Journal Pre-proof

Predicting monthly natural gas production in China using a novel grey seasonal model with particle swarm optimization

Nu Li, Jianliang Wang, Lifeng Wu, Yongmei Bentley

PII: S0360-5442(20)32225-8

DOI: <https://doi.org/10.1016/j.energy.2020.119118>

Reference: EGY 119118

To appear in: *Energy*

Received Date: 5 August 2019

Revised Date: 3 October 2020

Accepted Date: 17 October 2020

Please cite this article as: Li N, Wang J, Wu L, Bentley Y, Predicting monthly natural gas production in China using a novel grey seasonal model with particle swarm optimization, *Energy*, <https://doi.org/10.1016/j.energy.2020.119118>.

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2020 Elsevier Ltd. All rights reserved.

Nu LI: Data Curation, Software, Validation, Formal analysis. **Jiangliang WANG:** Conceptualization, Methodology, Investigation, Writing- Original draft preparation, Supervision. **Lifeng WU:** Conceptualization, Methodology. **Yongmei BENTLEY:** Conceptualization, Writing- Reviewing and Editing.

Predicting monthly natural gas production in China using a novel grey seasonal model with particle swarm optimization

Nu Li ^a, Jianliang Wang ^{a,*}, Lifeng Wu ^b, Yongmei Bentley ^c

^a School of Business Administration, China University of Petroleum, Beijing, 102249, China

^b College of Economics and Management, Hebei University of Engineering, Handan, 056038, China

^c Business School, University of Bedfordshire, Luton, UK

Corresponding Author: E-mail address: wangjianliang305@163.com.

Abstract: Accurate prediction of short and medium-term monthly natural gas production in a country is the basis for understanding the supply capacity of natural gas in different months, and for the timely adjustment of natural gas production and import strategies. In China the monthly production of natural gas has obvious seasonal and cyclical variations, thus the use of a traditional grey prediction model is not very effective. As a result, a novel grey seasonal model is proposed in this paper. This is the Particle swarm optimized Fractional-order-accumulation non-homogenous discrete grey Seasonal Model (PFSM(1,1) model). This model enhances the adaptability to seasonal fluctuation data in two ways: the seasonal adjustment of the original data, and improvement of model self-adaptability. We use monthly natural gas production data of China for the period 2013-2018 as samples to predict those for the period 2019-2023. To demonstrate the PFSM(1,1) model does indeed exhibit better predictive capability, we also use the Holt–Winters model and a seasonal GM(1,1) model to predict monthly natural gas production, and compare the results with the model proposed here. The prediction results show that monthly natural gas production in China will continue to increase throughout the 2019-2023 period, that the peak-to-valley differences in monthly production values will also increase, and that the seasonal variations in production will become increasingly pronounced. Moreover, although Chinese production of natural gas is increasing, it will still be difficult to meet future demand, and hence the gap between supply and demand will increase year by year. We conclude that China needs to develop a more complete import plan for gas to meet expected natural gas consumption.

Keywords: Grey seasonal model; Fractional order accumulation; Particle swarm optimization; Monthly production of natural gas; Gap between supply and demand

1. Introduction

1.1 Background

Currently, as China is strongly advocating low-carbon development, natural gas has become a key energy choice in China's future energy strategy. However, natural gas currently accounts for a much smaller proportion of China's energy mix than the global average, and although the country has implemented a series of measures to improve its natural gas infrastructure, pricing systems and monitoring mechanisms, China has yet to increase this proportion significantly because of limits to its natural gas production capacity.

The changes in China's natural gas production and consumption growth rates from 2007 to 2017 are shown in Fig. 1 (data from the National Bureau of Statistics of China <http://www.stats.gov.cn/>). It can be seen from Fig. 1 that in recent years the growth rate of natural gas production in China has nearly always been lower than its growth rate in consumption, which has worsened the gap between domestic production and demand. To fill this gap, in 2018 for example, China imported 121.3 billion cubic meters (bcm) of natural gas, which represented a year-on-year increase of 31.9%, and accounted for 12.9 % of global natural gas trade, making China the world's largest natural gas importer [1]. The global market for natural gas has long been one of oversupply, hence the increase in China's natural gas imports has effectively helped stabilize the global gas supply market. Moreover, with the significant increase of liquefied natural gas (LNG) trade, the regional natural gas market is expected to increasingly develop into an international market. As the world's largest natural gas buyer, China's import volume changes will to a certain extent affect the buyer's market initiative, and have an impact on the existing natural gas pricing model.

For these reasons, accurately predicting future production of natural gas in China is important for understanding the process of China's reform of its energy structures, including its natural gas production, utilization and import strategies, as well as for understanding the development prospects of the world's natural gas market as a whole.

Fig. 1 Growth rates of the natural gas production and consumption in China 2007-2016

Research on the prediction of natural gas production covers multiple time scales and prediction perspectives. In terms of time scales, the research can be divided into long-term prediction (usually more than five years) and medium- and short-term prediction (less than five years), while prediction perspectives reflect mainly differentiating annual forecasting from monthly or quarterly forecasting. Different time scales and prediction perspectives reflect different aspects of natural gas production, and are complementary, such that each research path is useful. Long-term production forecasts support long-term strategy making; while medium and short-term forecasts generally support more tactical decisions.

A review of the literature shows that current studies of China's future natural gas production are mainly long-term forecasts. Examples include the following: A multi-cyclic Gauss model was used to predict natural gas reserves and production in the Ordos Basin for the 2011-2030 period [2]. Generalized prediction models types I and II, and also the Weng model, have been used to make predictions of natural gas production in the Sichuan Basin for the 2021-2030 period [3]. A multicycle Hubbert model, and a rolling grey model, have been applied to forecast natural gas production of China up to 2050 [4]. Production of conventional gas (including tight gas), coalbed methane and shale gas in China up to 2050 have been predicted using multiple methods [5]. The multicycle Generalized Weng model has been chosen to predict the long-term trend of natural gas production in China up to 2100 [6]. And a geologic-resources supply-demand model was used to forecast production of unconventional gas in China to 2200 [7]. These forecasts are mainly of annual production, and the prediction periods are long-term, and hence the 'annual + long-term' perspective is the focus of a large proportion of the research into natural gas production in China.

These existing ‘annual + long term’ prediction models provide an important basis for the formulation of China’s natural gas development strategies, and help put the country’s natural gas development on a scientific basis with a view to avoiding imbalance between the ratios of different energy supplies that might follow uncontrolled exploitation of one energy source vs. another, and thus help ensure the overall sustainable development of the country’s energy supplies. However, this does not mean that medium and short-term predictions of gas production have less value, as specific medium and short-term plans for the country’s natural gas development can then be formulated, which, if combined with existing policies on environmental factors, technical conditions, energy market dynamics and other factors, can support the healthy development of the natural gas supply market.

In this context it is important to note that from a monthly or quarterly perspective, China's medium and short-term natural gas production is characterized by significant seasonal and cyclical fluctuations, as shown for example in Fig. 2 for the period January 2013 to December 2018. As can be seen, production of natural gas in China generally tends to be higher in the autumn and winter months, and lower in spring and summer months. This is mainly due to the impact of the consumer market, where winter heating in the north of the country increases the amount of gas required. However, this important feature of seasonal fluctuations in production seems to have received relatively little attention within the academic community, as currently there are only two academic studies known to us that focus on forecasting monthly gas production, and both apply to shale gas production in the United States. One used a hybrid ARIMA and metabolic nonlinear grey model to predict the U.S. shale gas monthly production [8], and the other a method that integrated linear and nonlinear forecasting techniques, based on grey theory and artificial intelligence, to forecast monthly shale gas production in Pennsylvania and Texas [9]. As far as we have been able to ascertain, predictions of monthly natural gas production in China have not been reported in the academic literature.

Fig. 2 Monthly data of natural gas production in China during the period 2013-2018

Note: Non-zero ordinate.

Data sources:

China's National Development and Reform Commission (<http://www.ndrc.gov.cn/>)

International Gas (<http://gas.in-en.com/html/gas-2399865.shtml>)

The China Reporting (<http://market.chinabaogao.com/nengyuan/041I302A2018.html>)

The National Bureau of Statistics (<http://www.stats.gov.cn/>).

With the increased implementation of air pollution mitigation and clean-heating policies in China, we expect that the gas usage peak in winter will continue to grow, and the difference between seasonal peaks and valleys in natural gas supply will correspondingly increase year-on-year. For this reason, this paper researches prediction models with a 'monthly + medium and short-term' perspective, to provide an improved understanding of China's natural gas production. In addition, the prediction results can provide the data to support investigation of the impact of environmental protection and energy policies on natural gas medium and short-term supply, and hence in the planning of China's natural gas development.

In this paper we select a grey prediction model for forecasting the production of natural gas in China. This is because such models are designed to reduce the randomness of input data, and mine the potential 'law of data' through the action of a grey generation or sequence operator, to predict the changing trend of a time series when the sample data set is small [10]. Such a model has the benefit that it has been widely used in energy forecasting. More specifically, this paper proposes a particular type of grey prediction model for monthly gas production of China, that of the 'particle swarm optimized fractional-order-accumulation non-homogeneous discrete grey seasonal model (PFSM(1,1)). Details of this model are set out in Section 2 of this paper.

1.2 Literature review

To set the above model into context, this section briefly reviews the application of grey prediction models for natural gas forecasting, and introduces some of the different ways that have been employed for seasonal improvement for grey prediction models.

1.2.1 Application of grey prediction models in natural gas forecasting

Up to now, application of grey prediction models for natural gas prediction has mainly focused on consumption and production. For different types of data, researchers have chosen different grey prediction models to solve the individual problems faced. The literature related to the prediction of natural gas by grey models over the last ten years is summarized in Table 1.

This table shows that the time-spans of grey models for natural gas prediction are all within 10 years, which supports the idea that the grey prediction is most applicable to small sample prediction. The literature shows that use of grey prediction models over the recent ten years has mainly been for the prediction of natural gas consumption, while the prediction of production has been less. Moreover, most models have been used for annual prediction, with only references [8] and [9] being used to forecast monthly production, and where, as mentioned, these have focused on the United States. So far there has been no use of grey models for the prediction of monthly natural gas production in China. This paper sets out to fill this gap, and hence better understand the developmental changes of natural gas production in China.

1.2.2 Seasonal improvement of grey prediction models

With the increasing application of grey prediction models, the forms of data involved are becoming more complex. The traditional grey prediction model that is applicable to the data of exponential change cannot meet the needs of many practical applications. Data with seasonal changes has become of increasing interest in the use of grey prediction models in recent years because of its wide application. In order to make the grey model suitable for the prediction of seasonal changes, numerous researchers have suggested improvements. The literature on seasonal improvements of grey models over the recent ten years is summarized in Table 2.

As the table shows, while there are many methods for making seasonal improvements to the grey prediction model, these mainly focus on two aspects. 1) The first is the improvement of model self-adaptability, and where improvements to the GM(1,1) power model based on oscillating data accounts for a large proportion of this class of improvement. Moreover, data with seasonal fluctuations are oscillating data, and the introduction of the GM(1,1) power model designed for oscillating data, and the improvement of its prediction performance, have broadened the application field of the grey prediction model, making this model more suitable for the prediction of data with different changing trends. Based on the improvement methods of other

models, the improvement of model self-adaptability is done either by improving the prediction model without changing the trends of the original data, or by the establishment of a combined model, so that the model is more suitable for the prediction when using seasonal data. 2) The second aspect is the seasonal adjustment of the original data, which is to process the original data to eliminate its seasonal change and then use the grey prediction model with these processed data in order to achieve better predictions.

Two studies [41,42] further improve the model by combining the above two improvement aspects. The research focus of reference [41] is the oscillation sequence. The basic idea is to pre-process the original data through an accelerated translation and weighted-mean generation operation, and then modify the background value of the GM(1,1) model to improve its prediction accuracy. Reference [42] instead adjusts the initial values by separating out the seasonal index of the original data, and then applies an adaptive parameter learning mechanism to the SFGM(1,1) model to improve its prediction performance. The model proposed in this paper combines the improvement concepts of references [41] and [42]. However, the data processing method and prediction model itself that we use are greatly different from those in these two references., and where our new method is more suitable for the prediction of monthly natural gas production in China. We call our method, the ‘Particle swarm optimized fractional-order-accumulation non-homogenous discrete grey seasonal model’, i.e., PFSM(1,1).

Table 1 Literature on the application of grey models for predicting natural gas production or consumption

Reference	Authors	Countries or areas covered	Research area	Time-span	Research perspective	Methods
[11]	Xunbo Shuai (2010)	China	Production	1995-2001	Year	Grey model based on exponential smoothing
[12]	Li and Wang (2014)	China	Production	2002-2011	Year	Grey prediction model based on Generalized Weng Model's logarithmic equation
[13]	Weibing Chen (2015)	Shanghai, Shanxi and Shaanxi, China	Consumption	2012	Year	Grey inhomogeneous model
[14]	Liu et al. (2016)	Spain	Consumption	2010-2015	Year	Grey neural network and input-output combined forecasting model
[15]	Zhang and Yang (2016)	29 provinces China	Consumption	2016-2020	Year	Grey metabolic GM(1,1) prediction model
[16]	Jingying Guo (2017)	China	Consumption	2006-2015	Year	Grey relational degree combination model
[17]	Faheemullah et al.(2017)	China	Consumption	2014-2020	Year	Grey Verhulst Model and the Nonlinear Grey Bernoulli Model
[18]	Ma and Liu (2017)	China	Consumption	2005-2013	Year	Time-delayed polynomial grey prediction model
[19]	Wu and Shen (2017)	China	Consumption	2013-2015	Year	Grey-related least squares support vector machine optimization model
[20]	Ye and Chen (2018)	China	Consumption	2007-2016	Year	Residual grey model
[21]	Zeng et al. (2018)	China	Production	2017-2025	Year	Unbiased grey prediction model
[22]	Song Ding (2018)	China	Consumption	2015-2020	Year	Self-adapting intelligent grey model
[23]	Wang et al. (2018)	U.S.	Production	2014-2016	Monthly	Hybrid ARIMA and metabolic nonlinear grey model
[24]	Ma et al.(2019)	Chongqing, China	Consumption	2007-2016	Year	Fractional time delayed grey model
[25]	Ma et al.(2019)	UAE,Brazil,Bolivia,Denmark,Netherlands,Qatar,Nigeria,Turkmenistan,Brunei,Italy,India	Production	2008-2016	Year	The conformable fractional grey system model
[26]	Hu et al.(2019)	Sichuan, China	Consumption	2018-2027	Year	Fractional order grey neural network combination model
[27]	Wang and Jiang (2019)	U.S.	Production	2017-2018	Monthly	Nonlinear metabolic grey model - autoregressive integrated moving average model

Table 2 Literature on the seasonal improvement of grey models

Reference	Authors	Methods	Improved aspects	Notes
[28]	Wang et al. (2011)	Modeling approach for oscillatory sequences based on GM(1, 1) power model		
[29]	Zhengxin Wang (2013)	The oscillating GM(1,1) power model with system latency and time-varying parameters	GM(1,1) power model is a nonlinear gray model, in which power exponents can flexibly determine the form of the model. These models improve and optimize the GM(1,1) power model in different ways to make it suitable for the prediction of small-sample oscillation sequences.	Improvement of model adaptability
[30]	Zhengxin Wang (2014)	Grey forecasting method for small sample oscillating sequences based on Fourier series		
[31]	Wang and Luo (2017)	Fractional order discrete grey GM(1, 1) power model based on oscillation sequences		
[32]	Liang Zeng (2019)	Grey GM(1, 1 sin) power model based on oscillation sequences		
[33]	Yang and Cheng (2010)	Grey-cycle epitaxial combination model	Based on GM(1,1) model, the residual periodic epitaxy model is established.	Improvement of model adaptability
[34]	Wu and Wang (2012)	Integrated optimum grey support vector machine model	GM(1,1) and support vector machine (SVM) are weighted by particle swarm optimization (PSO) to obtain a combined model that minimizes the prediction error and makes it more suitable for seasonal data prediction.	Improvement of model adaptability
[35]	Zhang et al. (2017)	Modified Grey model using Fourier series	Using the Fourier series to improve the prediction performance of the original grey model	Improvement of model adaptability
[36]	Rui Guo (2018)	Grey seasonal variation index model	The variable prediction equation is transformed by the variable period, and the prediction equation suitable for periodic data is obtained.	Improvement of model adaptability
[37]	Xiao et al. (2013)	An improved seasonal rolling grey forecasting model	The sequence of seasonal fluctuation is converted to a flat sequence using the cycle truncation accumulated generating operation	Seasonal adjustment of original data
[38]	Wang et al. (2017)	The data grouping approach based grey modelling method	By grouping the original data according to the seasons, a monotone increasing time series is obtained, so as to eliminate the seasonal fluctuation of the original data.	Seasonal adjustment of original data

[39]	Wang et al. (2017)	The seasonal GM(1,1) model	The seasonal fluctuation of the original data is separated by dividing the original data by the seasonal index.	Seasonal adjustment of original data
[40]	Qian and Dang (2009)	GM(1,1) model based on oscillation sequences	It turns the oscillation sequence into monotone sequence by means of accelerated translation theory, then using the weighted mean value generating operation.	Seasonal adjustment of original data
[41]	Fu and Gao (2010)	GM(1, 1) model with modified background values	After preprocessing the original data with by means of accelerated translation and weighted mean value generating operations, the background values of a GM(1,1) model was modified	Both
[42]	Wang et al. (2012)	APL-SFGM(1,1) model	Firstly, the seasonal index of the original data is separated to generate the SFGM(1,1) model, then an adaptive parameter learning mechanism is applied to the SFGM(1,1) model	Both

1.3 Aim, Contribution and Organization

As mentioned earlier, China's natural gas production has an important impact on the natural gas market of both China and the world, and the aim of this paper is to improve the accuracy of predictions of monthly natural gas production in China. Based on the literature review of existing natural gas grey prediction models summarized in Table 1, and the seasonal improvements shown in Table 2, the main contributions of this paper are shown in Fig. 3, and as follows:

Fig. 3 Contributions of this research

The first contribution of this paper is that we focus on China's medium- and short- term production (MSTP) of natural gas, as well as on seasonal fluctuations in MSTP by using monthly data. As stated in section 1.1, both long-term production (LTP) and MSTP analyses are necessary and important for better understanding China's future natural gas production. However, current studies mainly focus on 'annual+LTP' analysis, and only two studies focus on 'monthly+MSTP' analysis, and where the latter examine U.S. shale gas production. Therefore, in terms of natural gas prediction, 'China+MSTP+Monthly' analysis, which we focus on, is a new perspective, and constitutes a unique contribution to the research area.

The second contribution of this paper is that we propose a novel grey seasonal prediction model, the PFSM(1,1) model. This model combines two seasonal improvement methods: firstly, the monthly data (which exhibit seasonal fluctuations) is summed for each year to obtain the

annual data to realize the seasonal adjustment of original data. Secondly, a fractional order accumulation operator is introduced into the non-homogenous discrete grey model, and particle swarm optimization is used to optimize the model, so as to realize adaptive improvement of the model. Although there are many seasonal improvement methods about grey model, the modeling idea of the PFSM(1,1) model is novel. In particular, there is no similar research idea in the existing literature on the processing method of original data and the combination method with seasonal factors. Therefore, PFSM(1,1) model provides a new idea for the seasonal improvement of grey model.

Following the introduction and literature review given above, the remainder of this paper is organized as follows: The methods and specific model used in this study are described in Section 2. The prediction of the monthly natural gas production in China is presented in Section 3. ‘Out of sample’ prediction results and discussion are given in Section 4, and the conclusions and outlook are given in Section 5.

2. Methods

2.1 Model selection

Research on prediction can be divided into two main approaches. One is factor-driven structural prediction (multivariate). This kind of prediction seeks to first determine the influencing factors of the target problem, and then to predict this target problem based on future changes of these influencing factors. The other approach is data-driven extrapolation prediction (univariate). This starts from the historical data of the target problem, and forecasts the future by analyzing the data itself. Both methods have their advantages and disadvantages. The first considers the influence of external factors on the prediction results, but has the disadvantage that the prediction of the target problem depends on the future trend of the influencing factors. It is necessary to make corresponding prediction on the influencing factors first, and then use the prediction data to predict the target problem, which can cause accumulation of prediction errors and increase the uncertainty of the prediction results. The second approach only relies on the target data, which hence requires less data, and the prediction method can be simple, but this approach lacks the consideration of influencing factors of the prediction process.

Both kinds of prediction approaches are widely used in academic circles. For example, the literature summarized in Section 1.2 follows the second kind of prediction approach, and our research here also belongs to this. As for the first prediction approach, in the current prediction research of grey model, the data of known influencing factors are mostly taken as samples to study the prediction performance of the model [43-45], and it is rare to predict target problems with unknown influencing factor data. The research in this paper is to predict the unknown data in the

future, at the same time, it is difficult to obtain the influencing factors monthly data of natural gas production. Therefore, in order to ensure the validity of the prediction, this paper adopts the second approach. In this paper, a type of univariate grey prediction model, the non-homogenous discrete grey model (NDGM(1,1)), is used as the basic model on which we carry out both optimization and improvement.

2.2 Meanings of abbreviations, symbols and equations

In the following sections, we will abbreviate the models' name. Table 3 shows the abbreviations of the models and their corresponding meanings. In addition, the meanings represented by symbols and equations are shown in Table 4 and 5 respectively.

Table 3 The abbreviation of the model's name and their meanings

Index	Abbreviation	Meaning
1	NDGM(1,1)	Non-homogenous Discrete Grey Model
2	NDGM ^r (1,1)	Non-homogenous Discrete Grey Model with Fractional order accumulation
3	PFSM(1,1)	Particle swarm optimized Fractional-order-accumulation non-homogenous discrete grey Seasonal Model
4	PSO	Particle Swarm Optimization
5	SGM(1,1)	Seasonal GM(1,1) Model
6	HW	Holt-Winters model

Table 4 The meaning of symbols

Index	Symbol	Meaning
1	$X^{(0)}$	An original time sequence
2	$X^{(1)}$	First-order accumulation generation (AGO) sequence of $X^{(0)}$
3	$X^{(r)}$	r-order accumulation generation sequence of $X^{(0)}$
4	$\hat{X}^{(0)}$	The simulation time sequence of $X^{(0)}$
5	$Y^{(0)}$	The annual data sequence
6	$\hat{Y}^{(0)}$	The simulation time sequence of $Y^{(0)}$
7	r	Accumulation order
8	β	The parameter of NDGM(1,1) or NDGM ^r (1,1) model
9	B, Y	The intermediate variables in the calculation of parameter β with the least square method

- 10 $\alpha^{(r)} \hat{X}^{(r)}$ r-order inverse accumulated generating operator (IAGO) of $\hat{X}^{(r)}$
- 11 S_l The seasonal factor

Table 5 The meaning of equations

Index	Equation	Meaning
1	$x^{(1)}(k+1) = \beta_1 x^{(1)}(k) + \beta_2 k + \beta_3$	The original form of NDGM(1,1)
2	$\hat{x}^{(1)}(k+1) = \hat{\beta}_1^k \hat{x}^{(1)}(1) + \hat{\beta}_2 \sum_{j=1}^k j \hat{\beta}_1^{k-j} + \frac{1 - \hat{\beta}_1^k}{1 - \hat{\beta}_1} \hat{\beta}_3$	The time response function of NDGM(1,1)
3	$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$	The regressive reduction formula of NDGM(1,1)
4	$x^{(r)}(k+1) = \beta_1 x^{(r)}(k) + \beta_2 k + \beta_3$	The original form of NDGM ^r (1,1)
5	$\hat{x}^{(r)}(k+1) = \hat{\beta}_1^k \hat{x}^{(1)}(1) + \hat{\beta}_2 \sum_{j=1}^k j \hat{\beta}_1^{k-j} + \frac{1 - \hat{\beta}_1^k}{1 - \hat{\beta}_1} \hat{\beta}_3$	The time response function of NDGM ^r (1,1)
6	$\alpha^{(1)} \hat{x}^{(r)(1-r)}(k) = \hat{x}^{(r)(1-r)}(k) - \hat{x}^{(r)(1-r)}(k-1)$	The regressive reduction formula of NDGM ^r (1,1)
7	$S_l = \frac{\bar{x}_l^{(0)}(i)}{\bar{x}^{(0)}(i)}$	The expression of seasonal factor
8	$\hat{x}^{(0)}(i) = \hat{y}(\left\lfloor \frac{i-1}{L} \right\rfloor + 1) \cdot \frac{S_{i - \left\lfloor \frac{i-1}{L} \right\rfloor \cdot L}}{L}$	The restored value of sequence $X^{(0)}$

2.3 Modeling

This section introduces the modeling process of the three models, and the relationships among the three models are independent and progressive. The PFSM(1,1) model proposed in this paper is improved on the basis of NDGM(1,1) model. The improvement process and the correlation among models are shown in Fig.4. The modeling process of the three models is described in detail below.

Fig.4 Correlation between models

2.3.1 The NDGM(1,1) model

While traditional GM(1,1) model and the discrete GM(1,1) model are suitable for the prediction of original data sequences with a homogeneous index trend, in practical applications, only a few systems can meet this requirement, as most original data sequences are in accord with non-homogeneous index trends. However, the NDGM(1,1) model is seen as a prediction model suitable for a non-homogeneous index sequence [46]. Thus the NDGM(1,1) model can solve the problem that the traditional grey prediction model cannot, in terms of predicting a non-homogeneous index sequence, and extends grey model prediction to a wider application field. The modeling process is as follows:

Let the original time sequence be: $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\}$, where the first-order accumulation generation sequence of the original sequence is

$$X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \dots, x^{(1)}(n)\} \quad (1)$$

And where $x^{(1)}(k) = \sum_{i=1}^k x^{(0)}(i), k = 1, 2, \dots, n$.

The NDGM(1,1) model is defined by the equation:

$$x^{(1)}(k+1) = \beta_1 x^{(1)}(k) + \beta_2 k + \beta_3 \quad (2)$$

where β_1, β_2 and β_3 are parameters of the NDGM(1,1) model and can be obtained by using the least-squares method. The formulas for determining the β parameters are as follows:

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = (B^T B)^{-1} B^T Y, \quad (3)$$

where

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

The time response function of the NDGM(1,1) model is:

$$\hat{x}^{(1)}(k+1) = \hat{\beta}_1^k \hat{x}^{(1)}(1) + \hat{\beta}_2 \sum_{j=1}^k j \hat{\beta}_1^{k-j} + \frac{1-\hat{\beta}_1^k}{1-\hat{\beta}_1} \hat{\beta}_3, k = 1, 2, \dots, n-1 \quad (4)$$

The regressive reduction formula of NDGM (1,1) is:

$$\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1), k = 2, 3, \dots, n \quad (5)$$

2.3.2 The NDGM^r(1,1) model

The difference between grey prediction models and other models is that the data used by grey prediction models are the accumulated data, rather than the original data. Use of a fractional order accumulated generating operator enriches and improves the grey system theory, satisfies the principle of new information priority to a great extent, and improves the prediction accuracy and stability of the model [47]. It has been shown that when a fractional order accumulated generating operator is applied to the NDGM(1,1) model it much improves the accuracy and stability of the model [48]. The modeling process of the NDGM(1,1) model is as follows:

Let the original non-negative sequence $X^{(0)}$ have the r -order ($r \in R_+$) accumulated generating sequence as $X^{(r)}$,

$$X^{(r)} = \{x^{(r)}(1), x^{(r)}(2), \dots, x^{(r)}(n)\},$$

Set

$$x^{(r)}(k) = \sum_{j=1}^k \binom{k-j+r-1}{k-j} x^{(0)}(j), k=1, 2, \dots, n, \quad (6)$$

$$\text{where } \binom{r-1}{0} = 1, \binom{k-1}{k} = 0, \binom{k-j+r-1}{k-j} = \frac{(k-j+r-1)(k-j+r-2)\dots(r+1)r}{(k-j)!}$$

is the general Newton binomial coefficient.

The equation

$$x^{(r)}(k+1) = \beta_1 x^{(r)}(k) + \beta_2 k + \beta_3 \quad (7)$$

is called the NDGM^r(1,1) model; when $r=1$ it is the traditional NDGM(1,1) model.

Similarly, the parameters are obtained by the least square method:

$$\begin{pmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \hat{\beta}_3 \end{pmatrix} = (B^T B)^{-1} B^T Y, \quad (8)$$

where

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

The time response function of the NDGM^r(1,1) model is:

$$\hat{x}^{(r)}(k+1) = \hat{\beta}_1^k \hat{x}^{(1)}(1) + \hat{\beta}_2 \sum_{j=1}^k j \hat{\beta}_1^{k-j} + \frac{1 - \hat{\beta}_1^k}{1 - \hat{\beta}_1} \hat{\beta}_3 \quad (9)$$

For $\hat{X}^{(r)} = (\hat{x}^{(r)}(1), \hat{x}^{(r)}(2), \dots, \hat{x}^{(r)}(n), \dots)$, the r ($0 < r < 1$) order inverse accumulated generating operator is shown as follows:

$$\hat{X}^{(0)} = \alpha^{(r)} \hat{x}^{(r)}(k) = \{\alpha^{(1)} \hat{x}^{(r)(1-r)}(1), \alpha^{(1)} \hat{x}^{(r)(1-r)}(2), \dots, \alpha^{(1)} \hat{x}^{(r)(1-r)}(n), \dots\} \quad (10)$$

where $\alpha^{(1)} \hat{x}^{(r)(1-r)}(k) = \hat{x}^{(r)(1-r)}(k) - \hat{x}^{(r)(1-r)}(k-1)$.

The predicted values are: $\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \dots, \hat{x}^{(0)}(n), \hat{x}^{(0)}(n+1), \dots$

2.3.3 Optimization of the fractional order r based on PSO

For the NDGM^r(1,1) model, the value of order r has a great influence on the prediction accuracy of the model. The main idea of finding the optimal value of r is to minimize the error of the NDGM^r(1,1) model. The mean absolute percentage error (MAPE) is usually used to represent the model error (the calculation of MAPE is given in Section 2.4). However, the determination of the optimal order r requires a large number of repeated operations, which is difficult to achieve using traditional methods. PSO can provide a good method for the determination of the optimal order r .

PSO is a random search algorithm based on group cooperation developed by simulating the foraging behavior of birds [49-50]. It is often considered a form of swarm intelligence. The basic idea of PSO is to find the optimal solution through cooperation and information sharing among individuals in the group. The advantage of PSO is that it is easy to implement as it does not have many parameters to adjust. It has been widely used in function optimization, neural network training, fuzzy system control and other application fields. In this paper, PSO algorithm is used to determine the optimal order r .

The basic flow of PSO is as follows:

Assume that there are m particles in a D -dimensional target search space, and the position of each particle represents a potential solution. The position vector of the i th particle is $X_i = (x_i^1, x_i^2, \dots, x_i^D)$, the velocity vector of the i th particle is $V_i = (v_i^1, v_i^2, \dots, v_i^D)$. Its best position is recorded as p_{best} , the optimal position searched by the whole particle swarm, which is recorded as g_{best} . In each iteration, particle velocity is updated by p_{best} and g_{best} , and the formula to calculate the change of particle velocity is:

$$V_{i+1} = wv_i^d = c_1 r_1 (p_i^d - x_i^d) + c_2 r_2 (p_g^d - x_i^d) \quad (11)$$

where, V_{i+1} is the velocity of the updated particle, w is inertia vector, r_1 and r_2 are random numbers varying within $[0,1]$, c_1 and c_2 are accelerating constants (generally $c_1 = c_2 = 2$), v_i is limited by a maximum speed of $v_{\max}$.

In each iteration, the position of each particle is modified by velocity vector and position vector, and the formula to determine the particle position is

$$x_{i+1} = x_i + v_i \quad (12)$$

where, x_{i+1} is the updated particle position.

Iteration termination conditions are determined according to specific problems, generally selected as the optimal position searched by particle swarm to meet the preset minimum adaptive threshold or reach the maximum iteration number.

2.3.4 The PFSM(1,1) model

When there are obvious periodic fluctuations of the data, the NDGM(1,1) and NDGM^r(1,1) models discussed in the above sections are no longer applicable due to the inability to effectively identify and predict the periodic fluctuation of the time series. Therefore, in this paper we propose a new model, the PFSM(1,1) model, which is suitable for data showing periodic fluctuations. The basic idea of PFSM(1,1) model is as follows: The monthly data (which contains seasonal fluctuations) are summed to obtain the annual data for each year. Then the NDGM^r(1,1) model is used to predict the annual data, and PSO is used to calculate the optimal order. Finally, the seasonal factors are used to restore the prediction results of annual data to monthly data, which then contains the original data's seasonal trends. The modeling process of the PFSM(1,1) model is as follows:

The original time sequence with significant periodic (quarterly/monthly) changes is:

$$X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \dots, x^{(0)}(n)\} \quad (13)$$

Supposing $X^{(0)}$ has a period of L , the new sequence of annual data is:

$$Y^{(0)} = \{y^{(0)}(1), y^{(0)}(2), \dots, y^{(0)}(m)\}, m = \left\lfloor \frac{n}{L} \right\rfloor, \quad (14)$$

$$\text{where } y^{(0)}(1) = \sum_{i=1}^L x^{(0)}(i), y^{(0)}(2) = \sum_{i=L+1}^{2L} x^{(0)}(i), \dots, y^{(0)}(m) = \sum_{i=(m-1)L+1}^{mL} x^{(0)}(i).$$

Then using the **Eq.(6) - Eq.(10)** in Section 2.3.2 to calculate the following equation:

$$\hat{Y}^{(0)} = \{\hat{y}^{(0)}(1), \hat{y}^{(0)}(2), \dots, \hat{y}^{(0)}(m), \hat{y}^{(0)}(m+1), \dots\}. \quad (15)$$

where the optimal order is determined by PSO, as in Section 2.3.3.

The seasonal factor is a measure of seasonal change expressed in relative numbers, which

describes the typical seasonal characteristics of time series in different months or quarters.

The corresponding period average method [51] is used to calculate the seasonal factors $S_l (l=1,2,L, L)$:

$$S_l = \frac{\bar{x}_l^{(0)}(i)}{\bar{x}^{(0)}(i)} \quad (16)$$

where $\bar{x}_l^{(0)}(i)$ indicates the average of l -quarter (or month) and $\bar{x}^{(0)}$ represents the average of all quarters (or months).

The predicted value of the original time sequence is:

$$\hat{X}^{(0)} = \{ \hat{x}^{(0)}(1), \hat{x}^{(0)}(2), L, \hat{x}^{(0)}(n), L \} \quad (17)$$

Where

$$\hat{x}^{(0)}(i) = \hat{y}\left(\left[\frac{i-1}{L}\right] + 1\right) \cdot \frac{S_{i - \left[\frac{i}{L}\right] \cdot L}}{L}, i = 1, 2, L, n \quad (18)$$

A flowchart to illustrate the above modeling process is shown in Fig.5.

Fig. 5 The PFSM(1,1) modeling process

2.4 Metrics for evaluating accuracy of the models

The mean absolute percentage error (MAPE), root mean squares error (RMSE), mean absolute error (MAE) and Theil's U-statistic (U) which are used for evaluating the models' error are calculated as follows.

$$\text{MAPE} = \frac{1}{n} \sum_{k=1}^n \left| \frac{\hat{x}^{(0)}(k) - x^{(0)}(k)}{x^{(0)}(k)} \right| \times 100\% \quad (19)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{k=1}^n (\hat{x}^{(0)}(k) - x^{(0)}(k))^2} \quad (20)$$

$$\text{MAE} = \frac{1}{n} \sum_{k=1}^n |\hat{x}^{(0)}(k) - x^{(0)}(k)| \quad (21)$$

$$U = \sqrt{\frac{\sum_{k=1}^{n-1} (FRC_{k+1} - ARC_{k+1})^2}{\sum_{k=1}^{n-1} (ARC_{k+1})^2}} \quad (22)$$

In Eq.(19), Eq.(20) and Eq.(21), $\hat{x}^{(0)}(k)$ and $x^{(0)}(k)$ are the predicted values and the actual values respectively. In Eq.(22), $FRC_{k+1} = \frac{\hat{x}^{(0)}(k+1) - \hat{x}^{(0)}(k)}{\hat{x}^{(0)}(k)}$, the FRC_{k+1} is forecast relative change. $ARC_{k+1} = \frac{x^{(0)}(k+1) - x^{(0)}(k)}{x^{(0)}(k)}$, the ARC_{k+1} is the actual relative change. The smaller the calculation result of the above formulas, the higher the prediction accuracy of the model.

3. Prediction of monthly natural gas production in China

3.1 Model establishment

The monthly natural gas production data for the period 2013 - 2018 in China, drawn from the data sources shown in Fig.2, were used in the modeling to predict the monthly production data from 2019 to 2023. The aim of this section is to highlight the applicability and accuracy of the PFSM(1,1) model. To do so, the Holt-Winters model and the seasonal GM(1,1) model, which are both suitable for predicting seasonal data, were also used to predict the same data, and hence compare and analyze the accuracy of the different models. It should be noted that the prediction errors for each model are obtained by comparing the predicted values and the actual values over

the period 2013-2018. The forecast results for future monthly natural gas production (i.e., 2019-2023) are shown in section 4.

3.1.1 The PFSM(1,1) model

The modeling process of the PFSM(1,1) model is as follows.

Step 1. Firstly, monthly natural gas production data are processed according to **Eq.(14)**, and the annual natural gas production sequence $Y^{(0)}$ is obtained by summing.

$$Y^{(0)} = \{113.62, 124.13, 126.89, 136.64, 147.91, 158.38\}$$

Step 2. Using the NDGM^r(1,1) model to predict the $Y^{(0)}$, and using PSO to select the optimal order $r = 0.43$, was carried out in MATLAB2016b. Then, the parameters β_1, β_2 and β_3 are obtained by the least-squares method.

$$(\hat{\beta}_1, \hat{\beta}_2, \hat{\beta}_3) = (-0.0452, 44.7627, 133.3564)$$

Step 3. The corresponding time response function obtained from the above parameters is:

$$\begin{aligned} \hat{y}^{(0.43)}(k+1) &= \hat{\beta}_1^k \hat{x}^{(1)}(1) + \hat{\beta}_2 \sum_{j=1}^k j \hat{\beta}_1^{k-j} + \frac{1-\hat{\beta}_1^k}{1-\hat{\beta}_1} \hat{\beta}_3 \\ &= (-13.9693) \times (-0.0452)^k + 44.7627 \times \sum_{j=1}^k j (-0.0452)^{k-j} + 127.5893, k = 1, 2, 3, 4, 5. \end{aligned}$$

Step 4. By $\hat{Y}^{(0.43)}$, the prediction value $\hat{Y}^{(0)}$ can be obtained using 0.57-order IAGO. To demonstrate the accuracy of the NDGM^r(1,1) model, traditional NDGM(1,1) and GM(1,1) models with fractional-order accumulation (GM(1,1)^r) [52] were also used to predict $Y^{(0)}$. The fitted results of these different models are given in Table 6.

Table 6 The fitted results of different models (bcm natural gas/yr.)

Year	Actual value	GM ^{0.0004} (1,1)	NDGM(1,1)	NDGM ^{0.43} (1,1)
2013	113.62	113.62	113.62	113.62
2014	124.13	120.82	122.76	124.13
2015	126.89	128.80	128.95	126.76
2016	136.64	137.64	136.73	136.96
2017	147.91	147.42	146.53	147.71
2018	158.38	158.24	158.85	158.38
MAPE(%)		1.06	0.81	0.10

Table 6 shows that the MAPE of NDGM^{0.43}(1,1) model is much smaller than the other two models, indicating that its prediction performance is the best. Therefore, the prediction results of the NDGM^{0.43}(1,1) model are used to make subsequent calculations.

Step 5. The seasonal factors for monthly natural gas production were calculated by the corresponding period average method (**Eq.(16)**). The results are shown in Table 7.

Table 7 Seasonal factors for monthly natural gas production

Month	Seasonal factors	(Seasonal factors) / 12
1	1.0994	0.0916
2	0.9874	0.0823
3	1.0610	0.0884
4	0.9519	0.0793
5	0.9522	0.0793
6	0.9213	0.0768
7	0.9498	0.0792
8	0.9635	0.0803
9	0.9079	0.0757
10	0.9895	0.0825
11	1.0616	0.0885
12	1.1546	0.0962

Step 6. Finally, multiplying each item of $\hat{Y}^{(0)}$ by (seasonal factors)/12 gives the original sequence predicted value $\hat{X}^{(0)}$, where the fitted results of the PFSM(1,1) model are shown in Table 8.

3.1.2 The Seasonal GM(1,1) model

The SGM(1,1) model was established by defining a seasonal accumulating generation operator [39]. This is an effective prediction method for data having obvious seasonal changes.

Using the corresponding period average method to calculate the seasonal factors for monthly natural gas production (**Eq.(16)**), the results are also shown in Table 7. Then, the 1-SAGO is generated by seasonal factors and the SGM(1,1) model is established,

$$\text{where, the parameters are } \begin{bmatrix} \hat{a}_s \\ \hat{b}_s \end{bmatrix}^T = [-0.0054, 9.1550].$$

The corresponding time response function obtained from the above parameters is:

$$\hat{x}_s^{(1)}(k) = \left(\frac{x^{(0)}(1)}{f_s(1)} - \frac{\hat{b}_s}{\hat{a}_s} \right) e^{-\hat{a}_s(k-1)} + \frac{\hat{b}_s}{\hat{a}_s} = 1705.1681e^{0.0054(k-1)} + 1695.3704.$$

From $\hat{x}_s^{(1)}(k)$, the prediction value $\hat{x}_s^{(0)}(k)$ can be obtained through 1-order IAGO, and the results as shown in Table 8.

3.1.3 The Holt-Winters model

The basic idea of the HW model [53] is to decompose a time series having a linear trend, plus seasonal changes and random changes, and combine it with exponential smoothing to estimate the long-term trend, trend increment, and seasonal change of data respectively, and then establish a prediction model. The HW model can deal with the seasonal variations of data effectively and

filter out the effects of random fluctuations appropriately. Therefore, it is suitable for the prediction of time series with seasonal changes. The HW model's three basic formulas are:

$$St = \alpha \frac{x_t}{I_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}), 0 < \alpha < 1.$$

$$bt = \gamma(S_t - S_{t-1}) + (1-\gamma)b_{t-1}, 0 < \gamma < 1.$$

$$It = \beta \frac{X_t}{S_t} + (1-\beta)I_{t-L}, 0 < \beta < 1$$

where, L is the length of seasonality, I is the seasonal index, S is the stable composition, b is the linear component, α, β, γ are the smoothing parameters.

The HW model's prediction formula is $F_{t+m} = (s_t + b_t m)I_{t+m-L}$, where, m is the number of cycles.

The initial values are calculated as follows:

$$s_t = x_{L+1}, I_L = \frac{\overline{x_L}}{x}, b_{L+1} = \frac{(x_{L+1} - x_1) + (x_{L+2} - x_1) + (x_{L+3} - x_1)}{3 \times L}.$$

For the data prediction of the monthly natural gas production, use can be made of the grey wolf algorithm [54] to calculate the smoothing parameters of HW model, which are $\alpha = 0.9303, \beta = 0.01, \gamma = 0.01$ respectively. The results of HW model are shown in Table 8.

Table 8 The fitted values of the different models (bcm natural gas/month)

Time	Actual value	PFSM(1,1)	HW	SGM(1,1)	Time	Actual value	PFSM(1,1)	HW	SGM(1,1)
Jan-2013	11.20	10.41	11.20	10.57	Jan-2016	13.1	12.55	11.89	11.58
Feb-2013	9.02	9.35	9.02	9.50	Feb-2016	11.97	11.27	11.76	11.55
Mar-2013	9.93	10.05	9.93	9.86	Mar-2016	12.27	12.11	12.91	11.98
Apr-2013	9.21	9.01	9.21	8.89	Apr-2016	10.58	10.86	11.10	10.81
May-2013	8.88	9.02	8.88	8.88	May-2016	10.83	10.87	10.67	10.79
Jun-2013	8.85	8.72	8.85	8.70	Jun-2016	10.16	10.51	10.51	10.57
Jul-2013	8.98	8.99	8.98	9.02	Jul-2016	10.31	10.84	10.54	10.96
Aug-2013	8.90	9.12	8.90	9.20	Aug-2016	10.78	11.00	10.52	11.18
Sep-2013	8.50	8.60	8.50	8.71	Sep-2016	10.15	10.36	10.18	10.59
Oct-2013	9.23	9.37	9.23	9.63	Oct-2016	10.83	11.29	11.11	11.70
Nov-2013	10.01	10.05	10.01	10.38	Nov-2016	12.35	12.12	11.69	12.61
Dec-2013	10.91	10.93	10.91	11.30	Dec-2016	13.31	13.18	13.44	13.73
Jan-2014	10.99	11.37	10.99	10.17	Jan-2017	13.1	13.53	12.73	12.36
Feb-2014	10.40	10.21	10.91	10.14	Feb-2017	12	12.15	11.80	12.32
Mar-2014	10.95	10.97	11.27	10.52	Mar-2017	13.58	13.06	12.94	12.78
Apr-2014	9.78	9.85	9.90	9.49	Apr-2017	12.2	11.72	12.20	11.53
May-2014	9.89	9.85	9.84	9.47	May-2017	11.99	11.72	12.26	11.51
Jun-2014	9.33	9.53	9.61	9.28	Jun-2017	11.55	11.34	11.67	11.28

Jul-2014	9.83	9.82	9.69	9.62	Jul-2017	11.74	11.69	11.97	11.70
Aug-2014	9.89	9.97	10.01	9.82	Aug-2017	11.95	11.86	11.98	11.93
Sep-2014	9.54	9.39	9.37	9.30	Sep-2017	11.15	11.18	11.31	11.30
Oct-2014	10.31	10.24	10.44	10.27	Oct-2017	12.41	12.18	12.21	12.48
Nov-2014	11.07	10.98	11.12	11.08	Nov-2017	12.63	13.07	13.36	13.46
Dec-2014	12.15	11.94	12.10	12.05	Dec-2017	13.61	14.21	13.85	14.65
Jan-2015	12	11.61	11.61	10.85	Jan-2018	13.6	14.51	13.01	13.19
Feb-2015	10.47	10.43	10.82	10.82	Feb-2018	12.59	13.03	12.24	13.15
Mar-2015	11.15	11.21	11.33	11.23	Mar-2018	13.52	14.00	13.56	13.64
Apr-2015	9.4	10.05	10.06	10.13	Apr-2018	12.89	12.56	12.18	12.31
May-2015	9.87	10.06	9.49	10.11	May-2018	12.62	12.57	12.90	12.29
Jun-2015	9.93	9.73	9.57	9.91	Jun-2018	12.18	12.16	12.28	12.04
Jul-2015	10.1	10.03	10.26	10.27	Jul-2018	12.96	12.54	12.62	12.48
Aug-2015	10.42	10.18	10.30	10.47	Aug-2018	12.9	12.72	13.18	12.73
Sep-2015	9.58	9.59	9.86	9.92	Sep-2018	12.18	11.98	12.23	12.06
Oct-2015	10.39	10.45	10.51	10.96	Oct-2018	13.42	13.06	13.33	13.32
Nov-2015	11.11	11.21	11.20	11.82	Nov-2018	14.27	14.01	14.45	14.36
Dec-2015	12.47	12.20	12.14	12.86	Dec-2018	15.25	15.24	15.60	15.63

The degree of agreement between the different model fitted values and the actual monthly values is shown in Figs.6 and 7.

Fig. 6 The fitted values of different models for the period 2013-2015

Fig. 7 The fitted values of different models for the period 2016-2018

3.2 Comparison of prediction accuracy

To effectively compare the prediction accuracy of the three different models discussed above, the corresponding MAPE (Eq.(19)), RMSE (Eq.(20)), MAE (Eq.(21)) and U (Eq.(22)) values were calculated. The calculation results (i.e. the errors) are shown in Fig.8.

Fig. 8 The errors of the three models

The accuracy criteria of MAPE are shown as Table 9. This is the criterion to determine the acceptability of the prediction error. When MAPE value is less than 10%, the prediction accuracy of the model is considered as “excellent”, which is a strong argument for its subsequent use.

Table 9 Criteria of MAPE

MAPE (%)	Forecasting ability
< 10	Excellent
10–20	Good
20–50	Reasonable
> 50	Incorrect

The smaller the error, the higher the prediction accuracy. According to Fig.8, the values of MAPE, RMSE and MAE of the PFSM(1,1) model are the smallest among the three models, and MAPE value is 2.16%, lower than the target level of 10%. These results quantitatively support the proposition that the new model proposed here has a high prediction accuracy, and the ability to predict the data exhibiting significant seasonal variation.

4. Out of sample prediction results and discussion

4.1 The prediction results of natural gas monthly production during 2019-2023

The PFSM(1,1) model was finally used to predict monthly natural gas production in China for the period 2019 - 2023. The results are shown in Table 10. The variation trend of monthly and annual production are demonstrated in Fig.9 and 10.

Table 10 Predicted monthly natural gas production for the period 2019–2023 (bcm/month)

Months	Years				
	2019	2020	2021	2022	2023
Jan.	15.46	16.38	17.27	18.13	18.97
Feb.	13.89	14.71	15.52	16.29	17.04
Mar.	14.92	15.81	16.67	17.50	18.31
Apr.	13.39	14.19	14.95	15.70	16.42
May.	13.39	14.19	14.95	15.70	16.42
Jun.	12.96	13.73	14.48	15.20	15.90
Jul.	13.36	14.15	14.93	15.68	16.40
Aug.	13.55	14.36	15.14	15.90	16.63
Sep.	12.77	13.53	14.27	14.99	15.68
Oct.	13.92	14.75	15.56	16.33	17.08
Nov.	14.93	15.82	16.69	17.52	18.33
Dec.	16.24	17.21	18.14	19.04	19.92
Total	168.77	178.83	188.57	197.98	207.10

Fig.9 The actual and predicted month production of nature gas

Fig.10 The actual and predicted annual production of nature gas

As can be seen from Table 10, during the period 2019 - 2023, natural gas production in China for each month is forecast to increase year by year, and will be higher during the winter months and lower during the summer months. From Fig.9, the gap in natural gas production between the summer and winter months is expected to continue to grow; with the maximum peak-to-valley

difference expected to increase from 3.07 bcm during 2018 to 4.24 bcm during 2023. Therefore, the demands being imposed on China's capacity to supply natural gas in a stable manner are increasing.

From Fig.10, annual gas production will continue to grow in the next few years, but the rate of growth has slowed, and our predictions shows that the annual natural gas production in China will be about 180 bcm by 2020, which is significantly less than the target provided in the *13th Five Year Plan for Energy*, of 220 bcm by 2020, the gap will be 41.17 bcm. Therefore, under the current technological conditions and natural gas development level in China, we judge that it will be extremely challenging for China to achieve the goals of the *13th Five Year Plan for Energy*. Furthermore, the 'energy consumption structure' target in the *13th Five Year Plan for Energy* is to increase the share of natural gas consumption in the primary energy mix to 10% or greater by 2020. However, our predicted natural gas production of China in 2020 is much less than that required to meet this target. Imports will therefore need to increase in the future to meet the basic domestic demand for natural gas consumption.

4.2 Discussion of prediction results

4.2.1 Discussion on the rationality of the prediction results

(1) Quantitative analysis

In order to verify the rationality of the prediction results of PFSM(1,1) model, the projections of natural gas production in this paper were compared with those predicted by other research. However, as no comparable research has yet released data on the future monthly production of natural gas in China, only annual data, a direct comparison is not possible. For this reason, the forecast monthly production data in this paper are summarized into annual data, and these data used in the comparison. The data to make this comparison were selected from [4], [5], [6], *World Energy Outlook 2018* published by International Energy Agency (IEA-WEO-2018) and *BP Energy Outlook 2019 edition* published by British Petroleum (BP-EO-2019). These sources predict China's gas production over several decades, but here for a more useful comparison we compare predictions only up to 2030. The results of this comparison are shown in Fig.11.

Fig. 11 China's predicted domestic natural gas production

As can be seen from Fig.11, the predictions of natural gas production in China from 2019 to 2023 given in this paper are very close to those of other literature covering this period, indicating that the results of this paper are reasonable. Note that the other studies predicting natural gas production in China all focus on the long-term trend, while this paper focuses on the seasonal change in the medium and short-term to allow more accurate predictions over this time-frame.

(2) Qualitative analysis

The fluctuations of monthly production: as an efficient and clean energy, natural gas is favored by the government and the market. The project of changing fuel from coal to natural gas, as one of the important measures to improve air quality, has been widely promoted. In the future, this project will continue to be carried out on a large scale across the country, with the gradual implementation and advancement of various environmental protection policies. The demand for natural gas in winter will continue to rise, the seasonal difference of monthly production will also continue to expand, and the supply will fluctuate with the change of demand. Therefore, the prediction results of this paper show that the continuous expansion of peak-to-valley difference of natural gas production in the future is a normal development and change of natural gas supply market in the future, which conforms to the actual development situation.

The fluctuations of annual production: at present, natural gas is still in the early stage of exploration and development. In the next few years, with the increase of exploration and development efforts, natural gas production will increase significantly. However, due to the slow progress of upstream institutional reform and insufficient market competition, the production

potential and efficiency of natural gas are far from being released. In the medium and short term, the growth momentum is insufficient and the growth rate declines somewhat. In addition, the rich areas of oil and gas resources and important ecological function areas are objectively overlapped. The current laws and policies lack detailed regulations on hierarchical management and classification of production and construction activities in the important ecological function areas, so the oil and gas production and construction activities in the ecological function areas are limited and their output is also affected to some extent. Moreover, with the continuous progress of the project of changing fuel from coal to natural gas and other projects, the pressure of natural gas supply continues to increase. However, limited by the long construction period of underground gas storage and LNG storage tank, it is difficult to meet the demand of market regulation in the short and medium term, thus restricting the healthy development of natural gas market. Therefore, according to the prediction results of this paper, the annual production will continue to increase in the next few years, but the growth rate will decline, which is a normal change under the regulation of the policy and the market, in line with the actual development situation, is reasonable.

4.2.2 Discussion of nature gas import based on prediction results

(1) Analysis of natural gas import volume

If we know the gap between the supply and demand of natural gas in China during the period predicted in this paper, a better natural gas import strategy can be formulated. To do this we compare the projection of China's natural gas production in this paper with the corresponding consumption data predicted by other research. As with production forecasts, there is no research on monthly consumption forecasts, and therefore annual data are used for comparison. The data for comparison are mainly published by the mainstream institutions in the world. In addition to data from the IEA-WEO-2018 and BP-EO-2019 reports, we also use data from *World Oil Outlook 2018* published by Organization of the Petroleum Exporting Countries (OPEC-WOO-2018) and *Energy Outlook 2018* published by the China National Petroleum Corporation & Economic Technology Research Institute (CNPC-ETRI-EO-2018). The results of this comparison are given in Fig.12.

Fig. 12 Forecasts of China's domestic natural gas production and consumption from different sources

The average of forecasts from the four mainstream institutions is used as a benchmark for future gas consumption. The ratio between annual natural gas production and consumption from 2019 to 2023 is calculated and listed in Table 11.

Table 11 The ratio of production to consumption for natural gas

Year	Ratio
2019	60.67%
2020	58.84%
2021	58.23%
2022	57.36%
2023	56.29%

From Fig.12, China's natural gas consumption predicted by the above mainstream institutions are that demand will continue to grow until 2040 (and likely well beyond). By comparing this paper's production predictions for China from 2019 to 2023 with the consumption predicted by these mainstream institutions, there will still be a large gap between production and consumption, reaching some 160 bcm/yr. by 2023, that is, China will need to import some 160 bcm of natural gas per year to meet domestic demand. Taking global gas trade in 2018, of 943.4 bcm/yr. as a benchmark, China's imports will account for about 17% of global trade volume in 2023, up over 4% in absolute terms from 2018.

According to Table 11, natural gas production in China currently accounts for about 60% of the country's consumption, with this proportion forecast to decrease year by year, indicating that

China's dependence on external gas supply will continue to rise. As this increases, it seems likely that China will need to seek opportunities for energy cooperation in more countries, and hence greater economic benefits to a greater number of gas exporting countries, with a corresponding impact on the development of the world economy.

(2) Uncertainty analysis of natural gas import

From the above analysis, it can be seen that the gap between supply and demand of natural gas will continue to increase in the future, and imported natural gas will become an important source to maintain the normal operation of China's natural gas consumption market. However, affected by many factors, there are many uncertainties in natural gas import, which directly affects the sustainable and healthy development of China's natural gas market. Therefore, this section will analyze the uncertainty of natural gas import from the following aspects to provide a reference for making a more reasonable natural gas import strategy.

1) Natural gas production itself is constrained by infrastructure technology and process, so it is unlikely to produce large fluctuations in the short term, and there is less uncertainty, so it will show a trend of stable development in the short term. However, in terms of demand, due to its own uncertainty, it will indirectly lead to fluctuations in the gap between supply and demand, thus affecting China's import of natural gas. Therefore, the uncertainty of natural gas import is mainly caused by the uncertainty of demand. At present, the uncertainty of natural gas demand market mainly has two aspects: on the one hand is the implementation of environmental protection policies. China is in a critical period of air pollution control. Driven by environmental protection policies, the demand for natural gas continues to increase, but if policy drivers start to weaken in the future, gas demand growth will slow and imports will fall. On the other hand is the rise of renewable energy. Natural gas has long been considered as an important low-carbon transitional fossil energy. As the production of renewable energy increase, part of the demand for high-carbon fossil energy will be replaced by renewable energy, which will reduce the demand for natural gas and its import. Therefore, the formulation of China's import strategy should not blindly pursue increase, it should be reasonably adjusted according to the changes of policies and the market.

As China has become the world's largest natural gas buyer, how to ensure a stable supply of natural gas imported from overseas is a problem that needs to be paid attention to by all parties. One of the characteristics of international natural gas supply is unstable, the change of the international situation, the fluctuation of the international energy market will have a great impact on China's natural gas import. In addition, transnational natural gas import has never been a simple trade issue, and it is necessary to effectively connect the upstream natural gas reserves, intermediate transport pipelines and downstream consumer markets. Therefore, fluctuations in any

of the intermediate links will increase the uncertainty of China's natural gas import trade. It is based on this point that when developing the natural gas import strategy, we should diversify the import channels as much as possible, so as to reduce the impact of uncertainties in the import process.

5. Conclusions and outlook

5.1 Main conclusions

The two main conclusions of this paper are:

(1) This paper has proposed an innovative medium- and short-term forecasting model applicable to data having significant seasonal variation, which we term the PFSM(1,1) model. In this model, two approaches (i.e., the improvement of model self-adaptability and of seasonal adjustment of original data) are combined to provide an innovative PFSM(1,1) model for the prediction of seasonal data. We then compared this model using monthly natural gas production data for China with the HW and SGM(1,1) models, which are also suitable for predicting seasonal data. The results show that the predictive errors of the PFSM(1,1) model were substantially less than with the other two models.

(2) Prediction of natural gas production over the medium- and short-term, and including especially over monthly periods, has seen very little research, and constitutes something of a 'research gap'. In this paper we used the PFSM(1,1) model to predict monthly production of natural gas in China for the period 2019-2023. The results showed that China's production of natural gas will continue to increase in the short and medium term, and that the seasonal variations in this production will also continue to increase over this period. Furthermore, the results indicate that China is likely to produce about 180 bcm/yr. of natural gas by 2020, which is considerably less than the production target outlined in China's *13th Five Year Plan for Energy*, for this date of 220 bcm/yr. This difference results in a bigger gap than expected between supply and demand, and indeed our forecast is that by 2023, China will have to import some 160 bcm/yr. of natural gas to meet demand. It can be expected that this rapid growth in natural gas imports is likely to have a significant impact on the international natural gas market.

5.2 Limitations and future research directions

The category of predictive model used in this paper is univariate data-driven extrapolation prediction, which is based on historical and realistic trends of the system in a scientific manner. Therefore, the predicted results are the basic trend when the potential influencing factors in the future are not significantly different from the historical situation. It has been proved by a great deal of evidence that the development process of things, although it may sometimes involve certain leaps, is mainly gradual, so the predictions in this paper are a scientific and reasonable

judgement of future trends. However, if future external factors such as policies, market conditions and technologies change significantly, the actual situation may be greatly deviated from the forecast results. But the forecast results of this paper can still be used as the reference for the assumption that there is no change, so as to quantify the specific impact of the external factors change.

The forecasts in this paper are primarily from a data-driven perspective, and do not consider the possible influence of external factors. Future research could expand forecasts to include a range of external factor driven perspectives and consider the influence of these on the prediction results. In addition, while considering external factors, other prediction models, such as ARFIMA, ARIMA and VAR, can also be used to make corresponding predictions based on the same data of this paper, so as to reflect the accuracy and rationality of the prediction results in this paper.

This paper mainly focuses on the forecast of national natural gas production, however, due to the differences in resource distribution and economic development, energy consumption and energy supply in different regions of China can be quite different than for the country as a whole. Our future work will focus on the natural gas market in different regions within China, so as to better analyze these markets which face different economic and resource conditions. We can also compare the natural gas production of China with that of other parts of the world in order to better understand the development prospects and risks to China's natural gas supply market.

Acknowledgements

The relevant research projects carried out for this paper were supported by the National Natural Science Foundation of China (No.71874201; 71503264; 71871084), the Science Foundation of China University of Petroleum Beijing (No. ZX20200109), the Humanities and Social Sciences Youth Foundation of the Ministry of Education of China (Grant No. 19YJCZH106), the Excellent Young Scientist Foundation of Hebei Education Department (No. SLRC2019001), and the project of high-level talent in Hebei province. We also received helpful comments from Dr. Roger Bentley of the Petroleum Analysis Centre, Ireland, and from anonymous reviewers.

References

- [1] British Petroleum. BP Statistical Review of World Energy, <https://www.bp.com/>; 2019 [accessed 20 February 2020].
- [2] Wang WF, Liu P, Zheng L, Zhou XG, Jiang S. Natural Gas Reserves and Production Prediction of Ordos Basin. *Natural Gas Geoscience* 2014; 25(09):1483-1490.

- [3] Yu G, Fang YZ, Li HT, Chen YR. Application of prediction model in medium-and long-term development planning of natural gas production: A case study on conventional natural gas in the Sichuan Basin. *Natural Gas Exploration and Development* 2017; 40(03):90-95 .
- [4] Wang JZ, Jiang HY, Zhou QP, Wu J, Qin SS. Chinas natural gas production and consumption analysis based on the multicycle Hubbert model and rolling Grey model. *Renewable and Sustainable Energy Reviews* 2016;53:1149-1167.
- [5] Lu JL, Zhao SP, Sun YP, Tang HJ. Gas production peaks in China: Research and strategic proposals. *Natural Gas Industry B* 2018; 5(4):371-379.
- [6] Wang JL, Liu R. Prediction Study of the medium and long-term natural gas production in China. *Coal Economic Research*. 2019;39(10):41-47.
- [7] Wang JL, Steve M, Feng LY, Liu HH, Gail ET. Analysis of resource potential for China's unconventional gas and forecast for its long-term production growth. *Energy Policy* 2016; 88: 389-401.
- [8] Wang Q, Li SY, Li RR, Ma ML. Forecasting U.S. shale gas monthly production using a hybrid ARIMA and metabolic nonlinear grey model. *Energy* 2018; 160:378-387.
- [9] Wang Q, Jiang F. Integrating linear and nonlinear forecasting techniques based on grey theory and artificial intelligence to forecast shale gas monthly production in Pennsylvania and Texas of the United States. *Energy* 2019; 178:781-803
- [10] Wang, J.L., Li, N., 2020. Influencing factors and future trends of natural gas demand in the eastern, central and western areas of China based on the grey model. *Natural Gas Industry B*, in press, <https://doi.org/10.1016/j.ngib.2020.09.005>
- [11] Shuai XB. One Natural Gas Production Prediction Grey Model Based on Exponential Smoothing. *Computer Technology and Development* 2010;20(03):243-245+249.
- [12] Li HX, Wang HJ. China's Natural Gas Demand Forecast Based on the Improved Gray Model. *Henan Science* 2014; 32(05):872-876.
- [13] Chen WB. Establishment of a grey in-homogeneous gas demand forecasting model. *Oil & Gas Storage and Transportation* 2015; 34(03):288-292.
- [14] Liu XL, Blanca M, Ana SG. A grey neural network and input-output combined forecasting model. *Primary energy consumption forecasts in Spanish economic sectors*. *Energy* 2016; 115:1042-1054.
- [15] Zhang T, Yang JM. Trend prediction of natural gas consumption in different regions of China. *Natural Gas Industry* 2016; 36(09):135-140.
- [16] Guo JY. Research on Energy Consumption Prediction Based on Grey Combination Model. *Journal of Chongqing University of Technology(Natural Science)* 2017;31(06):95-98.
- [17] Faheemullah S, Qiang J, Pervez HS, Nayyar HM, Muhammad AU. Forecasting China's natural gas demand based on optimised nonlinear grey models. *Energy* 2017;140(1):941-951.
- [18] Ma X, Liu ZB. Application of a novel time-delayed polynomial grey model to predict the natural gas consumption in China. *Journal of Computational and Applied Mathematics* 2017; 324:17-24.
- [19] Wu YH, Shen H. Grey-related least squares support vector machine optimization model and its application in predicting natural gas consumption demand. *Journal of Computational and Applied Mathematics* 2018; 338:212-220.
- [20] Ye ZH, Chen EL. Study on the Prediction of Natural Gas Consumption Based on Residual Grey Model. *Journal of Chongqing University of Technology (Natural Science)* 2018;32(09):99-102.
- [21] Zeng B, Duan HM, Bai Y, Meng W. Forecasting the output of shale gas in China using an unbiased grey model and weakening buffer operator. *Energy* 2018; 151:238-249.

- [22] Song D. A novel self-adapting intelligent grey model for forecasting China's natural-gas demand. *Energy* 2018; 162:393-407.
- [23] Wang Q, Li SY, Li RR, Ma ML. Forecasting U.S. shale gas monthly production using a hybrid ARIMA and metabolic nonlinear grey model. *Energy* 2018; 160:378-387.
- [24] Ma X, Mei X, Wu WQ, Wu XX, Zeng B. A novel fractional time delayed grey model with Grey Wolf Optimizer and its applications in forecasting the natural gas and coal consumption in Chongqing China. *Energy* 2019; 178:487-507.
- [25] Ma X, Wu WQ, Zeng B, Wang Y, Wu XX. The conformable fractional grey system model. *ISA Transactions* 2020; 96:255-271.
- [26] Hu Y, Su LL, Huang ZM, Chen XZ, Tang GX. Forecast Analysis of Natural Gas Demand in Sichuan Province Based on Fractional Order Grey Neural Network Combination Model. *Computer Applications of Petroleum* 2018;26(Z1):60-64+2.
- [27] Wang Q, Jiang F. Integrating linear and nonlinear forecasting techniques based on grey theory and artificial intelligence to forecast shale gas monthly production in Pennsylvania and Texas of the United States. *Energy* 2019; 178:781-803
- [28] Wang ZX, Dang YG, Pei LL. Modeling approach for oscillatory sequences based on GM(1, 1) power model. *Systems Engineering and Electronics* 2011;33(11):2440-2444.
- [29] Wang ZX. Oscillating GM(1,1) power model and its application. *Control and Decision* 2013;28(10):1459-1464+1472.
- [30] Wang ZX. Grey forecasting method for small sample oscillating sequences based on Fourier series .*Control and Decision* 2014;29(02):270-274.
- [31] Wang JF, Luo D. Fractional order discrete grey GM(1, 1) power model based on oscillation sequences and its application. *Control and Decision* 2017;32(1):176-180.
- [32] Zeng L. Grey GM(1,1|sin) power model based on oscillation sequences and its application. *Journal of Zhejiang University(Science Edition)* 2019;46(06):697-704.
- [33] Yang JX, Cheng SF. Application of grey - cycle epitaxial model in coal demand forecasting. *Statistics & Decision* 2010; 13:162-163.
- [34] Wu Y, Wang J. Integrated Optimum Gray Support Vector Machine Model for the Seasonal Power Load Forecasting. *East China Electric Power* 2012;40(01):18-21.
- [35] Zhang LP, Wang L, Zheng YL, Wang K, Zhang XL, Zheng YJ. Time Prediction Models for Echinococcosis Based on Gray System Theory and Epidemic Dynamics. *International Journal of Environmental Research and Public Health* 2017;14:262.
- [36] Guo R. Application of grey seasonal variation index GSVI (1,1) model in urban water demand prediction. *Water Resources Planning and Design* 2018;01:55-57+162.
- [37] Xiao XP, Yang JW, Mao SH. An improved seasonal rolling grey forecasting model using a cycle truncation accumulated generating operation for traffic. *Applied Mathematical Modelling* 2013; 33:1324-1331.
- [38] Wang ZX, Li Q, Pei LL. Grey forecasting method of quarterly hydropower production in China based on a data grouping approach. *Applied Mathematical Modelling* 2017; 51:302-316.
- [39] Wang ZX, Li Q, Pei LL. A seasonal GM(1,1) model for forecasting the electricity consumption of the primary economic sectors. *Energy* 2018; 154(1):522-534.
- [40] Qian WY, Dang YG. GM(1,1) model based on oscillation Sequences. *Systems Engineering-Theory& Practice* 2009; 29(3):149-154.
- [41] Fu HY, Gao HQ. Application of GM(1,1) model with modified background value in oscillation sequence. *Statistics & Decision* 2010; 15:28-30.

- [42] Wang JZ, Ma XL, Wu J, Dong Y. Optimization models based on GM(1, 1) and seasonal fluctuation for electricity demand forecasting, *Electrical Power and Energy Systems* 2012; 43:109-117.
- [43] Liu CY, Yao J. Real estate price forecasting model based on multiple. *Statistics & Decision* 2017; (17):33-38.
- [44] Ma DJ. Application research of grey multivariable model MGM(1, n) in characteristic parameter analysis on electromagnetic relay. *Electronic Components and Materials* 2014;33(12):78-81.
- [45] Ding S, Dang YG, Xu N, Wang JJ. Multivariable grey forecasting model based on interaction effect and its application. *Systems Engineering and Electronics* 2018; 40(03):595-602.
- [46] Xie NM, Liu SF, Yang YJ, Yuan CQ. On novel grey forecasting model based on non-homogeneous index sequence. *Applied Mathematical Modelling* 2013; 37:5059-5068.
- [47] Wu LF, Liu SF, Yao LG, Yan SL, Liu DL. Grey system model with the fractional order accumulation. *Communications in Nonlinear Science and Numerical Simulation* 2013; 18(7):1775-1785.
- [48] Wu LF, Liu SF, Cui W, Liu DL, Yao TX. Non-homogenous discrete grey model with fractional-order accumulation. *Neural Computing & Applications* 2014; 25:1215-1221.
- [49] Kennedy J, Eberhart RC. Particle swarm optimization. *Proceedings of IEEE International Conference on Neural Networks*. Perth: IEEE Press; 1995.p.1942-1948.
- [50] Eberhart RC, Kennedy J. A new optimizer using particle swarm theory. *Proceedings of the Sixth International Symposium on Micromachine and Human Science*. Nagoya: IEEE Press; 1995. p. 39-43.
- [51] Xiang MR, Wang QH, Ma D. *Statistical*. Beijing: Machinery Industry Press; 2014.
- [52] Wu LF, Liu SF, Yao LG, Yan SL, Liu DL. Grey system model with the fractional order accumulation. *Communications in Nonlinear Science and Numerical Simulation* 2013;18(7):1775-1785.
- [53] Gaynor PE, Kirkpatrick RC. *Introduction to time-series modeling and forecasting in business and economics*. New York: McGraw-Hill; 1994.
- [54] Seyedali M, Seyed MM, Andrew L. Grey wolf optimizer. *Advances in Engineering Software* 2014; 69:46-61.

Highlights

- A novel grey seasonal model with particle swarm optimization (PFSM(1,1)) is proposed.
- The PFSM(1,1) model has the lowest prediction errors compared to other models.
- The proposed model is used to predict China's monthly natural gas production.
- The predict results show that China's gas production will continue to increase.
- The gas production can't meet the demand and the gap is forecasted to increase.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐ The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: