenization equation of the generalized GM(r, h) model, where the vectors \( \vec{a} = [-a_1, -a_2, \ldots, -a_r]^T \) and \( \vec{b} = [b_1, b_2, \ldots, b_h]^T \) are called developing and driving coefficients vectors, respectively. And there is one dependent variable \( F_1^1 \) and \( h - 1 \) independent variables \( X_1^0(t), X_2^1(t), \ldots, X_{h-1}^1(t) \) in (4).

2.2. Grey GM(1, 1) Model

Particularly when \( r = 1 \) and \( h = 1 \) in (4), the first-order differential equation with one dependent variable

\[ \frac{d F_1^1}{dt} + a F_1^1 = b \] (5)

is called the whitenization equation of the GM(1, 1) model, where the parameter \( a \) represents the developing coefficient, \( b \) represents the grey input coefficient, and \( F_1^1 \) is the dependent variable with AGO input value \( f_1^1 \). Meanwhile the equation

\[ f_1^1 + 0.5a (f_1^1 + f_{1-1}^0) = b, \quad t = 1, 2, \ldots, n \] (6)

is called the basic form of the GM(1, 1) model.
Let sampling time $\Delta t = 1$; then by applying the least square method with input data $f_0^n$, the parameters $a$ and $b$ in matrix $\hat{R}$ can be obtained as

$$\hat{R} = \left(\frac{\hat{a}}{\hat{b}}\right) = (B_1^TB_1)^{-1}B_1^TF_n^0,$$  \quad (7)

where

$$F_n^0 = \begin{bmatrix} f_2^0 \\ f_3^0 \\ \vdots \\ f_n^0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} -0.5(f_2^0 + f_1^0) & 1 \\ -0.5(f_3^0 + f_2^0) & 1 \\ \vdots & \vdots \\ -0.5(f_n^0 + f_{n-1}^0) & 1 \end{bmatrix}.$$  \quad (8)

By making the initial value $f_1^0 = f_0^1$, the time response sequence of GM(1,1) model is given by:

$$f_{t+1}^1 = f_t^1 - b + \frac{a}{\hat{a}} f_t^0, \quad \forall t \geq 1,$$  \quad (9)

and the simulated value $f_{t+1}^1$ of dependent variable $F_t^1$ can be obtained from (9) accordingly. Consider the inverse accumulated generation operation (abbreviated as IAGO)

$$f_{t+1}^0 = f_{t+1}^1 - f_t^1, \quad \forall t \geq 1;$$  \quad (10)

then the simulated value $f_{t+1}^0$ of IAGO variable $F_t^0$ can be obtained.

As mentioned above, GM(1,1) model can be considered as the most simple and special case of the generalized GM(r, h) model. Conversely, generalized GM(r, h) model can be treated as the extension of GM(1,1) model. As the basic and core part of grey systems theory, GM(1,1) model is one of the most frequently used grey prediction models for time series. It is based on the grey exponential law resulting from accumulated generating and moving averaging operation, which has an ideal predictive effect for the time series with the approximate homogeneous exponential characteristics. Meanwhile, there exist numerous literatures which play an important role in promoting the predictive performance of GM(1,1) model by means of various kinds of optimization techniques [8, 27, 28].

2.3. Grey NGM(1,1,k) Model. When the traditional GM(1,1) model is used for modeling analysis, it is assumed that the original data sequence must obey approximate homogeneous exponential growth law. However, in fact, the approximate homogeneous exponential data sequence is very limited. As for the time series with the characteristics of approximate nonhomogeneous exponential with stochastic fluctuation, Cui et al. first put forward a novel grey prediction model termed NGM(1,1,k) model [17, 29]. Aiming at some defects of parameter setting in original NGM(1,1,k) model, Cui and Lu [30] constructed a modified NGM(1,1,k) model by optimizing parameters of whitenization differential equation. And then Chen and Wei [31] further optimized the grey derivative of approximate nonhomogeneous index sequence GM(1,1) model, thus perfecting the model parameters.

The first-order differential equation

$$\frac{dF_1^1}{dt} + aF_1^1 = \gamma t + b$$  \quad (11)

is called the whitenization equation of the NGM(1,1,k) model, where the parameter $\gamma$ represents the control coefficient and $a, b$ are the same as mentioned above. Meanwhile, the equation

$$f_t^0 + 0.5a(f_t^0 + f_{t-1}^0) = \gamma t + b, \quad t = 1, 2, \ldots, n$$  \quad (12)

is called the basic form of the NGM(1,1,k) model.

Let sampling time $\Delta t = 1$; then the least square estimate of the parameters sequence $a, \gamma, b$ is given by

$$\hat{R} = \left(\frac{\hat{a}}{\hat{b}}\right) = (B_2^TB_2)^{-1}B_2^TF_n^0,$$  \quad (13)

where

$$F_n^0 = \begin{bmatrix} f_2^0 \\ f_3^0 \\ \vdots \\ f_n^0 \end{bmatrix}, \quad B_2 = \begin{bmatrix} -0.5(f_2^0 + f_1^0) & 2 & 1 \\ -0.5(f_3^0 + f_2^0) & 3 & 1 \\ \vdots & \vdots & \vdots \\ -0.5(f_n^0 + f_{n-1}^0) & n & 1 \end{bmatrix}.$$  \quad (14)

By making the initial value $f_1^0 = f_0^1$, the time response sequence of NGM(1,1,k) model is given by

$$f_{t+1}^1 = \left(\frac{f_0^0 - \frac{\gamma}{a} + \frac{\gamma t}{a} - \frac{b}{a}}{\hat{a}}\right) e^{-at} + \frac{\gamma}{a} (t + 1)$$  \quad (15)

and the simulated value $f_{t+1}^0$ of dependent variable $F_t^1$ can be obtained from (15) accordingly. As mentioned above, the simulated value $f_{t+1}^0$ of IAGO variable $F_t^0$ can be calculated by IAGO.

NGM(1,1,k) model is a sort of significant nonlinear grey prediction model in grey systems theory. It can reflect well the approximate nonhomogeneous exponential characteristic of real time series data, which possesses higher accuracy of simulation and prediction. And the NGM(1,1,k) model is still further superior to GM(1,1) model with respect to applied range and predictive performance for the approximate nonhomogeneous exponential law time series with stochastic fluctuation.

3. Novel NGM(1,1,k) Self-Memory Coupling Prediction Model

3.1. Fundamental Principles of Self-Memory Principle of Dynamic System. By introducing the memory concept into physics, the self-memory principle of dynamic system is
proposed on the basis that natural and social phenomena are all irreversible. The history information should be investigated fully if we want to realize present system and predict its future. Accordingly, the principle emphasizes the relationship between before and after of system status itself, particularly on the systematic evolution law per se. After the memory function which contains historical information is introduced into the system’s dynamic differential equation, it can be transformed into an appropriate difference-integral equation which is called a self-memoryization one by defining the inner product in Hilbert space. Because the systematic self-memoryization equation contains multiple time-point initial fields instead of only single-time-point initial field, the weakness as being sensitive to initial value of the original dynamic differential equation can be overcome. Then through studying systematic inner memorability, the superior of utilizing self-memory principle lies in that systematic evolutionary trend can be modeled and predicted. The advantage of self-memory principle lies in that the systematic predictive ability can be improved by means of not only combining dynamics calculations and estimating parameters of historical data, but also extracting systematic information from historical data in statistics.

Based on the above-mentioned literature analysis, the superior self-memory technique is introduced in this section to support the NGM(1,1,k) model so as to devise a novel NGM(1,1,k) self-memory coupling prediction model. Let the original time series and the l-AGO time series be $F^0_t = \{f^0_t \mid t \in 1,2,\ldots,n\}$ and $F^1_t = \{f^1_t \mid t \in 1,2,\ldots,n\}$, respectively. Let $dF^1_t/ dt$ in the whitenization equation of the NGM(1,1,k) model be $F(x,\lambda,t)$; then

$$F(x,\lambda,t) = -aF^1_t + yt + b. \quad (16)$$

The differential equation $dF^1_t/ dt$, which has been determined by (16), is considered to be the system self-memory dynamic equation of the NGM(1,1,k) self-memory coupling prediction model:

$$\frac{dx}{dt} = F(x,\lambda,t), \quad (17)$$

where $x$ is a variable, $\lambda$ is a parameter, $t$ is the time interval series, and $F(x,\lambda,t)$ is the dynamic kernel. Introduce a memory function $\beta(t)$ and define an inner product in the Hilbert space:

$$(f,g) = \int_{\xi=0}^{\xi_0} f(\xi) g(\xi) d\xi \quad (f,g \in L^2). \quad (18)$$

3.2 Coupling Modeling Process of NGM(1,1,k) Model and Self-Memory Principle. Then, the step-by-step modeling procedure of a novel NGM(1,1,k) prediction model coupled with self-memory principle is described as follows.

Step 1 (deducing the difference-integral equation). Let one time set $T = [t_{-p},t_{-p+1},\ldots,t_{-1},t_0,t]$, where $t_{-p},t_{-p+1},\ldots,t_{-1},t_0$ is historical observation time, $t_0$ is predicted initial time, $t$ is coming prediction time, the retrospective order of the equation is $p$, and time sampling interval is $\Delta t$.

Apply the above inner product operation into (17) and suppose that variables $x$, $\beta$ are continuous, differentiable, and integrable; the analytic formula of (17) is therefore obtained as

$$\int_{t_{-p}}^{t} \beta(t) \frac{dx}{dt} dt = \int_{t_{-p}}^{t} \beta(t) F(x,\lambda,t) dt; \quad (19)$$

that is,

$$\int_{t_{-p}}^{t_{-p+1}} \beta(t) \frac{dx}{dt} dt + \int_{t_{-p+1}}^{t_{-p+2}} \beta(t) \frac{dx}{dt} dt + \cdots + \int_{t_0}^{t} \beta(t) \frac{dx}{dt} dt = \int_{t_{-p}}^{t} \beta(t) F(x,\lambda,t) dt. \quad (20)$$

For every integral term in the left-hand side of (20), after integration by parts, applying the median theorem and performing algebra operation, a difference-integral equation is deduced as

$$\beta_i x_i - \beta_{-p} x_{-p} - \sum_{i=-p}^{0} x^m_i (\beta_{i+1} - \beta_i) - \int_{t_{-p}}^{t} \beta(t) F(x,\lambda,t) dt = 0, \quad (21)$$

where $\beta_i \equiv \beta(t), x_i \equiv x(t), \beta_i \equiv \beta(t_i), x_i \equiv x(t_i), i = -p, -p + 1,\ldots,0$, and midvalue $x^m_i \equiv x(t_m), t_i < t_m < t_{i+1}$.

Step 2 (discretizing the self-memory prediction equation). Let $x_{-p} \equiv x^{m}_{-p-1}$ and let $\beta_{-p-1} \equiv 0$; then (21) can be converted into

$$x_i = \frac{1}{\beta_i} \sum_{i=-p}^{0} x^m_i (\beta_{i+1} - \beta_i) + \frac{1}{\beta_i} \int_{t_{-p}}^{t} \beta(t) F(x,\lambda,t) dt = S_1 + S_2, \quad (22)$$

which is called the self-memory equation with the retrospective order $p$. As the first term $S_1$ in (22) denotes the relative contributions of historical data at $p + 1$ times to the value of variable $x_i$, it is defined as the self-memory term. The second term $S_2$ is the total contribution of the function $F(x,\lambda,t)$ in the retrospective time interval $[t_{-p},t_0]$, and it is defined as the exogenous effect term. Equation (22) emphasizes serial correlation of the system by itself, that is, the self-memory characteristic of the system. Therefore, it is the self-memory prediction equation of the system.

If integral operation is substituted by summation and differential is transformed into difference in (25), then the midvalue $x^m_i$ is replaced simply by two values of different times; namely,

$$x^m_i = \frac{1}{2} (x_{i+1} + x_i) \equiv y_i. \quad (23)$$
By taking equidistance time interval $\Delta t_i = t_{i+1} - t_i = 1$ and merging $\beta_i$ and $\beta_i$ together the self-memory equation of discrete form is shown as follows:

$$x_i = \sum_{i=-p}^{0} \alpha_i y_i + \sum_{i=-p}^{0} \theta_i F(x, \lambda, i), \quad (24)$$

where $\alpha_i = (\beta_{i+1} + \beta_i)/\beta_i$ and $\theta_i = \beta_i/\beta_i$, and $\alpha_i$ and $\theta_i$ are called memory coefficients, and $F(x, \lambda, t)$ is determined by the dynamic kernel $-AF^1_1 + bt + c$ of NGM(1,1,$k$) model.

**Step 3** (solving the self-memory prediction model). Assume that there are $L$ items of historical data; the memory coefficients $\alpha_i$ and $\theta_i$ can be estimated by the least square method. Let

$$X_{Lx} = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1L} \end{bmatrix},$$

$$Y_{Lx(p+1)} = \begin{bmatrix} y_{-p-1,1} & y_{-p-1,2} & \cdots & y_{-p-1,L} \\ y_{-p-2,1} & y_{-p-2,2} & \cdots & y_{-p-2,L} \\ \vdots & \vdots & \ddots & \vdots \\ y_{-p,L,1} & y_{-p,L,2} & \cdots & y_{-p,L,L} \end{bmatrix},$$

$$A_{(p+1)x1} = \begin{bmatrix} \alpha_{-p-1} \\ \vdots \\ \alpha_{-1} \end{bmatrix},$$

$$\Gamma_{Lx(p+1)} = \begin{bmatrix} F(x, \lambda, -p) & F(x, \lambda, -p+1) & \cdots & F(x, \lambda, 0) \\ F(x, \lambda, -p-1) & F(x, \lambda, -p) & \cdots & F(x, \lambda, 0) \\ \vdots & \vdots & \ddots & \vdots \\ F(x, \lambda, -p) & F(x, \lambda, -p+1) & \cdots & F(x, \lambda, 0) \end{bmatrix},$$

$$\Theta_{(p+1)x1} = \begin{bmatrix} \theta_{-p} \\ \vdots \\ \theta_{0} \end{bmatrix},$$

then (24) can be expressed as matrix form as follows:

$$X_i = YA + \Gamma \Theta.$$  \hspace{1cm} (26)

Let $Z = [Y, \Gamma]$ and $W = [A]$, then (26) turns into

$$X_i = ZW.$$  \hspace{1cm} (27)

**Step 4** (modeling simulation and prediction accuracy check). The absolute percentage error (abbreviated as APE) at time $t$ is denoted by

$$\text{APE}_t = \frac{|\hat{f}_t - f_t|}{f_t} \times 100\%,$$ \hspace{1cm} (30)

and the mean absolute percentage error (abbreviated as MAPE) at all times is defined as

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^{n} \text{APE}_t.$$ \hspace{1cm} (31)

Accordingly, the comparison analysis between actual values and simulative values derived from each prediction model can be analyzed using APE, and MAPE values.

At the same time, the established NGM(1,1,$k$) self-memory coupling prediction model must pass the simulation accuracy check before performing extrapolation and prediction. We usually check their accuracy by methods such as the "posterior variance ratio" and "small error probability" according to Table 1.

*Table 1: Test list of posterior variance ratio and small error probability.*

<table>
<thead>
<tr>
<th>Modeling accuracy class</th>
<th>Posterior variance ratio $C$</th>
<th>Small error probability $p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st level (superior)</td>
<td>$\leq 0.35$</td>
<td>$\geq 0.95$</td>
</tr>
<tr>
<td>2nd level (qualified)</td>
<td>$0.35 - 0.50$</td>
<td>$0.80 - 0.95$</td>
</tr>
<tr>
<td>3rd level (marginal)</td>
<td>$0.50 - 0.65$</td>
<td>$0.70 - 0.80$</td>
</tr>
<tr>
<td>4th level (disqualified)</td>
<td>$\geq 0.65$</td>
<td>$\leq 0.70$</td>
</tr>
</tbody>
</table>

$\hat{f}_t$ of the first-order accumulated generation sequence in NGM(1,1,$k$) self-memory coupling model, their inverse accumulated values $f_t$ can be obtained as follows:

$$\hat{f}_t = f_{t-1} - f_t,$$ \hspace{1cm} (29)

where $t = 1, 2, \ldots, n$ and $f_0 = 0$.

When the memory coefficient matrix $W$ is obtained, the simulating and predicting of original data sequence $f_t$ can be carried out. For the simulating and predicted values $\hat{f}_t$, the simulating and predicting of original data sequence $f_t$ can be carried out.

3.3. Programming Procedure of Coupling Prediction Model. The calculation process is carried out as mentioned above with the help of Matlab software in order to save the computational effort. And the programming flowchart for NGM(1,1,$k$) self-memory coupling prediction model is shown in Figure 1.
4. Illustrative Examples for Energy Consumption Prediction

China is one of the major countries in energy consumption. According to China’s energy consumption structure, coal is the main primary energy and electricity is the main secondhand energy. Therefore, the illustrative examples of total energy consumption, coal consumption, and electricity consumption in China are adopted to verify the effectiveness and practicability of the proposed NGM(1, 1, k) self-memory coupling prediction model (abbreviated as NGM + self-memory model). According to the China Energy Statistical Yearbook 2013, Figures 2, 3, and 4 show the annual consumption of total energy, coal, and electricity in China from 1999 to 2012, respectively. As it is seen in Figures 2–4, there has been a tremendous rise in energy consumption for each energy source, accompanied by irregular fluctuations due to the unstable changes occurring in the social and economic
The self-memorization equation can be established for coal consumption forecasting. The value of retrospective order is determined as \( p = 1 \) by trial calculation method under the principle of minimum error. After the differential equation is dealt with discretely, the memory coefficients can be solved by the least square method. Then the prediction equation of total energy consumption can be expressed as

\[
\frac{dF^1_t}{dt} = -0.0199F^1_t + 22066.6569t + 90668.0270. \tag{32}
\]

If the right-side terms of (32) are regarded as the dynamic kernel \( F(x, \lambda, t) \), then \( \frac{dF^1_t}{dt} = F(x, \lambda, t) \). The self-memorization equation can be established for total energy consumption forecasting. The value of retrospective order is determined as \( p = 1 \) by trial calculation method under the principle of minimum error. After the differential equation is

\[
\frac{dF^1_t}{dt} = -0.0123F^1_t + 21890.3477t + 86775.2490. \tag{33}
\]

If the right-side terms of (33) are regarded as the dynamic kernel \( F(x, \lambda, t) \), then \( \frac{dF^1_t}{dt} = F(x, \lambda, t) \). The self-memorization equation can be established for coal consumption forecasting. The value of retrospective order is

\[
\frac{dF^1_t}{dt} = -0.0161F^1_t + 20675.7221t + 81130.3477. \tag{34}
\]

In the statistical models, Markov chain can be used to explain the stochastic fluctuation phenomenon in which state transfer probability matrix is the basis of Markov prediction model. In general, we cannot determine the typical distribution of random variables and can only use the frequency instead of the probability. Probability theory points out that this replacement is meaningful only on the premise of large sampled values. On the contrary, the grey prediction models just possess an apparent superiority over a limited amount of statistical data without knowing their statistical distribution. Thereinto, the NGM(1, 1, \( k \)) model is especially appropriate for the approximate nonhomogeneous exponential law time series with stochastic fluctuation.

Following the coupling modeling process as mentioned above, the NGM(1, 1, \( k \)) self-memory coupling prediction models for total energy, coal, and electricity consumption are established to model and predict the consumption amounts for the upcoming three years, respectively. At the same time, these novel models are compared with the traditional GM(1, 1) model (abbreviated as GM model), the GM(1, 1) model with three-point moving average (abbreviated as GM + three-point model) [32], and the traditional NGM(1, 1, \( k \)) model (abbreviated as NGM model) to perform the error analysis. APE\(_t\) and MAPE are used to compare the actual values with simulative values to evaluate the predictive performance of novel NGM + self-memory model over other popular grey models.

4.1. Total Energy Consumption Forecasting in China. Based on statistical data from 1999 to 2012, the differential equation of NGM model is formulated as follows:

\[
\frac{dF^1_t}{dt} = -0.0199F^1_t + 22066.6569t + 90668.0270. \tag{32}
\]

The actual values and the simulative results for total energy consumption from 1999 to 2012 obtained by four different grey models are presented in Table 2, respectively. It is shown that the simulative MAPE of NGM model is lower than GM model and GM + three-point model. Moreover, the self-memory principle significantly further improves the prediction accuracy of NGM model. Consequently, the NGM + self-memory model yields the lowest MAPE compared with the other popular grey models.

The actual values and the simulative results for total energy consumption from 1999 to 2012 obtained by four different grey models are also presented in Figure 5. As can be seen from Table 2 and Figure 5, the NGM + self-memory model can better catch the development tendency of total energy consumption with the characteristics of nonhomogeneous exponential law. And the self-memory principle possesses an apparent advantage over other grey models when dealing with the stochastic fluctuation phenomenon.

4.2. Coal Consumption Forecasting in China. Based on statistical data from 1999 to 2012, the differential equation of NGM model is formulated as follows:

\[
\frac{dF^1_t}{dt} = -0.0123F^1_t + 21890.3477t + 86775.2490. \tag{34}
\]

The actual values and the simulative results for total energy consumption from 1999 to 2012 obtained by four different grey models are also presented in Table 2, respectively. It is shown that the simulative MAPE of NGM model is lower than GM model and GM + three-point model. Moreover, the self-memory principle significantly further improves the prediction accuracy of NGM model. Consequently, the NGM + self-memory model yields the lowest MAPE compared with the other popular grey models.

The actual values and the simulative results for total energy consumption from 1999 to 2012 obtained by four different grey models are also presented in Figure 5. As can be seen from Table 2 and Figure 5, the NGM + self-memory model can better catch the development tendency of total energy consumption with the characteristics of nonhomogeneous exponential law. And the self-memory principle possesses an apparent advantage over other grey models when dealing with the stochastic fluctuation phenomenon.
Table 2: Comparison of four different models for total energy consumption in China (unit: 10,000 tce).

<table>
<thead>
<tr>
<th>Years</th>
<th>Actual values</th>
<th>GM model Simulative values</th>
<th>GM + three-point model Simulative values</th>
<th>NGM model Simulative values</th>
<th>NGM + self-memory model Simulative values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>APE ( t ) (%)</td>
<td>APE ( t ) (%)</td>
<td>APE ( t ) (%)</td>
<td>APE ( t ) (%)</td>
</tr>
<tr>
<td>1999</td>
<td>140568.82</td>
<td>0.00</td>
<td>140578.39</td>
<td>3.21</td>
<td>140568.82</td>
</tr>
<tr>
<td>2000</td>
<td>145530.86</td>
<td>5.89</td>
<td>153607.37</td>
<td>5.55</td>
<td>121587.39</td>
</tr>
<tr>
<td>2001</td>
<td>150405.80</td>
<td>10.47</td>
<td>166470.28</td>
<td>10.68</td>
<td>142213.09</td>
</tr>
<tr>
<td>2002</td>
<td>159430.99</td>
<td>12.37</td>
<td>179496.73</td>
<td>12.59</td>
<td>162614.72</td>
</tr>
<tr>
<td>2003</td>
<td>183791.82</td>
<td>5.11</td>
<td>193176.71</td>
<td>5.31</td>
<td>182794.71</td>
</tr>
<tr>
<td>2004</td>
<td>213455.99</td>
<td>2.42</td>
<td>208687.18</td>
<td>2.23</td>
<td>202755.47</td>
</tr>
<tr>
<td>2005</td>
<td>235996.65</td>
<td>4.83</td>
<td>224591.93</td>
<td>4.65</td>
<td>222499.39</td>
</tr>
<tr>
<td>2006</td>
<td>258676.30</td>
<td>6.38</td>
<td>242591.93</td>
<td>6.45</td>
<td>242029.78</td>
</tr>
<tr>
<td>2007</td>
<td>280507.94</td>
<td>6.91</td>
<td>261116.02</td>
<td>6.83</td>
<td>261346.08</td>
</tr>
<tr>
<td>2008</td>
<td>291448.29</td>
<td>3.40</td>
<td>281548.49</td>
<td>3.21</td>
<td>280453.49</td>
</tr>
<tr>
<td>2009</td>
<td>306647.15</td>
<td>1.00</td>
<td>303579.81</td>
<td>0.81</td>
<td>299353.33</td>
</tr>
<tr>
<td>2010</td>
<td>324939.15</td>
<td>0.74</td>
<td>327335.10</td>
<td>0.93</td>
<td>318047.84</td>
</tr>
<tr>
<td>2011</td>
<td>348001.66</td>
<td>5.21</td>
<td>352949.25</td>
<td>5.11</td>
<td>343617.35</td>
</tr>
<tr>
<td>2012</td>
<td>361732.01</td>
<td>5.51</td>
<td>350567.72</td>
<td>5.21</td>
<td>348298.80</td>
</tr>
</tbody>
</table>

MAPE (%) 4.95

There are no simulative values from year 1999 to 2000 owing to the retrospective order \( p = 1 \).

Table 3: Comparison of four different models for coal consumption in China (unit: 10,000 ton).

<table>
<thead>
<tr>
<th>Years</th>
<th>Actual values</th>
<th>GM model Simulative values</th>
<th>GM + three-point model Simulative values</th>
<th>NGM model Simulative values</th>
<th>NGM + self-memory model Simulative values</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>APE ( t ) (%)</td>
<td>APE ( t ) (%)</td>
<td>APE ( t ) (%)</td>
<td>APE ( t ) (%)</td>
</tr>
<tr>
<td>1999</td>
<td>139336.46</td>
<td>0.00</td>
<td>142733.77</td>
<td>2.44</td>
<td>139336.46</td>
</tr>
<tr>
<td>2000</td>
<td>141091.70</td>
<td>5.98</td>
<td>150048.17</td>
<td>6.35</td>
<td>137190.05</td>
</tr>
<tr>
<td>2001</td>
<td>144528.11</td>
<td>11.59</td>
<td>161587.50</td>
<td>11.80</td>
<td>137509.12</td>
</tr>
<tr>
<td>2002</td>
<td>152282.66</td>
<td>14.23</td>
<td>174286.49</td>
<td>14.45</td>
<td>157579.06</td>
</tr>
<tr>
<td>2003</td>
<td>180587.04</td>
<td>4.90</td>
<td>187983.48</td>
<td>4.10</td>
<td>177402.94</td>
</tr>
<tr>
<td>2004</td>
<td>207561.29</td>
<td>6.74</td>
<td>202756.91</td>
<td>6.23</td>
<td>196983.77</td>
</tr>
<tr>
<td>2005</td>
<td>231851.07</td>
<td>6.76</td>
<td>218691.36</td>
<td>5.68</td>
<td>216324.53</td>
</tr>
<tr>
<td>2006</td>
<td>255065.45</td>
<td>7.70</td>
<td>235878.09</td>
<td>7.52</td>
<td>235428.16</td>
</tr>
<tr>
<td>2007</td>
<td>272745.88</td>
<td>6.90</td>
<td>254415.50</td>
<td>6.72</td>
<td>254297.57</td>
</tr>
<tr>
<td>2008</td>
<td>280105.92</td>
<td>9.26</td>
<td>274409.74</td>
<td>7.26</td>
<td>272935.63</td>
</tr>
<tr>
<td>2009</td>
<td>295833.08</td>
<td>0.14</td>
<td>295975.32</td>
<td>0.05</td>
<td>291345.17</td>
</tr>
<tr>
<td>2010</td>
<td>312236.50</td>
<td>2.05</td>
<td>319235.71</td>
<td>2.24</td>
<td>309529.01</td>
</tr>
<tr>
<td>2011</td>
<td>342950.24</td>
<td>0.21</td>
<td>344324.11</td>
<td>0.40</td>
<td>327489.91</td>
</tr>
<tr>
<td>2012</td>
<td>352647.07</td>
<td>5.11</td>
<td>361673.81</td>
<td>5.26</td>
<td>345230.59</td>
</tr>
</tbody>
</table>

MAPE (%) 4.95

There are no simulative values from year 1999 to 2001 owing to the retrospective order \( p = 2 \).

determined as \( p = 2 \) by trial calculation method under the principle of minimum error. After the differential equation is dealt with discretely, the memory coefficients can be solved by the least square method. Then the prediction equation of coal consumption can be expressed as

\[
x_t = \sum_{i=-3}^{-1} \alpha_i y_{t+i} + \sum_{i=-2}^{0} \theta_i F(x, \lambda, i),
\]

where \( \alpha_{-3} = 0.0717, \alpha_{-2} = -0.4124, \alpha_{-1} = 1.2960, \theta_{-2} = -49.6300, \theta_{-1} = 102.7246, \) and \( \theta_0 = -51.2947 \).

Through calculation, the actual and simulative values of four different models are presented in Table 3, respectively. It is shown that the simulative MAPE of NGM model is lower than GM model and GM + three-point model. Moreover, the self-memory principle significantly further improves the prediction accuracy of NGM model. Consequently, the NGM + self-memory model yields the lowest MAPE compared with the other popular grey models.

The actual values and the simulative results for coal consumption from 1999 to 2012 obtained by four different grey models are also presented in Figure 6. As can be
seen from Table 3 and Figure 6, the NGM + self-memory model can better catch the development tendency of coal consumption with the characteristics of nonhomogeneous exponential law. And the self-memory principle possesses an apparent advantage over other grey models when dealing with the stochastic fluctuation phenomenon.

4.3. Electricity Consumption Forecasting in China. Based on statistical data from 1999 to 2012, the differential equation of NGM model is formulated as follows:

\[
\frac{dF^1_t}{dt} = 0.0599F^1_t + 1390.9613t + 8703.5942. \quad (36)
\]

If the right-side terms of (36) are regarded as the dynamic kernel \( F(x, \lambda, t) \), then \( dF^1_t/dt = F(x, \lambda, t) \). The self-memorization equation can be established for electricity consumption forecasting. The value of retrospective order is determined as \( p = 2 \) by trial calculation method under the principle of minimum error. After the differential equation is dealt with discretely, the memory coefficients can be solved by the least square method. Then the prediction equation of electricity consumption can be expressed as

\[
x_t = \sum_{i=3}^{-1} \alpha_i y_{t-i} + \sum_{i=-2}^{0} \theta_i F(x, \lambda, i), \quad (37)
\]

where \( \alpha_{-3} = 0.2790, \alpha_{-2} = -0.5414, \alpha_{-1} = 1.1746, \theta_{-2} = 14.9812, \theta_{-1} = -21.8940, \) and \( \theta_0 = 9.6612 \).

Through calculation, the actual and simulative values of four different models are presented in Table 4, respectively. It is shown that the simulative MAPE of NGM model is lower than GM model and GM + three-point model. Moreover, the self-memory principle significantly further improves the prediction accuracy of NGM model. Consequently, the NGM model + self-memory model yield the lowest MAPE compared with the other popular grey models.

The actual values and the simulative results for electricity consumption from 1999 to 2012 obtained by four different grey models are also presented in Figure 7. As can be seen from Table 4 and Figure 7, the NGM + self-memory model can better catch the development tendency of electricity consumption with the characteristics of nonhomogeneous exponential law. And the self-memory principle possesses an apparent advantage over other grey models when dealing with the stochastic fluctuation phenomenon.

The MAPE of different prediction models for total energy, coal, and electricity consumption are all presented in Table 5, respectively. It is obvious that the prediction error of NGM(1,1,k) model is lower than GM(1,1) model and GM(1,1) model with three-point, and the self-memory principle further reduces the prediction error of NGM(1,1,k) model remarkably.

The “posterior error” method is used to perform the simulation accuracy check, with the results showing that the posterior variance ratio and small error probability of three NGM + self-memory prediction models are all up to the first precision level on the basis of Table 1. The simulation accuracies of the simulative values to the actual values are 97.70%, 98.27%, and 98.46%, respectively. Because three NGM + self-memory prediction models have all passed through the simulation accuracy check, they all could be used to carry out extrapolation and prediction, which could reasonably reflect the growth trend of the future energy consumption of China.

The total energy, coal, and electricity consumption in China from 2013 to 2015 is predicted according to the NGM + self-memory prediction models as mentioned above. The results show that, compared with the data of 2012, the total energy, coal energy, and electricity energy consumption will increase at the annual average rate of 3.44%, 2.53%, and 8.08%, respectively, in the next three years. In the energy consumption structure, the proportion of coal energy and electricity energy consumption will increase gradually, as shown in Table 6. This means that along with the rapid development of China’s economy, the need for energy is increasing.
Table 4: Comparison of four different models for electricity consumption in China (unit: 100 million kW-h).

<table>
<thead>
<tr>
<th>Years</th>
<th>Actual values</th>
<th>GM model</th>
<th>GM + three-point model</th>
<th>NGM model</th>
<th>NGM + self-memory model</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Simulative values</td>
<td>APE_{t} (%)</td>
<td>Simulative values</td>
<td>APE_{t} (%)</td>
</tr>
<tr>
<td>1999</td>
<td>12304.71</td>
<td>12304.71</td>
<td>0.00</td>
<td>13019.57</td>
<td>5.81</td>
</tr>
<tr>
<td>2000</td>
<td>13472.38</td>
<td>14449.29</td>
<td>7.25</td>
<td>14275.93</td>
<td>5.96</td>
</tr>
<tr>
<td>2001</td>
<td>14723.46</td>
<td>16073.79</td>
<td>9.17</td>
<td>16134.67</td>
<td>9.58</td>
</tr>
<tr>
<td>2002</td>
<td>16465.45</td>
<td>17880.94</td>
<td>8.60</td>
<td>17948.66</td>
<td>9.01</td>
</tr>
<tr>
<td>2003</td>
<td>19031.60</td>
<td>19891.26</td>
<td>4.52</td>
<td>19966.59</td>
<td>4.91</td>
</tr>
<tr>
<td>2004</td>
<td>21971.37</td>
<td>22127.59</td>
<td>0.71</td>
<td>22211.40</td>
<td>1.09</td>
</tr>
<tr>
<td>2005</td>
<td>24940.32</td>
<td>24615.35</td>
<td>1.31</td>
<td>24708.58</td>
<td>0.93</td>
</tr>
<tr>
<td>2006</td>
<td>28587.97</td>
<td>27382.80</td>
<td>4.22</td>
<td>27486.52</td>
<td>3.85</td>
</tr>
<tr>
<td>2007</td>
<td>32711.81</td>
<td>30461.39</td>
<td>6.68</td>
<td>30576.77</td>
<td>6.53</td>
</tr>
<tr>
<td>2008</td>
<td>34541.35</td>
<td>33886.11</td>
<td>1.90</td>
<td>34014.45</td>
<td>1.53</td>
</tr>
<tr>
<td>2009</td>
<td>37032.14</td>
<td>37695.85</td>
<td>1.79</td>
<td>37838.63</td>
<td>2.18</td>
</tr>
<tr>
<td>2010</td>
<td>41934.49</td>
<td>41933.92</td>
<td>0.00</td>
<td>42092.75</td>
<td>0.38</td>
</tr>
<tr>
<td>2011</td>
<td>47000.88</td>
<td>46648.47</td>
<td>0.75</td>
<td>46825.15</td>
<td>0.37</td>
</tr>
<tr>
<td>2012</td>
<td>49762.64</td>
<td>51893.06</td>
<td>4.28</td>
<td>50144.86</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>MAPE (%)</td>
<td>3.83</td>
<td></td>
<td>3.78</td>
<td></td>
</tr>
</tbody>
</table>

There are no simulative values from year 1999 to 2001 owing to the retrospective order \( p = 2 \).

Table 5: MAPE of four different models for energy consumption.

<table>
<thead>
<tr>
<th>Energy sources</th>
<th>GM model</th>
<th>GM + three-point model</th>
<th>NGM model</th>
<th>NGM + self-memory model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total energy</td>
<td>4.95%</td>
<td>4.74%</td>
<td>4.42%</td>
<td>2.30%</td>
</tr>
<tr>
<td>Coal</td>
<td>5.13%</td>
<td>4.93%</td>
<td>4.66%</td>
<td>1.73%</td>
</tr>
<tr>
<td>Electricity</td>
<td>3.83%</td>
<td>3.78%</td>
<td>3.66%</td>
<td>1.54%</td>
</tr>
</tbody>
</table>

Table 6: The predictive values of the NGM + self-memory model and the growth rates.

<table>
<thead>
<tr>
<th>Years</th>
<th>Total Energy (10,000 tce)</th>
<th>Coal (10,000 ton)</th>
<th>Electricity (100 million kW-h)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Quantity</td>
<td>Growth rate</td>
<td>Quantity</td>
</tr>
<tr>
<td>2013</td>
<td>373905.27</td>
<td>3.37%</td>
<td>357900.57</td>
</tr>
<tr>
<td>2014</td>
<td>385239.04</td>
<td>3.03%</td>
<td>364615.31</td>
</tr>
<tr>
<td>2015</td>
<td>400317.60</td>
<td>3.91%</td>
<td>380018.09</td>
</tr>
</tbody>
</table>

continuously. Moreover, with the increasing growth of coal energy consumption, the proportion of electricity energy consumption is also growing rapidly with a growth rate far higher than coal energy.

5. Conclusion

In this study, aiming at the approximate nonhomogeneous exponential data sequence with stochastic fluctuation emerging in the energy consumption, the predictive performance of the traditional NGM(1, 1, k) model has been markedly improved by using the self-memory principle of dynamic system. The illustrative examples show the superiority of NGM(1, 1, k) self-memory coupling prediction model over other popular grey models. This superiority results from the organic integration of the grey NGM(1, 1, k) model and the self-memory principle. The coupling prediction model can take full advantage of the systematic multitime historical data and tightly catch the stochastic fluctuation tendency. The future total energy, coal, and electricity consumption of China has been effectively predicted using the proposed NGM(1, 1, k) self-memory coupling prediction model. It is worth popularizing and applying in other relevant energy consumption predictions.

These results may guide China’s institutions related to energy production in implementing the energy planning studies and framing the suitable energy strategies. In the future, China’s energy consumption structure will still mainly depend on coal energy, together with an obvious upward trend of the consumption share of electricity energy. This conforms to the future sustainable multiple clean energy consumption strategy that is based mainly on coal. Therefore, the proposed coupling prediction model could provide a reference for other countries (especially developing countries) to establish and adjust the energy consumption structure and coordinate the relationship among energy, economy, and environment.
This study also supports that there is still room for improving the performance of the existing prediction methods for energy consumption. As a future work, grey prediction models based on various kinds of optimization techniques will be integrated with the self-memory principle in order to further improve the prediction accuracy and stability in energy consumption. Meanwhile, we have not found an ideal algorithm for the optimal retrospective order, only by means of the trial calculation method under the principle of minimum error. Therefore, whether there are certain intelligent optimization algorithms, such as nonlinear programming and particle swarm optimization, that could be introduced into the coupling model needs further exploration.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors are grateful to the editor and anonymous reviewers for their valuable comments and suggestions which improved the paper. This work was supported by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme (no. FP7-PIIF-GA-2013-629051); the National Natural Science Foundation of China (no. 7117113, no. 71363046); the Humanistic and Social Science Research Project of Ministry of Education of China (no. 10YJA790174, no. 13YJC790198); the Funding of Jiangsu Innovation Program for Graduate Education (no. CXZZ13_0184); the Fundamental Research Funds for the Central Universities; the Funding for Talent Introduction in NUAA (no. 1009-YAH14003); and the Nantong Science and Technology Planning Project (no. HS2013026).

References


