Predicting Natural Gas Production in Various Nations Using a Fractional Grey Bernoulli Approach

Abstract

Accurate forecasting of natural gas production is crucial for economic stability, environmental sustainability, and market investment. This study presents an advanced forecasting method using the fractional grey Bernoulli model, which combines fractional accumulation and Bernoulli processes to enhance the predictive accuracy for nonlinear datasets. The model's versatility and flexibility allow it to adapt to various data characteristics and complexities, thereby outperforming traditional grey models in forecasting performance. To optimize the model parameters, this study employs the Particle Swarm Optimization (PSO) algorithm, further improving the model's effectiveness. Empirical analysis of natural gas production data from Brazil, Italy, and Qatar demonstrates that the model exhibits significant advantages in both fitting and forecasting capabilities. The findings indicate that the fractional grey Bernoulli model achieves high accuracy and reliability in predicting natural gas production in these countries, providing a robust framework for strategic energy planning and investment decision-making. With average forecast errors of 1.9113%, 4.0353%, and 1.8902% for natural gas production in Brazil, Italy, and Qatar respectively, this study underscores the model's effectiveness in enhancing forecast reliability and minimizing risk, providing valuable insights for sustainable energy development.

keywords: Fractional order, Grey Bernoulli model, Particle Swarm Optimization, Natural gas production

1 Introduction

Accurate predictions of natural gas production are crucial for economic stability, environmental sustainability, and informed investment decisions on a global scale [1, 2]. As a vital energy source, natural gas significantly influences the energy landscape of countries like Brazil, Italy, and Qatar. These forecasts are essential not only for optimizing extraction strategies but also for aiding policymakers and businesses in addressing the environmental challenges linked to energy production, including carbon emissions. Understanding the dynamics of natural gas production is key to enhancing energy policies and investment strategies in these regions, ultimately contributing to a more sustainable and resilient energy future.

Grey system theory was first proposed by Deng in 1982 [3], with the grey model GM being the most important and fundamental model. It laid the foundation for the development of grey system theory and played a crucial role. Following the development of NGM(1,1) [4], DGM [5] and other univariate models were subsequently proposed. Moreover, beyond univariate models, grey system modeling has gradually expanded to include research on multivariate grey models. Based on the GM(1,1) model, Deng [6] proposed the GM(1,n) model in 1984, which is more general and widely applicable than the GM(1,1) model. However, the GM(1,n) model has some theoretical issues that were pointed out by Tien [7] in 2012. Tien noted that using the trapezoidal rule to discretize the convolution integral of the analytical solution to the whitening equation could lead to better prediction results. In 2005, Tien [8] proposed the GMC(1,n) multivariable grey model. Later,the models such as NGMC(1,n) [9], DGMC(1,n) [10] were proposed soon. In 2008 and 2010, Chen [11, 12] et al. proposed a nonlinear Bernoulli equation based on the GM(1,1) model, which they referred to as the nonlinear grey Bernoulli model(BernoulliGM), abbreviated as the NBGM model. The above models are all traditional first-order(1-AGO) accumulated grey models. They have played a crucial role in the subsequent expansion, improvement, and optimization of grey models.

The grey model, as an important tool for handling uncertainty and small sample data, has demonstrated its unique advantages in various fields [13]. Existing research indicates that traditional large-sample statistical models perform poorly when data is insufficient, failing to provide accurate predictions and analyses [14]. In contrast, grey models are particularly suited for small sample data because they can achieve high prediction accuracy even with limited sample data through inherent mathematical processing and model optimization [15]. This capability makes grey models especially important in practical applications, particularly in fields such as energy [16–18], where data collection is challenging or sample sizes are small. Further research and optimization of grey models' small sample prediction capabilities can enhance the accuracy of future trend predictions and the scientific basis for decision-making [19]. In practical applications, due to the difficulty in obtaining large, high-quality data samples, small sample prediction becomes particularly important. Especially in the energy sector, where data collection is often limited, grey models are widely used for energy forecasting. These models not only assist relevant departments in formulating accurate energy policies but also effectively assess energy demand, contributing to the stability and sustainable development of the energy market.

Various grey prediction models, such as the grey model(GM(1,1)) [6], have been widely used due to their simplicity and applicability to small samples. However, these models often struggle with complex, nonlinear datasets. Recent advancements have led to the development of more sophisticated grey models, including the fractional grey Bernoulli model(FBernoulliGM) [20], nonlinear grey Bernoulli model(BernoulliGM) [12, 21], hybrid acculation grey model(HAGM) [22], new information priority nonlinear grey Bernoulli model(NIPBernoulliGM) [23], fractional-order nonhomogeneous discrete grey model(FNDGM) [24], new information priority grey model(NIPGM) [25] ,and fractional grey model(FGM) [26] to improve prediction accuracy. The fractional grey Bernoulli model(FBernoulliGM) stands out among these models, demonstrating superior performance in handling nonlinear relationships and providing more reliable forecasts.

The fractional grey Bernoulli model(FBernoulliGM) [20] emerges as a powerful tool for this purpose, integrating the concepts of fractional accumulation and Bernoulli processes to enhance predictive accuracy in nonlinear scenarios [27, 28]. Unlike traditional grey models, the fractional grey Bernoulli model leverages fractional orders to better capture the intricacies of real-world data, accommodating the variability and uncertainty inherent in energy systems. Its adaptability allows it to effectively model both short-term fluctuations and long-term trends in natural gas production.

Studies have shown that the fractional grey Bernoulli model can significantly enhance forecasting accuracy across diverse contexts. By employing this model, we can obtain more precise estimates of natural gas production, which are crucial for strategic planning and policy formulation. The flexibility of the The fractional grey Bernoulli model not only improves prediction reliability but also facilitates a deeper understanding of energy dynamics in the studied regions.

This paper focuses on utilizing the fractional grey Bernoulli model to analyze and forecast natural gas production in Brazil, Italy and Qatar. The research aims to validate the model's effectiveness through empirical case studies, revealing its advantages over traditional forecasting methods. The subsequent sections will elaborate on the model's structure, present empirical findings, and draw conclusions regarding its implications for energy forecasting and policy.

The remainder of this article is structured as follows: Section 2 introduces the prediction model and the fractional order optimization method. Section 3 presents three prediction cases, while Section 4 concludes the discussion.

2 RESEARCH METHODOLOGY

2.1 The fractional accumulation Bernoulli grey model

From Ref. [26], consider a nonnegative original series $X^{(0)} = \left\{X^{(0)}(1), X^{(0)}(2), ..., X^{(0)}(t)\right\}$. The r-order fractional accumulation (denoted as r-FOA) of this series can be expressed as follows:

$$X^{(r)}(m) = \sum_{j=1}^{m} \binom{m-j+r-1}{m-j} X^{(0)}(j), \quad m = 1, 2, \ldots, t, \tag{1}$$

where r represents the fractional order, and the binomial coefficient $\binom{m-j+r-1}{m-j}$ is defined as:

$$\binom{m-j+r-1}{m-j} = \frac{(m-j+r-1)(m-j+r-2)\cdots(r+1)r}{(m-j)!}. \tag{2}$$

Similarly, the inverse fractional accumulation of order r (referred to as r-IFOA) is defined by the following expression:

$$X^{(-r)}(m) = \sum_{j=1}^{m} \binom{m-j-r-1}{m-j} X^{(0)}(j), \quad m = 1, 2, \ldots, t. \tag{3}$$

By utilizing these definitions, the relationship between r-FOA and r-IFOA can be established as:

$$(X^{(r)}(m))^{(-r)} = \sum_{j=1}^{m} \binom{m-j-r-1}{m-j} X^{(r)}(j) = X^{(0)}(m), \quad m = 1, 2, \ldots, t. \tag{4}$$

Eq.(4) highlights the intrinsic connection between the accumulated series $X^{(r)}(m)$, the inverse accumulation $X^{(-r)}(m)$, and the original series $X^{(0)}(m)$. This relationship is crucial for recovering the original series from predicted values $\hat{X}^{(r)}(m)$ obtained through the fractional grey model, a process that will be further explained in subsequent sections.

From Ref. [20], the fractional accumulation nonlinear Bernoulli grey model can be formulated as the following nonlinear differential equation:

$$\frac{\mathrm{d}X^{(r)}(t)}{\mathrm{d}t} + aX^{(r)}(t) = b(X^{(r)}(t))^{\gamma}, \tag{5}$$

where γ represents a nonlinear parameter. When $\gamma = 0$, this model simplifies to the traditional Bernoulli grey model (BernoulliGM). If $\gamma \neq 0$ and $\gamma \neq 1$, the equation can be linearized as:

$$\frac{\mathrm{d}Y^{(r)}(t)}{\mathrm{d}t} + (1-\gamma)aY^{(r)}(t) = (1-\gamma)b, \tag{6}$$

where the transformation $Y^{(r)}(t) = (X^{(r)}(t))^{1-\gamma}$ is applied. This linear Eq.(6) is known as the whitening equation of the fractional Bernoulli grey model (FBernoulliGM). The corresponding discrete form of this equation is:

$$Y^{(r)}(m) - Y^{(r)}(m-1) + (1-\gamma)aZ^{(r)}(m) = (1-\gamma)b, \quad m = 2, 3, \ldots, t, \tag{7}$$

where the background value $Z^{(r)}(m)$ is given by:

$$Z^{(r)}(m) = \frac{1}{2}\left[Y^{(r)}(m) + Y^{(r)}(m-1)\right]. \tag{8}$$

Given the fractional order γ. To estimate the parameters a and b in Eq.(6), the least squares method can be used. The estimation equation is:

$$[(1-\gamma)a, (1-\gamma)b]^T = (B^T B)^{-1} B^T Y, \tag{9}$$

where the matrices B and Y are constructed as follows:

$$B = \begin{bmatrix} -Z^{(r)}(2) & 1 \\ -Z^{(r)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(r)}(t) & 1 \end{bmatrix}_{(t-1)\times 2}, \quad Y = \begin{bmatrix} Y^{(r)}(2) - Y^{(r)}(1) \\ Y^{(r)}(3) - Y^{(r)}(2) \\ \vdots \\ Y^{(r)}(t) - Y^{(r)}(t-1) \end{bmatrix}_{(t-1)\times 1}. \tag{10}$$

Solving the differential Eq.(6) and substituting the values of $(1-\gamma)a$ and $(1-\gamma)b$, we derive the solution for $\hat{Y}^{(r)}(m)$ as:

$$\hat{Y}^{(r)}(m) = \left(Y^{(r)}(1) - \frac{b}{a}\right) e^{-a(1-\gamma)(m-1)} + \frac{b}{a}, \quad m = 1, 2, \ldots, t, \tag{11}$$

where $Y^{(r)}(1) = Y^{(0)}(1)$. Finally, using equation (4) and the transformation in equation (6), the predicted values for the original series can be expressed as:

$$\hat{X}^{(0)}(m) = \left((\hat{Y}^{(r)}(m))^{\frac{1}{1-\gamma}}\right)^{(-r)}, \quad m = 1, 2, \ldots, t. \tag{12}$$

2.2 Optimization algorithm solves nonlinear parameters

Based on the previous content, we can observe that the nonlinear parameter significantly affects the accuracy of the model's predictions to a considerable extent. This section provides a detailed discussion on optimizing the algorithm for solving the optimal nonlinear parameters.

To solve this optimization problem, we reformulate it as an error-minimization task by tuning the parameters within the constraints set by the proposed model's structure. Among various error metrics, we choose the Mean Absolute Percentage Error (MAPE) as the standard for assessing the model' s prediction accuracy. MAPE provides a straightforward representation of relative error, known for its robustness and interpretability, making it suitable for different types of datasets.

The Mean Absolute Percentage Error (MAPE) for model fitting, validation, and forecasting is defined as follows:

$$MAPE_{fit} = \frac{1}{n} \sum_{m=1}^{n} \left| \frac{\hat{Y}^{(r)}(m) - Y^{(r)}(m)}{Y^{(r)}(m)} \right| \times 100\% \tag{13}$$

$$MAPE_{valid} = \frac{1}{s} \sum_{m=n+1}^{n+s} \left| \frac{\hat{Y}^{(r)}(m) - Y^{(r)}(m)}{Y^{(r)}(m)} \right| \times 100\% \tag{14}$$

$$MAPE_{pred} = \frac{1}{p} \sum_{m=n+s+1}^{n+s+p} \left| \frac{\hat{Y}^{(r)}(m) - Y^{(r)}(m)}{Y^{(r)}(m)} \right| \times 100\% \tag{15}$$

where n is the number of fitting points, s refers to the number of valid points, and p refers to the number of predicted points.

Minimizing MAPE allows us to determine the optimal value of r, improving the accuracy of the model' s predictions. The optimization problem can be mathematically expressed as:

$$\min MAPE_{valid} = \frac{1}{s} \sum_{m=n+1}^{n+s} \left| \frac{\hat{Y}^{(r)}(m) - Y^{(r)}(m)}{Y^{(r)}(m)} \right| \times 100\% \tag{16}$$

$$s.t. \begin{cases} [(1-\gamma)a, (1-\gamma)b]^T = (B^T B)^{-1} B^T Y \\ B = \begin{bmatrix} -Z^{(r)}(2) & 1 \\ -Z^{(r)}(3) & 1 \\ \vdots & \vdots \\ -Z^{(r)}(t) & 1 \end{bmatrix}_{(t-1)\times 2} \\ Y = \begin{bmatrix} Y^{(r)}(2) - Y^{(r)}(1) \\ Y^{(r)}(3) - Y^{(r)}(2) \\ \vdots \\ Y^{(r)}(t) - Y^{(r)}(t-1) \end{bmatrix}_{(t-1)\times 1} \\ Z^{(r)}(m) = \frac{1}{2}\left[Y^{(r)}(m) + Y^{(r)}(m-1)\right] \\ \hat{Y}^{(r)}(m) = \left(Y^{(r)}(1) - \frac{b}{a}\right) e^{-a(1-\gamma)(m-1)} + \frac{b}{a}, \quad m = 1, 2, \ldots, t \\ \hat{X}^{(0)}(m) = \left((\hat{Y}^{(r)}(m))^{\frac{1}{1-\gamma}}\right)^{(-r)}, \quad m = 1, 2, \ldots, t \end{cases}$$

The optimization problem above is solved using the Particle Swarm Optimization (PSO) algorithm.

Particle Swarm Optimization (PSO) [29] [30] is a population-based optimization technique inspired by the coordinated behaviors seen in bird flocking and fish schooling. Each individual in the population, referred to as a particle, represents a candidate solution. By iteratively adjusting their positions, particles explore the search space, leveraging both their personal experiences and information shared within the swarm to approach the optimal or near-optimal solution.

To address the optimization challenge, we employ PSO, which utilizes both individual historical information and collective knowledge to iteratively improve solutions. Each particle adjusts its position in the search space based on its velocity, its own best-known position, and the best-known position found by the swarm as a whole.

The equations for updating each particle’s velocity and position are given as:

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot (p_i - x_i(t)) + c_2 \cdot r_2 \cdot (g - x_i(t)), \tag{17}$$

$$x_i(t+1) = x_i(t) + v_i(t+1), \tag{18}$$

where $v_i(t)$ and $x_i(t)$ denote the velocity and position of the i-th particle at iteration t, w is the inertia weight, and c_1 and c_2 are cognitive and social acceleration coefficients. r_1 and r_2 are random variables uniformly drawn from $[0, 1]$, p_i represents the particle’s personal best position, and g denotes the global best position discovered by the swarm.

This algorithm iteratively reduces the objective function, guiding particles toward optimal solutions. Particle Swarm Optimization (PSO) is chosen for its simplicity, adaptability, and strong global search performance. In this study, Particle Swarm Optimization (PSO) is used to tune model parameters, aiming to minimize the Mean Absolute Percentage Error (MAPE) during model validation.

To address the optimization problem mentioned above, we developed a solution process based on the Particle Swarm Optimization (PSO) algorithm. The procedure is outlined in Algorithm 1. This algorithm is implemented using Python.

Algorithm 1: PSO-based optimization of γ and least squares for a and b

input : The original series $X^{(0)} = \{X^{(0)}(1), X^{(0)}(2), \ldots, X^{(0)}(t)\}$, max iterations, population size

output: The optimal values of (a, b, γ)

Initialize: PSO parameters (population size, inertia weight, cognitive and social coefficients), swarm particles with random initial positions and velocities for γ

Set $(MAPE_{\min}) = \infty$

for each iteration up to max iterations **do**

  for each particle in the swarm **do**

    Step 1: Compute accumulative sequence $Z^{(r)}(m)$ using:

$$Z^{(r)}(m) = \frac{1}{2}[Y^{(r)}(m) + Y^{(r)}(m-1)]$$

    Step 2: Construct matrices $\mathbf{B}$ and $\mathbf{Y}$:

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

    Step 3: Solve for $(1-\gamma)a, (1-\gamma)b$ and compute a and b:

$$[(1-\gamma)a, (1-\gamma)b]^T = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T\mathbf{Y}$$

    Step 4: Predict $Y^{(r)}(m)$:

$$\hat{Y}^{(r)}(m) = \left(Y^{(r)}(1) - \frac{b}{a}\right) e^{-a(1-\gamma)(m-1)} + \frac{b}{a}$$

    Step 5: Compute $\hat{X}^{(1)}(m)$ and $\hat{X}^{(0)}(m)$ using:

$$\hat{X}^{(1)}(m) = \hat{Y}^{(r)}(m)^{\frac{1}{1-\gamma}}, \quad \hat{X}^{(0)}(m) = (\hat{X}^{(1)}(m))^{(-r)}$$

    Step 6: Evaluate the particle fitness using MAPE:

$$MAPE_{\text{valid}} = \frac{1}{n}\sum_{m=1}^{n}\left|\frac{X^{(0)}(m) - \hat{X}^{(0)}(m)}{X^{(0)}(m)}\right|$$

  end

  if $\text{MAPE}_{\text{valid}} < MAPE_{\min}$ **then**

    Update $MAPE_{\min} \leftarrow MAPE_{\text{valid}}$

    Update the best values of γ^*

  end

  `Update` particle velocities and positions using PSO update rules

end

3 Applications in forecasting natural gas production

Forecasting natural gas production trends in regions such as Qatar, Brazil, and Italy is essential for ensuring stability in the global energy landscape and promoting economic growth. These countries are significant players in the international natural gas market, influencing supply dynamics and global pricing mechanisms. Accurate predictions of their production levels allow policymakers and businesses to formulate effective energy strategies, ensuring market stability while reducing risks and supporting economic development. Furthermore, these insights contribute to energy security and guide investment decisions, ultimately fostering the sustainable growth of the global energy sector.

We meticulously collected annual natural gas production data from Qatar, Brazil, and Italy for the years 2008 to 2016. This data is sourced from reference [31], which clearly indicates that all figures are extracted from the authoritative BP Statistical Review of World Energy.

In our study, we utilized data from 2008 to 2012 for model fitting to ensure the accuracy and reliability of the constructed model and data from 2013 to 2014 for model validation. Subsequently, we selected data from 2015 to 2016 for prediction to applicate the model's effectiveness and practicality. Through this process, we aim to not only reveal the production trends of these three countries in natural gas but also provide robust data support and theoretical foundations for future energy policy formulation, thereby promoting the sustainable development of the global energy market.

3.1 Case 1: forecasting natural gas production in Brazil

The study of natural gas production in Brazil is essential for understanding the country's growing role in the global energy market. As Brazil emerges as a key player in South America, its natural gas resources are becoming increasingly significant for the region's energy security and economic development. By examining production trends in Brazil, we can assess the country's potential to reduce its reliance on imported energy, boost domestic industries, and contribute to a more stable and diversified regional energy supply. Moreover, Brazil's advancements in natural gas production could have a positive impact on its environmental policies, promoting a cleaner energy transition within the region.

In Table 1, the prediction results and errors of several models (FBernoulliGM, BernoulliGM, HAGM, NIPBernoulliGM, NIPGM, FNDGM, GM, and FGM) are compared against the original data from 2008 to 2016. Each model’s performance is measured by MAPE values during fitting (MAPE_{fit}), validation (MAPE_{valid}), and prediction (MAPE_{pred}).

Notably, the NIPGM model achieves the lowest fitting error with a MAPE_{fit} of 0.5771%, while HAGM performs the best overall in fitting with the second-lowest MAPE_{fit} of 0.595%. FBernoulliGM, however, stands out in the validation phase, achieving the best MAPE_{valid} of 0.1408%, indicating its superior capability in predicting validation data. In terms of prediction, FBernoulliGM also excels, showing the lowest $_{of}$1.9113%, outperforming other models such as NIPGM, which has a $_{of}$17.3919%. The GM model shows the worst prediction results with a high $_{of}$43.156%. Overall, FBernoulliGM demonstrates strong prediction capabilities, while the NIPGM and HAGM models excel in fitting performance.

As shown in Fig.1, the comparison of the predicted data for different models is thoroughly illustrated. The parameters were fine-tuned through the PSO (Particle Swarm Optimization) algorithm, with 1,000 iterations. The convergence curve of MAPE$_{valid}$ is displayed in Fig.2.

Table 1. Detailed results in Case 1

Year	Original Data	FBernoulliGM	BernoulliGM	HAGM	NIPBernoulliGM	NIPGM	FNDGM	GM	FGM
2008	14.00	14.00	14.00	14.00	14.00	14.00	14.00	14.00	14.00
2009	11.90	12.09	11.65	11.79	11.78	11.96	11.92	12.18	12.21
2010	14.60	14.23	14.43	14.60	14.46	14.43	14.43	14.22	14.23
2011	16.70	16.80	16.89	16.93	16.94	16.86	16.96	16.59	16.64
2012	19.30	19.23	19.07	19.17	19.15	19.25	19.19	19.37	19.35
2013	21.30	21.24	21.00	21.29	21.07	21.62	21.12	22.61	22.37
2014	22.70	22.70	22.70	23.30	22.70	23.94	22.81	26.39	25.75
2015	23.10	23.58	24.20	25.22	24.06	26.23	24.32	30.80	29.54
2016	23.50	23.91	25.53	27.04	25.18	28.49	25.69	35.95	33.81
MAPE$_{fit}$		1.0185	1.1189	0.595	0.8363	0.5771	**0.6919**	1.1954	1.1515
MAPE$_{valid}$		**0.1408**	0.7042	1.3451	0.5399	3.4825	0.6648	11.2029	9.2298
MAPE$_{pred}$		**1.9113**	6.7001	12.1207	5.6524	17.3919	7.3003	43.156	35.8756

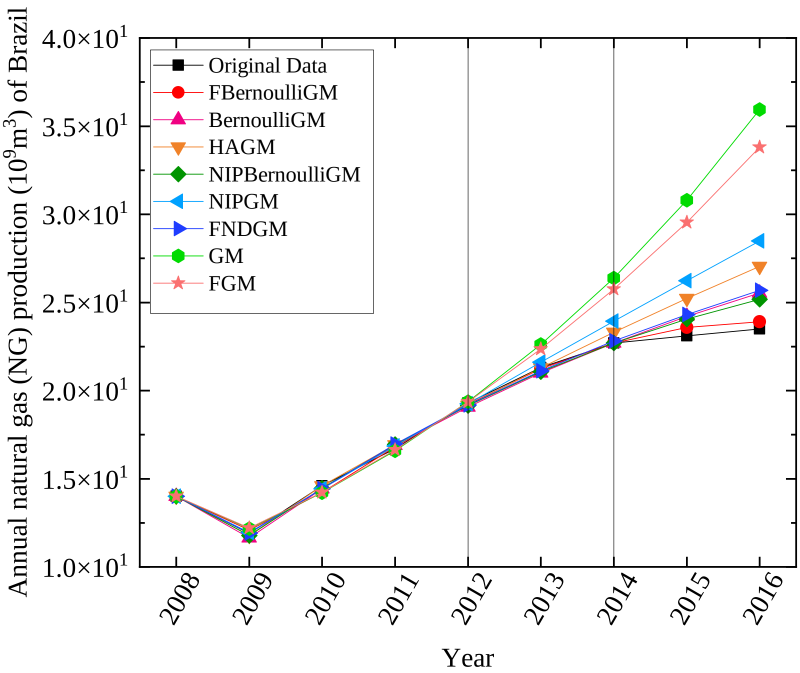

Fig. 1. Comparison of observed and fitted data in Case 1

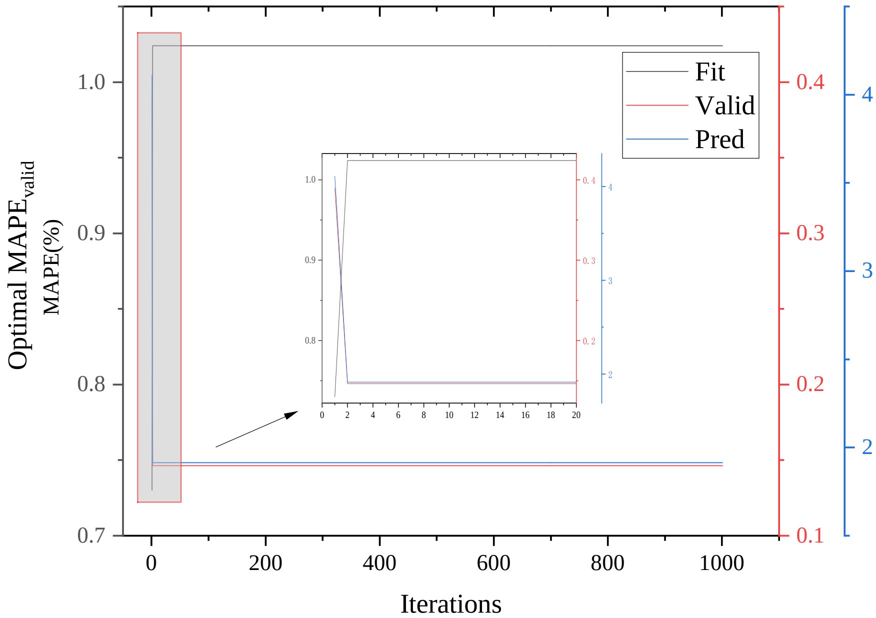

Fig. 2. MAPE$_{valid}$ convergence trend in Case 1

3.2 Case 2: forecasting natural gas production in Italy

Research into Italy' s natural gas production is vital for grasping its role in the European energy landscape. As a major consumer and hub for energy distribution in Europe, Italy' s production capabilities play a pivotal role in balancing the region' s energy supply. By analyzing Italy's natural gas output, we can evaluate its contribution to energy diversification, which is crucial for reducing Europe' s dependence on external suppliers, such as Russia, and strengthening energy resilience. The study also highlights Italy' s efforts in fostering sustainable energy practices and supporting Europe' s transition toward greener energy sources.

In Table 2, a comparison of the original data and predictions from the same models over the years 2008 to 2016 is presented. The fitting errors (MAPE$_{fit}$) show that FNDGM provides the best fitting performance with a MAPE$_{fit}$ of 0.3995%, significantly outperforming the other models. HAGM also performs well in fitting with a MAPE$_{fit}$ of 0.7401%. However, in the validation phase, NIPBernoulliGM performs the best with a perfect MAPE$_{valid}$ of 0%, making it the most accurate in predicting unseen data. Despite this, FBernoulliGM proves to be the best performer in the prediction phase, achieving a MAPE$_{pred}$ of 4.0353%, slightly outperforming BernoulliGM and HAGM, which have MAPE$_{pred}$ values of 4.3046% and 8.0386%, respectively. In contrast, NIPGM exhibits the highest error with a MAPE$_{pred}$ of 32.6385%, indicating weak predictive ability.

The comparison of the predicted results from each model is presented in detail in Fig.3. The parameter optimization is achieved using the PSO (Particle Swarm Optimization) algorithm, with the number of iterations set to 1000. The convergence curve of MAPE$_{valid}$ is shown in Fig.4.

Table 2. Detailed results in Case 2

Year	Original Data	FBernoulliGM	BernoulliGM	HAGM	NIPBernoulliGM	NIPGM	FNDGM	GM	FGM
2008	8.40	8.40	8.40	8.40	8.40	8.40	8.40	8.40	8.40
2009	7.30	6.94	6.99	7.21	6.86	7.73	7.35	7.36	7.96
2010	7.60	7.76	7.74	7.64	7.81	7.65	7.52	7.52	7.73
2011	7.70	7.87	7.88	7.81	7.91	7.61	7.72	7.68	7.59
2012	7.80	7.60	7.62	7.76	7.53	7.59	7.80	7.84	7.50
2013	7.00	7.11	7.12	7.51	7.00	7.58	7.71	8.01	7.44
2014	6.50	6.50	6.50	7.09	6.50	7.58	7.43	8.18	7.39
2015	6.20	5.84	5.83	6.53	6.10	7.58	6.93	8.35	7.35
2016	5.30	5.18	5.16	5.87	5.81	7.58	6.19	8.53	7.32
$\mathbf{MAPE}_{fit}$		2.3617	2.1468	0.7401	2.9959	2.0819	**0.3995**	0.5294	3.2053
$\mathbf{MAPE}_{valid}$		0.7857	0.8571	8.1813	**0**	12.4505	12.2253	20.1374	9.989
$\mathbf{MAPE}_{pred}$		**4.0353**	4.3046	8.0386	5.6178	32.6385	14.2833	47.8104	28.3308

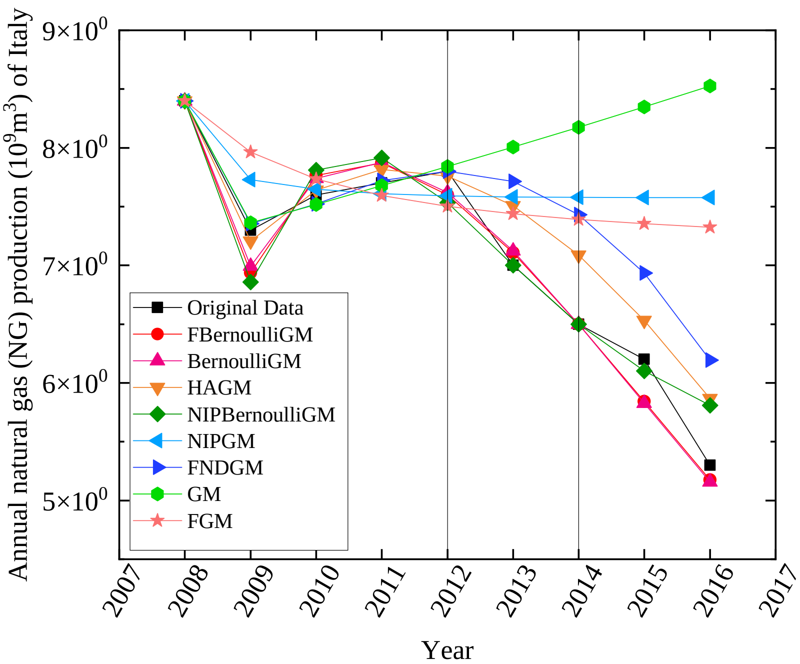

Fig. 3. Comparison of observed and fitted data in Case 2

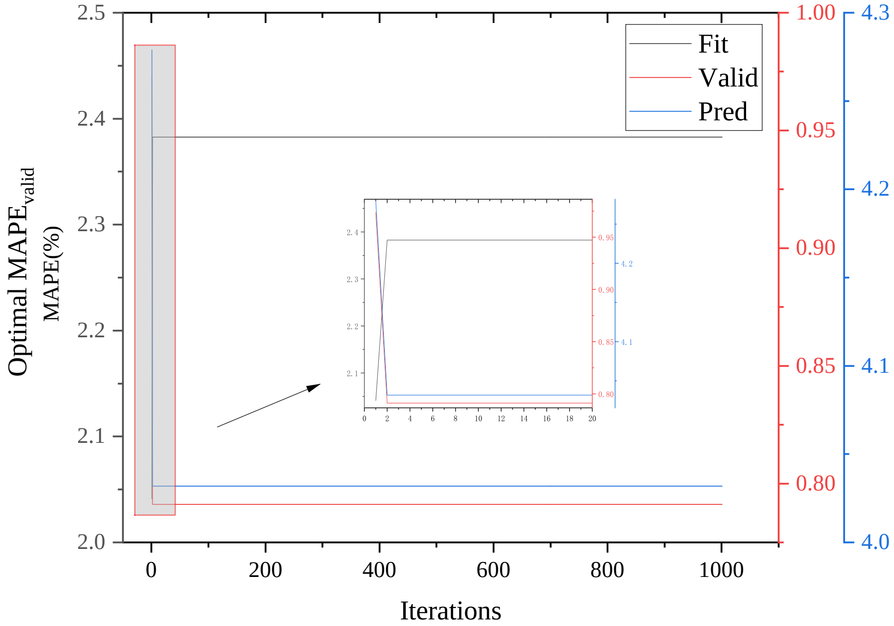

Fig. 4. MAPE$_{valid}$ convergence trend in Case 2

3.3 Case 3: forecasting natural gas production in Qatar

The significance of studying natural gas production in Qatar lies in its status as one of the world’s largest exporters of natural gas. Qatar’s production levels have a profound influence on the global energy market, particularly in meeting the energy demands of Asia and Europe. By focusing on Qatar’s production trends, we can understand its role in maintaining global energy security, stabilizing natural gas prices, and ensuring a steady supply to major importing countries. Additionally, Qatar’s strategies in managing its vast natural gas reserves provide valuable lessons for other nations in optimizing resource utilization and advancing environmental sustainability.

In Table 3, predictions and errors from various models are analyzed over the years 2008 to 2016. The FNDGM model delivers the most accurate fitting performance with a MAPE_{fit} of 0.0446%, showcasing its superior capacity to model the training data. In terms of validation performance, NIPBernoulliGM emerges as the best model with a MAPE_{valid} of 2.5704%, followed closely by FBernoulliGM, which also performs well in both validation (MAPE_{valid} of 2.5873%) and prediction ($_{of}$1.8902%), making it a top performer overall. Meanwhile, BernoulliGM achieves the lowest prediction error with a $_{of}$1.8652%. In comparison, HAGM’s $_{of}$91.4939% highlights significant shortcomings in its predictive ability. GM and FGM show poor predictive performance, with $_{val}$ues of 58.011% and 13.262%, respectively.

Fig.5 provides a detailed illustration of the prediction results for each model. Parameters were optimized using the PSO (Particle Swarm Optimization) algorithm, with 1,000 iterations. The MAPE_{valid} convergence curve is depicted in Fig.6.

Table 3. Detailed results in Case 3

Year	Original Data	FBernoulliGM	BernoulliGM	HAGM	NIPBernoulliGM	NIPGM	FNDGM	GM	FGM
2008	77.00	77.00	77.00	77.00	77.00	77.00	77.00	77.00	77.00
2009	89.30	97.14	95.72	98.63	95.67	97.89	89.30	101.50	99.92
2010	131.20	123.57	123.35	122.52	122.68	121.56	131.16	118.87	121.51
2011	145.30	143.99	143.96	143.25	143.28	141.62	145.45	139.21	140.78
2012	157.00	158.66	158.63	157.68	158.35	158.61	156.86	163.04	157.87
2013	177.60	168.41	168.38	160.24	168.47	172.99	166.15	190.95	173.04
2014	174.10	174.10	174.10	140.78	174.10	185.18	174.10	223.64	186.49
2015	178.50	176.52	176.56	81.07	175.59	195.51	181.11	261.92	198.43
2016	181.20	176.36	176.41	-51.47	173.25	204.25	187.40	306.75	209.03
MAPE$_{fit}$		3.3108	3.0266	3.7816	3.1755	4.105	**0.0446**	6.2196	4.5886
MAPE$_{valid}$		2.5873	2.5957	14.4566	**2.5704**	4.4799	3.2235	17.9859	4.8421
MAPE$_{pred}$		1.8902	**1.8652**	91.4939	3.0088	11.1251	2.4419	58.011	13.262

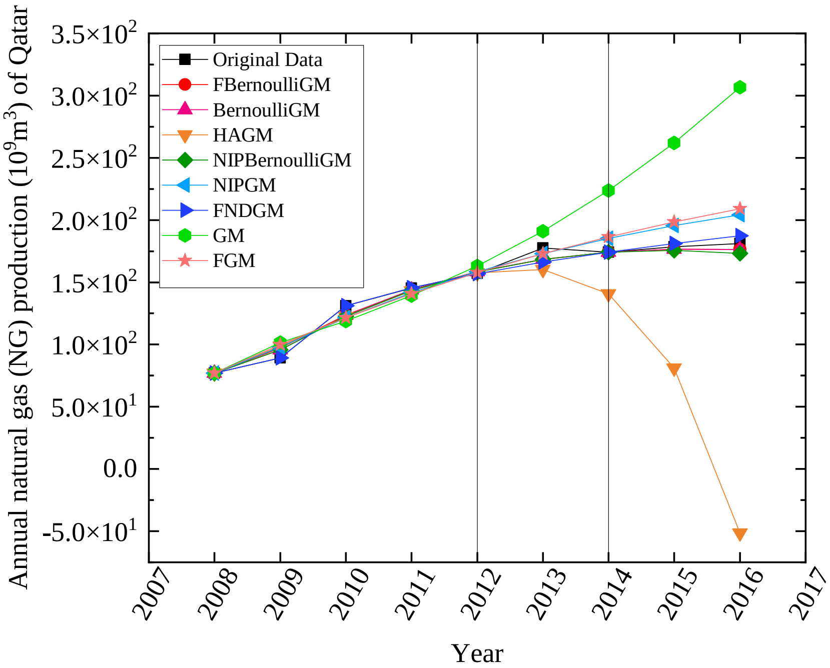

Fig. 5. Comparison of observed and fitted data in Case 3

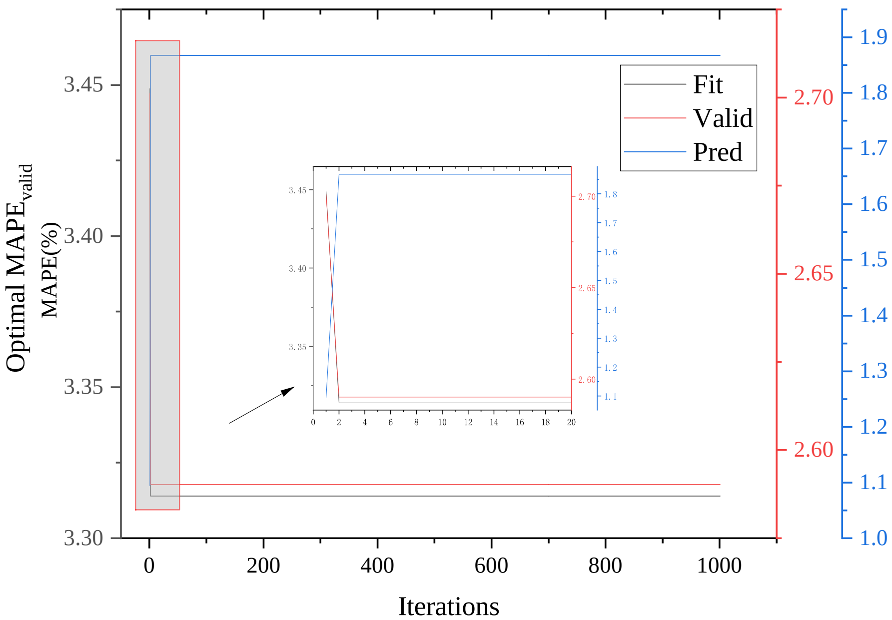

Fig. 6. MAPE_{valid} convergence trend in Case 3

3.4 Discussion

In this study, based on the analysis of natural gas production forecasts, we compared the FBernoulliGM model with various classical grey models, such as BernoulliGM, HAGM, NIPBernoulliGM, NIPGM, FNDGM, GM, and FGM. The experimental results demonstrate that FBernoulliGM excels in accuracy and predictive power, particularly when handling nonlinear and complex trends in the data.

Compared to traditional GM and FGM models, FBernoulliGM, through the introduction of fractional-order accumulation and the New Information Priority Accumulation method, enhances the model' s sensitivity to historical data, allowing it to better capture subtle fluctuations and trends. This advantage is especially evident when applied to natural gas production data with small sample sizes and significant fluctuations, significantly improving the model's forecasting accuracy.

Additionally, the FBernoulliGM model has a more simplified structure compared to more complex models like NIPBernoulliGM and HAGM, which offers advantages in computational efficiency. Especially when performing large-scale forecasting tasks, FBernoulliGM converges faster and has lower computational costs. This allows FBernoulliGM to strike a better balance between accuracy and efficiency.

However, FBernoulliGM also has certain limitations. For instance, when the data exhibits extreme nonlinearity or abrupt trends, the prediction error of this model may increase, making it less effective at handling all outliers. Moreover, while the fractional-order accumulation method enhances data adaptability, the selection of its parameters still relies on experience, which may affect the model' s generalizability.

Overall, the FBernoulliGM model demonstrates strong adaptability and stability in natural gas production forecasting, especially in small-sample scenarios, where its forecasting performance significantly surpasses other traditional grey models. This provides new ideas and tools for data forecasting in the energy sector, while also indicating that there is room for improvement in future work, particularly in optimizing model parameters and further enhancing the ability to handle nonlinearities.

4 Conclusions

This study demonstrates that the FBernoulliGM model exhibits significant effectiveness in predicting natural gas production, especially in situations involving small samples and fluctuating trends. By combining the fractional-order accumulation method with the New Information Priority accumulation method, the FBernoulliGM model significantly improves prediction accuracy and adaptability compared to traditional grey models (such as GM, FGM, and NIPBernoulliGM). Its ability to capture subtle data trends and fluctuations makes it particularly well-suited for forecasting natural gas production in different regions.

In this study, the particle swarm optimization (PSO) algorithm was utilized to finely tune the model parameters, and the Mean Absolute Percentage Error (MAPE) was employed as the evaluation criterion to systematically assess the performance of seven models. The analysis results further confirm that the FBernoulliGM model has significant advantages in both fitting and predictive capabilities, demonstrating its outstanding reliability in forecasting natural gas production data. These findings provide a solid theoretical foundation and practical reference for future natural gas production forecasts.

However, despite the model's excellent performance in most scenarios, its predictive effectiveness may decline in highly nonlinear conditions, and its dependence on parameter selection remains a limitation. Nonetheless, the FBernoulliGM model achieves a good balance between prediction accuracy and computational efficiency, providing a valuable tool for forecasting in the energy sector. Future research can focus on further optimizing parameter selection and enhancing the model's capability to handle extreme nonlinear data. Overall, this study contributes significantly to the application of grey system theory in the energy field, particularly in natural gas production forecasting.

References

[1] A. Safari, N. Das, O. Langhelle, J. Roy, M. Assadi, Natural gas: A transition fuel for sustainable energy system transformation?, Energy Science & Engineering 7 (4) (2019) 1075–1094.

[2] O. Yemelyanov, A. Symak, T. Petrushka, O. Vovk, O. Ivanytska, D. Symak, A. Havryliak, T. Danylovych, L. Lesyk, Criteria, indicators, and factors of the sustainable energy-saving economic development: the case of natural gas consumption, Energies 14 (18) (2021) 5999.

[3] J.-L. Deng, Control problems of grey systems., Systems Control Lett. 1 (5) (1982) 288–294.

[4] J. Cui, Y. Dang, S. Liu, A new grey prediction model and its modeling mechanism, Control and Decision 24 (2009) 1702–1706.

[5] N. Xie, S. Liu, The discrete gm(1,1) model and the modeling mechanism of grey prediction models, Systems Engineering Theory and Practice (2005) 93–99.

[6] J.-L. Deng, The theory and method of social and economic grey systems, Chinese Social Sciences (1984) 47–60.

[7] T.-L. Tien, A research on the grey prediction model gm (1, n), Applied mathematics and computation 218 (9) (2012) 4903–4916.

[8] T.-L. Tien, The indirect measurement of tensile strength of material by the grey prediction model GMC(1, n), Measurement Science and Technology 16 (6) (2005) 1322–1328.

[9] Z.-X. Wang, Nonlinear grey prediction model with convolution integral ngmc (1, n) and its application to the forecasting of china’s industrial so2 emissions, Journal of Applied Mathematics 2014 (1) (2014) 580161.

[10] X. Naiming, L. Sifeng, Multivariate discrete grey model and its properties, Systems Engineering Theory and Practice 28 (6) (2008) 143–150.

[11] C.-I. Chen, H. L. Chen, S.-P. Chen, Forecasting of foreign exchange rates of taiwan’s major trading partners by novel nonlinear grey bernoulli model ngbm (1, 1), Communications in Nonlinear Science and Numerical Simulation 13 (6) (2008) 1194–1204.

[12] C.-I. Chen, P.-H. Hsin, C.-S. Wu, Forecasting taiwan’s major stock indices by the nash nonlinear grey bernoulli model, Expert Systems with Applications 37 (12) (2010) 7557–7562.

[13] B. Zeng, H. Duan, Y. Zhou, A new multivariable grey prediction model with structure compatibility, Applied Mathematical Modelling 75 (2019) 385–397.

[14] K. Li, L. Liu, J. Zhai, T. M. Khoshgoftaar, T. Li, The improved grey model based on particle swarm optimization algorithm for time series prediction, Engineering Applications of Artificial Intelligence 55 (2016) 285–291.

[15] Z.-X. Wang, D.-J. Ye, Forecasting chinese carbon emissions from fossil energy consumption using non-linear grey multivariable models, Journal of Cleaner Production 142 (2017) 600–612.

[16] X. Ma, M. Xie, W. Wu, B. Zeng, Y. Wang, X. Wu, The novel fractional discrete multivariate grey system model and its applications, Applied Mathematical Modelling 70 (2019) 402–424.

[17] W. Wu, X. Ma, Y. Zhang, W. Li, Y. Wang, A novel conformable fractional non-homogeneous grey model for forecasting carbon dioxide emissions of brics countries, Science of the Total Environment 707 (2020) 135447.

[18] X. Ma, Z. Liu, Y. Wang, Application of a novel nonlinear multivariate grey bernoulli model to predict the tourist income of china, Journal of Computational and Applied Mathematics 347 (2019) 84–94.

[19] X. Ma, Z.-b. Liu, The kernel-based nonlinear multivariate grey model, Applied Mathematical Modelling 56 (2018) 217–238.

[20] W. Wu, X. Ma, B. Zeng, Y. Wang, W. Cai, Forecasting short-term renewable energy consumption of china using a novel fractional nonlinear grey bernoulli model, Renewable energy 140 (2019) 70–87.

[21] C.-I. Chen, Application of the novel nonlinear grey bernoulli model for forecasting unemployment rate, Chaos, Solitons & Fractals 37 (1) (2008) 278–287.

[22] X. Zhao, X. Ma, Y. Cai, H. Yuan, Y. Deng, Application of a novel hybrid accumulation grey model to forecast total energy consumption of southwest provinces in china, Grey Systems: Theory and Application 13 (4) (2023) 629–656.

[23] J. Xiao, X. Xiao, Forecast of clean energy generation in china based on new information priority nonlinear grey bernoulli model, Environmental Science and Pollution Research 30 (51) (2023) 110220–110239.

[24] L.-F. Wu, S.-F. Liu, W. Cui, D.-L. Liu, T.-X. Yao, Non-homogenous discrete grey model with fractional-order accumulation, Neural Computing and Applications 25 (2014) 1215–1221.

[25] L. Wu, S. Liu, H. Chen, N. Zhang, et al., Using a novel grey system model to forecast natural gas consumption in china, Mathematical Problems in Engineering 2015 (2015).

[26] L. Wu, S. Liu, J. Liu, Fractional accumulation model of grey gm(1,1) and its stability, Control and Decision 29 (2014) 919–924.

[27] C. Yin, S. Mao, Fractional multivariate grey bernoulli model combined with improved grey wolf algorithm: Application in short-term power load forecasting, Energy 269 (2023) 126844.

[28] U. Şahin, T. Şahin, Forecasting the cumulative number of confirmed cases of covid-19 in italy, uk and usa using fractional nonlinear grey bernoulli model, Chaos, Solitons & Fractals 138 (2020) 109948.

[29] F. Marini, B. Walczak, Particle swarm optimization (pso). a tutorial, Chemometrics and Intelligent Laboratory Systems 149 (2015) 153–165.

[30] D. Wang, D. Tan, L. Liu, Particle swarm optimization algorithm: an overview, Soft computing 22 (2) (2018) 387–408.

[31] X. Ma, W. Wu, B. Zeng, Y. Wang, X. Wu, The conformable fractional grey system model, ISA transactions 96 (2020) 255–271.