Scientia Iranica E (2021) 28(1), 386–394

Sharif University of Technology
Scientia Iranica
Transactions E: Industrial Engineering
http://scientiairanica.sharif.edu

Forecasting natural gas production and consumption using grey model with latent information function: The cases of China and USA

L. Wu*, **K. Zhang**, **and T. Zhao**

College of Management Engineering and Business, Hebei University of Engineering, Handan 056038, China.

Received 16 October 2017; received in revised form 1 February 2019; accepted 29 April 2019

KEYWORDS
Forecasting;
Natural gas consumption;
Natural gas production;
Grey model;
China.

Abstract. This paper aimed to develop a grey model for short-term forecasting of natural gas consumption and production in China and USA. To enhance the prediction accuracy of the proposed model, the outliers were determined by the error of the latent information function and then, they were corrected according to the test sample and the future trend. The sequence with corrected outliers was used to construct a grey model. The proposed model was employed to predict the natural gas consumption and production in China and USA. The results demonstrated that the proposed model could raise the forecast accuracy of the grey model. In addition, it was shown that China would inevitably face a massive expansion of natural gas imports.

1. Introduction

Given that natural gas emits much less CO_2 than coal and is the cleanest burning of all fossil fuels, it can be considered as an important adjunct to renewable energy sources as well as a bridge to the new energy economy [1]. It is playing an increasingly important role when the energy supply of the world is challenged by many significant changes such as the rapid change in geopolitical situations [2], a shift in the relationship between natural gas supply and demand, climate change and environmental pressures, breakthrough of unconventional gas in the United States, the impact of Japan's nuclear crisis, the economy and techniques of new energy, etc. The relationship between natural gas energy consumption and economic growth in Gulf Cooperation Council countries is investigated using the multivariate framework model [3]. The natural gas and electricity prices share common long-term dynamics in the Spanish market [4].

For the sake of a stabilized world economy, accurate forecasting of gas production has become ever more critical and important [5]. This has motivated many researchers to focus their research on natural gas forecasting. Mehmet Melikoglu reviewed natural gas demand forecasting in Turkey between 2013 and 2030 [6]. The projections indicate that world natural gas production will peak between 2025 and 2066 [7]. Lin and Wang predicted the production peak and import trends of natural gas in China [8]. Primo et al. compared static and adaptive models for short-term residential natural gas forecasting in Croatia [9]. The Bayesian method of vector autoregression and neural network quantile regression were used separately to predict the future natural gas consumption [10].

Among the various forecasting models [11–13], the grey model enjoys the potential to become a

*. *Corresponding author.*
E-mail address: wlf6666@126.com (L. Wu)

doi: 10.24200/sci.2019.5378.1240

very powerful tool for natural gas forecasting due to its superiority in limited data forecasting [14,15], particularly the grey model of the first-order and one variable (GM(1,1)) [16]. The GM(1,1) model achieved significant improvements in prediction and simulation accuracy using the following aspects: a generating operator [17], parameter estimation [18,19], differential equations [20], non-equigap grey model [21], hybrid models [22–24], etc. Wu et al. improved the grey model to give more weight to recently obtained information [25]. However, these models did not discuss the errors of the GM(1,1) model in detail. In this paper, the outlier was considered by analyzing the errors of the GM(1,1) model.

The rest of this paper is organized as follows. A novel grey model with corrected outliers is given in Section 2. The natural gas production and consumption of China is predicted by using this grey model in Section 3. The conclusion and discussion are presented in Section 4. Implication for practice is given in Section 5. The effectiveness of the proposed modeling is further demonstrated in Section 6.

2. The proposed model

Let $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$ be an original non-negative sequence and $x^{(0)}(k)$ be the value at time k. The first-order accumulated generating operator of $X^{(0)}$ is:

$$X^{(1)} = \left\{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right\},$$

where,

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i); \quad k = 1, 2, \cdots, n.$$

The original form of the GM(1,1) model is $x^{(0)}(k) + az^{(1)}(k) = b$, where:

$$z^{(1)}(k) = \frac{x^{(1)}(k) + x^{(1)}(k-1)}{2}, \quad k = 2, 3, \cdots, n.$$

The ordinary least squares estimate of the GM(1,1) parameters can be obtained by:

$$\begin{bmatrix} \hat{a} \\ \hat{b} \end{bmatrix} = (A^T A)^{-1} A^T Y,$$

where:

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, \quad A = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}.$$

The solution of the whitenization equation $\frac{dx^{(1)}(t)}{dt} + ax^{(1)}(t) = b$, is given by:

$$\hat{x}^{(1)}(t) = \left[x^{(0)}(1) - \frac{\hat{b}}{\hat{a}}\right] e^{-\hat{a}t} + \frac{\hat{b}}{\hat{a}}.$$

The inverse accumulated generating operator is:

$$\begin{aligned} \hat{x}^{(0)}(k+1) &= \hat{x}^{(1)}(k+1) - \hat{x}^{(1)}(k) \\ &= \left[x^{(0)}(1) - \frac{\hat{b}}{\hat{a}}\right](1 - e^{\hat{a}})e^{-\hat{a}k}, \end{aligned}$$

$$k = 1, 2, \cdots.$$

The error $e(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$, thus the error sequence is $E = \{e(1), e(2), \cdots, e(n)\}$. If all of $e(k)$'s are positive or negative, there are system errors in the sequence. The sequence E is used to construct another GM(1,1), that is to say, a residual correction model is constructed [23]. If all of $e(k)$'s are not positive or negative, there will be $e(k)$'s irregular errors with possible outliers.

Hence, a novel model for identification of outliers via the Latent Information (LI) function is proposed. The LI function is developed to analyze the likelihood of occurrence for the potential data [26]. It utilizes four statistical indexes (Central Tendency (CT), dispersion, skewness, and kurtosis) to describe the data feature, and can help extract hidden information. In this paper, the LI function is used to analyze the likelihood of occurrence of the above error. The complete constructing procedure of the LI function is shown below:

Step 1. Consider the error sequence $E = \{e(1), e(2), \cdots, e(n)\}$, to calculate the range (R): $R = e_{\max} - e_{\min}$, where $e_{\max}$ is the maximal error, and $e_{\min}$ is the minimal error.

Step 2. Determine the central tendency (CT):

$$CT = \frac{\sum_{k=1}^{n} ie(i)}{\sum_{k=1}^{n} i}, \qquad i = 1, 2, \cdots, n.$$

Step 3. Find the Central Location (CL) of the existing error:

$$CL = \frac{e_{\min} + e_{\max}}{2}.$$

Step 4. Count the number of elements which is larger than CL to be written as $|E^+|$ and the number of elements which is smaller than CL to be written as $|E^-|$.

Step 5. Compute the Decreasing Tendency (DT) and the Increasing Tendency (IT):

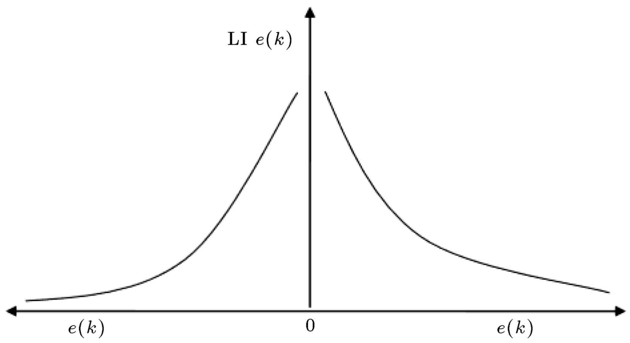

Figure 1. The relationship between error and the $LI\ e(k)$.

$$IT = \frac{|E^+|}{n}, \qquad DT = \frac{|E^-|}{n}.$$

Step 6. Extend the Upper Bound (UB) and Lower Bound (LB):

$$UB = e_{\max} + IT \times \frac{R}{n}, \qquad LB = e_{\min} - DT \times \frac{R}{n}.$$

Step 7. The LI value of the existing error $e(k)$ can be obtained by the following formulas:

$$LI\ e(k) = \begin{cases} \frac{e(k)-LB}{\mathrm{CT-LB}}, & if\ e(k) < CT \\ 1 & if\ e(k) = CT \\ \frac{UB-e(k)}{UB-CT}, & if\ e(k) > CT \end{cases}$$

$LI\ e(k)$ indicates the likelihood of occurrence of the error $e(k)$. Simply, put the relationship between $LI\ e(k)$ and $e(k)$ is considered as the normality, as shown in Figure 1. Larger $e(k)$ with smaller $LI\ e(k)$ and smaller $e(k)$ with larger $LI\ e(k)$ are rational. Larger $e(k)$ with larger $LI\ e(k)$ and smaller $e(k)$ with smaller $LI\ e(k)$ are not logical. These unreasonable errors are caused by outliers; therefore these outliers must be corrected to improve the prediction accuracy [27].

Actually, in some cases, it is difficult to select the outlier. The value $e(k) \times LI\ e(k)$ is another index. The larger $e(k) \times LI\ e(k)$ is, the more likelihood that $x(k)$ is an outlier. The smaller $e(k) \times LI\ e(k)$ is, the more likelihood that $x(k)$ is an outlier. Outliers appear in this model, because some abnormal factors often produce outliers [28]. For example, the values of tourism-based foreign exchange earnings over the years 2000 to 2006 in China are chosen as the in-sample data. This value in 2003 was known to be an outlier due to the outbreak of severe acute respiratory syndrome, which greatly affected China's tourism industry. The errors of GM(1,1) and their LI values are shown in Table 1. $e(k) \times LI\ e(2003)$ is the smallest among all the $e(k) \times LI\ e(k)$ $(k = 2000, 2001, \cdots, 2006)$. The other $e(k) \times LI\ e(k)$ values are very approximate. Therefore, this value in 2003 is considered as the outlier. This conforms with the actual situation; it is indicated that the proposed model is effective in the identification of outlier.

The outlier is corrected to increase prediction accuracy as far as possible, although by doing so, the fitting accuracy may decrease. The strategy for the correction of an outlier is to firstly increase the prediction accuracy as far as possible with acceptable fitting accuracy. Then, the sequence with the corrected outliers is used to construct a GM(1,1). The proposed model is called GM(1,1) with Corrected Outlier (GMCO(1,1)). The detailed flow chart is shown in Figure 2.

3. Empirical results

3.1. China

China's natural gas consumption reached 167.6 billion cubic meters in 2013, a year-on-year rise of 13.9%. China has become the world's third largest natural gas consumer. Natural gas's share of total energy is 4.8%. China's natural gas consumption has recently maintained a double-digit growth trend. The proportion of natural gas to the primary energy consumption will continue to increase by more than 10% with the advancement of controlling haze by government. Thus, it is essential to accurately forecast the consumption and production of natural gas. In this section, firstly, the natural gas production is predicted

Table 1. The errors of GM(1,1) and their Latent Information (LI) values.

Year	Actual value	GM(1,1)	Error	LI value	$e(k) \times LI\ e(k)$
2000	162.24	162.24	0	–	–
2001	177.92	162.23	15.72	0.37	5.9
2002	203.85	187.33	16.55	0.34	5.6
2003	174.06	216.32	-42.24	0.04	-1.5
2004	257.39	249.79	7.59	0.71	5.4
2005	292.96	288.44	4.56	0.84	4.0
2006	339.49	333.07	6.39	0.76	4.8
MAPE	–	7.4	–	–	–

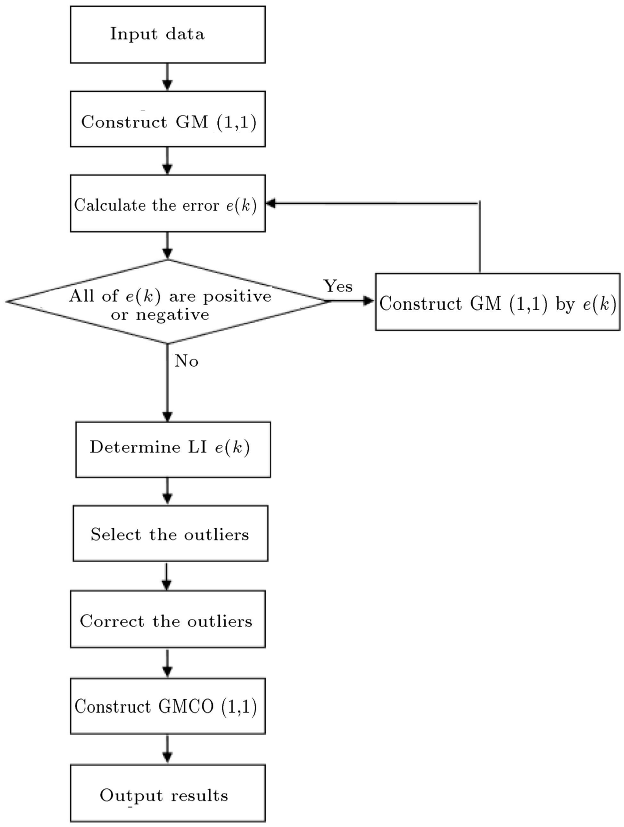

Figure 2. Flowchart of forming GMCO(1,1).

using traditional GM(1,1). Secondly, the natural gas consumption is predicted using GMCO(1,1). The data are obtained from the BP statistical review of World Energy in 2016, which can be found at the following link: http://www.bp.com/statisticalreview. Natural gas production (consumption) is measured in tonnes.

3.1.1. Natural gas production forecasting

In order to examine the current situation of natural gas production and predict the future trend, the 2007–2012 natural gas production value is set as the in-sample data. The natural gas production values from 2013 to 2015 are kept to verify the prediction accuracy. Mean Absoluten Percentage Error (MAPE): ($MAPE = 100\% \frac{1}{n} \sum_{k=1}^{n} |\frac{x^{(0)}(k)-\hat{x}^{(0)}(k)}{x^{(0)}(k)}|$), compares actual values with forecasted values to evaluate the precision. The results are reported in Table 2.

Both in-sample and out-of-sample prediction performance results show that the GM(1,1) enjoys higher accuracy; thus, it is suitable to forecast natural gas production. The short-term forecasting results are reported in Table 3.

3.1.2. Natural gas consumption forecasting

In order to examine the recent status of natural gas consumption and predict the trend, the 2007–2012 natural gas consumption values for China are set raw data. The errors of GM(1,1) and their LI values are shown in Table 4.

Table 2. The fitting values of GM(1,1).

Year	Actual value	GM(1,1)
2007	64.5	64.5
2008	74.8	75.0
2009	79.4	81.1
2010	89.2	87.8
2011	98.1	95.0
2012	100.7	102.9
MAPE	–	1.6
2013	110.0	111.3
2014	118.4	120.5
2015	124.2	130.4
MAPE	–	2.6

Table 3. The forecasting values of GM(1,1).

Year	GM(1,1)
2016	113.7
2017	122.6
2018	132.1

Table 4. The errors of GM(1,1) and their Latent Information (LI) values.

Year	Actual value	GM(1,1)	Error	LI value
2007	65.7	65.7	0	–
2008	75.7	73.9	1.96	0.61
2009	83.3	86.4	−2.53	0.21
2010	100.1	100.9	−0.95	0.64
2011	123.4	117.9	3.98	0.13
2012	135.8	137.7	−0.93	0.65
MAPE	–	2.5	–	–
2013	154.7	160.9	–	–
2014	169.6	188.0	–	–
2015	177.6	219.7	–	–
MAPE	–	12.9	–	–

As shown in Table 4, the forecasting accuracy from 2013 to 2015 is not acceptable. The outlier needs to be found and corrected. By means of the LI function, the natural gas consumption values in 2008 and 2011 are the outliers with the largest likelihood. This is because the natural gas consumption has been hit hard by the financial crisis of 2007–2008. Different experienced consumption values can correct the outliers: 65.7, 83.1,

Table 5. The forecasting value and Mean Absolution Percentage Error (MAPE) of GMCO(1,1).

Year	Actual value	GMCO(1,1)
2013	154.7	150.6
2014	169.6	172.3
2015	177.6	197.1
MAPE	–	5.1
2016	–	225.6
2017	–	258.1
2018	–	295.3

Table 6. The import values of natural gas in China.

Year	Import value
2016	111.9
2017	135.5
2018	163.2

83.3, 100.1, 111.4, 135.8. Only 83.1 and 111.4 may obtain the best forecasting accuracy from 2013 to 2015. Therefore, the values are corrected. The corrected sequence is {65.7, 83.1, 83.3, 100.1, 111.4, 135.8}; then, the corrected sequence is used to construct a GMCO(1,1). The short-term forecasting results of GMCO(1,1) are given in Table 5. As shown in Table 5, the forecasting value and MAPE of GMCO(1,1) are acceptable. This validation highlights the capacity of the proposed model to give a future outlook with better accuracy. It should be taken into account that the year (2013) model validation represents a tougher test, because strong changes in the trend are present. The proposed model demonstrates its capability to capture the rate of variations in gas consumption correctly, proving its strong prediction ability.

The short-term forecasting results given in Table 5 are considered reasonable. From Tables 3 and 5, the import values that China needs are seen in Table 6. According to Table 6, China will inevitably face a massive expansion of natural gas imports. As the largest developing country, China's massive imports of natural gas will affect the international natural gas price trend in the coming years.

3.2. USA

3.2.1. Natural gas production forecasting

From 2007 to 2012, the USA increased its natural gas production at a compound annual growth rate of 4.55%. In order to examine the current status of natural gas production and predict the future trend, the 2007–2012 natural gas production value is set as the in-sample data. The natural gas production values in 2013–2015 are kept to verify the prediction accuracy. The results are reported in Table 7.

Table 7. The fitting values of GM(1,1).

Year	Actual value	GM(1,1)
2007	498.6	498.6
2008	521.7	512.5
2009	532.7	536.4
2010	549.5	561.5
2011	589.8	587.7
2012	620.2	615.1
MAPE	–	1.1
2013	626.4	643.8
2014	669.1	673.8
2015	705.3	705.3
MAPE	–	1.1

Table 8. The forecasting values of GM(1,1).

Year	GM(1,1)
2016	738.2
2017	772.6
2018	808.7

Table 9. The errors of GM(1,1) and their Latent Information (LI) values.

Year	Actual value	GM (1,1)	Error	LI value
2007	596.3	596.3	0	–
2008	600.8	589.0	11.8	0.20
2009	590.1	603.7	–13.6	0.13
2010	619.3	618.9	0.37	0.97
2011	628.8	634.4	–5.6	0.64
2012	657.4	650.3	7.12	0.52
MAPE	–	1.3	–	–
2013	675.5	666.6	–	–
2014	692.7	683.3	–	–
2015	713.6	700.4	–	–
MAPE	–	1.6	–	–

Both in-sample and out-of-sample prediction performance results show that the GM(1,1) enjoys higher accuracy; thus, it is suitable to forecast natural gas production. The short-term forecasting results are reported in Table 8.

3.2.2. Natural gas consumption forecasting

To examine the recent status of natural gas consumption and predict the trend, the 2007–2012 natural gas consumption values for China were taken as the sample data. The errors of GM(1,1) and their LI values are shown in Table 9.

Table 10. The forecasting value and Mean Absolution Percentage Error (MAPE) of GMCO(1,1).

Year	Actual value	GMCO(1,1)
2013	675.5	672.7
2014	692.7	692.7
2015	713.6	713.3
MAPE	–	0.2
2016	–	734.5
2017	–	756.3
2018	–	778.8

Table 11. The export values of natural gas in the USA.

Year	Export values
2016	3.7
2017	16.3
2018	29.9

As shown in Table 9, the forecasting accuracy in 2013–2015 is not acceptable. The outlier needs to be corrected. By means of the LI function, the natural gas consumption values in 2008–2009 are the outliers with the smallest likelihood, because the natural gas consumption has been hit hard by the financial crisis of 2007–2008. The different experienced values of consumption are meant to correct the outlier and they are 596.3, 585.8, 592.1, 619.3, 628.8, 657.4. Only this corrected sequence can obtain the best forecasting accuracy for 2013–2015. Then, the corrected sequence is used to construct a GMCO(1,1). The short-term forecasting results of this model are presented in Table 10. As shown in Table 10, the forecasting value and MAPE of GMCO(1,1) are acceptable. This validation highlights the capacity of the proposed model to give a future outlook with better accuracy. These results given in Table 10 are considered reasonable. From Tables 8 and 10, the export values of the USA are seen in Table 11.

The U.S. has the ability to export 5.6 billion cubic feet of natural gas per day; however, the first volume of gas consumption moved out of the U.S. until sometime in 2015 [29]. The results shown in Table 11 imply that the export values of natural gas will increase in short term in the U.S. Thus, the results are consistent with the actual situation.

4. Conclusion and discussion

In the field of time series prediction, some of the data may be corrupted by outliers that may cause large estimation errors. Therefore, it is important to identify those outliers in order to remove or correct the corrupted data. Although known methods for outlier detection achieve good results, their complexity is usually high and their theoretical basis involves large samples. However, small data problems are considered as important issues in the early prediction. It is more difficult to detect the outliers from limited data. The proposed method can detect the outlier. The outlier can be corrected by a better predictive value. That is to say, the aim of correcting the outlier is to produce a better predictive value.

To increase the prediction and simulation accuracy of GM(1,1) model, the outliers were found by the Latent Information (LI) function. The sequence with corrected outliers was used to construct GM(1,1) model in this paper. The GMCO(1,1) could ensure better forecasting results, as demonstrated in this paper. It may be applied to other real prediction problems to further confirm the model effectiveness. This paper was subject to drawbacks, given that the method used to correct the outliers was rather a heuristic or a trial-and-error procedure. The grey model was proved suitable for the limited sample. LI function can detect the outlier of the limited sample. Linear regression as well as autoregressive and exponential smoothing models are suitable for a large sample. LI function may not be suitable for these models. In the future, the method of detecting the outliers from the large sample is worth researching.

5. Implication for practice

China has become the second largest economy in the world with a rapid economic growth. Restricted by the limited domestic natural gas resource, a large portion of natural gas needs to be imported from overseas. In terms of energy security, the Chinese government must advocate saving energy by adjusting the prices of natural gas, because China's natural gas price is still controlled by the government and has remained low so far. The massive imports will exert great pressure on China's gas price reform. Additionally, considering the low-carbon economic development in China, developing renewable energy should be the most feasible way for the future.

According to analysts at Goldman Sachs, the U.S. can realistically sustain about 7.7 billion cubic feet of Liquefied Natural Gas (LNG) exports by 2020 without significantly affecting natural gas prices. The USA, now nearly energy independent, could export LNG to both Europe and Asia. There are many good reasons for this. The USA has 100 to 120 years of proven reserves and could significantly boost its economy with billions of dollars a year that would pour in from a significant export program. In this paper, from a short forecasting viewpoint, the export values of natural gas will increase. This may cause significant impact on natural gas prices and energy policy.

Acknowledgements

This research was supported by the National Natural Science Foundation of China (No. 71871084, 71401051) and the project of high-level talent in Hebei province.

References

1. Karacaer-Ulusoy, M. and Kapusuzoglu, A. "The dynamics of financial and macroeconomic determinants in natural gas and crude oil markets: Evidence from organization for economic cooperation and development/gulf cooperation council/organization of the petroleum exporting countries", *International Journal of Energy Economics and Policy*, **7**(3), pp. 167–187 (2017).

2. Azadeh, A. Asadzadeh, S.M. Mirseraji, G.H., et al. "An emotional learning-neuro-fuzzy inference approach for optimum training and forecasting of gas consumption estimation models with cognitive data", *Technological Forecasting and Social Change*, **91**, pp. 47–63 (2015).

3. Ozturk, I. and Al-Mulali, U. "Natural gas consumption and economic growth nexus: panel data analysis for GCC countries", *Renewable and Sustainable Energy Reviews*, **51**, p. 9981003 (2015).

4. Furio, D. and Poblacion, J. "Electricity and natural gas prices sharing the long-term trend: some evidence from the Spanish market", *International Journal of Energy Economics and Policy*, **8**(5), pp. 173–180 (2018).

5. Tamba, J.G., Essiane, S.N., et al. "Forecasting natural gas: a literature survey", *International Journal of Energy Economics and Policy*, **8**(3), pp. 216–249 (2018).

6. Melikoglu, M. "Vision 2023: Forecasting Turkey's natural gas demand between 2013 and 2030", *Renewable and Sustainable Energy Reviews*, **22**, pp. 393–400 (2013).

7. Mohr, S.H. and Evans, G.M. "Long term forecasting of natural gas production", *Energy Policy*, **39**(9), pp. 5550–5560 (2011).

8. Boqiang, L. and Wang, T. "Forecasting natural gas supply in China: production peak and import trends", *Energy Policy*, **49**, pp. 225–233 (2012).

9. Primo, P., Boidar, S., Goran, I., et al. "Comparison of static and adaptive models for short-term residential natural gas forecasting in Croatia", *Applied Energy*, **129**, pp. 94–103 (2014).

10. Zhu, Q., Lu, Q.Y., Zhou, X.Y., et al. "A driving force analysis and forecast for gas consumption demand in China", *Mathematical Problems in Engineering*, **2014**, pp. 1–11 (2014).

11. Palivonaite, R., Lukoseviciute, K., and Ragulskis, M. "Short-term time series algebraic forecasting with mixed smoothing", *Neurocomputing*, **171**, pp. 854–865 (2016).

12. Palivonaite, R., Lukoseviciute, K., and Ragulskis, M. "Algebraic segmentation of short nonstationary time series based on evolutionary prediction algorithms", *Neurocomputing*, **121**(9), pp. 354-364 (2013).

13. Spoladore, A., Borelli, D., Devia, F., et al. "Model for forecasting residential heat demand based on natural gas consumption and energy performance indicators", *Applied Energy*, **182**, pp. 488–499 (2016).

14. Soldo, B. "Forecasting natural gas consumption", *Applied Energy*, **92**, pp. 26–37 (2012).

15. Wu, L., Liu, S., Yao, L., et al. "The effect of sample size on the grey system model", *Applied Mathematical Modelling*, **37**(9), pp. 6577–6583 (2013).

16. Mao, S., He, Q., Xiao, X., et al. "Study of the correlation between oil price and exchange rate under the new state of the economy", *Scientia Iranica*, **26**(4), pp. 2472–2483 (2019).

17. Li, G., Masuda, Sh., and Nagai, M. "The prediction for Japan's domestic and overseas automobile production", *Technological Forecasting and Social Change*, **87**, pp. 224–231 (2014).

18. Hamzacebi, C. and Es, H.A. "Forecasting the annual electricity consumption of Turkey using an optimized grey model", *Energy*, **70**(1), pp. 165–171 (2014).

19. Hashem-Nazari, M., Esfahanipour, A., and Fatemi Ghomi, S.M.T. "A basic form-focused modeling and a modified parameter estimation technique for grey prediction models", *Scientia Iranica, E.*, **25**(5), pp. 2867–2880 (2018).

20. Yi, Y., Yanhua, Ch., Jun, S., et al. "An improved grey neural network forecasting method based on genetic algorithm for oil consumption of China", *Journal of Renewable and Sustainable Energy*, **8**(2), p. 024104 (2016).

21. Chang, C., Li, D., Chen, C., et al. "A forecasting model for small non-equigap data sets considering data weights and occurrence possibilities", *Computers & Industrial Engineering*, **67**, pp. 139–145 (2014).

22. Li, G., Masuda, S., and Nagai, M. "Predicting the subscribers of fixed-line and cellular phone in Japan by a novel prediction model", *Technological Forecasting and Social Change*, **81**, pp. 321–330 (2014).

23. Xiuli, L., Blanca, M., and Salom, G. "A grey neural network and input-output combined forecasting model", *Primary Energy Consumption Forecasts in Spanish Economic Sectors, Energy*, **115**, pp. 1042–1054 (2016).
24. Wang, J., Jiang, H., Zhou, Q., et al. "Chinas natural gas production and consumption analysis based on the multicycle Hubbert model and rolling grey model", *Renewable and Sustainable Energy Reviews*, **53**, pp. 1149–1167 (2016).
25. Wu, L., Liu, S., Chen, H., et al. "Using a novel grey system model to forecast natural gas consumption in China", *Mathematical Problems in Engineering*, **2015**, pp. 1–7 (2015).
26. Chang, C., Li, D., Dai, W., et al. "A latent information function to extend domain attributes to improve the accuracy of small-data-set forecasting", *Neurocomputing*, **129**, pp. 343–349 (2014).
27. Vaghefi, M., Mahmoodi, K., and Akbari, M. "A comparison among data mining algorithms for outlier detection using flow pattern experiments", *Scientia Iranica, A.*, **25**(2), pp. 590–605 (2018).
28. Li, Z. and Tian, X. "The data outlier elimination method based on grey relation", *Proceedings of 2011 Cross Strait Quad-Regional Radio Science and Wireless Technology Conference*, Harbin, China (2011).
29. http://www.fool.com/investing/general/2013/08/11/how-much-natural-gas-can-americaexport.aspx
30. Lianfu, H., Changfeng, F., Jun, W., et al. "Outlier detection and correction for the deviations of tooth profiles of gears", *Measurement Science Review*, **13**(2), pp. 56–62 (2013).
31. Lee, Y.S. and Tong, L.I. "Forecasting energy consumption using a grey model improved by incorporating genetic programming", *Energy Conversion and Management*, **52**(1), pp. 147–152 (2011).
32. Tsai, C.F. "Dynamic grey platform for efficient forecasting management", *Journal of Computer and System Sciences*, **81**(6), pp. 966–980 (2015).

Appendix

$LI\ e(k)$ is proper for detecting outliers. The real case (deviations of tooth) is considered to demonstrate the effectiveness of detecting outliers. The data were taken from [30]. $e(k)$ and $LI\ e(k)$ are listed in Table A.1.

The data (2.1) has larger $e(k)$ and smaller $LI\ e(k)$. Therefore, 2.1 is an outlier. This result is consistent with that in [30].

Four real cases are considered to demonstrate the effectiveness of GMCO(1,1). They are given in the following:

Case 1. The benchmarking data set was taken from [31]. These models estimated Chinese energy consumption over the years 1990–2003. The MAPE values from 2004 to 2007 (model testing) are 26.2%, 27.8%, and 5.1%, respectively, as shown in Table A.2. The results indicate that the GMCO(1,1) has the smallest MAPE values among these models;

Case 2. The benchmarking data set was taken from [32]. These models estimated tourism demand during the period of 1989–2000. The MAPE values of the year 2001 (model testing) are 10.6%, 6.0%, and 0%, as shown in Table A.3. The results indicate that the GMCO(1,1) has the smallest MAPE values among these models;

Case 3. These models estimated the output of opto-electronics industry during the period of 1990–2005. The MAPE values from the years 2006 to 2008 (model testing) are shown in Table A.4. The results indicate that the GMCO(1,1) has the smallest MAPE values among these models;

Case 4. The benchmarking data set is borrowed from [32]. These models estimated tourism demand during the period of 1990–2001. The MAPE values from 2002 to 2005 (model testing) are shown in Table A.5. The results indicate that the GMCO(1,1) has the smallest MAPE values among these models.

Table A.1. The forecasting value and Mean Absolution Percentage Error (MAPE) of different models.

Order	Modeling data	$e(k)$	$LI\ e(k)$
1	3.0145	0.24	0.40
2	3.2297	0.29	0.29
3	3.3106	0.20	0.51
4	3.5751	0.28	0.33
5	2.1	−1.39	0.007
6	3.8370	0.14	0.66
7	3.9965	0.08	0.81
8	4.2424	0.08	0.80
9	4.3983	−0.01	0.99
10	4.7465	0.08	0.81
11	4.9828	0.03	0.92
12	5.3216	0.07	0.82
13	5.5961	0.04	0.91

Table A.2. The forecasting value and Mean Absolution Percentage Error (MAPE) of different models.

Year	Actual value	GM(1,1)	Linear regression [31]	GMCO(1,1)
2004	203227	166600	164572	209220
2005	224682	172163	169059	222183
2006	264270	177911	173546	235948
2007	265583	183851	178032	250566
MAPE	–	26.2	27.8	5.1

Table A.3. The forecasting value and Mean Absolution Percentage Error (MAPE) of different models.

Year	Actual value	GM(1,1)	ARIMA (0,1,0) [32]	GMCO(1,1)
2001	33716	37301	35722	33747
MAPE	–	10.6	6.0	0

Table A.4. The forecasting value and Mean Absolution Percentage Error (MAPE) of different models.

Year	Actual value	GM(1,1)	EGM(1,1)	ARIMA (0,1,0) [32]	GMCO(1,1)
2006	12774	16096	19368	11770	15173
2007	20665	20709	25389	12488	18521
2008	20095	33282	26645	13207	22607
MAPE	–	43.5	20.0	27.2	13.9

Table A.5. The forecasting value and Mean Absolution Percentage Error (MAPE) of different models.

Year	Actual value	GM(1,1)	EGM(1,1)	ARIMA (0,1,0) [32]	GMCO(1,1)
2002	6529	8085	7097	5690	6456
2003	8188	10010	8786	6111	7777
2004	10990	12395	10877	6531	9368
2005	11141	15347	13466	6952	11284
MAPE	–	24.2	9.5	29.1	5.5

Biographies

Lifeng Wu is an Associate Professor of the College of Management Engineering and Business, Hebei University of Engineering. His main research direction is grey system modeling. He has authored 40 papers that appeared in journals including Technological Forecasting & Social Change, IEEE Transactions on Systems, Man, and Cybernetics: Systems, Information Sciences, Applied Soft Computing, Energy, Applied Mathematical Modelling, Communications in Nonlinear Science and Numerical Simulation, Applied Mathematics and Computation, and International Journal of Intelligent Systems and Soft Computing.

Kai Zhang is a Master's student of the College of Management Engineering and Business, Hebei University of Engineering. His main research direction is grey system modeling.

Ting Zhao is a Lecturer of the College of Management Engineering and Business, Hebei University of Engineering. Her main research direction is grey forecasting modeling.