

A novel energy consumption forecasting model combining an optimized DGM (1, 1) model with interval grey numbers

Jing Ye¹, Yaoguo Dang², Song Ding^{3*}, Yingjie Yang⁴

1 School of Management Science and Engineering, Nanjing University of Finance & Economics, Nanjing, Jiangsu, 210023, China

2 College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, 210016, China

3 School of Economics, Zhejiang University of Finance & Economics, Hangzhou, Zhejiang, 310018, China

4 Centre for Computational Intelligence, De Montfort University, Leicester, LE1 9BH, UK

Abstract: Since energy consumption (EC) is becoming an important issue for sustainable development in the world, it has a practical significance to predict EC effectively. However, there are two main uncertainty factors affecting the accuracy of a region's EC prediction. Firstly, with the ongoing rapid changes in society, the consumption amounts can be non-smooth or even fluctuating during a long time period, which makes it difficult to investigate the sequence's trend in order to forecast. Secondly, in a given region, it is difficult to express the consumption amount as a real number, as there are different development levels in the region, which would be more suitably described as interval numbers. Most traditional prediction models for energy consumption forecasting deal with long-term real numbers. It is seldom found to discover research that focuses specifically on uncertain EC data. To this end, a novel energy consumption forecasting model has been established by expressing ECs in a region as interval grey numbers combining with the optimized discrete grey model (DGM(1,1)) in Grey System Theory (GST). To prove the effectiveness of the method, per capita annual electricity consumption in southern Jiangsu of China is selected as an example. The results show that the proposed model reveals the best accuracy for the short data sequences (the average fitting error is only 2.19% and the average three-step forecasting error is less than 4%) compared with three GM models and four classical statistical models. By extension, any fields of EC, such as petroleum consumption, natural gas consumption, can also be predicted using this novel model. As the sustained growth in EC of China's, it is of great significance to predict EC accurately to manage serious energy security and environmental pollution problems, as well as formulating relevant energy policies by the government.

Key Words: energy consumption; Grey system theory; interval number; prediction; electricity

* Corresponding author : dingsong1129@163.com (S. Ding)

E-mail address: yejingjenson@163.com (J. Ye); iamdangyg@163.com (Y.G. Dang); dingsong1129@163.com (S. Ding); yyang@dmu.ac.uk (Y.J. Yang)

1 Introduction

Today, energy remains the main engine of the economy. Energy consumption has been one of the hottest issues in energy research over the years. In China, as rapid and diversified society developments continue, EC forecasting has become more and more complicated and full of uncertainty.

To be specific, with the economic growth and transformation of industrial structures, per capita household EC sequence (1993-2014) in China show a non-smooth rising trend and is difficult to fit its curve trend (Data source: National Statistical Yearbook of China, 2016), thus it is not suitable for forecasting by using the whole sequence to establish models. In fact, energy data capacity is often very limited (15 observations or less) due to changes of energy policy or updates within the industrial structure. However, classical prediction models (e.g. exponential smoothing, Holt-Winters method, linear method, exponential model, regression, autoregressive integrated moving average (ARIMA)) or modern prediction techniques (e.g. genetic algorithm (GA), artificial neural networks (ANN), Particle Swarm Optimization (PSO), fuzzy system) often require larger sample sizes, based on mathematical statistics principles or machine learning rules and computational intelligence (Xia and Wong, 2014). To be specific, for EC forecasting, Yukseltan et al. (2017) developed a linear model, as an expansion in Fourier series, supplemented with a modulation by seasonal harmonics, to forecast electricity demand for Turkey. Al-mulali et al. (2014) conducted panel data (labour, gross fixed capital formation, and total trade) analysis, based on statistics to describe their relationships with renewable EC and non-renewable EC. Zahedi et al. (2013) predicted the electricity demand of Canada by using a neuro-fuzzy model. Felice et al. (2015) presented an assessment of the use of seasonal climate forecasts of temperature for medium-term electricity demand prediction, by using both deterministic and probabilistic forecasting approaches. The classical models and the modern methods mentioned above are applied to cases with sufficient sample data which cannot capture the sensitive latest changes of data trends.

In reality, for energy data samples where the trends show significant changes, their newest or most recent information play the most important role in the modeling process rather than the amount of data. Owing to these reasons, Grey Model (GM) in GST (Liu et al., 2010) is introduced to forecast EC with small amounts of newest data. GM is suitable for small samples which emphasizes on new information, and it has been verified by many scholars: according to the principle of new information priority, Hamzacebi and Es (2014) used Optimized Grey Modeling (1,1) to forecast the annual electricity consumption of Turkey; Xiong et al. (2014) proposed a novel GM (1,1) model based on optimizing initial condition and applied in China's energy consumption and production forecasting. Moreover, it has been proved that grey prediction models can predict short-term data with the most recent information more efficiently compared with using long term old data (Wu et al., 2013a). Various papers on energy forecasting, using Grey Models or hybrid models with Grey Models, have now been disseminated. Ding (2018) designed a self-adapting grey prediction model having a nonlinear optimized initial value by utilizing an ant lion optimizer (ALO) algorithm to forecast China's natural-gas demand. Xu et al. (2017) proposed a novel grey model with optimal time response, based on particle swarm algorithm to forecast the electricity consumption data of China and compared with three alternative grey models. Truong et al. (2012) established a modified grey model MGM(1,1) for real-time control of wave energy

1 converters in irregular waves and the traditional grey model and a typical autoregressive (AR)
2 model were carried out for comparisons. Ding designed a new multivariable model that considers
3 the accumulative effects of the influencing factors, and used this proposed model to accurately
4 forecast the future output value of high-tech industries in China. Lee and Tong (2012) combined a
5 dynamic grey model with genetic programming to forecast energy in the US and compared with
6 ARIMA, GP and other grey models. Erdal et al. (2010) investigated the accuracies of different
7 grey models such as GM(1,1), Grey Verhulst model, modified grey models using Fourier series,
8 and the fitting and forecasting results of the modified grey models revealed higher performances.
9 Ding et al. (2018) analyzed the properties of a multiple transformation and optimized background
10 value in the GM(1,N) model to simulate and predict CO₂ emissions. In summary, none of the
11 above papers considered regional unbalances in a certain country or region, and the research
12 objects of them are all described as single real numbers. For model comparisons, most of the
13 papers only took into account grey models, because of the short-term sequences, instead of using
14 forecasting models, such as statistical models, which is neither comprehensive nor convincing.
15
16
17
18

19 Due to the regional characteristics of EC, real values are difficult to ascertain when
20 deciphering the different levels of per capita annual ECs in one region. The interval numbers are
21 introduced to characterize the annual ECs which can cover the full information of this energy data.
22 Focusing on the feature of the uncertainty, interval grey numbers have been discussed in many
23 other fields, like decision-making, information retrieval, knowledge discovery and so on
24 (Yamaguchi et al., 2007). Professor Deng Julong (1982) originally proposed interval grey numbers,
25 whereby the interval grey number is a typical form of grey number which refers to uncertain
26 values of the numbers in an interval or a general number set. In terms of grey prediction of
27 interval grey numbers, researchers have made a series of attempts to transform interval grey
28 number sequences into real number sequences, to construct grey prediction models by means of
29 using geometric characteristics, formations of operators and establishments of combination
30 models etc., and then restore back to the interval grey number sequences. Zeng et al. (2014)
31 converted an interval grey number sequence into the kernel and area sequences to establish grey
32 forecasting models. This method avoids extreme data errors by generating average value and grey
33 degree enlargement scenarios. But the characteristic of non-homogeneous sequences is not taken
34 into account. Wu et al. (2013b) converted an interval grey number sequence into a real number
35 sequence by calculating the areas of grey number layers and the cognition degree of interval grey
36 numbers. In this method, the surface areas are introduced to accurately express the areas of grey
37 number layers. The solution of the equations is quite complex in the process of calculation. In
38 reality, most information distributions of interval grey numbers are not equal. The above grey
39 prediction models, based on interval grey numbers can not reflect this outstanding feature. Thus,
40 Zeng Bo et al. (2010) introduced the typical whitenization weight function to establish the grey
41 prediction model. In this model, the upper and lower bounds of interval grey numbers and two
42 turning points in whitenization weight function are all used without omitting any information of
43 the grey number. It is usually difficult to obtain the typical whitenization weight function which
44 needs to determine two turning points. And there is no doubt of the high costs taken by
45 synthesizing experts' opinions or experiences as well as collecting and analyzing previous data to
46 obtain this function, which also leads to more deviations. To simplify the typical whitenization
47 weight function, a central point which represents the maximum possibility of an interval can be
48 considered to replace two turning points in the typical whitenization weight function. By these
49 means, this paper utilizes the central-point sequence to establish a novel energy consumption
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

1 forecasting model, which is easier to acquire, in describing uneven distributions of information of
2 interval grey numbers.

3
4 To optimize grey model, the initial condition of DGM(1, 1) is reset by minimizing the fitting
5 values' errors in this paper. Instead of using the first data set as an initial condition in the original
6 GM (Liu et al., 2010), it ignores the extraction of new information. In terms of other initial
7 condition optimizations, Dang et al. (2004) set the last in-sample data's accumulated generating
8 value($x^{(1)}(n)$) as the initial condition which put too much emphasis on the latest information and
9 became vulnerable to abnormal fluctuations; Liu et al. (2011) discussed three initial conditions:
10 the starting-point fixed initial value, the middle-point fixed initial value and the end-point fixed
11 initial value, though there was no indication of the suitability of each setting; Wang et al. (2010)
12 used the weighted combination of the first and last data set from the 1-AGO sequence as the initial
13 condition, but how to determine the weight had not been given; Xiong et al. (2014) expanded
14 Wang's method by weighting all the in-shape's accumulated generating data as the initial
15 condition. The weights were fixed which meant the weights could not adjust according to
16 sequences. To this end, this paper optimizes the initial condition of the grey forecasting model by
17 setting a parameter which makes the fitting accuracy minimum.
18
19
20
21
22

23 To sum up, in-sample data sets in this paper are characterized as interval grey numbers to
24 establish the novel energy consumption forecasting model to forecast EC. The innovation and
25 contribution of this research compared with other references mainly lies in the following three
26 aspects: Firstly, the uncertain information of regional energy consumption is expressed by interval
27 grey numbers, in which the upper bound and the lower bound are used to characterize the highest
28 and the lowest levels of this region. Secondly, short-term most recent energy consumption data are
29 used to establish the novel forecasting model, and long-term data are applied as comparisons.
30 Thirdly, the accuracy of the novel forecasting model is improved by full information
31 transformation of energy consumption data, described by interval grey numbers and the model's
32 initial value optimization.
33
34
35
36

37 The reminder of this paper is structured as follows: in Section 2, information of the interval
38 grey numbers with central points is transformed fully to prepare for forecasting. Section 3
39 optimizes the DGM(1, 1) of interval grey numbers with central points and the initial condition to
40 improve the energy consumption forecasting model's precision. Section 4 discusses a case of per
41 capita annual EC from five cities in Southern Jiangsu to forecast near future trends and verifies the
42 practicability of the novel model by comparing with other forecasting methods. Finally, conclusions
43 are drawn in Section 5.
44
45
46

47 **2 The effective information conversion of interval grey numbers with** 48 **central points**

49 **2.1 The basic concepts**

50
51
52
53
54
55
56 **Definition 1** (Liu et al., 2010). Let $\otimes \in [a, b]$, $a \leq b$, $a, b \in R$, then \otimes is called an
57 interval grey number; if $a = b$, then \otimes is a real number and $\otimes = a$.
58
59
60
61
62
63
64
65

Definition 2. Let $\otimes = [a, b, c]$ be an interval grey number, $a \leq b \leq c$. Among them, a , c are the lower bound and the upper bound of the interval grey number, and b is the maximum possibility point of this interval which is called the central point. This kind of interval grey number is named as an interval grey number with a central point. The sequence of interval grey numbers with central points can be described as $X(\otimes) = (\otimes_1, \otimes_2, \dots, \otimes_n)$ in which $\otimes_m \in [a_m, b_m, c_m]$, $a_m \leq b_m \leq c_m$, $m = 1, 2, \dots, n$.

Definition 3. Let the sequence of interval grey numbers with central points as $X(\otimes) = (\otimes_1, \otimes_2, \dots, \otimes_n)$, and map all elements of $X(\otimes)$ into a two-dimensional Cartesian coordinate system. Then, sequentially connect the lower bounds, the upper bounds and the central points of adjacent interval grey numbers respectively to constitute the lower-bound line, the upper-bound line and the central-point line, which can be seen in Fig. 1. Among them, the sequence of the lower-bound line can be denoted as $X(\underline{\otimes}) = (a_1, a_2, \dots, a_n)$, the sequence of the central-point line can be denoted as $X(\tilde{\otimes}) = (b_1, b_2, \dots, b_n)$, and the sequence of the upper-bound line can be denoted as $X(\overline{\otimes}) = (c_1, c_2, \dots, c_n)$.

Definition 4 (Zeng et al., 2010). The graphics of the lower and upper bounds sequences of interval grey numbers are known as grey numbers' belts; the sections of the adjacent interval grey numbers are called grey numbers' layers. According to positions of grey numbers' layers in grey numbers' belts, they are marked by $1, 2, \dots, n$ (See in Fig. 1).

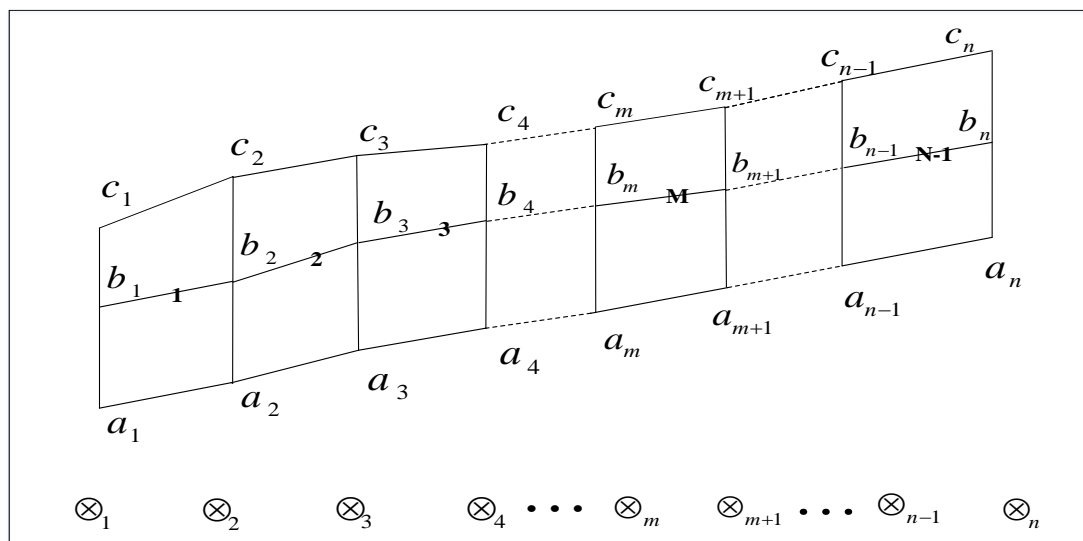


Fig.1. The grey numbers' belt of the interval grey number $\otimes = [a, b, c]$

Definition 5. If the grey numbers' layers (GL) are marked by $1, 2, \dots, n$, each GL can be divided into two parts by the central-point sequence. The two parts can be called: the upper grey layer and the lower grey layer.

2.2 Establish the effective information conversion sequences

Let the sequence of interval grey numbers with the central point as $X(\otimes) = ([a_1, b_1, c_1], [a_2, b_2, c_2], \dots, [a_n, b_n, c_n])$ in which a_m and c_m are the lower and upper bound values, b_m is the central point. The effective information conversion sequences can be established by the sequences of trapezoid areas made by upper bound values and central-point values; trapezoid areas made by lower bound values and central-point values and middle values calculated by upper bound values and lower bound values.

Lemma 1 (Zeng and Liu, 2011). In the grey number layer m (see in Fig. 1), the longitudinal coordinate of the middle value calculated by upper bound values and lower bound values ($a_m, c_m, a_{m+1}, c_{m+1}$) is:

$$L_m = \frac{a_m + c_m + a_{m+1} + c_{m+1}}{4}, \quad m = 1, 2, \dots, n-1.$$

According to Lemma 1, the sequence of middle values can be denoted as $X(L) = (L_1, L_2, \dots, L_{n-1})$.

Theorem 1 (Trapezoid areas of the lower grey layer). In the grey number layer m (see in Fig. 1), the trapezoid area of lower bound values (a_m, a_{m+1}) and central-point values (b_m, b_{m+1}) is:

$$\underline{S}_m = \frac{(b_{m+1} - a_{m+1}) + (b_m - a_m)}{2}, \quad m = 1, 2, \dots, n-1.$$

Proof: The differences of longitudinal coordinates between the lower bounds and the central points are $(b_{m+1} - a_{m+1})$, $(b_m - a_m)$ which represent the trapezoid's top and bottom. And the height of the trapezoid can be obtained by the differences of the adjacent grey number layers, which is $(m+1) - m = 1$. By this means, the trapezoid area is $\underline{S}_m = \frac{(b_{m+1} - a_{m+1}) + (b_m - a_m)}{2}$.

Theorem 2 (Trapezoid areas of the upper grey layer). In the grey number layer m (see in Fig. 1), the trapezoid area of upper bound values (c_m, c_{m+1}) and central-point values (b_m, b_{m+1}) is:

$$\bar{S}_m = \frac{(c_{m+1} - b_{m+1}) + (c_m - b_m)}{2}, \quad m = 1, 2, \dots, n-1.$$

Proof: Similar to Theorem 1.

According to Theorem 1 and Theorem 2, here are the sequence of trapezoid areas of the lower grey layer $X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1})$ and the sequence of trapezoid areas of the upper grey layer $X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1})$.

In addition, from Lemma 1, Theorem 1 and Theorem 2, we can obtain:

$$a_{m+1} + a_m = 2L_m - \bar{S}_m - \underline{S}_m$$

$$b_{m+1} + b_m = 2L_m - \bar{S}_m + \underline{S}_m$$

$$c_{m+1} + c_m = 2L_m + \bar{S}_m + \underline{S}_m.$$

That is to say,

$$X(\otimes) = ([a_1, b_1, c_1], [a_2, b_2, c_2], \dots, [a_n, b_n, c_n]) \Leftrightarrow \begin{cases} X(L) = (L_1, L_2, \dots, L_{n-1}) \\ X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1}) \\ X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1}) \end{cases} .$$

The information contained in the interval grey numbers' sequence with central points are retained completely by the information conversion which means the conversion preserves the characteristics of information equivalence and data integrity.

3 The optimized DGM(1, 1) model

3.1. DGM(1,1) of interval grey numbers with central points

In the modeling process of discrete grey models, parameters estimation, simulation and prediction are all adopted in the form of discrete equations. The DGM model does not differentiate approximate substitution between the discrete equation and the continuous equation compared with GM(1, 1) model, in which the method of estimating parameters is the discrete equation and the method of simulating and forecasting is the continuous equation, so it has higher fitting precision by calculating the average relative percentage error (ARPE) (Xie and Liu, 2009). Let $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ be the original data sequence, $x^{(0)}(k) \geq 0 (k=1, 2, \dots, n)$. Based on the traditional DGM(1, 1) model (Xie and Liu, 2009), here are: the recursive function of DGM(1,1) (set $x^{(1)}(1) = x^{(0)}(1)$)

$$\hat{x}^{(1)}(k+1) = \beta_1^k x^{(0)}(1) + \frac{1-\beta_1^k}{1-\beta_1} \cdot \beta_2, (k=1, 2, \dots, n-1) \quad (1)$$

and the restored value of the recursive function of DGM(1,1) at (k+1)

$$\hat{x}^{(0)}(k+1) = \alpha^{(1)} \hat{x}^{(1)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), (k=1, 2, \dots, n-1) \quad (2)$$

Similarly, for the sequence of middle values $X(L) = (L_1, L_2, \dots, L_{n-1})$, the time response sequence of its DGM(1,1) is proposed by Zeng et al. (2010):

$$\hat{L}_{k+1}^{(1)} = \alpha_1^k L_1 + \frac{1-\alpha_1^k}{1-\alpha_1} \alpha_2 \quad (3)$$

Note: α_1, α_2 are parameters of DGM(1,1).

The restored value of the prediction model is:

$$\hat{L}_{k+1} = \alpha_1^{k-1} [L_1 (\alpha_1 - 1) + \alpha_2] \quad (4)$$

From Formula (4), we can obtain the parameters' forecasting values of the sequence of middle values made up by the upper and lower bounds values. According to Lemma 1, it contains the "plus" information of the lower and upper bounds of interval grey numbers.

From $L_m = \frac{a_m + c_m + a_{m+1} + c_{m+1}}{4}$, here is:

$$\hat{a}_{m+1} + \hat{c}_{m+1} = 4\hat{L}_m - (\hat{a}_m + \hat{c}_m) \quad (5)$$

When $m = k - 1$,

$$\hat{a}_k + \hat{c}_k = 4\hat{L}_{k-1} - 4\hat{L}_{k-2} + \cdots + (-1)^{k-1}(a_1 + c_1) \quad (6)$$

According to Formula (4), the first $(k - 2)$ items in Formula (6) consist of a geometric series whose common ratio is $q = -\alpha_1^{-1}$. $\hat{a}_k + \hat{c}_k$, which can be calculated by the summation formula of geometric series:

$$\hat{a}_k + \hat{c}_k = \frac{4\alpha_1^{k-3}[L_1(\alpha_1 - 1) + \alpha_2][1 - (-\alpha_1^{-1})^{k-2}]}{1 + \alpha_1^{-1}} + (-1)^k(a_2 + c_2) \quad (7)$$

For the sequence of trapezoid areas of the lower grey layer $X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1})$, the time response sequence of its DGM(1,1) is:

$$\hat{\underline{S}}_{k+1}^{(1)} = \beta_1^k \underline{S}_1 + \frac{1 - \beta_1^k}{1 - \beta_1} \beta_2 \quad (8)$$

Note: β_1, β_2 are parameters of DGM(1,1).

The restored value of the prediction model is:

$$\hat{\underline{S}}_{k+1} = \beta_1^{k-1}[\underline{S}_1(\beta_1 - 1) + \beta_2] \quad (9)$$

From Formula (9), we can obtain parameters' forecasting values of the sequence of trapezoid areas made by lower bound values and central-point values.

For the sequence of trapezoid areas of the upper grey layer $X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1})$, the time response sequence of its DGM(1,1) is:

$$\hat{\bar{S}}_{k+1}^{(1)} = \gamma_1^k \bar{S}_1 + \frac{1 - \gamma_1^k}{1 - \gamma_1} \gamma_2 \quad (10)$$

Note: γ_1, γ_2 are parameters of DGM(1,1).

The restored value of the prediction model is:

$$\hat{\bar{S}}_{k+1} = \gamma_1^{k-1}[\bar{S}_1(\gamma_1 - 1) + \gamma_2] \quad (11)$$

From Formula (11), we can obtain the parameters' forecasting values of the sequence of trapezoid areas made by upper bound values and central-point values.

According to Theorem 1 and Theorem 2, they contain the "minus" information of the lower and upper bounds of interval grey numbers.

Similarly, from $\underline{S}_m = \frac{(b_{m+1} - a_{m+1}) + (b_m - a_m)}{2}$ and $\bar{S}_m = \frac{(c_{m+1} - b_{m+1}) + (c_m - b_m)}{2}$,

here are:

$$\hat{c}_{m+1} - \hat{a}_{m+1} = 2(\hat{\underline{S}}_m + \hat{\bar{S}}_m) - (\hat{c}_m - \hat{a}_m) \quad (12)$$

When $m = k - 1$,

$$\hat{c}_k - \hat{a}_k = 2(\hat{\underline{S}}_{k-1} + \hat{\bar{S}}_{k-1}) - 2(\hat{\underline{S}}_{k-2} + \hat{\bar{S}}_{k-2}) + \cdots + (-1)^{k-1}(c_1 - a_1) \quad (13)$$

According to Formula (9) and Formula (11), the first $(k - 2)$ items in Formula (13) consist of two geometric series whose common ratio are $q' = -\beta_1^{-1}$ and $q'' = -\gamma_1^{-1}$. $\hat{c}_k - \hat{a}_k$ can be calculated by the summation formula of geometric series:

$$\hat{c}_k - \hat{a}_k = \frac{2\beta_1^{k-3}[\underline{S}_1(\beta_1 - 1) + \beta_2][1 - (-\beta_1^{-1})^{k-2}]}{1 + \beta_1^{-1}} + \frac{2\gamma_1^{k-3}[\bar{S}_1(\gamma_1 - 1) + \gamma_2][1 - (-\gamma_1^{-1})^{k-2}]}{1 + \gamma_1^{-1}} + (-1)^k(c_2 - a_2) \quad (14)$$

Among them,

$$\hat{b}_k - \hat{a}_k = \frac{2\beta_1^{k-3}[\underline{S}_1(\beta_1 - 1) + \beta_2][1 - (-\beta_1^{-1})^{k-2}]}{1 + \beta_1^{-1}} + (-1)^k(b_2 - a_2) \quad (15)$$

$$\hat{c}_k - \hat{b}_k = \frac{2\gamma_1^{k-3}[\bar{S}_1(\gamma_1 - 1) + \gamma_2][1 - (-\gamma_1^{-1})^{k-2}]}{1 + \gamma_1^{-1}} + (-1)^k(c_2 - b_2) \quad (16)$$

Combine with Formula (7), Formula (15) and Formula (16),

$$\begin{cases} \hat{a}_k + \hat{c}_k = 4F_L\alpha_1^{k-3}[1 - (-\alpha_1^{-1})^{k-2}] + (-1)^k(a_2 + c_2) \\ \hat{c}_k - \hat{a}_k = 2F_{\underline{S}}\beta_1^{k-3}[1 - (-\beta_1^{-1})^{k-2}] + 2F_{\bar{S}}\gamma_1^{k-3}[1 - (-\gamma_1^{-1})^{k-2}] + (-1)^k(c_2 - a_2) \\ \hat{b}_k - \hat{a}_k = 2F_{\underline{S}}\beta_1^{k-3}[1 - (-\beta_1^{-1})^{k-2}] + (-1)^k(b_2 - a_2) \end{cases}$$

Among them,

$$F_L = \frac{[L_1(\alpha_1 - 1) + \alpha_2]}{1 + \alpha_1^{-1}}, \quad F_{\underline{S}} = \frac{[\underline{S}_1(\beta_1 - 1) + \beta_2]}{1 + \beta_1^{-1}}, \quad F_{\bar{S}} = \frac{[\bar{S}_1(\gamma_1 - 1) + \gamma_2]}{1 + \gamma_1^{-1}}.$$

Note: F_L , $F_{\underline{S}}$, $F_{\bar{S}}$ are all seen as constant terms due to the confirmation of all the parameters.

Finally, we obtain the grey prediction model of interval grey numbers with central points:

$$\begin{cases} \hat{a}_k = 2F_L\alpha_1^{k-3}[1-(-\alpha_1^{-1})^{k-2}] - F_{\underline{S}}\beta_1^{k-3}[1-(-\beta_1^{-1})^{k-2}] - F_{\bar{S}}\gamma_1^{k-3}[1-(-\gamma_1^{-1})^{k-2}] + (-1)^k a_2 \\ \hat{b}_k = 2F_L\alpha_1^{k-3}[1-(-\alpha_1^{-1})^{k-2}] + F_{\underline{S}}\beta_1^{k-3}[1-(-\beta_1^{-1})^{k-2}] - F_{\bar{S}}\gamma_1^{k-3}[1-(-\gamma_1^{-1})^{k-2}] + (-1)^k b_2 \\ \hat{c}_k = 2F_L\alpha_1^{k-3}[1-(-\alpha_1^{-1})^{k-2}] + F_{\underline{S}}\beta_1^{k-3}[1-(-\beta_1^{-1})^{k-2}] + F_{\bar{S}}\gamma_1^{k-3}[1-(-\gamma_1^{-1})^{k-2}] + (-1)^k c_2 \end{cases} \quad (*)$$

In summary, the establishment steps of grey prediction models of interval grey numbers with central points are:

Step1 : From the sequence of interval grey numbers with central points $X(\otimes) = ([a_1, b_1, c_1], [a_2, b_2, c_2], \dots, [a_n, b_n, c_n])$, calculate the sequence of middle values made up by the upper and lower bound values $X(L) = (L_1, L_2, \dots, L_{n-1})$, the sequence of trapezoid areas of the lower grey layer $X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1})$ and the sequence of trapezoid areas of the lower grey layer $X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1})$.

Step2: Establish the DGM(1,1) for the above 3 sequences, and obtain parameters $(\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2)$.

Step3: Calculate $F_L, F_{\underline{S}}, F_{\bar{S}}$.

Step4: Calculate Equation Group (*) to obtain the forecasting values of the lower bounds, the upper bounds and central points.

3.2. Optimization of the initial conditions in the DGM(1,1)

As it mentioned in Section 3.1, the traditional DGM model reveals higher accuracy because there is no approximate substitution of continuous equations (Xie and Liu, 2009). However, the traditional DGM model seeks to align the exponential sequences (ca^k) modeling, ignoring many non-homogeneous exponential sequences $(ca^k + b)$ in reality. The interval grey number sequences mostly present non-homogeneous characteristics for their upper and lower bounds sequences, because of the uncertainties of their internal mechanism. To this end, Xie et al. (2013) proposed an approximate non homogeneous discrete grey model (NDGM(1,1)).

$$\begin{cases} \hat{x}^{(1)}(k+1) = \beta_1 \hat{x}^{(1)}(k) + \beta_2 \cdot k + \beta_3 \\ \hat{x}^{(1)}(1) = x^{(1)}(1) + \beta_4 \end{cases} \quad (17)$$

Based on the least square method, the parameters $\beta_1, \beta_2, \beta_3$ can be calculated:

$$\hat{\beta} = [\beta_1, \beta_2, \beta_3]^T = (B^T B)^{-1} B^T Y,$$

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

Moreover, the accuracy of this model is not only determined by the structural parameters $\hat{\beta}$, but also the initial condition. For the initial conditions of the traditional DGM, the starting-point value, middle-point value and end-point value are all discussed for the iterations respectively, to explain the influences for exponential sequence' fitting. The three forms of initial conditions of the traditional DGM are (Liu and Xie, 2011):

The starting-point fixed DGM(SDGM) is:

$$\begin{cases} \hat{x}^{(1)}(k+1) = \beta_1 \hat{x}^{(1)}(k) + \beta_2 \\ \hat{x}^{(1)}(1) = x^{(1)}(1) = x^{(0)}(1) \end{cases} \quad (18)$$

The middle-point fixed DGM(MDGM) is:

$$\begin{cases} \hat{x}^{(1)}(k+1) = \beta_1 \hat{x}^{(1)}(k) + \beta_2 \\ \hat{x}^{(1)}(1) = x^{(1)}(m) = \sum_{i=1}^m x^{(0)}(i), 1 < m < n \end{cases} \quad (19)$$

The end-point fixed DGM(EDGM) is:

$$\begin{cases} \hat{x}^{(1)}(k+1) = \beta_1 \hat{x}^{(1)}(k) + \beta_2 \\ \hat{x}^{(1)}(1) = x^{(1)}(n) = \sum_{i=1}^n x^{(0)}(i) \end{cases} \quad (20)$$

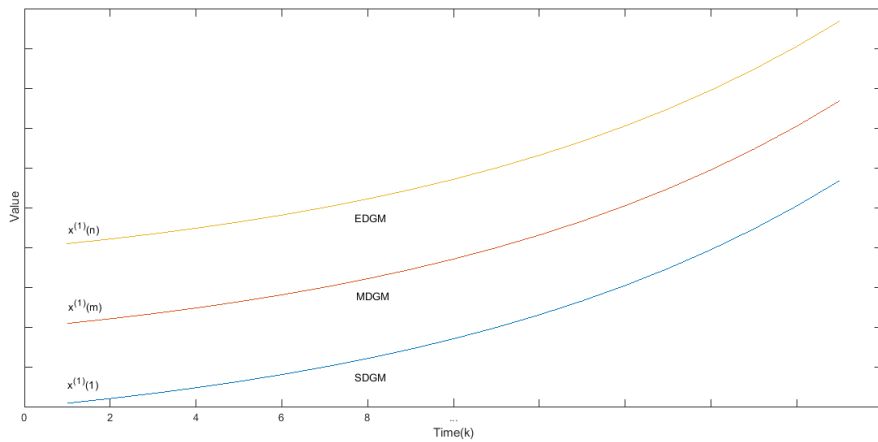


Fig.2. The effect of different initial conditions

The choice of the initial value will directly affect the first point and the trend of the values in

the model (shown in Fig. 2). If the initial condition of the model can be processed properly, there will be a positive effect on the accuracy of the model. For the SDGM and EDGM, the initial conditions are fixed (the starting-point value and the end-point value determine the first point) which may not be the optimal solution. For the MDGM, the middle-point value is hard to identify and the results may not be so accurate. In order to better explore the initial conditions, the optimization method of the initial condition should be put forward. In NDGM(1,1) (Xie et al., 2013), the initial condition is set by the first data instance as a reference and a corrective parameter β_4 ($\hat{x}^{(1)}(1) = x^{(1)}(1) + \beta_4$) which can be optimized by establishing a parameter in an overall consideration.

According to NDGM(1,1) ($\hat{x}^{(1)}(m+1) = \beta_1 \hat{x}^{(1)}(m) + \beta_2 \cdot m + \beta_3$), the recurrence model is obtained:

$$\hat{x}^{(1)}(m+1) = \beta_1^m \hat{x}^{(1)}(1) + \beta_2 \sum_{j=1}^m j \beta_1^{m-j} + \frac{1-\beta_1^m}{1-\beta_1} \beta_3, m=1,2,\dots \quad (21)$$

In order to establish a simple and effective optimization method, here the solution to establish the initial condition can be expressed as C to replace the first data and a corrective parameter β_4 :

$$\hat{x}^{(1)}(m+1) = C \cdot \beta_1^m + \beta_2 \sum_{j=1}^m j \beta_1^{m-j} + \frac{1-\beta_1^m}{1-\beta_1} \beta_3 \quad (22)$$

where C is an undetermined parameter.

To calculate the value of C by using a similar method like the least square principle, an unconstrained optimization model is established to minimize the sum of squares of errors between $\hat{x}^{(1)}(k)$ and $x^{(1)}(k)$: $\min_C \sum_{k=1}^n [\hat{x}^{(1)}(k) - x^{(1)}(k)]^2$ where the number of in-shape values is n .

We then, bring the obtained parameters $\beta_1, \beta_2, \beta_3$ into

$\hat{x}^{(1)}(k+1) = C \cdot \beta_1^k + \beta_2 \sum_{j=1}^k j \beta_1^{k-j} + \frac{1-\beta_1^k}{1-\beta_1} \beta_3$ and let $\hat{x}^{(1)}(1) = x^{(1)}(1)$, we set:

$$F(C) = \sum_{k=1}^n [\hat{x}^{(1)}(k) - x^{(1)}(k)]^2 = 0 + \sum_{k=1}^{n-1} [C \cdot \beta_1^k + \beta_2 \sum_{j=1}^k j \beta_1^{k-j} + \frac{1-\beta_1^k}{1-\beta_1} \beta_3 - x^{(1)}(k+1)]^2 \quad (23)$$

Here is

$$\frac{\partial F(C)}{\partial C} = 2 \sum_{k=1}^{n-1} [C \cdot \beta_1^k + \beta_2 \sum_{j=1}^k j \beta_1^{k-j} + \frac{1-\beta_1^k}{1-\beta_1} \beta_3 - x^{(1)}(k+1)] \cdot \beta_1^k = 0 \quad (24)$$

C can be solved:

$$C = \frac{\sum_{k=1}^{n-1} [x^{(1)}(k+1) - \beta_2 \sum_{j=1}^k j\beta_1^{k-j} - \frac{1-\beta_1^k}{1-\beta_1} \beta_3] \cdot \beta_1^k}{\sum_{k=1}^{n-1} \beta_1^{2k}} \quad (25)$$

The optimized NDGM(1,1) is obtained which can be denoted as CNDGM(1,1):

$$\hat{x}^{(1)}(m+1) = \frac{\sum_{k=1}^{n-1} [x^{(1)}(k+1) - \beta_2 \sum_{j=1}^k j\beta_1^{k-j} - \frac{1-\beta_1^k}{1-\beta_1} \beta_3] \cdot \beta_1^k}{\sum_{k=1}^{n-1} \beta_1^{2k}} \cdot \beta_1^m + \beta_2 \sum_{j=1}^m j\beta_1^{m-j} + \frac{1-\beta_1^m}{1-\beta_1} \beta_3 \quad (26)$$

The restored value of the recursive function of CNDGM(1,1) at (k+1) is:

$$\hat{x}^{(0)}(m+1) = \alpha^{(1)} \hat{x}^{(1)}(m+1) = \hat{x}^{(1)}(m+1) - \hat{x}^{(1)}(m), m=1,2,\dots \quad (27)$$

The novel DGM (CNDGM(1,1)) with an optimized initial condition under overall consideration has been established.

3.3. CNDGM(1,1) of interval grey number with central points

For the sequence of middle values $X(L) = (L_1, L_2, \dots, L_{n-1})$, the time response sequence of its CNDGM(1,1) is:

$$\hat{L}_{m+1}^{(1)} = \frac{\sum_{k=1}^{n-1} [L_{k+1}^{(1)} - \alpha_2 \sum_{j=1}^k j\alpha_1^{k-j} - \frac{1-\alpha_1^k}{1-\alpha_1} \alpha_3] \cdot \alpha_1^k}{\sum_{k=1}^{n-1} \alpha_1^{2k}} \cdot \alpha_1^m + \alpha_2 \sum_{j=1}^m j\alpha_1^{m-j} + \frac{1-\alpha_1^m}{1-\alpha_1} \alpha_3 \quad (28)$$

$$\hat{L}_{m+1}^{(0)} = \hat{L}_{m+1}^{(1)} - \hat{L}_m^{(1)} \quad (29)$$

Note: $\alpha_1, \alpha_2, \alpha_3$ are parameters of CNDGM(1,1).

Similarly, the CNDGM(1,1) models of the sequences of $X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1})$ and $X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1})$ are:

$$\hat{\underline{S}}_{m+1}^{(1)} = \frac{\sum_{k=1}^{n-1} [\underline{S}_{k+1}^{(1)} - \beta_2 \sum_{j=1}^k j\beta_1^{k-j} - \frac{1-\beta_1^k}{1-\beta_1} \beta_3] \cdot \beta_1^k}{\sum_{k=1}^{n-1} \beta_1^{2k}} \cdot \beta_1^m + \beta_2 \sum_{j=1}^m j\beta_1^{m-j} + \frac{1-\beta_1^m}{1-\beta_1} \beta_3 \quad (30)$$

$$\hat{\underline{S}}_{m+1}^{(0)} = \hat{\underline{S}}_{m+1}^{(1)} - \hat{\underline{S}}_m^{(1)} \quad (31)$$

$$\hat{\bar{S}}_{m+1}^{(1)} = \frac{\sum_{k=1}^{n-1} [\bar{S}_{k+1}^{(1)} - \gamma_2 \sum_{j=1}^k j\gamma_1^{k-j} - \frac{1-\gamma_1^k}{1-\gamma_1} \gamma_3] \cdot \gamma_1^k}{\sum_{k=1}^{n-1} \gamma_1^{2k}} \cdot \gamma_1^m + \gamma_2 \sum_{j=1}^m j\gamma_1^{m-j} + \frac{1-\gamma_1^m}{1-\gamma_1} \gamma_3 \quad (32)$$

$$\hat{\bar{S}}_{m+1}^{(0)} = \hat{\bar{S}}_{m+1}^{(1)} - \hat{\bar{S}}_m^{(1)} \quad (33)$$

Among them, $\hat{L}_m^{(0)}, \hat{\underline{S}}_m^{(0)}, \hat{\bar{S}}_m^{(0)}$ are the forecasting values of the m_{th} values of $X(L)$,

$$X(\underline{S}), X(\bar{S}).$$

Combine with Lemma 1, Theorem 1 and Theorem 2,

$$\begin{cases} \hat{a}_{m+1} + a_m = 2\hat{L}_m^{(0)} - \hat{S}_m^{(0)} - \hat{S}_m^{(0)} \\ \hat{b}_{m+1} + b_m = 2\hat{L}_m^{(0)} + \hat{S}_m^{(0)} - \hat{S}_m^{(0)} \\ \hat{c}_{m+1} + c_m = 2\hat{L}_m^{(0)} + \hat{S}_m^{(0)} + \hat{S}_m^{(0)} \end{cases}$$

Finally, we obtain grey a prediction model of interval grey numbers with central points:

$$\begin{cases} \hat{a}_{m+1} = 2\hat{L}_m^{(0)} - \hat{S}_m^{(0)} - \hat{S}_m^{(0)} - a_m \\ \hat{b}_{m+1} = 2\hat{L}_m^{(0)} + \hat{S}_m^{(0)} - \hat{S}_m^{(0)} - b_m \quad (\#) \\ \hat{c}_{m+1} = 2\hat{L}_m^{(0)} + \hat{S}_m^{(0)} + \hat{S}_m^{(0)} - c_m \end{cases}$$

The procedures involved for using the CNDGM(1,1) can be summarized as follows.

Step 1: Similar with Step 1 in the process of DGM(1,1) of interval grey numbers with central points.

Step2: Establish the CNDGM(1, 1) for the above 3 sequences, and obtain parameters($\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \gamma_3$).

Step3: Calculate the optimized initial condition C , then the forecasting values ($\hat{L}_m^{(0)}, \hat{S}_m^{(0)}, \hat{S}_m^{(0)}$) are obtained.

Step 4: Calculate Equation Group (#) to obtain the forecasting values of the lower bounds, the upper bounds and central points.

Step 5: Models' error analysis: The relative percentage error (RPE) describes the percentage of difference between the real and the fitting or forecasting values to evaluate the precision at a certain time instance k . The RPE can be defined as

$$RPE(k) = \left| \frac{\hat{x}^{(0)}(k) - x^{(0)}(k)}{x^{(0)}(k)} \right| \cdot 100\%$$

The total precision of a model can be described by calculating the average relative percentage error (ARPE):

$$ARPE = \frac{1}{m-k+1} \sum_{i=k}^m RPE(i), m \geq k$$

1 Note: Given θ , when $ARPE < \theta$ and $RPE(k) < \theta$, the model can be seen as a
2
3 satisfactory residual model.
4
5

6 **4 Numerical examples analysis**

7

8
9 As the most important basic industrial energy source for economic development, electricity is
10 one of the main priorities in the economic development strategy of all countries across the world.
11 Specific to China, with the society's sustained development, electricity consumption (EC)
12 continues to show a growth trend in recent years. Focusing on the southern Jiangsu area, it is
13 located in the core area of the Yangtze River Delta which is one of the most developed and
14 modernized areas in China. Therefore, the predictions of per capita annual ECs of southern
15 Jiangsu can be seen as a benchmark for the Yangtze River Delta's economic trend to provide a
16 strong reference for the local government to set future energy policy. In this section, per capita
17 annual electricity consumptions of southern Jiangsu are selected as practical examples. Owing to
18 the unbalanced development levels of the five cities (Nanjing, Wuxi, Changzhou, Suzhou and
19 Zhenjiang), single real numbers are difficult to decipher in expressing the actual situation of
20 electricity consumption in this area. The data sets are described as interval grey numbers. Among
21 them, the per capita annual electricity consumption is defined as the amount of annual electricity
22 consumption divided by the local population. To be specific, the upper and lower bounds of the
23 interval grey numbers are characterized by the highest and lowest levels in the region, and the
24 central-point values are obtained by the ratio of total amount of annual electricity consumption of
25 southern Jiangsu and total population. Due to the revision of statistical standards, the data of per
26 capita annual electricity consumption can now be collected from 1993 onwards. Hence, the
27 expressions of interval grey numbers for the data can be obtained in Table 1 below, the trend of
28 data are shown in Fig. 3.
29
30
31
32
33
34
35

36 To investigate the applicability of the novel model in more depth, three GM models and four
37 classical statistical models are introduced. For the three grey models and the novel grey model
38 proposed in this paper, the original DGM(1,1) (DGM) is established by using DGM(1,1) to
39 forecast the lower bound sequence, the central-point sequence and the upper bound sequence
40 respectively, which consider the three sequences independently; DGM(1,1) of interval grey
41 numbers with central points (DGMC) is set up in Section 3.1; CNDGM(1,1) of interval grey
42 numbers with central points (CNDGM) can be processed as steps set in Section 3.3;
43 GA-GM(1,1)(GAGM) uses the Genetic Algorithm to calculate the parameters of GM(1,1) (Xiong
44 et al., 2014). On the other hand, the four classical statistical models are as follows: the double
45 exponential smoothing method (DES) applies to the double exponential smoothing method instead
46 of single exponential smoothing, as it can better predict sequences with a certain linear trend; for
47 the linear model (LM), three sequences are modeled as $y(t) = a + bt$ (a,b are parameters, $y(t)$
48 is the t_{th} value of sequences) to fit their linear trends respectively; Similarly, in the exponential
49 model (EM), three sequences are modeled as $y(t) = a * b^t$ (a,b are parameters, $y(t)$ is the t_{th}
50 value of sequences) to fit their exponential trends; for the autoregressive moving average (ARMA),
51 AR(2) or ARMA (2,1) models are established by adopting autocorrelation and partial correlation
52 tests. Considering the features of grey models and classical statistical models, long data sequences
53
54
55
56
57
58
59
60
61
62
63
64
65

(1993-2012) and short data sequences (2001-2012) are applied to establish forecasting models. Take the data of 1993-2012 or 2001-2012 as in-sample fitting data, and the data of 2013, 2014 and 2015 as out-sample comparative data. According to the long data sequences and the short data sequences, the fitting errors of each model are shown in Fig.4.

Table 1

Per capita annual electricity consumption of southern Jiangsu in 1993-2015

Unit: kiloWatt-hour (kWh)/person

| Year | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 |
|-------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| The lower bound | 1164.42 | 1288.68 | 1384.50 | 1433.19 | 1457.82 | 1476.95 | 1577.41 | 1822.07 |
| The central point | 1467.51 | 1643.68 | 1795.24 | 1940.74 | 2029.73 | 2128.57 | 2348.22 | 2718.08 |
| The upper bound | 1795.47 | 2030.82 | 2244.95 | 2400.04 | 2505.86 | 2522.46 | 2716.36 | 3286.99 |
| Year | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
| The lower bound | 1997.52 | 2270.43 | 2653.92 | 3104.67 | 3631.40 | 4112.38 | 4717.54 | 4976.98 |
| The central point | 3064.88 | 3569.16 | 4408.90 | 5283.90 | 6283.43 | 7295.38 | 8264.12 | 8529.37 |
| The upper bound | 3891.62 | 4685.71 | 6051.24 | 7589.04 | 9320.45 | 11142.91 | 12928.09 | 13472.38 |
| Year | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | |
| The lower bound | 5352.03 | 5908.48 | 6281.66 | 6655.81 | 7194.48 | 7252.74 | 7578.51 | |
| The central point | 8895.40 | 10150.63 | 11034.49 | 11443.45 | 12196.43 | 12115.71 | 12405.57 | |
| The upper bound | 13894.80 | 16060.33 | 17617.58 | 18368.50 | 19319.86 | 19182.55 | 19665.67 | |

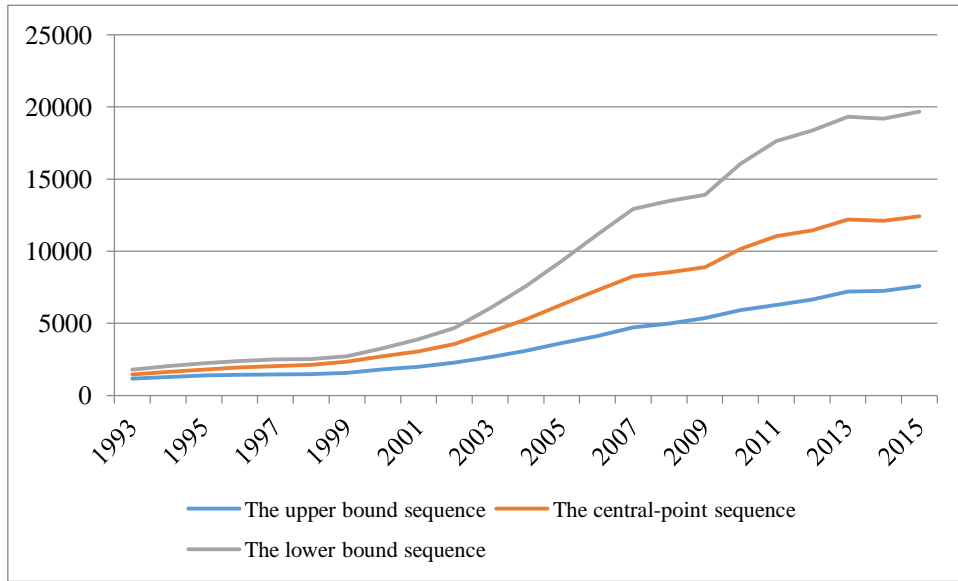


Fig.3. Interval grey numbers with central points of per capita annual electricity consumptions of southern Jiangsu

Special attention should be paid to the data trends. Before 2000, the trends of the lower bound, the central point and the upper bound sequences of interval grey numbers were all moderate. Then, from 2001 to 2005, three data sequences demonstrated steep rises. After experiencing a stable period from 2006-2009, the data displays a sharper growth trend which was ahead of the stable period of data in China from 2010. From an overall perspective, the data trends are complicated, and the valid data for recent information modelling are limited. It is difficult for forecasting models to capture the features of the sequences to make accurate predictions.

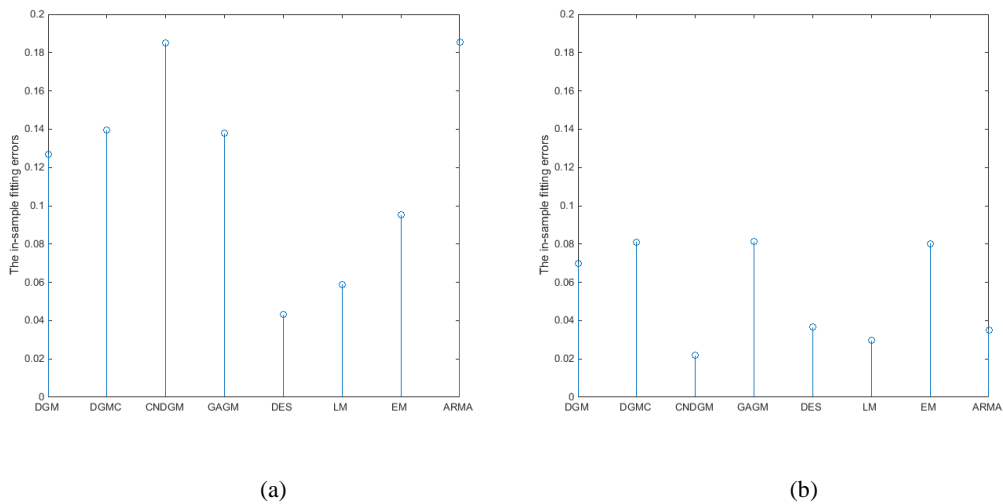


Fig. 4. (a) The in-sample fitting errors of eight models, based on the data of 1993-2012

(b) The in-sample fitting errors of eight models, based on the data of 2001-2012

It can be clearly seen that DES & LM obtain the smallest simulating errors for the long data sequences modelling, the ARPEs being 4.32% and 5.9% respectively. On the contrary, the fitting results of CNDGM for the long data sequences are less effective with an ARPE of 18.51% which is almost the highest ARPE, just lower than ARMA's 18.53%. For other models' results, EM also

shows high performance in fitting errors and the average relative error is 9.52%; the ARPEs of DGM, DGMC and GAGM are all around 13% (12.7%, 13.96% and 13.8% respectively). For the upper bound sequence, the fitting errors of DGM, DGMC, GAGM, ARMA and CNDGM are all above 18% which means the simulation effects are not effective and unacceptable. Similarly, for the central-point sequence, the fitting error of ARMA reaches 19.35%. These fitting outcomes may lead to poor forecasting results.

In the modelling process of the short data sequences, it is obvious that the ARPE of CNDGM which is proposed in this paper is the smallest among the other seven models, which is 2.19%. To be specific, the simulating errors of the lower bound, the central point and the upper bound sequences of interval grey numbers by using CNDGM are all below 3% and all the RPEs are smallest compared to each corresponding item from other compared methods. This means CNDGM is accurately fitted original data. In terms of other models, the ARPEs of LM, ARMA and DES are all below 4% (2.97%, 3.52% and 3.65% respectively) and their fitting effects are just outside CNDGM's. As to EM, the ARPE is 8.03%, and the RPE of the upper bound sequence reaches 10.52% which is an unacceptable fitting accuracy. The ARPEs of other three grey models are 6.99% for DGM, 8.11% for DGMC and 8.13% for GAGM which are all much higher than the optimized grey model in this paper. Similarly, the RPEs of DGM, GAGM and DGMC for the upper bound sequence achieve 9.08%, 11.78% and 10.59% which are also not exceptional.

After analyzing the results of simulating, the CNDGM(1,1) model of interval grey numbers with central points proposed in this paper, demonstrates a great fitting performance in the short data sequences which surpasses any other fitting models. It verifies that grey models can predict short-term more effectively compared with using long term data and further improves the original Grey Model of interval grey numbers.

When it comes to forecasting values, eight models are proposed by using the long data sequences and the short data sequences. And the results are shown in Fig.5.

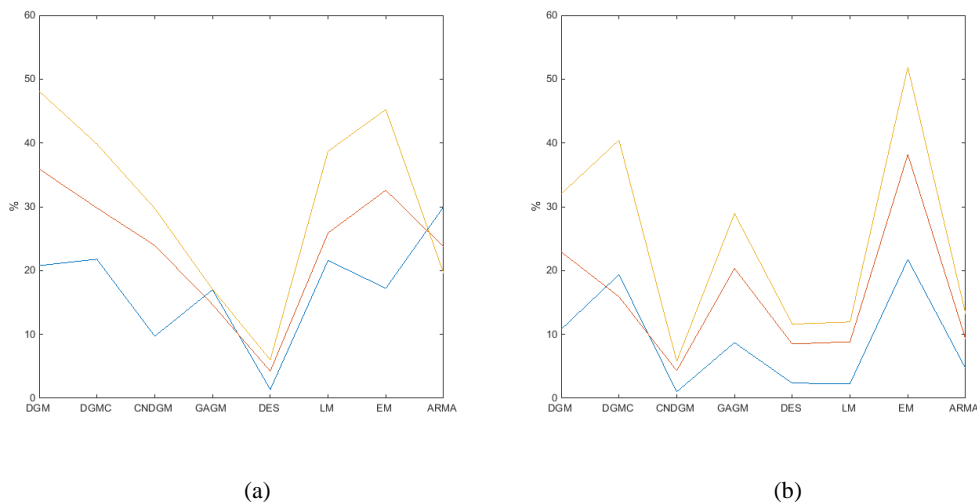


Fig. 5. (a) The out-sample forecasting errors of eight models, based on the data of 1993-2012

(b) The out-sample forecasting errors of eight models, based on the data of 2001-2012

From the perspective of the long data sequences, DES and LM should be discussed first because of their high fitting accuracy. For DES, the prediction results of 2013-2015 are considerably close to the actual values, the forecasting errors being 1.36%, 4.24% and 6.01% for 2013, 2014 and 2015. As to LM, the forecasting trends of the upper bound sequence, the central-point sequence and the lower bound sequence do not match the actual lines, which result in high forecasting errors all above 20% and for the year of 2015, the value is 38.69%. In terms of other models, the majority of the forecasting errors are over 20% which are not efficacious.

Moving on to the short data sequences, the novel CNDGM demonstrates extremely outstanding consequences of prediction. More concretely, the forecasting error of 2013 is only 1.01% and shows comparatively high accuracy towards the ARPEs of other models. In addition, the LM's, DES's and ARMA's are the second, third and fourth lowest (2.30%, 2.35% and 4.64%). The results show high fitting efforts lead to high first step forecasting. As for GAGM and DGM, the ARPEs are 8.71% and 10.79%, and the other two models' (DGMC and EM) are all around 20% which reveal disappointing forecasting performances. For the ARPEs of forecasting values in 2014 and 2015, similar consequences are proposed. Specifically, the lowest APEs are still from CNDGM which are 4.31% in 2014 and 5.77% in 2015. Followed by the DES's, LM's and ARMA's in 2014 and 2015, the ARPEs of DES are 8.49% and 11.61%, the ARPEs of LM are 8.85% and 11.95%, and the ARPEs of ARMA are 9.26% and 13.19%. As to GAGM, the ARPE of 2014 and 2015 are 20.34% and 28.94%. For DGMC, the average forecasting error of 2014 is 15.91% which is lower than the APE in 2013, but goes up to 40.44% in 2015. In terms of DGM, the ARPEs are 22.92% and 32.02% in 2014 and 2015. As to EM's, the forecasting errors reach to 38.19% in 2014 and 51.85% in 2015. These forecasting errors are up to four times greater than CNDGM's.

In summarizing, taking all the results of the long and short sequences into account, the CNDGM(1,1) model of interval grey numbers with central points proposed in this paper for the short data sequences reveals the best accuracy among the other seven models, not only in the aspect of fitting errors but also forecasting errors. It shows great performance and effectiveness of this model when applied to complicated and short data sequences.

In regards to the long-term forecasting of electricity consumption, this can be obtained by using average values of every three years. By this means, the smoothness of data sequences can be enhanced which may facilitate the models' simulations. Additionally, the average data of every three years are applicable for medium-long term forecasting. Following this up, the per capita annual electricity consumption in southern Jiangsu area up to 2031 can be obtained using the eight models above by extrapolating the in-sample data from 1993-2013. The fitting errors are displayed in Table 2.

Table 2

The simulation errors of eight models, based on every-three-year average values of the long data sequences

| RPE | The lower bound (a) | The central point (b) | The upper bound(c) | ARPE | a | b | c | ARPE |
|-----|--------------------------|----------------------------|-------------------------|------|---|---|---|------|
|-----|--------------------------|----------------------------|-------------------------|------|---|---|---|------|

| | | | | | | | | | |
|-------|-------|--------|--------|--------|------|--------|--------|--------|--------|
| DGM | 6.88% | 11.53% | 19.08% | 12.50% | DES | 12.89% | 16.07% | 21.82% | 16.93% |
| DGMC | 3.90% | 8.11% | 13.05% | 8.35% | LM | 19.09% | 22.33% | 32.93% | 24.78% |
| CNDGM | 6.71% | 8.90% | 12.45% | 9.35% | EM | 7.69% | 7.87% | 11.20% | 8.92% |
| GAGM | 5.65% | 11.34% | 20.82% | 12.61% | ARMA | 10.21% | 13.46% | 31.06% | 18.24% |

As seen from Table 2, the fitting errors of average values of every three year period by using DGMC, CNDGM and EM, show relatively high fitting precision (all under 10%). To compare the development trends by using these three models, the forecasting values are then shown in Table 3 and the future trends of the per capita annual electricity consumption in southern Jiangsu area are displayed in Fig.5.

Table 3

Future per capita annual electricity consumption in southern Jiangsu area (2014-2031) Unit: kWh/person

| Year | DGMC | | | CNDGM | | | EM | | |
|-----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| | a | b | c | a | b | c | a | b | c |
| 2014-2016 | 9202.47 | 16985.07 | 31354.11 | 9079.98 | 15393.69 | 25505.23 | 9346.82 | 17936.57 | 30775.46 |
| 2017-2019 | 11953.30 | 23972.33 | 48123.93 | 10640.75 | 17448.51 | 29005.66 | 12632.76 | 25518.83 | 46200.78 |
| 2020-2022 | 15445.66 | 31949.13 | 69553.84 | 13842.37 | 22551.72 | 38573.41 | 17073.91 | 36306.32 | 69357.60 |
| 2023-2025 | 19081.87 | 43662.56 | 105065.6 | 16223.21 | 25568.21 | 44060.33 | 23076.38 | 51653.95 | 104121.1 |
| 2026-2028 | 22720.72 | 57318.13 | 154644.3 | 20372.75 | 31769.31 | 55928.48 | 31189.07 | 73489.44 | 156308.8 |
| 2029-2031 | 24756.22 | 75399.64 | 232548.2 | 23848.46 | 36039.51 | 64080.12 | 42153.84 | 104555.3 | 234654.1 |

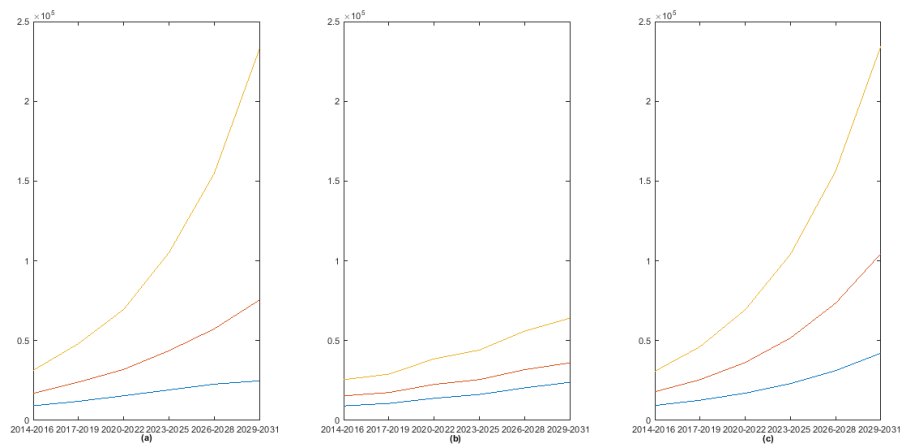


Fig.6. The forecasting values of the per capita annual electricity consumption in southern Jiangsu area by using DGMC (a), CNDGM (b) and EM (c)

1 According to forecasting values, the per capita annual electricity consumption in the southern
2 Jiangsu area will keep increasing. The growth rates of DGMC, CNDGM and EM are considerably
3 different. Specially, the trends of DGMC and EM (in Fig. 6(a) and (c)) totally exceed that of
4 CNDGM, and the growth rates of each sequence by using the former two models are all around
5 30%-50%. Furthermore, the magnitude of data will reach 2,000,000 for the upper sequences by
6 DGMC and EM which means the per capita annual electricity consumption in the southern
7 Jiangsu area will expand 10 times in the next 18 years. On the other hand, the forecasting values
8 by CNDGM proposed in this paper will experience a more moderate rise and the growth rates will
9 be between 13% and 33% for every three years. Additionally, for the upper bound sequence, by
10 2030 the value will reach to 64080.12 kWh/person which is 1.5 times more than 2014-2016's. It
11 shows that the growth rates of the highest levels in southern Jiangsu area will decrease slowly in
12 general, although the values of the per capita annual electricity consumption will continue to rise.
13 For the central-point sequence, the value of 2029-2031 will be 36039.51 kWh/person - a 1.34
14 times' increase compared with the value of 2014-2016. The growth rates' trend of the central-point
15 sequence during 2014-2031 is similar with the upper bound sequence. For the lower bound
16 sequence, by 2030 it will reach to 23848.46 kWh/person, which is more than 1.6 times compared
17 to that of 2014-2016. It is notable that the growth rates of the lower bound sequence are slightly
18 higher than the other two sequences'.

24 Focusing on the recent years (2013-2015) seen in Fig.3, the trends of per capita annual
25 electricity consumption in the southern Jiangsu area are tending to stabilize, that is to say, the
26 future growth rates of electricity consumption may not be as high as the last few years
27 (2001-2012). As China has implemented a series of measures to cut production capacity and adjust
28 industrial structure, the low growth rates of per capita annual electricity consumption reveal the
29 effects of these energy policies. Accordingly, the forecasting values of CNDGM show the
30 probable future trends considering the continuous energy policy on reducing energy consumption
31 in China. With respect to the other two models (DGMC and EM), the forecasting values
32 demonstrate that the energy consumption will experience high rises along with the increasing
33 trend from 2001, which ignore the effects of energy policy in China. In other words, taking into
34 account the energy policy of China in recent years, there is no difference for the development
35 trend of the per capita annual electricity consumption in southern Jiangsu in the future by using
36 DGMC and EM. Such a conclusion seriously underestimates the impact of the relevant energy
37 policy. In fact, the impact of these energy policies is clearly visible in Fig. 3: the steep curve
38 changes its direction in 2013. The data for this new trend is overwhelmed by the dominant
39 historical data for previous years, this is the reason why this trend is ignored by DGMC and EM.
40 For CNDGM, its association with the recent data is further enforced by our new method, and it
41 captures the new trends ignored by the other two models. In short, CNDGM provides a possible
42 way forward for forecast deviations, which reveals the hidden effectiveness of China's measures to
43 reduce energy consumption. The forecasting values by 2031 can be obtained by using the values
44 from CNDGM as the lower bound, and the values from DGMC and EM as the upper bound. The
45 ranges between the upper and the lower bound notify the importance of a series of energy
46 optimization policy.

56 To sum up, the continuous growth of the per capita annual electricity consumption in the
57 southern Jiangsu area will lead to sustained pressure on energy planning and configuration, as well
58 as restrict economic development.
59
60
61
62
63
64
65

5 Conclusions

A novel energy consumption forecasting model is proposed by optimizing the initial conditions of DGM (1, 1) model and establishing models of interval grey numbers with central points, in which the data sequences are fully transformed as the sequence of middle values, the sequence of trapezoid areas of the lower grey layer and the sequence of trapezoid areas of the upper grey layer. Based on the empirical analysis of per capita annual electricity consumptions of southern Jiangsu, the following conclusions are obtained:

- (1) The fitting and prediction results exhibits higher accuracy in short-term sequence modelling (the fitting and forecasting ARPEs are only 2.19% and 3.7%) compared with other forecasting models (DGM, DGMC, GAGM, DES, LM, EM, ARMA), regardless if the results stem from long-term or short-term sequences. It provides a feasible method for the prediction of interval grey numbers with central points for regional data of energy consumption and expands the research fields for different types of EC data.
- (2) For the long-term forecasting of electricity consumption (to the year 2031), the prediction values have been obtained by establishing three-year average values and the outcomes display high fitting precision by the model proposed in this paper compared with the other three grey models (DGM, DGMC, GAGM). The future trends of per capita annual electricity consumption of southern Jiangsu will continue to rise in the near future, but the growth rate will probably slow to more moderate levels due to the effects of sustainable energy policy in China.

Although the high precision results are obtained in the novel model, there are huge challenges involved in the forecasting of energy consumption in practice because of its trend variability and regional imbalance revealed in this paper. Taking into account the expansion of absolute amounts in more advanced cities and the increase in the growth rates in less advanced cities, it is still a massive pressure in controlling the electricity consumption in the southern Jiangsu area.

The case in this paper is simply to clarify the process of this novel model and shows the effectiveness of this novel model. Actually, this model can be applied in many other aspects of energy forecasting. For further research in the future, sufficient data should be collected for the models' verification compared with classic statistical models and intelligent algorithms. While most of the forecasting methods need more than 30 data sets for modeling, the data in this paper total just 23 data sets (due to the changes in statistical standards in the Statistical Yearbook of Jiangsu Province), which is more suitable for grey models.

6 Acknowledgments

The authors are grateful to anonymous referees for their helpful and constructive comments on this paper. This work was supported by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme entitled "Grey Systems and Its Application to Data Mining and Decision Support" Grant No. FP7-PIIF-GA-2013-629051, a project of the Leverhulme Trust International Network entitled "Grey Systems and Its Applications"(IN-2014-020). The authors would also like to acknowledge the support of the

National Natural Science Foundation of China (71771119, 71701101, 91746202).

REFERENCES

- Al-mulali U., Fereidouni H.G., Lee J.Y.M., 2014. Electricity consumption from renewable and non-renewable sources and economic growth: Evidence from Latin American countries. *Renewable and Sustainable Energy Reviews*. 30, 290–298.
- Dang Y.G., Liu S.F., 2004. The GM models that $x(n)$ be taken as initial value. *Kybernetes*. 33, 247–255.
- Deng J.L., 1982. The Control problem of grey systems. *Syst. Control Lett.* 1 (5), 288–294.
- Ding S., 2019. A novel discrete grey multivariable model and its application in forecasting the output value of China's high-tech industries. *Industrial Engineering*. 127, 749-760.
- Ding S., 2018. A novel self-adapting intelligent grey model for forecasting China's natural-gas demand. *Energy*. 162, 393-407.
- Ding S., Dang Y.G., Xu N., Wang J.J., Xu Z.D., 2018. A Novel Grey Model Based on the Trends of Driving Factors and Its Application. *The Journal of Grey System*. 30(3):105-126.
- Erdal, K., Baris, U., Okyay, K., 2010. Grey system theory-based models in time series prediction. *Expert Syst Appl.* 37(2), 1784–1789.
- Felice M.D., Alessandri A., Catalano F., 2015. Seasonal climate forecasts for medium-term electricity demand forecasting. *Applied Energy*. 137, 435–444.
- Hamzacebi C., Es H.A., 2014. Forecasting the annual electricity consumption of Turkey using an optimized grey model. *Energy*. 70, 165–171.
- Lee Y.S., Tong L.I., 2012. Forecasting nonlinear time series of energy consumption using a hybrid dynamic model. *Applied Energy*. 94, 251–256.
- Liu S.F., Dang Y.G., Fang Z.G., Xie N.M., 2010. *Grey System Theory and Its Application* (5th ed.). Science Press, Beijing. 15,122–123 & 146–149.
- Liu S.F., Xie N.M., 2011. A new grey evaluation method based on reformative triangular whitenization weight function. *Journal of Systems Engineering*. 31(2), 72–78.
- Truong D.Q., Ahn K.K., 2012. Wave prediction based on a modified grey model MGM(1,1) for real-time control of wave energy converters in irregular waves. *Renewable Energy*. 43, 242–255.
- Wang Y.H., Dang Y.G., Liu S.F., 2010. An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition. *Expert Syst Appl.* 37, 5640–5644.
- Wu, L., Liu, S., Yao, L., Yan, S., Liu, D., 2013a. Grey system model with the fractional order accumulation. *Commun Nonlinear Sci.* 18(7), 1775–1785.
- Wu L.F., Liu S.F., Yan S.L., 2013b. Grey Prediction Model for Hybrid Sequence. *Control and Decision*. 28(12), 1912–1914, 1920.
- Xia M., Wong W.K., 2014. A seasonal discrete grey forecasting model for fashion retailing. *Knowledge-Based Systems*. 57, 119–126.
- Xie N.M., Liu S.F., 2009. Discrete grey forecasting model and its optimization. *Applied Mathematical Modeling*, 33, 1173–1186.
- Xie N.M., Liu S.F., Yang Y.J., Yuan C.Q., 2013. On novel grey forecasting model based on non-homogeneous index sequence. *Applied Mathematical Modeling*. 37, 5059–5068.
- Xiong P.P., Dang Y.G., Yao T.X., Wang Z.X., 2014. Optimal modeling and forecasting of the energy consumption

and production in China. *Energy*. 77, 623–634.

Xu N., Dang Y.G., Gong Y.D., 2017. Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China. *Energy*. 118, 473–480.

Yamaguchi D., Li G.D., Nagai M., 2007. A grey-based rough approximation model for interval data processing. *Information Sciences*. 177, 4727–4744.

Yukseltan E., Yucekay A., Bilge A.H., 2017. Forecasting electricity demand for Turkey: Modeling periodic variations and demand segregation. *Applied Energy*. 193, 287–296.

Zahedi G., Azizi S., Bahadori A., Elkamel A., Alwi S.R.W., 2013. Electricity demand estimation using an adaptive neuro-fuzzy network: a case study from the Ontario province-Canada. *Energy*. 49, 323–328.

Zeng B., Li C., Long X.J., Xiong Y., Zhou X.Y., 2014. A Novel Interval Grey Number Prediction Model Given Kernel and Grey Number Band. *Journal of Grey System*. 26(3), 69–84.

Zeng B., Liu S.F., Cui J., 2010. Prediction model for interval grey number with known whitenization weight function. *Control and Decision*. 25(12), 1815–1820.

Zeng B., Liu S.F., 2011. Prediction model of interval grey number based on its geometrical characteristics. *Journal of System Engineering*. 26(2), 174–180.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65

A novel energy consumption forecasting model combining an optimized DGM (1, 1) model with interval grey numbers

Jing Ye¹, Yaoguo Dang², Song Ding^{3*}, Yingjie Yang⁴

1 School of Management Science and Engineering, Nanjing University of Finance & Economics, Nanjing, Jiangsu, 210023, China

2 College of Economics and Management, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu, 210016, China

3 School of Economics, Zhejiang University of Finance & Economics, Hangzhou, Zhejiang, 310018, China

4 Centre for Computational Intelligence, De Montfort University, Leicester, LE1 9BH, UK

Abstract: Since energy consumption (EC) is becoming an important issue for sustainable development in the world, it has a practical significance to predict EC effectively. However, there are two main uncertainty factors affecting the accuracy of a region's EC prediction. Firstly, with the ongoing rapid changes in society, the consumption amounts can be non-smooth or even fluctuating during a long time period, which makes it difficult to investigate the sequence's trend in order to forecast. Secondly, in a given region, it is difficult to express the consumption amount as a real number, as there are different development levels in the region, which would be more suitably described as interval numbers. Most traditional prediction models for energy consumption forecasting deal with long-term real numbers. It is seldom found to discover research that focuses specifically on uncertain EC data. To this end, a novel energy consumption forecasting model has been established by expressing ECs in a region as interval grey numbers combining with the optimized discrete grey model (DGM(1,1)) in Grey System Theory (GST). To prove the effectiveness of the method, per capita annual electricity consumption in southern Jiangsu of China is selected as an example. The results show that the proposed model reveals the best accuracy for the short data sequences (the average fitting error is only 2.19% and the average three-step forecasting error is less than 4%) compared with three GM models and four classical statistical models. By extension, any fields of EC, such as petroleum consumption, natural gas consumption, can also be predicted using this novel model. As the sustained growth in EC of China's, it is of great significance to predict EC accurately to manage serious energy security and environmental pollution problems, as well as formulating relevant energy policies by the government.

Key Words: energy consumption; Grey system theory; interval number; prediction; electricity

* Corresponding author : dingsong1129@163.com (S. Ding)

E-mail address: yejingjenson@163.com (J. Ye); iamdangyg@163.com (Y.G. Dang); dingsong1129@163.com (S. Ding); yyang@dmu.ac.uk (Y.J. Yang)

1 Introduction

Today, energy remains the main engine of the economy. Energy consumption has been one of the hottest issues in energy research over the years. In China, as rapid and diversified society developments continue, EC forecasting has become more and more complicated and full of uncertainty.

To be specific, with the economic growth and transformation of industrial structures, per capita household EC sequence (1993-2014) in China show a non-smooth rising trend and is difficult to fit its curve trend (Data source: National Statistical Yearbook of China, 2016), thus it is not suitable for forecasting by using the whole sequence to establish models. In fact, energy data capacity is often very limited (15 observations or less) due to changes of energy policy or updates within the industrial structure. However, classical prediction models (e.g. exponential smoothing, Holt-Winters method, linear method, exponential model, regression, autoregressive integrated moving average (ARIMA)) or modern prediction techniques (e.g. genetic algorithm (GA), artificial neural networks (ANN), Particle Swarm Optimization (PSO), fuzzy system) often require larger sample sizes, based on mathematical statistics principles or machine learning rules and computational intelligence (Xia and Wong, 2014). To be specific, for EC forecasting, Yukseltan et al. (2017) developed a linear model, as an expansion in Fourier series, supplemented with a modulation by seasonal harmonics, to forecast electricity demand for Turkey. Al-mulali et al. (2014) conducted panel data (labour, gross fixed capital formation, and total trade) analysis, based on statistics to describe their relationships with renewable EC and non-renewable EC. Zahedi et al. (2013) predicted the electricity demand of Canada by using a neuro-fuzzy model. Felice et al. (2015) presented an assessment of the use of seasonal climate forecasts of temperature for medium-term electricity demand prediction, by using both deterministic and probabilistic forecasting approaches. The classical models and the modern methods mentioned above are applied to cases with sufficient sample data which cannot capture the sensitive latest changes of data trends.

In reality, for energy data samples where the trends show significant changes, their newest or most recent information play the most important role in the modeling process rather than the amount of data. Owing to these reasons, Grey Model (GM) in GST (Liu et al., 2010) is introduced to forecast EC with small amounts of newest data. GM is suitable for small samples which emphasizes on new information, and it has been verified by many scholars: according to the principle of new information priority, Hamzacebi and Es (2014) used Optimized Grey Modeling (1,1) to forecast the annual electricity consumption of Turkey; Xiong et al. (2014) proposed a novel GM (1,1) model based on optimizing initial condition and applied in China's energy consumption and production forecasting. Moreover, it has been proved that grey prediction models can predict short-term data with the most recent information more efficiently compared with using long term old data (Wu et al., 2013a). Various papers on energy forecasting, using Grey Models or hybrid models with Grey Models, have now been disseminated. Ding (2018) designed a self-adapting grey prediction model having a nonlinear optimized initial value by utilizing an ant lion optimizer (ALO) algorithm to forecast China's natural-gas demand. Xu et al. (2017) proposed a novel grey model with optimal time response, based on particle swarm algorithm to forecast the electricity consumption data of China and compared with three alternative grey models. Truong et al. (2012) established a modified grey model MGM(1,1) for real-time control of wave energy

converters in irregular waves and the traditional grey model and a typical autoregressive (AR) model were carried out for comparisons. Ding designed a new multivariable model that considers the accumulative effects of the influencing factors, and used this proposed model to accurately forecast the future output value of high-tech industries in China. Lee and Tong (2012) combined a dynamic grey model with genetic programming to forecast energy in the US and compared with ARIMA, GP and other grey models. Erdal et al. (2010) investigated the accuracies of different grey models such as GM(1,1), Grey Verhulst model, modified grey models using Fourier series, and the fitting and forecasting results of the modified grey models revealed higher performances. Ding et al. (2018) analyzed the properties of a multiple transformation and optimized background value in the GM(1,N) model to simulate and predict CO₂ emissions. In summary, none of the above papers considered regional unbalances in a certain country or region, and the research objects of them are all described as single real numbers. For model comparisons, most of the papers only took into account grey models, because of the short-term sequences, instead of using forecasting models, such as statistical models, which is neither comprehensive nor convincing.

Due to the regional characteristics of EC, real values are difficult to ascertain when deciphering the different levels of per capita annual ECs in one region. The interval numbers are introduced to characterize the annual ECs which can cover the full information of this energy data. Focusing on the feature of the uncertainty, interval grey numbers have been discussed in many other fields, like decision-making, information retrieval, knowledge discovery and so on (Yamaguchi et al., 2007). Professor Deng Julong (1982) originally proposed interval grey numbers, whereby the interval grey number is a typical form of grey number which refers to uncertain values of the numbers in an interval or a general number set. In terms of grey prediction of interval grey numbers, researchers have made a series of attempts to transform interval grey number sequences into real number sequences, to construct grey prediction models by means of using geometric characteristics, formations of operators and establishments of combination models etc., and then restore back to the interval grey number sequences. Zeng et al. (2014) converted an interval grey number sequence into the kernel and area sequences to establish grey forecasting models. This method avoids extreme data errors by generating average value and grey degree enlargement scenarios. But the characteristic of non-homogeneous sequences is not taken into account. Wu et al. (2013b) converted an interval grey number sequence into a real number sequence by calculating the areas of grey number layers and the cognition degree of interval grey numbers. In this method, the surface areas are introduced to accurately express the areas of grey number layers. The solution of the equations is quite complex in the process of calculation. In reality, most information distributions of interval grey numbers are not equal. The above grey prediction models, based on interval grey numbers can not reflect this outstanding feature. Thus, Zeng Bo et al. (2010) introduced the typical whitenization weight function to establish the grey prediction model. In this model, the upper and lower bounds of interval grey numbers and two turning points in whitenization weight function are all used without omitting any information of the grey number. It is usually difficult to obtain the typical whitenization weight function which needs to determine two turning points. And there is no doubt of the high costs taken by synthesizing experts' opinions or experiences as well as collecting and analyzing previous data to obtain this function, which also leads to more deviations. To simplify the typical whitenization weight function, a central point which represents the maximum possibility of an interval can be considered to replace two turning points in the typical whitenization weight function. By these means, this paper utilizes the central-point sequence to establish a novel energy consumption

forecasting model, which is easier to acquire, in describing uneven distributions of information of interval grey numbers.

To optimize grey model, the initial condition of DGM(1, 1) is reset by minimizing the fitting values' errors in this paper. Instead of using the first data set as an initial condition in the original GM (Liu et al., 2010), it ignores the extraction of new information. In terms of other initial condition optimizations, Dang et al. (2004) set the last in-sample data's accumulated generating value($x^{(1)}(n)$) as the initial condition which put too much emphasis on the latest information and became vulnerable to abnormal fluctuations; Liu et al. (2011) discussed three initial conditions: the starting-point fixed initial value, the middle-point fixed initial value and the end-point fixed initial value, though there was no indication of the suitability of each setting; Wang et al. (2010) used the weighted combination of the first and last data set from the 1-AGO sequence as the initial condition, but how to determine the weight had not been given; Xiong et al. (2014) expanded Wang's method by weighting all the in-shape's accumulated generating data as the initial condition. The weights were fixed which meant the weights could not adjust according to sequences. To this end, this paper optimizes the initial condition of the grey forecasting model by setting a parameter which makes the fitting accuracy minimum.

To sum up, in-sample data sets in this paper are characterized as interval grey numbers to establish the novel energy consumption forecasting model to forecast EC. The innovation and contribution of this research compared with other references mainly lies in the following three aspects: Firstly, the uncertain information of regional energy consumption is expressed by interval grey numbers, in which the upper bound and the lower bound are used to characterize the highest and the lowest levels of this region. Secondly, short-term most recent energy consumption data are used to establish the novel forecasting model, and long-term data are applied as comparisons. Thirdly, the accuracy of the novel forecasting model is improved by full information transformation of energy consumption data, described by interval grey numbers and the model's initial value optimization.

The reminder of this paper is structured as follows: in Section 2, information of the interval grey numbers with central points is transformed fully to prepare for forecasting. Section 3 optimizes the DGM(1, 1) of interval grey numbers with central points and the initial condition to improve the energy consumption forecasting model's precision. Section 4 discusses a case of per capita annual EC from five cities in Southern Jiangsu to forecast near future trends and verifies the practicability of the novel model by comparing with other forecasting methods. Finally, conclusions are drawn in Section 5.

2 The effective information conversion of interval grey numbers with central points

2.1 The basic concepts

Definition 1 (Liu et al., 2010). Let $\otimes \in [a, b]$, $a \leq b$, $a, b \in R$, then \otimes is called an interval grey number; if $a = b$, then \otimes is a real number and $\otimes = a$.

Definition 2. Let $\otimes = [a, b, c]$ be an interval grey number, $a \leq b \leq c$. Among them, a , c are the lower bound and the upper bound of the interval grey number, and b is the maximum possibility point of this interval which is called the central point. This kind of interval grey number is named as an interval grey number with a central point. The sequence of interval grey numbers with central points can be described as $X(\otimes) = (\otimes_1, \otimes_2, \dots, \otimes_n)$ in which $\otimes_m \in [a_m, b_m, c_m]$, $a_m \leq b_m \leq c_m$, $m = 1, 2, \dots, n$.

Definition 3. Let the sequence of interval grey numbers with central points as $X(\otimes) = (\otimes_1, \otimes_2, \dots, \otimes_n)$, and map all elements of $X(\otimes)$ into a two-dimensional Cartesian coordinate system. Then, sequentially connect the lower bounds, the upper bounds and the central points of adjacent interval grey numbers respectively to constitute the lower-bound line, the upper-bound line and the central-point line, which can be seen in Fig. 1. Among them, the sequence of the lower-bound line can be denoted as $X(\underline{\otimes}) = (a_1, a_2, \dots, a_n)$, the sequence of the central-point line can be denoted as $X(\tilde{\otimes}) = (b_1, b_2, \dots, b_n)$, and the sequence of the upper-bound line can be denoted as $X(\overline{\otimes}) = (c_1, c_2, \dots, c_n)$.

Definition 4 (Zeng et al., 2010). The graphics of the lower and upper bounds sequences of interval grey numbers are known as grey numbers' belts; the sections of the adjacent interval grey numbers are called grey numbers' layers. According to positions of grey numbers' layers in grey numbers' belts, they are marked by $1, 2, \dots, n$ (See in Fig. 1).

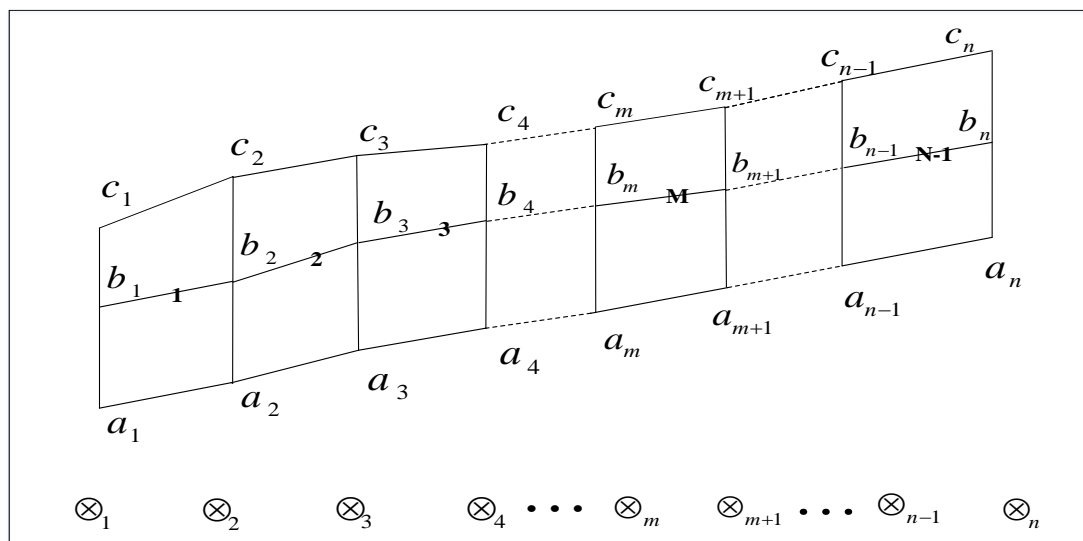


Fig.1. The grey numbers' belt of the interval grey number $\otimes = [a, b, c]$

Definition 5. If the grey numbers' layers (GL) are marked by $1, 2, \dots, n$, each GL can be divided into two parts by the central-point sequence. The two parts can be called: the upper grey layer and the lower grey layer.

2.2 Establish the effective information conversion sequences

Let the sequence of interval grey numbers with the central point as $X(\otimes) = ([a_1, b_1, c_1], [a_2, b_2, c_2], \dots, [a_n, b_n, c_n])$ in which a_m and c_m are the lower and upper bound values, b_m is the central point. The effective information conversion sequences can be established by the sequences of trapezoid areas made by upper bound values and central-point values; trapezoid areas made by lower bound values and central-point values and middle values calculated by upper bound values and lower bound values.

Lemma 1 (Zeng and Liu, 2011). In the grey number layer m (see in Fig. 1), the longitudinal coordinate of the middle value calculated by upper bound values and lower bound values ($a_m, c_m, a_{m+1}, c_{m+1}$) is:

$$L_m = \frac{a_m + c_m + a_{m+1} + c_{m+1}}{4}, \quad m = 1, 2, \dots, n-1.$$

According to Lemma 1, the sequence of middle values can be denoted as $X(L) = (L_1, L_2, \dots, L_{n-1})$.

Theorem 1 (Trapezoid areas of the lower grey layer). In the grey number layer m (see in Fig. 1), the trapezoid area of lower bound values (a_m, a_{m+1}) and central-point values (b_m, b_{m+1}) is:

$$\underline{S}_m = \frac{(b_{m+1} - a_{m+1}) + (b_m - a_m)}{2}, \quad m = 1, 2, \dots, n-1.$$

Proof: The differences of longitudinal coordinates between the lower bounds and the central points are $(b_{m+1} - a_{m+1}), (b_m - a_m)$ which represent the trapezoid's top and bottom. And the height of the trapezoid can be obtained by the differences of the adjacent grey number layers, which is $(m+1) - m = 1$. By this means, the trapezoid area is $\underline{S}_m = \frac{(b_{m+1} - a_{m+1}) + (b_m - a_m)}{2}$.

Theorem 2 (Trapezoid areas of the upper grey layer). In the grey number layer m (see in Fig. 1), the trapezoid area of upper bound values (c_m, c_{m+1}) and central-point values (b_m, b_{m+1}) is:

$$\bar{S}_m = \frac{(c_{m+1} - b_{m+1}) + (c_m - b_m)}{2}, \quad m = 1, 2, \dots, n-1.$$

Proof: Similar to Theorem 1.

According to Theorem 1 and Theorem 2, here are the sequence of trapezoid areas of the lower grey layer $X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1})$ and the sequence of trapezoid areas of the upper grey layer $X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1})$.

In addition, from Lemma 1, Theorem 1 and Theorem 2, we can obtain:

$$a_{m+1} + a_m = 2L_m - \bar{S}_m - \underline{S}_m$$

$$b_{m+1} + b_m = 2L_m - \bar{S}_m + \underline{S}_m$$

$$c_{m+1} + c_m = 2L_m + \bar{S}_m + \underline{S}_m.$$

That is to say,

$$X(\otimes) = ([a_1, b_1, c_1], [a_2, b_2, c_2], \dots, [a_n, b_n, c_n]) \Leftrightarrow \begin{cases} X(L) = (L_1, L_2, \dots, L_{n-1}) \\ X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1}) \\ X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1}) \end{cases} .$$

The information contained in the interval grey numbers' sequence with central points are retained completely by the information conversion which means the conversion preserves the characteristics of information equivalence and data integrity.

3 The optimized DGM(1, 1) model

3.1. DGM(1,1) of interval grey numbers with central points

In the modeling process of discrete grey models, parameters estimation, simulation and prediction are all adopted in the form of discrete equations. **The DGM model does not differentiate approximate substitution between the discrete equation and the continuous equation compared with GM(1, 1) model, in which the method of estimating parameters is the discrete equation and the method of simulating and forecasting is the continuous equation, so it has higher fitting precision by calculating the average relative percentage error (ARPE) (Xie and Liu, 2009).** Let $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n))$ be the original data sequence, $x^{(0)}(k) \geq 0 (k=1, 2, \dots, n)$. Based on the traditional DGM(1, 1) model (Xie and Liu, 2009), here are: the recursive function of DGM(1,1) (set $x^{(1)}(1) = x^{(0)}(1)$)

$$\hat{x}^{(1)}(k+1) = \beta_1^k x^{(0)}(1) + \frac{1-\beta_1^k}{1-\beta_1} \cdot \beta_2, (k=1, 2, \dots, n-1) \quad (1)$$

and the restored value of the recursive function of DGM(1,1) at (k+1)

$$\hat{x}^{(0)}(k+1) = \alpha^{(1)} \hat{x}^{(1)}(k+1) = \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k), (k=1, 2, \dots, n-1) \quad (2)$$

Similarly, for the sequence of middle values $X(L) = (L_1, L_2, \dots, L_{n-1})$, the time response sequence of its DGM(1,1) is proposed by Zeng et al. (2010):

$$\hat{L}_{k+1}^{(1)} = \alpha_1^k L_1 + \frac{1-\alpha_1^k}{1-\alpha_1} \alpha_2 \quad (3)$$

Note: α_1, α_2 are parameters of DGM(1,1).

The restored value of the prediction model is:

$$\hat{L}_{k+1} = \alpha_1^{k-1} [L_1 (\alpha_1 - 1) + \alpha_2] \quad (4)$$

From Formula (4), we can obtain the parameters' forecasting values of the sequence of middle values made up by the upper and lower bounds values. According to Lemma 1, it contains the "plus" information of the lower and upper bounds of interval grey numbers.

From $L_m = \frac{a_m + c_m + a_{m+1} + c_{m+1}}{4}$, here is:

$$\hat{a}_{m+1} + \hat{c}_{m+1} = 4\hat{L}_m - (\hat{a}_m + \hat{c}_m) \quad (5)$$

When $m = k - 1$,

$$\hat{a}_k + \hat{c}_k = 4\hat{L}_{k-1} - 4\hat{L}_{k-2} + \cdots + (-1)^{k-1}(a_1 + c_1) \quad (6)$$

According to Formula (4), the first $(k - 2)$ items in Formula (6) consist of a geometric series whose common ratio is $q = -\alpha_1^{-1}$. $\hat{a}_k + \hat{c}_k$, which can be calculated by the summation formula of geometric series:

$$\hat{a}_k + \hat{c}_k = \frac{4\alpha_1^{k-3}[L_1(\alpha_1 - 1) + \alpha_2][1 - (-\alpha_1^{-1})^{k-2}]}{1 + \alpha_1^{-1}} + (-1)^k(a_2 + c_2) \quad (7)$$

For the sequence of trapezoid areas of the lower grey layer $X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1})$, the time response sequence of its DGM(1,1) is:

$$\hat{\underline{S}}_{k+1}^{(1)} = \beta_1^k \underline{S}_1 + \frac{1 - \beta_1^k}{1 - \beta_1} \beta_2 \quad (8)$$

Note: β_1, β_2 are parameters of DGM(1,1).

The restored value of the prediction model is:

$$\hat{\underline{S}}_{k+1} = \beta_1^{k-1}[\underline{S}_1(\beta_1 - 1) + \beta_2] \quad (9)$$

From Formula (9), we can obtain parameters' forecasting values of the sequence of trapezoid areas made by lower bound values and central-point values.

For the sequence of trapezoid areas of the upper grey layer $X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1})$, the time response sequence of its DGM(1,1) is:

$$\hat{\bar{S}}_{k+1}^{(1)} = \gamma_1^k \bar{S}_1 + \frac{1 - \gamma_1^k}{1 - \gamma_1} \gamma_2 \quad (10)$$

Note: γ_1, γ_2 are parameters of DGM(1,1).

The restored value of the prediction model is:

$$\hat{\bar{S}}_{k+1} = \gamma_1^{k-1}[\bar{S}_1(\gamma_1 - 1) + \gamma_2] \quad (11)$$

From Formula (11), we can obtain the parameters' forecasting values of the sequence of trapezoid areas made by upper bound values and central-point values.

According to Theorem 1 and Theorem 2, they contain the "minus" information of the lower and upper bounds of interval grey numbers.

Similarly, from $\underline{S}_m = \frac{(b_{m+1} - a_{m+1}) + (b_m - a_m)}{2}$ and $\bar{S}_m = \frac{(c_{m+1} - b_{m+1}) + (c_m - b_m)}{2}$,

here are:

$$\hat{c}_{m+1} - \hat{a}_{m+1} = 2(\hat{\underline{S}}_m + \hat{\bar{S}}_m) - (\hat{c}_m - \hat{a}_m) \quad (12)$$

When $m = k - 1$,

$$\hat{c}_k - \hat{a}_k = 2(\hat{\underline{S}}_{k-1} + \hat{\bar{S}}_{k-1}) - 2(\hat{\underline{S}}_{k-2} + \hat{\bar{S}}_{k-2}) + \cdots + (-1)^{k-1}(c_1 - a_1) \quad (13)$$

According to Formula (9) and Formula (11), the first $(k - 2)$ items in Formula (13) consist of two geometric series whose common ratio are $q' = -\beta_1^{-1}$ and $q'' = -\gamma_1^{-1}$. $\hat{c}_k - \hat{a}_k$ can be calculated by the summation formula of geometric series:

$$\hat{c}_k - \hat{a}_k = \frac{2\beta_1^{k-3}[\underline{S}_1(\beta_1 - 1) + \beta_2][1 - (-\beta_1^{-1})^{k-2}]}{1 + \beta_1^{-1}} + \frac{2\gamma_1^{k-3}[\bar{S}_1(\gamma_1 - 1) + \gamma_2][1 - (-\gamma_1^{-1})^{k-2}]}{1 + \gamma_1^{-1}} + (-1)^k(c_2 - a_2) \quad (14)$$

Among them,

$$\hat{b}_k - \hat{a}_k = \frac{2\beta_1^{k-3}[\underline{S}_1(\beta_1 - 1) + \beta_2][1 - (-\beta_1^{-1})^{k-2}]}{1 + \beta_1^{-1}} + (-1)^k(b_2 - a_2) \quad (15)$$

$$\hat{c}_k - \hat{b}_k = \frac{2\gamma_1^{k-3}[\bar{S}_1(\gamma_1 - 1) + \gamma_2][1 - (-\gamma_1^{-1})^{k-2}]}{1 + \gamma_1^{-1}} + (-1)^k(c_2 - b_2) \quad (16)$$

Combine with Formula (7), Formula (15) and Formula (16),

$$\begin{cases} \hat{a}_k + \hat{c}_k = 4F_L\alpha_1^{k-3}[1 - (-\alpha_1^{-1})^{k-2}] + (-1)^k(a_2 + c_2) \\ \hat{c}_k - \hat{a}_k = 2F_{\underline{S}}\beta_1^{k-3}[1 - (-\beta_1^{-1})^{k-2}] + 2F_{\bar{S}}\gamma_1^{k-3}[1 - (-\gamma_1^{-1})^{k-2}] + (-1)^k(c_2 - a_2) \\ \hat{b}_k - \hat{a}_k = 2F_{\underline{S}}\gamma_1^{k-3}[1 - (-\beta_1^{-1})^{k-2}] + (-1)^k(b_2 - a_2) \end{cases}$$

Among them,

$$F_L = \frac{[L_1(\alpha_1 - 1) + \alpha_2]}{1 + \alpha_1^{-1}}, \quad F_{\underline{S}} = \frac{[\underline{S}_1(\beta_1 - 1) + \beta_2]}{1 + \beta_1^{-1}}, \quad F_{\bar{S}} = \frac{[\bar{S}_1(\gamma_1 - 1) + \gamma_2]}{1 + \gamma_1^{-1}}.$$

Note: F_L , $F_{\underline{S}}$, $F_{\bar{S}}$ are all seen as constant terms due to the confirmation of all the parameters.

Finally, we obtain the grey prediction model of interval grey numbers with central points:

$$\begin{cases} \hat{a}_k = 2F_L\alpha_1^{k-3}[1-(-\alpha_1^{-1})^{k-2}] - F_{\underline{S}}\beta_1^{k-3}[1-(-\beta_1^{-1})^{k-2}] - F_{\bar{S}}\gamma_1^{k-3}[1-(-\gamma_1^{-1})^{k-2}] + (-1)^k a_2 \\ \hat{b}_k = 2F_L\alpha_1^{k-3}[1-(-\alpha_1^{-1})^{k-2}] + F_{\underline{S}}\beta_1^{k-3}[1-(-\beta_1^{-1})^{k-2}] - F_{\bar{S}}\gamma_1^{k-3}[1-(-\gamma_1^{-1})^{k-2}] + (-1)^k b_2 \\ \hat{c}_k = 2F_L\alpha_1^{k-3}[1-(-\alpha_1^{-1})^{k-2}] + F_{\underline{S}}\beta_1^{k-3}[1-(-\beta_1^{-1})^{k-2}] + F_{\bar{S}}\gamma_1^{k-3}[1-(-\gamma_1^{-1})^{k-2}] + (-1)^k c_2 \end{cases} \quad (*)$$

In summary, the establishment steps of grey prediction models of interval grey numbers with central points are:

Step1 : From the sequence of interval grey numbers with central points $X(\otimes) = ([a_1, b_1, c_1], [a_2, b_2, c_2], \dots, [a_n, b_n, c_n])$, calculate the sequence of middle values made up by the upper and lower bound values $X(L) = (L_1, L_2, \dots, L_{n-1})$, the sequence of trapezoid areas of the lower grey layer $X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1})$ and the sequence of trapezoid areas of the lower grey layer $X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1})$.

Step2: Establish the DGM(1,1) for the above 3 sequences, and obtain parameters $(\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2)$.

Step3: Calculate $F_L, F_{\underline{S}}, F_{\bar{S}}$.

Step4: Calculate Equation Group (*) to obtain the forecasting values of the lower bounds, the upper bounds and central points.

3.2. Optimization of the initial conditions in the DGM(1,1)

As it mentioned in Section 3.1, the traditional DGM model reveals higher accuracy because there is no approximate substitution of continuous equations (Xie and Liu, 2009). However, the traditional DGM model seeks to align the exponential sequences (ca^k) modeling, ignoring many non-homogeneous exponential sequences $(ca^k + b)$ in reality. The interval grey number sequences mostly present non-homogeneous characteristics for their upper and lower bounds sequences, because of the uncertainties of their internal mechanism. To this end, Xie et al. (2013) proposed an approximate non homogeneous discrete grey model (NDGM(1,1)).

$$\begin{cases} \hat{x}^{(1)}(k+1) = \beta_1 \hat{x}^{(1)}(k) + \beta_2 \cdot k + \beta_3 \\ \hat{x}^{(1)}(1) = x^{(1)}(1) + \beta_4 \end{cases} \quad (17)$$

Based on the least square method, the parameters $\beta_1, \beta_2, \beta_3$ can be calculated:

$$\hat{\beta} = [\beta_1, \beta_2, \beta_3]^T = (B^T B)^{-1} B^T Y,$$

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

Moreover, the accuracy of this model is not only determined by the structural parameters $\hat{\beta}$, but also the initial condition. For the initial conditions of the traditional DGM, the starting-point value, middle-point value and end-point value are all discussed for the iterations respectively, to explain the influences for exponential sequence' fitting. The three forms of initial conditions of the traditional DGM are (Liu and Xie, 2011):

The starting-point fixed DGM(SDGM) is:

$$\begin{cases} \hat{x}^{(1)}(k+1) = \beta_1 \hat{x}^{(1)}(k) + \beta_2 \\ \hat{x}^{(1)}(1) = x^{(1)}(1) = x^{(0)}(1) \end{cases} \quad (18)$$

The middle-point fixed DGM(MDGM) is:

$$\begin{cases} \hat{x}^{(1)}(k+1) = \beta_1 \hat{x}^{(1)}(k) + \beta_2 \\ \hat{x}^{(1)}(1) = x^{(1)}(m) = \sum_{i=1}^m x^{(0)}(i), 1 < m < n \end{cases} \quad (19)$$

The end-point fixed DGM(EDGM) is:

$$\begin{cases} \hat{x}^{(1)}(k+1) = \beta_1 \hat{x}^{(1)}(k) + \beta_2 \\ \hat{x}^{(1)}(1) = x^{(1)}(n) = \sum_{i=1}^n x^{(0)}(i) \end{cases} \quad (20)$$

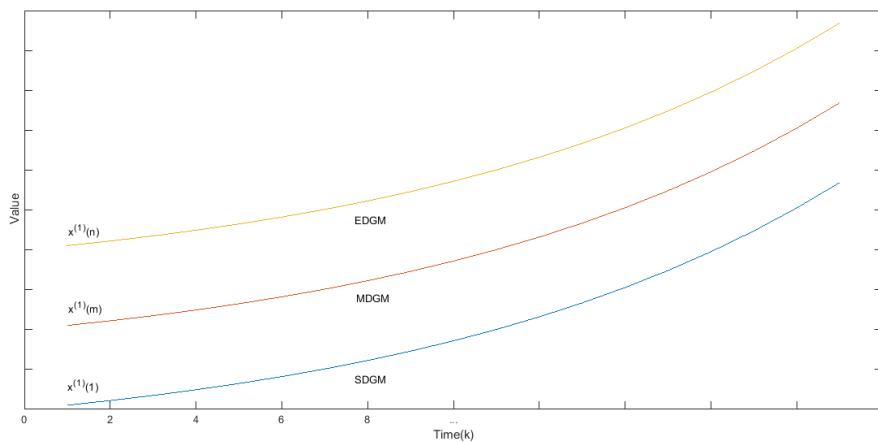


Fig.2. The effect of different initial conditions

The choice of the initial value will directly affect the first point and the trend of the values in

the model (shown in Fig. 2). If the initial condition of the model can be processed properly, there will be a positive effect on the accuracy of the model. For the SDGM and EDGM, the initial conditions are fixed (the starting-point value and the end-point value determine the first point) which may not be the optimal solution. For the MDGM, the middle-point value is hard to identify and the results may not be so accurate. In order to better explore the initial conditions, the optimization method of the initial condition should be put forward. In NDGM(1,1) (Xie et al., 2013), the initial condition is set by the first data instance as a reference and a corrective parameter β_4 ($\hat{x}^{(1)}(1) = x^{(1)}(1) + \beta_4$) which can be optimized by establishing a parameter in an overall consideration.

According to NDGM(1,1) ($\hat{x}^{(1)}(m+1) = \beta_1 \hat{x}^{(1)}(m) + \beta_2 \cdot m + \beta_3$), the recurrence model is obtained:

$$\hat{x}^{(1)}(m+1) = \beta_1^m \hat{x}^{(1)}(1) + \beta_2 \sum_{j=1}^m j \beta_1^{m-j} + \frac{1-\beta_1^m}{1-\beta_1} \beta_3, m=1,2,\dots \quad (21)$$

In order to establish a simple and effective optimization method, here the solution to establish the initial condition can be expressed as C to replace the first data and a corrective parameter β_4 :

$$\hat{x}^{(1)}(m+1) = C \cdot \beta_1^m + \beta_2 \sum_{j=1}^m j \beta_1^{m-j} + \frac{1-\beta_1^m}{1-\beta_1} \beta_3 \quad (22)$$

where C is an undetermined parameter.

To calculate the value of C by using a similar method like the least square principle, an unconstrained optimization model is established to minimize the sum of squares of errors between $\hat{x}^{(1)}(k)$ and $x^{(1)}(k)$: $\min_C \sum_{k=1}^n [\hat{x}^{(1)}(k) - x^{(1)}(k)]^2$ where the number of in-shape values is n .

We then, bring the obtained parameters $\beta_1, \beta_2, \beta_3$ into

$\hat{x}^{(1)}(k+1) = C \cdot \beta_1^k + \beta_2 \sum_{j=1}^k j \beta_1^{k-j} + \frac{1-\beta_1^k}{1-\beta_1} \beta_3$ and let $\hat{x}^{(1)}(1) = x^{(1)}(1)$, we set:

$$F(C) = \sum_{k=1}^n [\hat{x}^{(1)}(k) - x^{(1)}(k)]^2 = 0 + \sum_{k=1}^{n-1} [C \cdot \beta_1^k + \beta_2 \sum_{j=1}^k j \beta_1^{k-j} + \frac{1-\beta_1^k}{1-\beta_1} \beta_3 - x^{(1)}(k+1)]^2 \quad (23)$$

Here is

$$\frac{\partial F(C)}{\partial C} = 2 \sum_{k=1}^{n-1} [C \cdot \beta_1^k + \beta_2 \sum_{j=1}^k j \beta_1^{k-j} + \frac{1-\beta_1^k}{1-\beta_1} \beta_3 - x^{(1)}(k+1)] \cdot \beta_1^k = 0 \quad (24)$$

C can be solved:

$$C = \frac{\sum_{k=1}^{n-1} [x^{(1)}(k+1) - \beta_2 \sum_{j=1}^k j\beta_1^{k-j} - \frac{1-\beta_1^k}{1-\beta_1} \beta_3] \cdot \beta_1^k}{\sum_{k=1}^{n-1} \beta_1^{2k}} \quad (25)$$

The optimized NDGM(1,1) is obtained which can be denoted as CNDGM(1,1):

$$\hat{x}^{(1)}(m+1) = \frac{\sum_{k=1}^{n-1} [x^{(1)}(k+1) - \beta_2 \sum_{j=1}^k j\beta_1^{k-j} - \frac{1-\beta_1^k}{1-\beta_1} \beta_3] \cdot \beta_1^k}{\sum_{k=1}^{n-1} \beta_1^{2k}} \cdot \beta_1^m + \beta_2 \sum_{j=1}^m j\beta_1^{m-j} + \frac{1-\beta_1^m}{1-\beta_1} \beta_3 \quad (26)$$

The restored value of the recursive function of CNDGM(1,1) at (k+1) is:

$$\hat{x}^{(0)}(m+1) = \alpha^{(1)} \hat{x}^{(1)}(m+1) = \hat{x}^{(1)}(m+1) - \hat{x}^{(1)}(m), m=1,2,\dots \quad (27)$$

The novel DGM (CNDGM(1,1)) with an optimized initial condition under overall consideration has been established.

3.3. CNDGM(1,1) of interval grey number with central points

For the sequence of middle values $X(L) = (L_1, L_2, \dots, L_{n-1})$, the time response sequence of its CNDGM(1,1) is:

$$\hat{L}_{m+1}^{(1)} = \frac{\sum_{k=1}^{n-1} [L_{k+1}^{(1)} - \alpha_2 \sum_{j=1}^k j\alpha_1^{k-j} - \frac{1-\alpha_1^k}{1-\alpha_1} \alpha_3] \cdot \alpha_1^k}{\sum_{k=1}^{n-1} \alpha_1^{2k}} \cdot \alpha_1^m + \alpha_2 \sum_{j=1}^m j\alpha_1^{m-j} + \frac{1-\alpha_1^m}{1-\alpha_1} \alpha_3 \quad (28)$$

$$\hat{L}_{m+1}^{(0)} = \hat{L}_{m+1}^{(1)} - \hat{L}_m^{(1)} \quad (29)$$

Note: $\alpha_1, \alpha_2, \alpha_3$ are parameters of CNDGM(1,1).

Similarly, the CNDGM(1,1) models of the sequences of $X(\underline{S}) = (\underline{S}_1, \underline{S}_2, \dots, \underline{S}_{n-1})$ and $X(\bar{S}) = (\bar{S}_1, \bar{S}_2, \dots, \bar{S}_{n-1})$ are:

$$\underline{\hat{S}}_{m+1}^{(1)} = \frac{\sum_{k=1}^{n-1} [\underline{S}_{k+1}^{(1)} - \beta_2 \sum_{j=1}^k j\beta_1^{k-j} - \frac{1-\beta_1^k}{1-\beta_1} \beta_3] \cdot \beta_1^k}{\sum_{k=1}^{n-1} \beta_1^{2k}} \cdot \beta_1^m + \beta_2 \sum_{j=1}^m j\beta_1^{m-j} + \frac{1-\beta_1^m}{1-\beta_1} \beta_3 \quad (30)$$

$$\underline{\hat{S}}_{m+1}^{(0)} = \underline{\hat{S}}_{m+1}^{(1)} - \underline{\hat{S}}_m^{(1)} \quad (31)$$

$$\bar{\hat{S}}_{m+1}^{(1)} = \frac{\sum_{k=1}^{n-1} [\bar{S}_{k+1}^{(1)} - \gamma_2 \sum_{j=1}^k j\gamma_1^{k-j} - \frac{1-\gamma_1^k}{1-\gamma_1} \gamma_3] \cdot \gamma_1^k}{\sum_{k=1}^{n-1} \gamma_1^{2k}} \cdot \gamma_1^m + \gamma_2 \sum_{j=1}^m j\gamma_1^{m-j} + \frac{1-\gamma_1^m}{1-\gamma_1} \gamma_3 \quad (32)$$

$$\bar{\hat{S}}_{m+1}^{(0)} = \bar{\hat{S}}_{m+1}^{(1)} - \bar{\hat{S}}_m^{(1)} \quad (33)$$

Among them, $\hat{L}_m^{(0)}, \underline{\hat{S}}_m^{(0)}, \bar{\hat{S}}_m^{(0)}$ are the forecasting values of the m_{th} values of $X(L)$,

$$X(\underline{S}), X(\bar{S}).$$

Combine with Lemma 1, Theorem 1 and Theorem 2,

$$\begin{cases} \hat{a}_{m+1} + a_m = 2\hat{L}_m^{(0)} - \hat{\underline{S}}_m^{(0)} - \hat{\bar{S}}_m^{(0)} \\ \hat{b}_{m+1} + b_m = 2\hat{L}_m^{(0)} + \hat{\underline{S}}_m^{(0)} - \hat{\bar{S}}_m^{(0)} \\ \hat{c}_{m+1} + c_m = 2\hat{L}_m^{(0)} + \hat{\underline{S}}_m^{(0)} + \hat{\bar{S}}_m^{(0)} \end{cases}$$

Finally, we obtain grey a prediction model of interval grey numbers with central points:

$$\begin{cases} \hat{a}_{m+1} = 2\hat{L}_m^{(0)} - \hat{\underline{S}}_m^{(0)} - \hat{\bar{S}}_m^{(0)} - a_m \\ \hat{b}_{m+1} = 2\hat{L}_m^{(0)} + \hat{\underline{S}}_m^{(0)} - \hat{\bar{S}}_m^{(0)} - b_m \quad (\#) \\ \hat{c}_{m+1} = 2\hat{L}_m^{(0)} + \hat{\underline{S}}_m^{(0)} + \hat{\bar{S}}_m^{(0)} - c_m \end{cases}$$

The procedures involved for using the CNDGM(1,1) can be summarized as follows.

Step 1: Similar with Step 1 in the process of DGM(1,1) of interval grey numbers with central points.

Step2: Establish the CNDGM(1, 1) for the above 3 sequences, and obtain parameters($\alpha_1, \alpha_2, \alpha_3, \beta_1, \beta_2, \beta_3, \gamma_1, \gamma_2, \gamma_3$).

Step3: Calculate the optimized initial condition C , then the forecasting values ($\hat{L}_m^{(0)}, \hat{\underline{S}}_m^{(0)}, \hat{\bar{S}}_m^{(0)}$) are obtained.

Step 4: Calculate Equation Group (#) to obtain the forecasting values of the lower bounds, the upper bounds and central points.

Step 5: Models' error analysis: The relative percentage error (RPE) describes the percentage of difference between the real and the fitting or forecasting values to evaluate the precision at a certain time instance k . The RPE can be defined as

$$RPE(k) = \left| \frac{\hat{x}^{(0)}(k) - x^{(0)}(k)}{x^{(0)}(k)} \right| \cdot 100\%$$

The total precision of a model can be described by calculating the average relative percentage error (ARPE):

$$ARPE = \frac{1}{m-k+1} \sum_{i=k}^m RPE(i), m \geq k$$

Note: Given θ , when $ARPE < \theta$ and $RPE(k) < \theta$, the model can be seen as a satisfactory residual model.

4 Numerical examples analysis

As the most important basic industrial energy source for economic development, electricity is one of the main priorities in the economic development strategy of all countries across the world. Specific to China, with the society's sustained development, electricity consumption (EC) continues to show a growth trend in recent years. Focusing on the southern Jiangsu area, it is located in the core area of the Yangtze River Delta which is one of the most developed and modernized areas in China. Therefore, the predictions of per capita annual ECs of southern Jiangsu can be seen as a benchmark for the Yangtze River Delta's economic trend to provide a strong reference for the local government to set future energy policy. In this section, per capita annual electricity consumptions of southern Jiangsu are selected as practical examples. Owing to the unbalanced development levels of the five cities (Nanjing, Wuxi, Changzhou, Suzhou and Zhenjiang), single real numbers are difficult to decipher in expressing the actual situation of electricity consumption in this area. The data sets are described as interval grey numbers. Among them, the per capita annual electricity consumption is defined as the amount of annual electricity consumption divided by the local population. To be specific, the upper and lower bounds of the interval grey numbers are characterized by the highest and lowest levels in the region, and the central-point values are obtained by the ratio of total amount of annual electricity consumption of southern Jiangsu and total population. Due to the revision of statistical standards, the data of per capita annual electricity consumption can now be collected from 1993 onwards. Hence, the expressions of interval grey numbers for the data can be obtained in Table 1 below, the trend of data are shown in Fig. 3.

To investigate the applicability of the novel model in more depth, three GM models and four classical statistical models are introduced. For the three grey models and the novel grey model proposed in this paper, the original DGM(1,1) (DGM) is established by using DGM(1,1) to forecast the lower bound sequence, the central-point sequence and the upper bound sequence respectively, which consider the three sequences independently; DGM(1,1) of interval grey numbers with central points (DGMC) is set up in Section 3.1; CNDGM(1,1) of interval grey numbers with central points (CNDGM) can be processed as steps set in Section 3.3; GA-GM(1,1)(GAGM) uses the Genetic Algorithm to calculate the parameters of GM(1,1) (Xiong et al., 2014). On the other hand, the four classical statistical models are as follows: the double exponential smoothing method (DES) applies to the double exponential smoothing method instead of single exponential smoothing, as it can better predict sequences with a certain linear trend; for the linear model (LM), three sequences are modeled as $y(t) = a + bt$ (a,b are parameters, $y(t)$ is the t_{th} value of sequences) to fit their linear trends respectively; Similarly, in the exponential model (EM), three sequences are modeled as $y(t) = a * b^t$ (a,b are parameters, $y(t)$ is the t_{th} value of sequences) to fit their exponential trends; for the autoregressive moving average (ARMA), AR(2) or ARMA (2,1) models are established by adopting autocorrelation and partial correlation tests. Considering the features of grey models and classical statistical models, long data sequences

(1993-2012) and short data sequences (2001-2012) are applied to establish forecasting models. Take the data of 1993-2012 or 2001-2012 as in-sample fitting data, and the data of 2013, 2014 and 2015 as out-sample comparative data. According to the long data sequences and the short data sequences, the fitting errors of each model are shown in Fig.4.

Table 1

Per capita annual electricity consumption of southern Jiangsu in 1993-2015

Unit: kiloWatt-hour (kWh)/person

| Year | 1993 | 1994 | 1995 | 1996 | 1997 | 1998 | 1999 | 2000 |
|-------------------|----------|----------|----------|----------|----------|----------|----------|----------|
| The lower bound | 1164.42 | 1288.68 | 1384.50 | 1433.19 | 1457.82 | 1476.95 | 1577.41 | 1822.07 |
| The central point | 1467.51 | 1643.68 | 1795.24 | 1940.74 | 2029.73 | 2128.57 | 2348.22 | 2718.08 |
| The upper bound | 1795.47 | 2030.82 | 2244.95 | 2400.04 | 2505.86 | 2522.46 | 2716.36 | 3286.99 |
| Year | 2001 | 2002 | 2003 | 2004 | 2005 | 2006 | 2007 | 2008 |
| The lower bound | 1997.52 | 2270.43 | 2653.92 | 3104.67 | 3631.40 | 4112.38 | 4717.54 | 4976.98 |
| The central point | 3064.88 | 3569.16 | 4408.90 | 5283.90 | 6283.43 | 7295.38 | 8264.12 | 8529.37 |
| The upper bound | 3891.62 | 4685.71 | 6051.24 | 7589.04 | 9320.45 | 11142.91 | 12928.09 | 13472.38 |
| Year | 2009 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | |
| The lower bound | 5352.03 | 5908.48 | 6281.66 | 6655.81 | 7194.48 | 7252.74 | 7578.51 | |
| The central point | 8895.40 | 10150.63 | 11034.49 | 11443.45 | 12196.43 | 12115.71 | 12405.57 | |
| The upper bound | 13894.80 | 16060.33 | 17617.58 | 18368.50 | 19319.86 | 19182.55 | 19665.67 | |

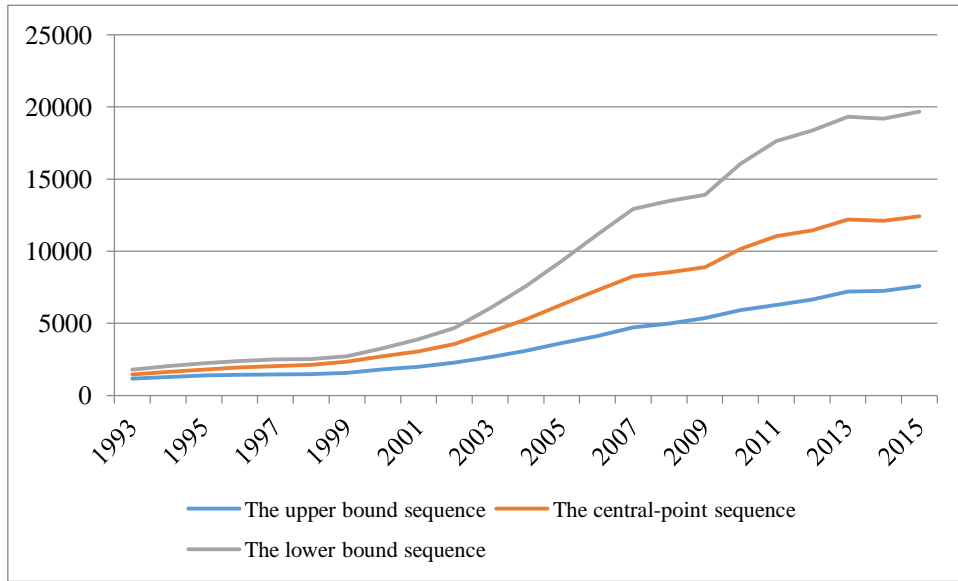


Fig.3. Interval grey numbers with central points of per capita annual electricity consumptions of southern Jiangsu

Special attention should be paid to the data trends. Before 2000, the trends of the lower bound, the central point and the upper bound sequences of interval grey numbers were all moderate. Then, from 2001 to 2005, three data sequences demonstrated steep rises. After experiencing a stable period from 2006-2009, the data displays a sharper growth trend which was ahead of the stable period of data in China from 2010. From an overall perspective, the data trends are complicated, and the valid data for recent information modelling are limited. It is difficult for forecasting models to capture the features of the sequences to make accurate predictions.

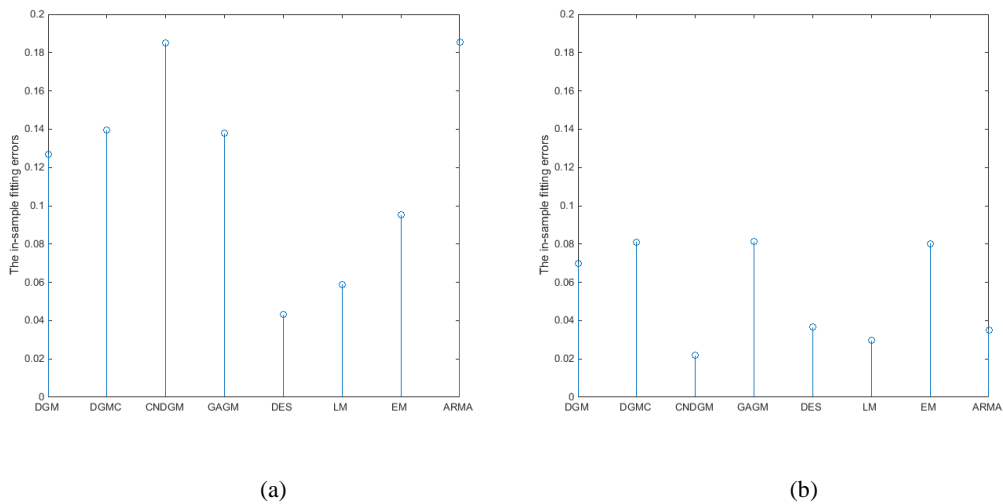


Fig. 4. (a) The in-sample fitting errors of eight models, based on the data of 1993-2012

(b)The in-sample fitting errors of eight models, based on the data of 2001-2012

It can be clearly seen that DES & LM obtain the smallest simulating errors for the long data sequences modelling, the ARPEs being 4.32% and 5.9% respectively. On the contrary, the fitting results of CNDGM for the long data sequences are less effective with an ARPE of 18.51% which is almost the highest ARPE, just lower than ARMA's 18.53%. For other models' results, EM also

shows high performance in fitting errors and the average relative error is 9.52%; the ARPEs of DGM, DGMC and GAGM are all around 13% (12.7%, 13.96% and 13.8% respectively). For the upper bound sequence, the fitting errors of DGM, DGMC, GAGM, ARMA and CNDGM are all above 18% which means the simulation effects are not effective and unacceptable. Similarly, for the central-point sequence, the fitting error of ARMA reaches 19.35%. These fitting outcomes may lead to poor forecasting results.

In the modelling process of the short data sequences, it is obvious that the ARPE of CNDGM which is proposed in this paper is the smallest among the other seven models, which is 2.19%. To be specific, the simulating errors of the lower bound, the central point and the upper bound sequences of interval grey numbers by using CNDGM are all below 3% and all the RPEs are smallest compared to each corresponding item from other compared methods. This means CNDGM is accurately fitted original data. In terms of other models, the ARPEs of LM, ARMA and DES are all below 4% (2.97%, 3.52% and 3.65% respectively) and their fitting effects are just outside CNDGM's. As to EM, the ARPE is 8.03%, and the RPE of the upper bound sequence reaches 10.52% which is an unacceptable fitting accuracy. The ARPEs of other three grey models are 6.99% for DGM, 8.11% for DGMC and 8.13% for GAGM which are all much higher than the optimized grey model in this paper. Similarly, the RPEs of DGM, GAGM and DGMC for the upper bound sequence achieve 9.08%, 11.78% and 10.59% which are also not exceptional.

After analyzing the results of simulating, the CNDGM(1,1) model of interval grey numbers with central points proposed in this paper, demonstrates a great fitting performance in the short data sequences which surpasses any other fitting models. It verifies that grey models can predict short-term more effectively compared with using long term data and further improves the original Grey Model of interval grey numbers.

When it comes to forecasting values, eight models are proposed by using the long data sequences and the short data sequences. And the results are shown in Fig.5.

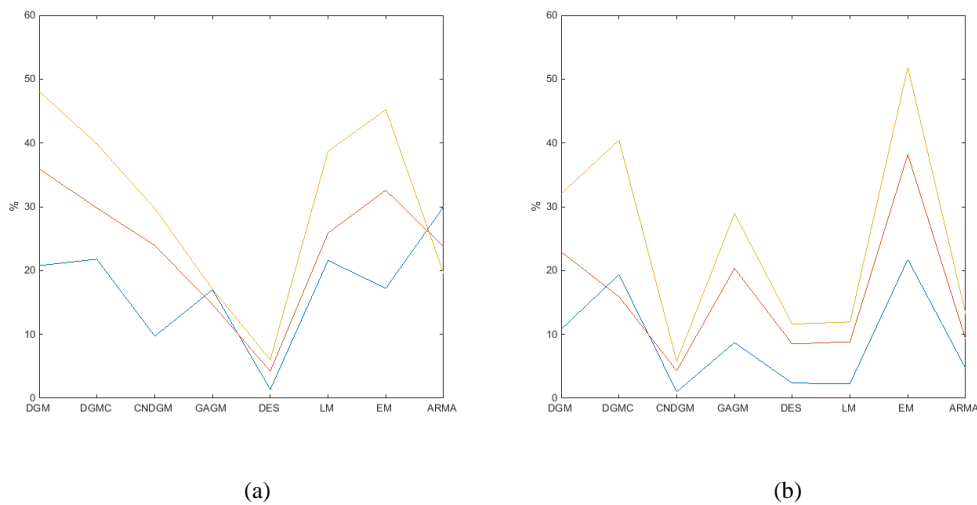


Fig. 5. (a) The out-sample forecasting errors of eight models, based on the data of 1993-2012

| | | | | | | | | | |
|-------|-------|--------|--------|--------|------|--------|--------|--------|--------|
| DGM | 6.88% | 11.53% | 19.08% | 12.50% | DES | 12.89% | 16.07% | 21.82% | 16.93% |
| DGMC | 3.90% | 8.11% | 13.05% | 8.35% | LM | 19.09% | 22.33% | 32.93% | 24.78% |
| CNDGM | 6.71% | 8.90% | 12.45% | 9.35% | EM | 7.69% | 7.87% | 11.20% | 8.92% |
| GAGM | 5.65% | 11.34% | 20.82% | 12.61% | ARMA | 10.21% | 13.46% | 31.06% | 18.24% |

As seen from Table 2, the fitting errors of average values of every three year period by using DGMC, CNDGM and EM, show relatively high fitting precision (all under 10%). To compare the development trends by using these three models, the forecasting values are then shown in Table 3 and the future trends of the per capita annual electricity consumption in southern Jiangsu area are displayed in Fig.5.

Table 3

Future per capita annual electricity consumption in southern Jiangsu area (2014-2031) Unit: kWh/person

| Year | DGMC | | | CNDGM | | | EM | | |
|-----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
| | a | b | c | a | b | c | a | b | c |
| 2014-2016 | 9202.47 | 16985.07 | 31354.11 | 9079.98 | 15393.69 | 25505.23 | 9346.82 | 17936.57 | 30775.46 |
| 2017-2019 | 11953.30 | 23972.33 | 48123.93 | 10640.75 | 17448.51 | 29005.66 | 12632.76 | 25518.83 | 46200.78 |
| 2020-2022 | 15445.66 | 31949.13 | 69553.84 | 13842.37 | 22551.72 | 38573.41 | 17073.91 | 36306.32 | 69357.60 |
| 2023-2025 | 19081.87 | 43662.56 | 105065.6 | 16223.21 | 25568.21 | 44060.33 | 23076.38 | 51653.95 | 104121.1 |
| 2026-2028 | 22720.72 | 57318.13 | 154644.3 | 20372.75 | 31769.31 | 55928.48 | 31189.07 | 73489.44 | 156308.8 |
| 2029-2031 | 24756.22 | 75399.64 | 232548.2 | 23848.46 | 36039.51 | 64080.12 | 42153.84 | 104555.3 | 234654.1 |

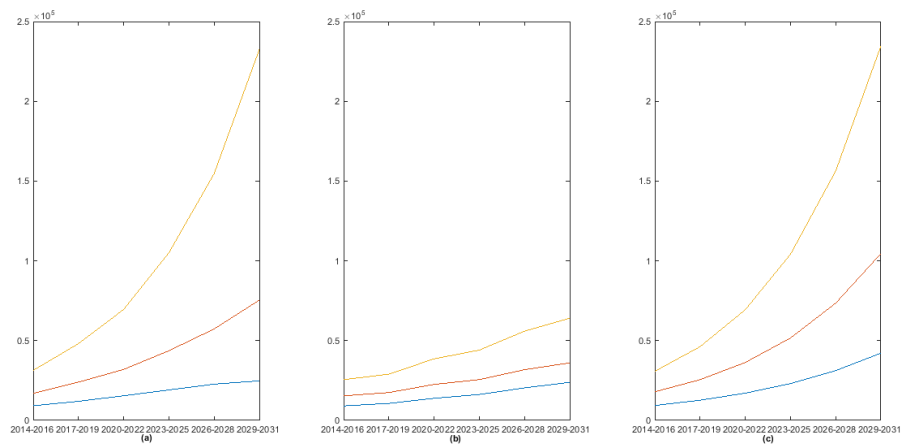


Fig.6. The forecasting values of the per capita annual electricity consumption in southern Jiangsu area by using DGMC (a), CNDGM (b) and EM (c)

According to forecasting values, the per capita annual electricity consumption in the southern Jiangsu area will keep increasing. The growth rates of DGMC, CNDGM and EM are considerably different. Specially, the trends of DGMC and EM (in Fig. 6(a) and (c)) totally exceed that of CNDGM, and the growth rates of each sequence by using the former two models are all around 30%-50%. Furthermore, the magnitude of data will reach 2,000,000 for the upper sequences by DGMC and EM which means the per capita annual electricity consumption in the southern Jiangsu area will expand 10 times in the next 18 years. On the other hand, the forecasting values by CNDGM proposed in this paper will experience a more moderate rise and the growth rates will be between 13% and 33% for every three years. Additionally, for the upper bound sequence, by 2030 the value will reach to 64080.12 kWh/person which is 1.5 times more than 2014-2016's. It shows that the growth rates of the highest levels in southern Jiangsu area will decrease slowly in general, although the values of the per capita annual electricity consumption will continue to rise. For the central-point sequence, the value of 2029-2031 will be 36039.51 kWh/person - a 1.34 times' increase compared with the value of 2014-2016. The growth rates' trend of the central-point sequence during 2014-2031 is similar with the upper bound sequence. For the lower bound sequence, by 2030 it will reach to 23848.46 kWh/person, which is more than 1.6 times compared to that of 2014-2016. It is notable that the growth rates of the lower bound sequence are slightly higher than the other two sequences'.

Focusing on the recent years (2013-2015) seen in Fig.3, the trends of per capita annual electricity consumption in the southern Jiangsu area are tending to stabilize, that is to say, the future growth rates of electricity consumption may not be as high as the last few years (2001-2012). As China has implemented a series of measures to cut production capacity and adjust industrial structure, the low growth rates of per capita annual electricity consumption reveal the effects of these energy policies. Accordingly, the forecasting values of CNDGM show the probable future trends considering the continuous energy policy on reducing energy consumption in China. With respect to the other two models (DGMC and EM), the forecasting values demonstrate that the energy consumption will experience high rises along with the increasing trend from 2001, which ignore the effects of energy policy in China. In other words, taking into account the energy policy of China in recent years, there is no difference for the development trend of the per capita annual electricity consumption in southern Jiangsu in the future by using DGMC and EM. Such a conclusion seriously underestimates the impact of the relevant energy policy. In fact, the impact of these energy policies is clearly visible in Fig. 3: the steep curve changes its direction in 2013. The data for this new trend is overwhelmed by the dominant historical data for previous years, this is the reason why this trend is ignored by DGMC and EM. For CNDGM, its association with the recent data is further enforced by our new method, and it captures the new trends ignored by the other two models. In short, CNDGM provides a possible way forward for forecast deviations, which reveals the hidden effectiveness of China's measures to reduce energy consumption. The forecasting values by 2031 can be obtained by using the values from CNDGM as the lower bound, and the values from DGMC and EM as the upper bound. The ranges between the upper and the lower bound notify the importance of a series of energy optimization policy.

To sum up, the continuous growth of the per capita annual electricity consumption in the southern Jiangsu area will lead to sustained pressure on energy planning and configuration, as well as restrict economic development.

5 Conclusions

A novel energy consumption forecasting model is proposed by optimizing the initial conditions of DGM (1, 1) model and establishing models of interval grey numbers with central points, in which the data sequences are fully transformed as the sequence of middle values, the sequence of trapezoid areas of the lower grey layer and the sequence of trapezoid areas of the upper grey layer. Based on the empirical analysis of per capita annual electricity consumptions of southern Jiangsu, the following conclusions are obtained:

- (1) The fitting and prediction results exhibits higher accuracy in short-term sequence modelling (the fitting and forecasting ARPEs are only 2.19% and 3.7%) compared with other forecasting models (DGM, DGMC, GAGM, DES, LM, EM, ARMA), regardless if the results stem from long-term or short-term sequences. It provides a feasible method for the prediction of interval grey numbers with central points for regional data of energy consumption and expands the research fields for different types of EC data.
- (2) For the long-term forecasting of electricity consumption (to the year 2031), the prediction values have been obtained by establishing three-year average values and the outcomes display high fitting precision by the model proposed in this paper compared with the other three grey models (DGM, DGMC, GAGM). The future trends of per capita annual electricity consumption of southern Jiangsu will continue to rise in the near future, but the growth rate will probably slow to more moderate levels due to the effects of sustainable energy policy in China.

Although the high precision results are obtained in the novel model, there are huge challenges involved in the forecasting of energy consumption in practice because of its trend variability and regional imbalance revealed in this paper. Taking into account the expansion of absolute amounts in more advanced cities and the increase in the growth rates in less advanced cities, it is still a massive pressure in controlling the electricity consumption in the southern Jiangsu area.

The case in this paper is simply to clarify the process of this novel model and shows the effectiveness of this novel model. Actually, this model can be applied in many other aspects of energy forecasting. For further research in the future, sufficient data should be collected for the models' verification compared with classic statistical models and intelligent algorithms. While most of the forecasting methods need more than 30 data sets for modeling, the data in this paper total just 23 data sets (due to the changes in statistical standards in the Statistical Yearbook of Jiangsu Province), which is more suitable for grey models.

6 Acknowledgments

The authors are grateful to anonymous referees for their helpful and constructive comments on this paper. This work was supported by a Marie Curie International Incoming Fellowship within the 7th European Community Framework Programme entitled "Grey Systems and Its Application to Data Mining and Decision Support" Grant No. FP7-PIIF-GA-2013-629051, a project of the Leverhulme Trust International Network entitled "Grey Systems and Its Applications"(IN-2014-020). The authors would also like to acknowledge the support of the

National Natural Science Foundation of China (71771119, 71701101, 91746202).

REFERENCES

- Al-mulali U., Fereidouni H.G., Lee J.Y.M., 2014. Electricity consumption from renewable and non-renewable sources and economic growth: Evidence from Latin American countries. *Renewable and Sustainable Energy Reviews*. 30, 290–298.
- Dang Y.G., Liu S.F., 2004. The GM models that $x(n)$ be taken as initial value. *Kybernetes*. 33, 247–255.
- Deng J.L., 1982. The Control problem of grey systems. *Syst. Control Lett.* 1 (5), 288–294.
- Ding S., 2019. A novel discrete grey multivariable model and its application in forecasting the output value of China's high-tech industries. *Industrial Engineering*. 127, 749–760.
- Ding S., 2018. A novel self-adapting intelligent grey model for forecasting China's natural-gas demand. *Energy*. 162, 393–407.
- Ding S., Dang Y.G., Xu N., Wang J.J., Xu Z.D., 2018. A Novel Grey Model Based on the Trends of Driving Factors and Its Application. *The Journal of Grey System*. 30(3):105-126.
- Erdal, K., Baris, U., Okyay, K., 2010. Grey system theory-based models in time series prediction. *Expert Syst Appl.* 37(2), 1784–1789.
- Felice M.D., Alessandri A., Catalano F., 2015. Seasonal climate forecasts for medium-term electricity demand forecasting. *Applied Energy*. 137, 435–444.
- Hamzacebi C., Es H.A., 2014. Forecasting the annual electricity consumption of Turkey using an optimized grey model. *Energy*. 70, 165–171.
- Lee Y.S., Tong L.I., 2012. Forecasting nonlinear time series of energy consumption using a hybrid dynamic model. *Applied Energy*. 94, 251–256.
- Liu S.F., Dang Y.G., Fang Z.G., Xie N.M., 2010. *Grey System Theory and Its Application* (5th ed.). Science Press, Beijing. 15,122–123 & 146–149.
- Liu S.F., Xie N.M., 2011. A new grey evaluation method based on reformative triangular whitenization weight function. *Journal of Systems Engineering*. 31(2), 72–78.
- Truong D.Q., Ahn K.K., 2012. Wave prediction based on a modified grey model MGM(1,1) for real-time control of wave energy converters in irregular waves. *Renewable Energy*. 43, 242–255.
- Wang Y.H., Dang Y.G., Liu S.F., 2010. An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition. *Expert Syst Appl.* 37, 5640–5644.
- Wu, L., Liu, S., Yao, L., Yan, S., Liu, D., 2013a. Grey system model with the fractional order accumulation. *Commun Nonlinear Sci.* 18(7), 1775–1785.
- Wu L.F., Liu S.F., Yan S.L., 2013b. Grey Prediction Model for Hybrid Sequence. *Control and Decision*. 28(12), 1912–1914, 1920.
- Xia M., Wong W.K., 2014. A seasonal discrete grey forecasting model for fashion retailing. *Knowledge-Based Systems*. 57, 119–126.
- Xie N.M., Liu S.F., 2009. Discrete grey forecasting model and its optimization. *Applied Mathematical Modeling*, 33, 1173–1186.
- Xie N.M., Liu S.F., Yang Y.J., Yuan C.Q., 2013. On novel grey forecasting model based on non-homogeneous index sequence. *Applied Mathematical Modeling*. 37, 5059–5068.
- Xiong P.P., Dang Y.G., Yao T.X., Wang Z.X., 2014. Optimal modeling and forecasting of the energy consumption

and production in China. *Energy*. 77, 623–634.

Xu N., Dang Y.G., Gong Y.D., 2017. Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China. *Energy*. 118, 473–480.

Yamaguchi D., Li G.D., Nagai M., 2007. A grey-based rough approximation model for interval data processing. *Information Sciences*. 177, 4727–4744.

Yukseltan E., Yucekay A., Bilge A.H., 2017. Forecasting electricity demand for Turkey: Modeling periodic variations and demand segregation. *Applied Energy*. 193, 287–296.

Zahedi G., Azizi S., Bahadori A., Elkamel A., Alwi S.R.W., 2013. Electricity demand estimation using an adaptive neuro-fuzzy network: a case study from the Ontario province-Canada. *Energy*. 49, 323–328.

Zeng B., Li C., Long X.J., Xiong Y., Zhou X.Y., 2014. A Novel Interval Grey Number Prediction Model Given Kernel and Grey Number Band. *Journal of Grey System*. 26(3), 69–84.

Zeng B., Liu S.F., Cui J., 2010. Prediction model for interval grey number with known whitenization weight function. *Control and Decision*. 25(12), 1815–1820.

Zeng B., Liu S.F., 2011. Prediction model of interval grey number based on its geometrical characteristics. *Journal of System Engineering*. 26(2), 174–180.

The supplementary information

In Section 4, according to the long data sequences and the short data sequences, the fitting errors of each model are listed in Table 1 and Table 2 which has been shown in Fig.4.

Table 1

The simulation RPE and ARPE of eight models, based on the data of 1993-2012

| RPE | The lower bound (<i>a</i>) | The central point (<i>b</i>) | The upper bound (<i>c</i>) | ARPE | | <i>a</i> | <i>b</i> | <i>c</i> | ARPE |
|-------|---------------------------------|-----------------------------------|---------------------------------|--------|------|----------|----------|----------|--------|
| DGM | 7.63% | 11.31% | 19.16% | 12.70% | DES | 3.14% | 4.18% | 5.65% | 4.32% |
| DGMC | 14.87% | 8.31% | 18.71% | 13.96% | LM | 9.62% | 4.84% | 3.25% | 5.90% |
| CNDGM | 13.36% | 16.74% | 25.43% | 18.51% | EM | 8.25% | 8.31% | 11.99% | 9.52% |
| GAGM | 7.53% | 12.68% | 21.21% | 13.80% | ARMA | 13.76% | 19.35% | 22.49% | 18.53% |

Table 2

The simulation RPE and ARPE of eight models based on the data of 2001-2012

| RPE | The lower bound (<i>a</i>) | The central point (<i>b</i>) | The upper bound (<i>c</i>) | ARPE | | <i>a</i> | <i>b</i> | <i>c</i> | ARPE |
|-------|---------------------------------|-----------------------------------|---------------------------------|-------|------|----------|----------|----------|-------|
| DGM | 5.20% | 6.68% | 9.08% | 6.99% | DES | 2.22% | 3.79% | 4.94% | 3.65% |
| DGMC | 5.46% | 8.26% | 10.59% | 8.11% | LM | 2.00% | 2.96% | 3.95% | 2.97% |
| CNDGM | 1.95% | 1.92% | 2.69% | 2.19% | EM | 5.62% | 7.94% | 10.52% | 8.03% |
| GAGM | 5.86% | 6.74% | 11.78% | 8.13% | ARMA | 6.09% | 2.20% | 2.27% | 3.52% |

When it comes to forecasting values, eight models are proposed by using the long data sequences and the short data sequences respectively. And the results are shown Table 3 and Table 4 which has been displayed in Fig.5.

Table 3

The out-sample results, RPE and ARPE of eight models, based on the data of 1993-2012

| | 2013 | Values | RPE (%) | ARPE (%) | 2014 | Values | RPE (%) | ARPE (%) | 2015 | Values | RPE (%) | ARPE (%) |
|-------|------|----------|---------|----------|------|----------|---------|----------|------|----------|---------|----------|
| DGM | a | 7948.50 | 10.48 | | a | 8813.74 | 21.52 | | a | 9773.18 | 28.96 | |
| | b | 14681.72 | 20.38 | 20.78 | b | 16437.20 | 35.67 | 35.92 | b | 18402.59 | 48.34 | 48.09 |
| | c | 25403.67 | 31.49 | | c | 28880.82 | 50.56 | | c | 32833.91 | 66.96 | |
| DGMC | a | 6252.97 | - | | a | 7438.92 | 2.57 | | a | 7370.68 | - | |
| | b | 13719.07 | 12.48 | 21.78 | b | 15399.07 | 27.10 | 29.81 | b | 16980.00 | 36.87 | 39.82 |
| | c | 27005.26 | 39.78 | | c | 30648.05 | 59.77 | | c | 35365.67 | 79.83 | |
| CNDGM | a | 7659.32 | 6.46 | | a | 8752.75 | 20.68 | | a | 9301.71 | 22.74 | |
| | b | 13370.83 | 9.63 | 9.71 | b | 14864.08 | 22.68 | 23.94 | b | 16103.46 | 29.81 | 29.74 |
| | c | 21838.50 | 13.04 | | c | 24641.47 | 28.46 | | c | 26876.47 | 36.67 | |
| GAGM | a | 7845.97 | 9.06 | | a | 8697.19 | 19.92 | | a | 9640.76 | 27.21 | |
| | b | 10857.18 | 10.98 | 17.00 | b | 11999.04 | 0.96 | 14.67 | b | 13260.99 | 6.90% | 17.08 |
| | c | 13340.01 | 30.95 | | c | 14742.99 | 23.14 | | c | 16293.52 | 17.15 | |
| DES | a | 7042.70 | - | | a | 7428.14 | 2.42 | | a | 7813.57 | 3.10% | |
| | b | 12014.28 | - | 1.36 | b | 12572.29 | 3.77 | 4.24 | b | 13130.31 | 5.84% | 6.01 |
| | c | 19413.62 | 0.49 | | c | 20433.60 | 6.52 | | c | 21453.59 | 9.09% | |
| LM | a | 10955.16 | - | | a | 11260.72 | - | | a | 11566.28 | - | |
| | b | 19625.07 | - | 21.58 | b | 20190.19 | - | 25.89 | b | 20755.30 | - | 38.69 |
| | c | 31738.81 | - | | c | 32680.48 | - | | c | 33622.15 | - | |
| EM | a | 7699.89 | 7.02 | | a | 8519.54 | 17.47 | | a | 9426.44 | 24.38 | |
| | b | 14538.42 | 19.20 | 17.22 | b | 16386.37 | 35.25 | 32.59 | b | 18469.20 | 48.88 | 45.24 |
| | c | 24233.89 | 25.44 | | c | 27825.65 | 45.06 | | c | 31949.74 | 62.46 | |
| ARMA | a | 18597.72 | 19.34 | | a | 20099.23 | 13.53 | | a | 21721.96 | 10.56 | |
| | b | 27541.12 | 34.85 | 29.99 | b | 29920.47 | 28.74 | 23.72 | b | 32505.37 | 24.40 | 19.55 |
| | c | 51657.48 | 35.79 | | c | 56811.31 | 28.88 | | c | 62479.34 | 23.70 | |

Table 4

The out-sample results, RPE and ARPE of eight models, based on the data of 2001-2012

| | 2013 | Values | RPE (%) | ARPE (%) | 2014 | Values | RPE (%) | ARPE (%) | 2015 | Values | RPE (%) | ARPE (%) |
|-------|----------|----------|---------|----------|----------|----------|---------|----------|----------|----------|---------|----------|
| | <i>a</i> | 7739.98 | 7.58 | | <i>a</i> | 8521.34 | 17.49 | | <i>a</i> | 9381.57 | 23.79 | |
| DGM | <i>b</i> | 13447.64 | 10.26 | 10.79 | <i>b</i> | 14837.77 | 22.47 | 22.92 | <i>b</i> | 16371.61 | 31.97 | 32.02 |
| | <i>c</i> | 22128.93 | 14.54 | | <i>c</i> | 24709.37 | 28.81 | | <i>c</i> | 27590.71 | 40.30 | |
| | <i>a</i> | 8155.49 | 13.36 | | <i>a</i> | 8144.77 | 12.30 | | <i>a</i> | 9767.44 | 28.88 | |
| DGMC | <i>b</i> | 14499.67 | 18.88 | 19.37 | <i>b</i> | 13926.48 | 14.95 | 15.91 | <i>b</i> | 17405.45 | 40.30 | 40.44 |
| | <i>c</i> | 24314.65 | 25.85 | | <i>c</i> | 23111.60 | 20.48 | | <i>c</i> | 29916.55 | 52.13 | |
| | <i>a</i> | 7088.06 | 1.48 | | <i>a</i> | 7418.37 | 2.28 | | <i>a</i> | 7830.48 | 3.32 | |
| CNDGM | <i>b</i> | 12038.91 | 1.29 | 1.01 | <i>b</i> | 12645.32 | 4.37 | 4.31 | <i>b</i> | 13185.17 | 6.28 | 5.77 |
| | <i>c</i> | 19268.31 | 0.27 | | <i>c</i> | 20387.15 | 6.28 | | <i>c</i> | 21182.17 | 7.71 | |
| | <i>a</i> | 8172.55 | 13.59 | | <i>a</i> | 9032.07 | 24.53 | | <i>a</i> | 9981.98 | 31.71 | |
| GAGM | <i>b</i> | 13679.69 | 12.16 | 8.71 | <i>b</i> | 15118.39 | 24.78 | 20.34 | <i>b</i> | 16708.41 | 34.68 | 28.94 |
| | <i>c</i> | 19390.10 | 0.36 | | <i>c</i> | 21429.38 | 11.71 | | <i>c</i> | 23683.12 | 20.43 | |
| | <i>a</i> | 7125.77 | 0.96 | | <i>a</i> | 7559.63 | 4.23 | | <i>a</i> | 7993.48 | 5.48 | |
| DES | <i>b</i> | 12420.59 | 1.84 | 2.35 | <i>b</i> | 13207.52 | 9.01 | 8.49 | <i>b</i> | 13994.45 | 12.81 | 11.61 |
| | <i>c</i> | 20143.62 | 4.26 | | <i>c</i> | 21530.23 | 12.24 | | <i>c</i> | 22916.83 | 16.53 | |
| | <i>a</i> | 7168.93 | 0.36 | | <i>a</i> | 7609.50 | 4.92 | | <i>a</i> | 8050.07 | 6.22 | |
| LM | <i>b</i> | 12472.57 | 2.26 | 2.30 | <i>b</i> | 13260.36 | 9.45 | 8.85 | <i>b</i> | 14048.15 | 13.24 | 11.95 |
| | <i>c</i> | 20148.84 | 4.29 | | <i>c</i> | 21517.60 | 12.17 | | <i>c</i> | 22886.36 | 16.38 | |
| | <i>a</i> | 8263.44 | 14.86 | | <i>a</i> | 9235.48 | 27.34 | | <i>a</i> | 10321.87 | 36.20 | |
| EM | <i>b</i> | 14721.53 | 20.70 | 21.76 | <i>b</i> | 16587.62 | 36.91 | 38.19 | <i>b</i> | 18690.26 | 50.66 | 51.85 |
| | <i>c</i> | 25063.98 | 29.73 | | <i>c</i> | 28836.49 | 50.33 | | <i>c</i> | 33176.83 | 68.70 | |
| | <i>a</i> | 6394.21 | 11.12 | | <i>a</i> | 6394.20 | 11.84 | | <i>a</i> | 6394.19 | 15.63 | |
| ARMA | <i>b</i> | 12090.95 | 0.86 | 4.64 | <i>b</i> | 12866.82 | 6.20 | 9.26 | <i>b</i> | 13643.16 | 9.98 | 13.19 |
| | <i>c</i> | 19693.33 | 1.93 | | <i>c</i> | 21049.37 | 9.73 | | <i>c</i> | 22413.69 | 13.97 | |