Hindawi Publishing Corporation
Mathematical Problems in Engineering
Volume 2015, Article ID 686501, 7 pages
http://dx.doi.org/10.1155/2015/686501

Research Article

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

Lifeng Wu,[1] Sifeng Liu,[1] Haijun Chen,[2] and Na Zhang[1]

[1] *College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*
[2] *Department 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

Academic Editor: Massimo Scalia

Accurate prediction of the future energy needs is crucial for energy management. This work presents a novel grey forecasting model that integrates the principle of new information priority into accumulated generation. This grey model can better reflect the priority of the new information theoretically. The results of two practical examples demonstrate that this grey model provides very remarkable short-term predication performance compared with traditional grey forecasting model for limited data set forecasting. It is applied to Chinese gas consumption forecasting to show its superiority and applicability.

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]. Ünler 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 $n-1$th component, respectively, and the nth component is more prior than the $n-1$th 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

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

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

$$\begin{bmatrix} \hat{\alpha} \\ \hat{\beta} \end{bmatrix} = \left(B^T B\right)^{-1} B^T Y, \tag{2}$$

where

$$Y = \begin{bmatrix} x(2) \\ x(3) \\ \vdots \\ x(n-1) \\ x(n) \end{bmatrix}, \qquad B = \begin{bmatrix} x(1) & 1 \\ x(2) & 1 \\ \vdots & \vdots \\ x(n-2) & 1 \\ x(n-1) & 1 \end{bmatrix}. \tag{3}$$

Theorem 1. *Assume that $\widehat{\alpha}$, $\widehat{\beta}$ are the same as* (1). *Set $\widehat{x}(1) = x(1)$; then recursive function is given by*

$$\widehat{x}(k+1) = \widehat{\alpha}^k \left[x(1) - \frac{\widehat{\beta}}{1-\widehat{\alpha}} \right] + \frac{\widehat{\beta}}{1-\widehat{\alpha}}, \qquad k = 1, 2, \ldots, n-1. \tag{4}$$

Proof. Substitute $\widehat{\alpha}$, $\widehat{\beta}$ into (1); then

$$\begin{aligned} \widehat{x}(k+1) &= \widehat{\alpha}\widehat{x}(k+1) + \widehat{\beta} = \widehat{\alpha}\left(\widehat{x}(k) + \widehat{\beta}\right) + \widehat{\beta} \\ &= \cdots = \widehat{\alpha}^k \widehat{x}(1) + \left(\widehat{\alpha}^{k-1} + \widehat{\alpha}^{k-2} + \cdots + \widehat{\alpha} + 1\right)\widehat{\beta}. \end{aligned} \tag{5}$$

Let $\widehat{x}(1) = x(1)$; then

$$\widehat{x}(k+1) = \widehat{\alpha}^k x(1) + \frac{1-\widehat{\alpha}^k}{1-\widehat{\alpha}}\widehat{\beta}. \tag{6}$$

We also obtain a nonhomogenous exponential model as follows:

$$\widehat{x}(k) = \widehat{\alpha}^{k-1} \left[x(1) - \frac{\widehat{\beta}}{1-\widehat{\alpha}} \right] + \frac{\widehat{\beta}}{1-\widehat{\alpha}}, \qquad k = 1, 2, \ldots, n-1. \tag{7}$$

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:

$$\text{Minimize}: \sum_{k=2}^{n} w_k \left[x(k) - \alpha x(k-1) - \beta \right]^2, \qquad w_k = \frac{k-1}{n-1}. \tag{8}$$

Then we obtain a nonhomogenous exponential model as follows:

$$\widehat{x}(k) = \widehat{\alpha}^{k-1} \left[x(1) - \frac{\widehat{\beta}}{1-\widehat{\alpha}} \right] + \frac{\widehat{\beta}}{1-\widehat{\alpha}}, \qquad k = 1, 2, \ldots, n-1. \tag{9}$$

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

$$X = \begin{bmatrix} x(2) & x(1) & 1 \\ x(3) & x(2) & 1 \\ \vdots & \vdots & \vdots \\ x(n-1) & x(n-2) & 1 \\ x(n) & x(n-1) & 1 \end{bmatrix}. \tag{10}$$

The data with the principle of new information priority is

$$\begin{bmatrix} 1 & 1 & \cdots & 1 & 1 \\ 0 & 1 & \cdots & 1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 1 \\ 0 & 0 & \cdots & 0 & 1 \end{bmatrix} \begin{bmatrix} x(2) & x(1) & 1 \\ x(3) & x(2) & 1 \\ \vdots & \vdots & \vdots \\ x(n-1) & x(n-2) & 1 \\ x(n) & x(n-1) & 1 \end{bmatrix}$$

$$= AX = \begin{bmatrix} \sum_{k=2}^{n} x(k) & \sum_{k=2}^{n} x(k-1) & n-1 \\ \sum_{k=3}^{n} x(k) & \sum_{k=3}^{n} x(k-1) & n-2 \\ \vdots & \vdots & \vdots \\ \sum_{k=n-1}^{n} x(k) & \sum_{k=n-1}^{n} x(k-1) & 2 \\ x(n) & x(n-1) & 1 \end{bmatrix}, \tag{11}$$

which is called the treatment of new information priority. The equation

$$\sum_{i=k}^{n} x(i) = \alpha \sum_{i=k}^{n} x(i-1) + (n-k+1)\beta, \qquad k = 2, 3, \ldots, n, \tag{12}$$

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

Theorem 2. *Assume that $X = \{x(1), x(2), \ldots, x(n)\}$ is certain. Actually, if the error is μ, then*

$$x(k) = \alpha x(k-1) + \beta + \mu_k, \quad k = 2, 3, \ldots, n. \tag{13}$$

Via the treatment of new information priority, the mean of the error is

$$E(A\mu) = \mu_n + \frac{(n-2)\mu_{n-1}}{n-1} + \cdots + \frac{2\mu_3}{n-1} + \frac{\mu_2}{n-1}, \tag{14}$$

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

$$\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} = \left(B^T B\right)^{-1} B^T Y, \tag{15}$$

where

$$Y = \begin{bmatrix} \sum_{k=2}^{n} x(k) \\ \sum_{k=3}^{n} x(k) \\ \vdots \\ \sum_{k=n-1}^{n} x(k) \\ x(n) \end{bmatrix}, \quad B = \begin{bmatrix} \sum_{k=2}^{n} x(k-1) & n-1 \\ \sum_{k=3}^{n} x(k-1) & n-2 \\ \vdots & \vdots \\ \sum_{k=n-1}^{n} x(k-1) & 2 \\ x(n-1) & 1 \end{bmatrix}. \tag{16}$$

The variable of

$$\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} \tag{17}$$

is $\operatorname{Var}\left[\begin{smallmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{smallmatrix}\right] = (B^T A^T AB)^{-1} B^T A^T E(A\mu\mu^T A^T) AB (B^T A^T AB)^{-1}$, *where*

$$E\left(A\mu\mu^T A^T\right) = \begin{bmatrix} \left[\sum_{i=2}^{n}\mu_i\right]^2 & \mu_3\left[\sum_{i=2}^{n}\mu_i\right] & \cdots & \mu_n\left[\sum_{i=2}^{n}\mu_i\right] \\ \mu_3\left[\sum_{i=2}^{n}\mu_i\right] & \left[\sum_{i=3}^{n}\mu_i\right]^2 & \cdots & \mu_n\left[\sum_{i=3}^{n}\mu_i\right] \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n\left[\sum_{i=2}^{n}\mu_i\right] & \mu_n\left[\sum_{i=3}^{n}\mu_i\right] & \cdots & \mu_n^2 \end{bmatrix}. \tag{18}$$

Proof. Via the treatment of new information priority, (11) becomes

$$AY = AB\begin{bmatrix}\alpha\\ \beta\end{bmatrix} + A\mu. \tag{19}$$

Obviously,

$$E(A\mu) = \mu_n + \frac{(n-2)\mu_{n-1}}{n-1} + \cdots + \frac{2\mu_3}{n-1} + \frac{\mu_2}{n-1}. \tag{20}$$

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

$$\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} = \left(B^T A^T AB\right)^{-1} B^T A^T AY. \tag{21}$$

Substituting

$$AY = AB\begin{bmatrix}\alpha\\ \beta\end{bmatrix} + A\mu \tag{22}$$

into (21), we obtain

$$\begin{aligned}\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} &= \left(B^T A^T AB\right)^{-1} B^T A^T \left[AB\begin{bmatrix}\alpha\\ \beta\end{bmatrix} + A\mu\right] \\ &= \begin{bmatrix}\alpha\\ \beta\end{bmatrix} + \left(B^T A^T AB\right)^{-1} B^T A^T A\mu.\end{aligned} \tag{23}$$

Then

$$\begin{aligned}E\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} &= \begin{bmatrix}\alpha\\ \beta\end{bmatrix} + E\left(\left[B^T A^T AB\right]^{-1} B^T A^T A\mu\right) \\ &= \begin{bmatrix}\alpha\\ \beta\end{bmatrix} + \left[B^T A^T AB\right]^{-1} B^T A^T E(A\mu).\end{aligned} \tag{24}$$

Actually, by

$$E\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} \tag{25}$$

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

$$\begin{gathered}\lim_{n\to\infty} \frac{n-2}{n-1} - \frac{n-3}{n-1} = 0, \\ \cdots, \\ \lim_{n\to\infty} \frac{2}{n-1} - \frac{1}{n-1} = 0.\end{gathered} \tag{26}$$

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

We can obtain

$$\operatorname{Var}\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} = \begin{bmatrix}\operatorname{Var}(\widehat{\alpha}) & \operatorname{Cov}\left(\widehat{\beta},\widehat{\alpha}\right) \\ \operatorname{Cov}\left(\widehat{\alpha},\widehat{\beta}\right) & \operatorname{Var}\left(\widehat{\beta}\right)\end{bmatrix}. \tag{27}$$

By

$$\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} - \begin{bmatrix}\alpha\\ \beta\end{bmatrix} = \left(B^T A^T AB\right)^{-1} B^T A^T A\mu, \tag{28}$$

thus

$$\begin{aligned}\operatorname{Var}\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} &= \left[B^T A^T AB\right]^{-1} B^T A^T E \\ &\quad \times \left(A\mu\mu^T A^T\right) AB\left[B^T A^T AB\right]^{-1}.\end{aligned} \tag{29}$$

In most cases, we pay attention to the $\mathrm{Var}(\widehat{\alpha})$ and $\mathrm{Var}(\widehat{\beta})$ and overlook $\mathrm{Cov}(\widehat{\beta},\widehat{\alpha})$; thus we denote (27) as

$$\mathrm{Var}\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} = \left[B^T A^T AB\right]^{-1} E\left(A\mu\mu^T A^T\right). \tag{30}$$

So

$$E\left(A\mu\mu^T A^T\right) = \begin{bmatrix} \left[\sum_{i=2}^{n}\mu_i\right]^2 & \mu_3\left[\sum_{i=2}^{n}\mu_i\right] & \cdots & \mu_n\left[\sum_{i=2}^{n}\mu_i\right] \\ \mu_3\left[\sum_{i=2}^{n}\mu_i\right] & \left[\sum_{i=3}^{n}\mu_i\right]^2 & \cdots & \mu_n\left[\sum_{i=3}^{n}\mu_i\right] \\ \vdots & \vdots & \vdots & \vdots \\ \mu_n\left[\sum_{i=2}^{n}\mu_i\right] & \mu_n\left[\sum_{i=3}^{n}\mu_i\right] & \cdots & \mu_n^2 \end{bmatrix}. \tag{31}$$

□

Let $E(A\mu\mu^T A^T) = D$; then the components from D_{11} to D_{nn} all contain μ_n^2. The components from D_{11} to D_{n-1n-1} all contain μ_{n-1}^2. The components from D_{11} to D_{n-2n-2} all contain μ_{n-2}^2. The components from D_{11} to D_{n-1n-1} all contain μ_{n-1}^2. Analogizing successively, the components D_{11} and D_{22} both contain μ_2^2. Only D_{11} contains μ_1^2. So

$$\mathrm{Var}\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} \tag{32}$$

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 α and β according to Theorem 2 is as follow.

$$\begin{bmatrix}\widehat{\alpha}\\ \widehat{\beta}\end{bmatrix} = \left(B^T A^T AB\right)^{-1} B^T A^T AY. \tag{33}$$

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

$$\widehat{x}(k) = \widehat{\alpha}^{k-1}\left[x(1) - \frac{\widehat{\beta}}{1-\widehat{\alpha}}\right] + \frac{\widehat{\beta}}{1-\widehat{\alpha}}, \quad k = 1, 2, \ldots, n-1. \tag{34}$$

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

$$x(k) = c\widehat{\alpha}^k + b. \tag{35}$$

Use the ordinary least squares method to estimate the parameters

$$\begin{bmatrix}\widehat{c}\\ \widehat{b}\end{bmatrix} = \left(B^T B\right)^{-1} B^T Y, \tag{36}$$

where

$$Y = \begin{bmatrix} x(1) \\ x(2) \\ \vdots \\ x(n-1) \\ x(n) \end{bmatrix}, \qquad B = \begin{bmatrix} \widehat{\alpha} & 1 \\ \widehat{\alpha}^2 & 1 \\ \vdots & \vdots \\ \widehat{\alpha}^{n-1} & 1 \\ \widehat{\alpha}^n & 1 \end{bmatrix}; \tag{37}$$

thus, we obtain

$$\widehat{x}(k) = \widehat{c}\widehat{\alpha}^k + \widehat{b}, \quad k = 1, 2, \ldots, n, \tag{38}$$

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 $= 100\%(1/n)\sum_{k=1}^{n}|(x(k)-\widehat{x}(k))/x(k)|$) 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),

$$x(k) = 22.478 * 1.276^k + 144.995. \tag{39}$$

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):

$$x(k) = 11.315 * 1.168^k + 32.49. \tag{40}$$

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

Table 1: The fitted values (unit: 100 million cubic meters) and MAPE of five grey models.

Year	Actual value	Model (39)	ARGM(1, 1)	Reference [14]	WLSGM(1, 1)	DGM(1, 1)
1995	177.53143	173.677	177.53143	190.24323	177.53143	177.53143
1996	188.04992	181.593	187.961	183.65849	190.990	149.31
1997	176.13278	191.694	201.029	184.90875	207.706	169.38
1998	218.69985	204.583	217.403	193.99401	228.466	192.15
1999	215.3371	221.030	237.920	210.91426	254.251	217.98
2000	244.81169	242.015	263.628	235.66952	286.276	247.29
2001	274.55447	268.793	295.840	268.25976	326.051	280.53
2002	292.18069	302.961	336.202	308.68501	375.451	318.25
2003	339.45429	346.559	386.775	356.94525	436.806	361.03
2004	397.28586	402.191	450.143	413.04049	513.010	409.57
2005	486.93617	473.177	529.543	476.97073	607.654	464.63
2006	561.05120	563.756	629.031	548.73596	725.202	527.10
MAPE		3.092	8.657	4.406	18.173	6.96
2007	698.903	679.333	753.690	679.955	871.197	597.96
2008	807.857	826.811	909.888	774.429	1052.523	678.35
MAPE		2.573	10.234	10.75	27.469	15.236

Table 2: Fitted values and MAPE of different models.

Year	Actual value	Model (40)	GM(1, 1)	Reference [5]	WLSGM(1, 1)	DGM(1, 1)
1990	45.41	44.71	45.41	44.07	45.41	45.41
1991	47.56	47.93	42.18	46.40	47.14	42.26
1992	50.58	50.52	46.32	49.48	49.19	46.40
1993	51.85	53.55	50.85	53.16	51.62	50.94
1994	56.25	57.09	55.83	57.49	54.50	55.93
1995	58.63	61.22	61.30	60.83	57.91	61.41
1996	59.94	66.04	67.31	60.62	61.97	67.42
1997	75.43	71.68	73.90	77.07	66.77	74.02
1998	82.45	78.27	81.14	82.15	72.48	81.27
1999	93.40	85.96	89.08	90.44	79.24	89.23
2000	100.67	94.94	97.81	98.72	87.26	97.96
2001	103.63	105.43	107.39	107.7	96.77	107.55
2002	111.71	117.68	117.91	116.93	108.05	118.09
2003	128.19	131.99	129.45	123.54	121.42	129.65
2004	151.04	148.71	142.13	146.05	137.29	142.34
MAPE		3.7	4.82	2.6	5.88	4.76
2005	166.29	168.24	156.06	164.21	156.11	156.28
2006	185.83	191.05	171.34	179.51	178.42	171.58
MAPE		1.99	3.48	2.3	5.06	3.42

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):

$$x(k) = 122.2219 * 1.14736^{k} - 30.8291. \tag{41}$$

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

Year	Actual value	Model (41)
2012	1463.0	1420.5
2013	1616.1	1634.3
MAPE		2.0

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

Year	Model (41)
2014	1879.7
2015	2161.2
2016	2484.3
2017	2854.9
2018	3280.2

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. 71171113,71271226, 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

[1] L. Hernández, J. M. Baladrón, J. M. Aguiar et al., "Artificial neural network for short-term load forecasting in distribution systems," *Energies*, vol. 7, no. 3, pp. 1576–1598, 2014.

[2] C. Jian, G. Jue, and L. Hu, "Forecasting energy demand of China using bayesian combination model," *China Population Resources and Environment*, vol. 18, no. 4, pp. 50–55, 2008.

[3] L.-Q. Ji, "Forecasting petroleum consumption in China: comparison of three models," *Journal of the Energy Institute*, vol. 84, no. 1, pp. 34–37, 2011.

[4] M. Meng and D. Niu, "Annual electricity consumption analysis and forecasting of China based on few observations methods," *Energy Conversion and Management*, vol. 52, no. 2, pp. 953–957, 2011.

[5] M. Zhang, H. Mu, G. Li, and Y. Ning, "Forecasting the transport energy demand based on PLSR method in China," *Energy*, vol. 34, no. 9, pp. 1396–1400, 2009.

[6] A. Ünler, "Improvement of energy demand forecasts using swarm intelligence: the case of Turkey with projections to 2025," *Energy Policy*, vol. 36, no. 6, pp. 1937–1944, 2008.

[7] M. Meng, D. Niu, and W. Sun, "Forecasting monthly electric energy consumption using feature extraction," *Energies*, vol. 4, no. 10, pp. 1495–1507, 2011.

[8] V. Bianco, O. Manca, S. Nardini, and A. A. Minea, "Analysis and forecasting of nonresidential electricity consumption in Romania," *Applied Energy*, vol. 87, no. 11, pp. 3584–3590, 2010.

[9] G.-D. Li, S. Masuda, and M. Nagai, "The prediction model for electrical power system using an improved hybrid optimization model," *International Journal of Electrical Power & Energy Systems*, vol. 44, no. 1, pp. 981–987, 2013.

[10] W. Niu, J. Cheng, and G. Wang, "Applications of extension grey prediction model for power system forecasting," *Journal of Combinatorial Optimization*, vol. 26, no. 3, pp. 555–567, 2013.

[11] H. Pao and C. Tsai, "Modeling and forecasting the CO_2 emissions, energy consumption, and economic growth in Brazil," *Energy*, vol. 36, no. 5, pp. 2450–2458, 2011.

[12] B. Wang, X. J. Liang, H. Zhang, L. Wang, and Y. M. Wei, "Vulnerability of hydropower generation to climate change in China: results based on grey forecasting model," *Energy Policy*, vol. 65, pp. 701–707, 2014.

[13] S. H. Mao, M. Y. Gao, J. H. Wen, and X. W. Nie, "Generalized admissible region of class ratio for fractional order GM(1, 1) model," *The Journal of Grey System*, vol. 26, no. 3, pp. 55–68, 2014.

[14] G. Xu and W. Wang, "Forecasting China's natural gas consumption based on a combination model," *Journal of Natural Gas Chemistry*, vol. 19, no. 5, pp. 493–496, 2010.

[15] M. Forouzanfar, A. Doustmohammadi, M. B. Menhaj, and S. Hasanzadeh, "Modeling and estimation of the natural gas consumption for residential and commercial sectors in Iran," *Applied Energy*, vol. 87, no. 1, pp. 268–274, 2010.

[16] R. Gutiérrez, A. Nafidi, and R. G. Sánchez, "Forecasting total natural-gas consumption in Spain by using the stochastic Gompertz innovation diffusion model," *Applied Energy*, vol. 80, no. 2, pp. 115–124, 2005.

[17] M. Brabec, O. Konár, E. Pelikán, and M. Malý, "A nonlinear mixed effects model for the prediction of natural gas consumption by individual customers," *International Journal of Forecasting*, vol. 24, no. 4, pp. 659–678, 2008.

[18] J. Li, X. Dong, J. Shangguan, and M. Hook, "Forecasting the growth of China's natural gas consumption," *Energy*, vol. 36, no. 3, pp. 1380–1385, 2011.

[19] A. Azadeh, S. M. Asadzadeh, M. Saberi, V. Nadimi, A. Tajvidi, and M. Sheikalishahi, "A Neuro-fuzzy-stochastic frontier analysis approach for long-term natural gas consumption forecasting

and behavior analysis: the cases of Bahrain, Saudi Arabia, Syria, and UAE," *Applied Energy*, vol. 88, no. 11, pp. 3850–3859, 2011.

[20] K. Sabo, R. Scitovski, I. Vazler, and M. Zekić-Sušac, "Mathematical models of natural gas consumption," *Energy Conversion and Management*, vol. 52, no. 3, pp. 1721–1727, 2011.

[21] E. F. Sánchez-Úbeda and A. Berzosa, "Modeling and forecasting industrial end-use natural gas consumption," *Energy Economics*, vol. 29, no. 4, pp. 710–742, 2007.

[22] V. Bianco, F. Scarpa, and L. A. Tagliafico, "Scenario analysis of nonresidential natural gas consumption in Italy," *Applied Energy*, vol. 113, pp. 392–403, 2014.

[23] Z. Wadud, H. S. Dey, M. A. Kabir, and S. I. Khan, "Modeling and forecasting natural gas demand in Bangladesh," *Energy Policy*, vol. 39, no. 11, pp. 7372–7380, 2011.