Research Article

Ying Huang, Weilong Huang, and Song Ding

1Industrial and Commercial Bank of China, Hangzhou Branch, Hangzhou 310003, China
2School of Economics, Zhejiang University of Finance and Economics, Hangzhou 310018, China

Correspondence should be addressed to Weilong Huang; hwl@zufe.edu.cn

Received 11 May 2022; Revised 2 June 2022; Accepted 11 June 2022; Published 6 July 2022

Copyright © 2022 Ying Huang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Since the limitation of carbon emissions, China’s photovoltaic (PV) industry has developed vigorously, while some traditional heavy industries have been violently hit. Therefore, the industrial production data exhibits significant nonlinear and complexity characteristics, which may affect prediction accuracy, thus hindering the corresponding department’s decision-making. Consequently, a novel structure-adaptive fractional Bernoulli grey model is presented in this paper to surmount this toughie, and the core innovations can be summarized as follows. Initially, a novel time function term is utilized to depict the accumulative time effect, which can smoothly represent the dynamic variations and significantly strengthen the robustness of the new model. Besides, the fractional-order accumulation technique, which could effectively improve the predicting accuracy, is employed in the proposed model. Furthermore, the adaptability and generalizability of the proposed model can be enhanced by the self-adaptive parameters optimized by the Particle Swarm Optimization. For illustration and verification purposes, experiments on forecasting the annual output of Photovoltaic modules in China and the annual output of steel in Beijing are compared with a range of benchmarks, including the classic GM (1, 1), conventional econometric technology, and machine learning methods. And the results confirmed that the proposed model is superior to all benchmark models, which indicates that the novel model is indeed suitable for industrial production forecasting.

1. Introduction

As the largest carbon emitter, China has announced that it will aim for emissions to peak before 2030 and be carbon neutral by 2060. Moreover, the carbon neutrality trajectory will be rooted in Five-Year Plans, connecting current policies to 2035 development goals and continuing the transition to 2060 [1]. To achieve such goals, the development and utilization of renewable energy have become increasingly crucial for the energy transition from carbon-based fossil fuels to clean energy [2, 3]. Specifically, as Figure 1 shows, China has set a goal of achieving 40% of electricity consumption from nonfossil generation by 2030. Moreover, the present has announced that solar PV and wind capacity will reach 1200 GW by 2030. Therefore, solar PV has large untapped potential and competitiveness and is expected to have a promising prospect in the following years.

PV energy is a representative of renewable energy that has been developed and used on a large scale due to its essential role in decarbonization. Additionally, the PV systems comprise one or more solar modules, which affect the electricity production amount [5]. Thus, it is imperative to predict PV modules for achieving the decarbonization goals accurately. However, the nonlinearity, volatility, and uncertainty of the PV system to solar irradiance, locations, and public emergency (for instance, COVID-19) pose significant challenges to producing precise forecasts [6, 7]. To overcome these challenges, accurate estimates of solar modules are indispensable for meeting the requirements of PV’s promising prospects. Moreover, it can also be helpful in grid management and power dispatch [8].

1.1. Previous Studies on Solar Photovoltaic Forecasts. Currently, researchers have conducted a range of studies on PV forecasts, and their adopted methods are typically divided into three categories, including physical approaches,
Figure 1: China renewable capacity additions, 2009–2026 (a) and nonfossil energy target proposal for 2030 (b). Note: 2026 values are estimated according to the 2030 trajectory, and the data source comes from international energy agency [4].

statistical models, deep learning methods, and grey prediction models [9]. Therefore, the appropriate and correct selection of a prediction technique is a necessary first step, and this action must be undertaken reasonably based on the rational analysis of the existing models and their forecasting performance.

Physical models normally use coupled thermodynamic techniques with initial and boundary value conditions to generate forecasts [10]. The advantages of the physical models are that they have no requirements for historical data; instead, they need geographical information, meteorological data, and information on PV accumulators that suffice [11]. However, it costs up to six hours of intensive computation to calculate the meteorological parameters for producing predictions [12]. Therefore, it is not suitable for ultra-short-term forecasts. Additionally, PV parameters usually suffer from the deviations from the manufacturers’ specifications during operation, thereby reducing the forecasting accuracy of physical methods [13, 14].

In addition to the physical approaches, statistical models play an alternative role in PV forecasting. The representative technique is the autoregressive moving average (ARIMA), which can effectively extract the statistical properties and adopt the Box-Jenkins method [15], as well as the time-series ensemble methods [16] and regression models [17]. However, there exist several limitations for these models. For instance, they highly rely on probabilistic autoregression and utilized probability mass bias for probabilistic predictions [18]. Moreover, these techniques may lose power in long-term forecast horizons owning to issues with adaptability, learning capability, and dependency on sufficient sample sizes for statistical modeling [19, 20].

To address nonlinearity and volatility in PV samples, deep learning methods have been implemented, which can describe nonlinear correlations and self-learning and self-adapting. The merits of these techniques are that no governing mathematical methods are required for specifying the correlations between inputs and outputs because they highly rely on pattern predictions to PV generation [19, 21]. Artificial neural networks (ANNs) are representatives for predicting PV generations [22]. Notable works relate that the ANN method was utilized to estimate PV parameters for seven cities from Turkey. Experimental results revealed that the ANN-based model performed the best to predict solar radiation in the specified areas [23]. Additionally, Long Short-Term Memory (LSTM) has shown superiority in PV forecasting. This model can extract more essential information from raw data and avoid the pitfalls of overfitting and fading, owing to its special gate structures [24]. Further, to mitigate the shortcomings of the methods discussed before, some researchers have used hybrid configurations of algorithms, for instance, combining LSTM and CNN [25]. Evidence shows that these hybrid models improved forecasting performance significantly compared to single-architecture models [26]. Ahmed et al. [9] employed an ensemble-based LSTM algorithm constituted of ten component LSTM models. Then, they used different data segmentations from three months to one data to compare the effects based on the various forecasting horizons from fourteen days to five minutes. Two case studies validated the efficacy and robustness of the proposed model with the MAPE of 6.02%. Li et al. [27] employed the CNN-LSTM to predict PV output. Their results revealed that the proposed hybrid model outperformed the single CNN and LSTM models. However, these hybrid models usually suffer from an excess of redundant information and strict requirements on the selected variables, which constrains the application field [28].

Although some statistical models and deep learning models are used in this field, they might lose power in predicting China’s PV modules featured by limited samples, volatility, and nonlinearity. Thus, the grey prediction model can be an alternative to dealing with such situations.
1.2. Previous Studies on the Grey Models. The grey system models put forward by Deng in 1982 [29] have been confirmed to be suitable methodologies for forecasting with limited information and much uncertainty. Compared with the traditional econometric prediction methods [30, 31] and machine learning methods [32–35], the most noticeable advantage of the grey theory is that it can still achieve pinpoint prediction when confronted with scarce data. Nevertheless, restricted by some inherent defects, such as the linear structure of the model formula and the mean value of the consecutive neighbors served as an estimate of the background value, the forecasting results of the classical models typically have weak adaptability. To overcome the defects, many scholars have modified the original GM (1, 1) from several perspectives, which can be summarized as follows: (1) the data preprocessing technique; (2) the initial condition; (3) the calculating method of background value; (4) the model structure.

The avenues to improve the performance of grey models through data preprocessing can be divided into various categories. For instance, Wang et al. introduced the data grouping approach to handle the seasonal attribute of the raw data [36] and made use of the seasonal factors to deseasonalize the original sequences [37], and the exponential smoothing method was utilized to erase the irregular disturbance contained in the raw data for the forecasting of solar energy consumption [38]. In addition, many scholars have investigated the buffer operator, such as the fractional-order accumulation generation operator proposed by Wu et al. [39], the damping accumulated generating operator presented by Liu et al. [40], and a novel weakening buffer operator put forward by Chen and Wu [41]. In general, compared with the first-order accumulation generating operation, the above-improved buffer operators can allocate the weights of old and new information more rationally, thus increasing the predicting accuracy.

From the time response function of the classic grey model, one can discover that the first component of the raw time series \( x^{(0)} (1) \) is served as the initial condition. Undoubtedly, this practice violates the new information principles and may result in serious biases. Considering the defect, Dang et al. [42] proposed the \( n \)th component \( x^{(b)} (n) \) as the initial condition, verified to improve the prediction performance effectively. Nevertheless, this approach completely ignores the role of early information. Referring to previous studies, Wang et al. [43] further optimized the initial condition of the grey model by using the linear combination of \( x^{(0)} (1) \) and \( x^{(0)} (n) \), which considers both new information and old information but still fails to take full advantage of intermediate data. Based on the above works, a novel initial condition determined by a linear combination of all collected observations was proposed by Ding et al. [44], which has been proved to be superior to those models mentioned above.

As for the background value, the methodology based on the mean value of the consecutive neighbors generating operator has been demonstrated to be one of the sources of error. To surmount the intrinsic faultiness, various mathematical methods are employed to calculate the background value, such as the compound trapezoid formula [45] and the compound Simpson quadrature formula [46]. As the theoretical analysis shows, the prediction accuracy of these optimized models is significantly higher than that of traditional models.

For the model structure, many studies have been conducted in recent years. To enhance the forecasting precision for the nonlinear sample data, Wang et al. [47] proposed \( NGBM(1, 1) \) model, and a method of estimating the power exponent based on the basic principle of information overlap is given [19]. Additionally, Zhou and Ding [3] put forward a discrete grey model incorporated with the seasonal dummy variables to describe the seasonal characteristic. Moreover, for better adaptability and robustness, the time-varying parameters are introduced into the discrete grey model for long-term power generation forecasting [48]. In addition to the above three models, other structural optimization can also obtain excellent prediction performance, for simplicity purposes, listed at the end of this paper [49–52].

1.3. Summarization and Motivation. It can be discovered from the above that although structural optimization is still the most potent aspect in comparison with the other three aspects, it does not mean that the majorization of data preprocessing, initial condition, and background value serves no purpose. Furthermore, although the existing research has improved significantly compared with the conventional GM (1, 1) model, these previous models cannot accurately predict the time series with prominent nonlinear characteristics. Thus, a novel self-adaptive Bernoulli grey model, abbreviated as the SFBGM model, which combines structural optimization with data preprocessing, is put forward to overcome this issue. Specifically, the core contributions are summarized as follows:

(1) A novel structure-adaptive fractional Bernoulli grey model combined is proposed in this paper. In particular, the grey action quantity included in the classic grey model is replaced by the time function term, which can effectively describe the cumulative time effect. Accordingly, the novel model's robustness, adaptability, and generalization ability are significantly improved.

(2) The PSO algorithm is employed to determine the generating parameters in the proposed model. Detailed procedures for utilizing this new model have been elaborated on in graphical manners. Accordingly, the predictive performance of the novel model is significantly enhanced.

(3) In the experimental part, the flexibility and efficacy of the proposed model have been verified by investigating several experiments with different development trends. Experimental results demonstrate that the new model can precisely predict issues with different developing characteristics. In contrast with the benchmarks, the SFBGM model performs the best. Thus, it can be demonstrated that the proposed model is indeed a valuable forecasting method.
The remaining parts are arranged as follows: Section 2 provides the theoretical analysis of the SFBGM model. After illustrating the mechanism of the novel model, two practical cases are selected to empirically demonstrate the efficacy of the proposed model in Section 3. Finally, conclusions and future work are summarized in Section 4.

2. Methodology

In this chapter, the novel self-adaptive fractional Bernoulli grey model, abbreviated as the SFBGM model, will be illustrated in detail. To illustrate the proposed model’s innovations and principles, this section is divided into the following three subsections. Initially, the construction process and the mathematical derivation of the proposed model are shown in Section 2.1. Subsequently, in Section 2.2, optimizing the self-adaptive parameters by the Particle Swarm Optimization (PSO) is elaborated on. Ultimately, in Section 2.3, the evaluation criteria are introduced to evaluate the model’s efficacy objectively.

2.1. The Establishment of the Proposed Model. To generate precise predictions for nonlinear time series, the first-order accumulation generating operation (1-AGO), which can effectively excavate the inherent laws of data, is utilized to smooth the original sequences. However, the above operator may not reflect that recent data are more crucial than previous data in forecasting, namely, the principle of new information priority. Therefore, the fractional-order accumulation is employed to resolve this defect.

**Definition 1.** Suppose that \( X^{(0)} \) is the raw data, and then its \( \gamma \)-order fractional-order accumulation (\( \gamma \)-FOA) series is presented as follows:

\[
x^{(\gamma)}(t) = \sum_{i=1}^{t} \left( \frac{t - i + \gamma - 1}{t - i} \right) x^{(0)}(i),
\]

where \( \frac{(\gamma + 1)t}{(t)!} \), and the value of \( \gamma \) is limited to the interval \((0,1)\) abided by the principle of new information priority. The closer the \( \gamma \) approaches to 0, the larger the weights assigned to the late information in the \( \gamma \)-FOA sequence are. Obviously, it will degenerate to the 1-AGO when \( \gamma \) equals one, i.e., \( x^{(1)}(t) = \sum_{i=1}^{t} x^{(0)}(i) \).

For restoration purposes, the \( \gamma \)-order inverse fractional-order accumulation (\( \gamma \)-IFOA) can be defined as follows:

\[
x^{(-\gamma)}(t) = \sum_{i=1}^{t} \left( \frac{t - i - \gamma - 1}{t - i} \right) x^{(0)}(i).
\]

**Definition 2.** Based on the \( \gamma \)-FOA sequence, the novel self-adaptive fractional Bernoulli grey model (SFBGM) is put forward. The whitened equation of the proposed model is presented as follows:

\[
\frac{dx^{(\gamma)}(t)}{dt} + a x^{(\gamma)}(t) = \left( \beta t^{(\gamma)} + \lambda \right) \left( x^{(\gamma)}(t) \right)^L,
\]

where \( L \) indicates the \( (\beta t^{(\gamma)} + \lambda) \) power index. \( y_1 \) and \( y_2 \) represent the accumulation order for the original sequences and time function items, respectively. The above parameters optimized by the PSO algorithm (as illustrated in Section 2.2) can be determined according to the characteristics of the different datasets with diverse evolution trends and oscillation states, thus generating more precise predictions. Therefore, the three parameters \( y_1, y_2, \) and \( L \) are called the self-adaptive parameters, which can significantly enhance the forecasting performance of the novel model. Moreover, the parameter \( \alpha \) in equation (3) is called the development coefficient describing the time sequences’ development trend, and \( (\beta t^{(\gamma)} + \lambda) \) is served as the grey action quantity. The parameters \( \alpha, \beta, \) and \( \lambda \) are linear parameters, which can be estimated by the ordinary least square method (as presented in Theorem 2).

In comparison with the conventional fractional Bernoulli model, the most significant modification of the novel model is that the traditional grey action quantity (a constant) is replaced by the time function \( (\beta t^{(\gamma)} + \lambda) \). Thereinto, \( f^{(\gamma)} \) is the \( y_2 \)-order fractional accumulation of time item, which can be served as a polynomial function, as presented as follows:

\[
f^{(\gamma)} = \sum_{i=1}^{t} \left( \frac{t - i + y_2 - 1}{t - i} \right) i.
\]

The advantages of the function \( (\beta t^{(\gamma)} + \lambda) \) can be concluded as the following two aspects. On the one hand, this function can feasibly describe the features of the equation itself developing with time, which leads to stronger generalizability of the proposed model. On the other hand, the value of the aforementioned self-adaptive parameter \( y_2 \) can be arbitrarily selected within the interval \((0, 1)\) to obtain more precise forecasting performance for the time series featured by nonlinearity.

Subsequently, the time response function of the presented model will be deduced in the following procedures.

**Theorem 1.** Let \( G(t) = [x^{(\gamma)}(t)]^{1-L}, \) \( a_1 = (1 - L)\alpha, \beta_1 = (1 - L)\beta, \lambda_1 = (1 - L)\lambda, \) and denote \( (\beta_1 t^{(\gamma)} + \lambda_1) \) as \( f(u) \). Substitute them into equation (3), and the time response function of the proposed model can be given by

\[
\tilde{x}^{(\gamma)}(t) = \tilde{G}(t)^{1/L},
\]

where...
\[
\tilde{G}(t) = \left[ x^{(0)}(1) \right]^{1-L} \cdot e^{-\alpha_1(t-1)} + \sum_{u=2}^{t} \left\{ e^{-\alpha_1(t-u+1/2)} \times \frac{1}{2} \left[ f(u) + f(u-1) \right] \right\}.
\]

Proof. Based on the whitened equation, the time response function can be inferred. For brevity purposes, the original equation needs to be simplified. Firstly, multiply both sides of equation (3) by \( [x^{(y)}(t)]^{-L} \), and one can obtain the following:
\[
\frac{dx^{(y)}(t)}{dt} - \left[ x^{(y)}(t) \right]^{1-L} = \alpha_1 x^{(y)}(t) = \beta \tilde{G}(t) + \lambda.
\]

Multiplying both sides of equation (7) by \((1-L)\), then one can get
\[
\frac{dG(t)}{dt} + \alpha(1-L)G(t) = \beta(1-L)t^{(y)} + \lambda(1-L).
\]

Subsequently, substitute \( \alpha_1, \beta_1, \) and \( \lambda_1 \) into equation (8), which can be simplified as
\[
\frac{dG(t)}{dt} + \alpha_1 G(t) = \beta_1 t^{(y)} + \lambda_1.
\]

By virtue of the above intermediate coefficients, equation (3) has been transformed into equation (9). Next, substitute the initial condition \( \tilde{G}(t)|_{t=1} = G(1) \) into equation (6), and the general solution of equation (9) can be deduced as follows:
\[
G(t) = P(t) \cdot G(1) \cdot e^{-\alpha_1(t-1)} = P(t) \left[ x^{(y)}(1) \right]^{1-L} \cdot e^{-\alpha_1(t-1)}.
\]

where \( P(t) \) represents a time function to be determined. Soon afterward, take equations (10) into (9), and integrate both sides over the interval \([1, t]\), and then the formula of \( P(t) \) can be given by
\[
P(t) = \int_1^t \frac{\beta_1 u^{(y)} + \lambda_1}{x^{(y)}(1)} \cdot e^{-\alpha_1(u-1)} du.\]

Substitute the formula of \( P(t) \) into equation (10), and it can be rewritten as

\[
G(t) = \left[ x^{(y)}(1) \right]^{1-L} \cdot e^{-\alpha_1(t-1)} + e^{-\alpha_1(t-1)} \int_1^t \left[ \beta_1 u^{(y)} + \lambda_1 \right] e^{-\alpha_1(u-1)} du.
\]

Definition 3. Integrating both sides of equation (3) on the interval \([k-1, k]\), the discrete form of the proposed model can be expressed as
\[
x^{(y)}(k) = x^{(y)}(k-1) + \alpha z^{(y)}(k)
\]

where \( z^{(y)}(k) \) is defined as the background value; actually, \( z^{(y)}(k) = \int_{k-1}^{k} x^{(y)}(t) dt \). However, the \( y^1 \)-FOA generating sequence \( x^{(y)}(t) \) cannot be integrated directly as a continuous function. Therefore, the background values of the novel model are approximated by the mean value of the consecutive neighbors of \( x^{(y)}(k) \). The expression is shown as follows:
\[
z^{(y)}(k) = \frac{1}{2} \left[ x^{(y)}(k) + x^{(y)}(k-1) \right], k = 2, 3, \ldots, n.
\]

Theorem 2. Supposing that \( \gamma_1, \gamma_2, \) and \( L \) are already known, the linear parameters \( \phi = (\alpha, \beta, \lambda) \) of the SFBGM model can be computed by the least squares method as follows:
\[
\tilde{\phi} = (B'B)^{-1}B'Y = (\tilde{\alpha}, \tilde{\beta}, \tilde{\lambda})',
\]

However, because the pinpoint results of the integral formula \( e^{-\alpha_1(t-1)} \int_1^t \left[ b_1 u^{(y)} + c_1 \right] e^{-\alpha_1(u-1)} du \) in equation (12) may not be computed straightforward, a mathematical method, which has been verified to be effective [34], should be utilized to estimate its true value. The procedure is expressed as follows:
\[
\int_1^t e^{-\alpha_1(t-u)} f(u) du = \sum_{u=2}^{t} \left\{ e^{-\alpha_1(t-u+1/2)} \times \frac{1}{2} \left[ f(u) + f(u-1) \right] \right\},
\]

where \( f(u) = (\beta_1 u^{(y)} + \lambda_1) \).

Utilizing equation (13), the ultimate formula of \( \tilde{G}(t) \) (shown as equation (9)) can be generated. Based on the above, the time response function can be deduced as follows:
\[
\tilde{x}^{(y)}(t) = \tilde{G}(t)^{1/L}, \quad \tilde{x}^{(0)}(t) = \tilde{x}^{(y)}(t)^{-y}.
\]

So far, the proof has been completed.

For elaboration purposes, the pseudocode about the modeling procedure of the SFBGM model is exhibited in Algorithm 1.

After obtaining the time response function formula, it is essential to estimate the modeling parameters that directly determine the forecasting accuracy. Therefore, in the below subsection, the estimation procedures of the linear parameters are elaborated on. □
concluded that the self-adaptive parameters
From the above derivation procedure, it can be
2.2. Determination of the Self-Adaptive Parameters by the PSO
Optimal values of damping accumulation parameters
While \( t < n + h \) do
Output the forecasting results \( \hat{X}_1^{(0)} = [\hat{x}_1^{(0)}(1), \hat{x}_1^{(0)}(2), \ldots, \hat{x}_1^{(0)}(n + h)] \).

\[
\text{Input:} \quad \text{The original sequence } X_i^{(0)} \text{ and the forecasting horizon } h.
\]
\[
\text{Output:} \quad \text{The forecasting results } \hat{X}_1^{(0)} = [\hat{x}_1^{(0)}(1), \hat{x}_1^{(0)}(2), \ldots, \hat{x}_1^{(0)}(n + h)].
\]
Construct the B and Y by equations (17) to (39);
Calculate the linear parameters \( \varphi \) by equation (33);
Obtain the optimal values of damping accumulation parameters \( \gamma_1, \gamma_2, L \) using PSO;
Initialize \( \hat{X}_1^{(0)}(1) = x_1^{(0)}(1) \);
\[
\begin{align*}
\text{While} \quad t & < n + h \\
\text{do} & \\
\text{Compute the } \hat{x}_1^{(0)}(t) & \text{ using equations (11) to (14)}; \\
\text{end} & \\
\end{align*}
\]
Output the forecasting results \( \hat{X}_1^{(0)} = [\hat{x}_1^{(0)}(1), \hat{x}_1^{(0)}(2), \ldots, \hat{x}_1^{(0)}(n + h)] \).

\[ \begin{align*}
\text{where} \\
B &= \begin{bmatrix}
-z(\gamma_2)(2) & \frac{1}{2}(z(\gamma_2) - 1(\gamma_2)) \cdot [z(\gamma_2)(2)]^L & [z(\gamma_2)(2)]^L \\
-z(\gamma_3)(3) & \frac{1}{2}(z(\gamma_3) - 2(\gamma_3)) \cdot [z(\gamma_3)(3)]^L & [z(\gamma_3)(3)]^L \\
\vdots & \vdots & \vdots \\
-z(\gamma_n)(n) & \frac{1}{2}(z(\gamma_n) - (n - 1)(\gamma_n)) \cdot [z(\gamma_n)(n)]^L & [z(\gamma_n)(n)]^L \\
\end{bmatrix} \\
\text{and} \quad Y &= \begin{bmatrix}
x(\gamma_2)(2) - x(\gamma_2)(1) \\
x(\gamma_3)(3) - x(\gamma_3)(2) \\
\vdots \\
x(\gamma_n)(n) - x(\gamma_n)(n - 1) \\
\end{bmatrix}.
\end{align*} \]

\[
\text{2.2. Determination of the Self-Adaptive Parameters by the PSO Algorithm.} \quad \text{From the above derivation procedure, it can be concluded that the self-adaptive parameters } \gamma_1, \gamma_2, \text{ and } L, \text{ which have complex nonlinear relationships with predicting errors, play a decisive role in the prognosis performance of the novel model. To optimize the parameter values, Particle Swarm Optimization (PSO) is adopted to seek the most appropriate solutions [53]. In addition, the mean absolute percentage error (MAPE) is employed as the objective function [54], and the optimization problem could be illustrated mathematically as follows:}
\[
\text{min}_{\gamma_1, \gamma_2, L} \text{MAPE} = \frac{1}{n - 1} \sum_{k=2}^{n} \left| \frac{x(0)(k) - \hat{x}(0)(k)}{x(0)(k)} \right| \times 100\% ,
\]
\[
\begin{align*}
\gamma_1 & \in (0, 1), \gamma_2 \in (0, 1), \\
\varphi & = (\tilde{\alpha}, \tilde{\beta}, \tilde{\lambda})^T = (B^T B)^{-1} B^T Y, \\
\alpha_1 & = (1 - L)\alpha, \beta_1 = (1 - L)\beta, \lambda_1 = (1 - L)\lambda, \\
f(u) & = \beta_1 u(\gamma_2) + \lambda_1, \\
\hat{x}(\gamma_1)(t) & = \tilde{G}(t)^{1/\gamma_1} \hat{x}(0)(t) = \left[ \hat{x}(0)(t) \right]^{(-1)} , \\
\tilde{G}(t) & = \left[ x(0)(1) \right]^{1 - L} e^{-\alpha_1 (t-1)} + \sum_{u=2}^{t} e^{-\alpha_1 (t-u+1/2)} \times \frac{1}{2} \left[ f(u) + f(u - 1) \right].
\end{align*}
\]
PSO was put forward by Eberhart and Kennedy [55]. The inspiration for this algorithm originated from the foraging of birds. Currently, the PSO algorithm has been widely applied in multi-constrained optimization problems [56]. During the operation of this algorithm, each particle updates its velocity and direction according to its own experience, which includes the best position of itself (hbest) and other particles (fbest). Moreover, the objective function is called the fitness function, and the ultimate optimization effect is determined by some preset parameters, which are usually considered below. Specifically, population size I should not be too large, generally not more than 300. Parameters $s_1$ and $s_2$ are called learning factors, and their values are usually equal to 2. The parameter $\omega$ is called inertia weight, which has a value between 0 and 1. Because there are three independent variables, the search dimension should be set to 3. Generally, the max iteration is between 200 and 1000. The procedure of PSO is provided as follows.

**Step 1.** Input the numerical value of the preset parameters, where $s_1 = s_2 = 2$, $\omega = 0.3$, $I = 50$, and the iteration $k = 1$.

**Step 2.** Generate the initial population randomly. Each individual particle can obtain an initial position as $q_i = (q_{i1}, q_{i2}, q_{i3}), i = 1, 2, \ldots, I$ and an initial velocity as $v_i = (v_{i1}, v_{i2}, v_{i3}), i = 1, 2, \ldots, I$.

**Step 3.** Compute the fitness of each particle $\text{MAPE}_i = \text{MAPE}(q_i)$, and search for the best particle $h_{bk}$.

Then, set $h_{best}^k = q_{h_{bk}}^k$, and $f_{best}^k = d_{h_{bk}}^k$.

**Step 4.** Update the velocity and position of all particles through the following formula:

$$
\begin{align*}
    v_{ij}^{k+1} &= \omega \times v_{ij}^k + s_1 \times \text{rand} \times (h_{best}^k - v_{ij}^k) \\
    &+ s_2 \times \text{rand} \times (f_{best}^k - v_{ij}^k), \quad j = 1, 2, 3, \\
    q_{ij}^{k+1} &= q_{ij}^k + v_{ij}^{k+1}, \quad j = 1, 2, 3.
\end{align*}
$$

**Step 5.** Update the $h_{best}$ and the $f_{best}$.

If $\text{MAPE}^{k+1} < \text{MAPE}(q_i)$, one can get that $h_{best}^k = q_i^{k+1}$, else $h_{best}^k = q_{h_{bk}}^k$ and search for the best particle $h_{h_{bk}}$. In addition, if $\text{MAPE}^{k+1} < \text{MAPE}(q_i)$, one can obtain that $g_{best}^{k+1} = h_{best}^{k+1}$, else $g_{best}^{k+1} = g_{best}^k$.

**Step 6.** Repeat steps 3–5 until $k$ reaches the max iteration.

The process of the SFBGM model can be expressed as a flowchart shown in Figure 2.

### 3. Experimental Study

For verification purposes, a range of representative prediction methods are employed to compete with the presented model, and the structural design of Section 3 is as follows. Experiment design and the contrastive analysis of the results are shown in Section 3.1 and Section 3.2, respectively. Furthermore, the advantages and limitations of the proposed model are summarized in Section 3.3.

#### 3.1. Experiment Design

The traditional GM (1, 1) model can generate an accurate prediction for the sequence close to exponential growth, but restricted by the fixed structure, the conventional method may lose power when confronted with the datasets with nonlinear characteristics. Hence, considering that nonlinearity generally exists in practical problems, the novel fractional Bernoulli grey model is put forward. For validation purposes, the annual output of Photovoltaic modules in China [57] and the annual output of steel in Beijing [58] (denoted as Case 1 and Case 2, respectively) are implemented.

For Case 1, as revealed in Figure 3(a), the annual output of PV modules in China has increased from 21 GW in 2011 to 124.6 GW in 2020. In recent years, the growth rate has been significantly accelerated, which is closely related to the vigorous promotion of clean energy.

After 2016, the raw data shows significant uncertainty. Overall, the original datasets of Case 1 and Case 2 show an increasing and decreasing trend, respectively. However, the nonnegligibly nonlinear variations, which seriously interfere with the prediction accuracy, are hidden in the trends. Thus, the SFBGM model is chosen to address the challenging issues.

For validation purposes, the classic GM (1, 1) model, Long Short-Term Memory (LSTM), and Autoregressive integrated moving average model (ARIMA) serve as the benchmarks [59]. To check out the robustness of the presented model, two different forecasting horizons are arranged in both Case 1 and Case 2. The designation is presented in Table 1. Specifically, take the Case 1 as an
Figure 2: The procedure of the SFBGM model.

Figure 3: The annual output of photovoltaic modules in China (a) and the annual output of steel in Beijing (b).
example, for the two-year horizon predicting, the original data from 2010 to 2018 are served as the training set, while those from 2019 to 2020 are regarded as the testing set. Moreover, for the one-year forecasting, the raw data from 2010 to 2019 are served as the training set, while the data of 2020 is served as the testing set.

3.2. The Comparative Analysis of the Experimental Results. The MAPE, RMSE, and MAE in the training period and testing period generated by the four models under two different forecasting horizons are listed in Tables 2 and 3, respectively. Meanwhile, the results in Tables 2 and 3 are shown in Figure 4 in the form of a histogram intuitively.

As observed from Tables 2, 3, and Figure 4, several conclusions can be drawn as follows:

(1) The performance of the SFBGM model is excellent in both two cases. Initially, for Case 1, since the MAPE values of the two grey models (including the in-sample and out-of-sample periods) are less than 11%, and the RMSE yielded by them are no more than 12, thus, it can be demonstrated that the two grey methods have achieved high-precision prediction. Thereinto, the MAPE and RMSE values incurred by the SFBGM model are the second lowest among the two forecasting horizons in the training stage, next only to the GM (1, 1) model. Moreover, the proposed model generated the most accurate forecasting performance with a two-year horizon and the second only to the traditional GM model with a one-year horizon. From the above, one can obtain that the grey models have the excellent fitting ability (both in-sample stage and out-of-sample stage) for exponential growth sequence. Subsequently, by investigating Case 2, it can be concluded from Table 2 that the MAPE values incurred by the novel model are the lowest, while the classic GM (1, 1) model lost power in the testing stage. Therefore, it has been verified that the SFBGM model has more robust adaptability and generalizability. Further, according to the stable performance of the model for different prediction horizons, the excellent robustness of the proposed model can be realized.

(2) The conventional GM (1, 1) model may be unsuitable for forecasting the time series with irregular fluctuations. In particular, the overall performance of the GM model in Case 1 is the best among all models, while the errors incurred by it in the testing period of Case 2 are unacceptable. As shown in Table 3, the predicting RMSE values both exceed 35. In general, the conventional model’s linear structure can only accurately describe the steady sequences, which blocks its application [60].

(3) As for the two non-grey benchmarks, LSTM and ARIMA, Figure 3 shows that the fitting accuracy of the LSTM model is almost the worst in these two cases. Primarily, MAPE values yielded by LSTM even exceed 40% in the in-sample period with a one-year horizon. Moreover, the prediction performance of the LSTM method is also poor (shown in Figure 3). As one of the econometric methods, the ARIMA method may be unsuitable for predicting the above sequences. It can be demonstrated from Tables 1 and 2 that, whether it is from the perspective of MAPE or RMSE, the error generated by the ARIMA model is relatively large compared with that of other models. The reason for the above phenomenon can be attributed to the inadequate datasets. Specifically, in the circumstances of insufficient data, the parameters of the LSTM method cannot be fully trained, and the assumption of the ARIMA method can be satisfied, which leads to poor performance.

3.3. The Advantages and Limitations of the SFBGM Model. As demonstrated by the two cases, the SFBGM model has achieved outstanding predicting performances in the annual output of Photovoltaic modules and steel in China and Beijing, respectively. The reasons for the excellent characteristics of the novel model are summarized as follows:

(1) Due to the principle of new information priority, the $\gamma$-order fractional-order accumulation generation operator ($\gamma$-FOA) replaces the first-order accumulation generating operation (1-AGO) to preprocess the raw sequences. This improved accumulation generation operator could raise the weight of new, thereby enhancing the generalizability of the presented model.

(2) The self-adaptive parameters $\gamma_1$, $\gamma_2$, and $L$ are constructed to play a pivotal role to cater the nonlinear features hidden in the original time series. The outstanding adaptability of the novel model, which is shown in the experiment, has proved the effectiveness of these self-adaptive parameters. Besides, the time function term $(\beta^{\gamma_2} + \lambda)$ is designed to describe the accumulative time effect, which significantly strengthens the robustness of the presented model.
Table 2: Error comparisons of the competing models with different horizons in Case 1.

<table>
<thead>
<tr>
<th>Models indicator</th>
<th>Horizon</th>
<th>GM (1,1)</th>
<th>SFBGM</th>
<th>LSTM</th>
<th>ARIMA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Two years</td>
<td>3.137</td>
<td>3.841</td>
<td>37.041</td>
<td>15.480</td>
</tr>
<tr>
<td>Training period</td>
<td>One year</td>
<td>5.085</td>
<td>5.710</td>
<td>42.157</td>
<td>15.014</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>2.410</td>
<td>2.926</td>
<td>13.327</td>
<td>7.619</td>
</tr>
<tr>
<td></td>
<td>One year</td>
<td>3.268</td>
<td>3.662</td>
<td>15.294</td>
<td>7.525</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>1.741</td>
<td>1.987</td>
<td>13.109</td>
<td>5.752</td>
</tr>
<tr>
<td></td>
<td>One year</td>
<td>2.668</td>
<td>2.742</td>
<td>14.897</td>
<td>5.763</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>6.473</td>
<td>8.638</td>
<td></td>
<td>14.448</td>
</tr>
<tr>
<td>Testing period</td>
<td>One year</td>
<td>1.642</td>
<td>3.844</td>
<td>13.340</td>
<td>15.677</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>5.710</td>
<td>12.741</td>
<td>19.820</td>
<td></td>
</tr>
<tr>
<td></td>
<td>One year</td>
<td>4.789</td>
<td>16.622</td>
<td>19.533</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>1.987</td>
<td>10.372</td>
<td>17.077</td>
<td></td>
</tr>
<tr>
<td></td>
<td>One year</td>
<td>2.742</td>
<td>16.622</td>
<td>19.533</td>
<td></td>
</tr>
</tbody>
</table>

Note. The optimum values of the above metrics are in bold type.

Table 3: Error comparisons of the competing models with different horizons in Case 2.

<table>
<thead>
<tr>
<th>Models indicator</th>
<th>Horizon</th>
<th>GM (1,1)</th>
<th>SFBGM</th>
<th>LSTM</th>
<th>ARIMA</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Two years</td>
<td>7.533</td>
<td>5.322</td>
<td>9.949</td>
<td>7.908</td>
</tr>
<tr>
<td>Training period</td>
<td>One year</td>
<td>7.427</td>
<td>4.743</td>
<td>5.693</td>
<td>7.170</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>15.596</td>
<td>14.202</td>
<td>25.496</td>
<td>18.285</td>
</tr>
<tr>
<td></td>
<td>One year</td>
<td>15.915</td>
<td>12.590</td>
<td>12.093</td>
<td>17.261</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>14.161</td>
<td>10.645</td>
<td>20.883</td>
<td>16.818</td>
</tr>
<tr>
<td></td>
<td>One year</td>
<td>13.865</td>
<td>9.415</td>
<td>11.056</td>
<td>15.190</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>18.952</td>
<td>5.332</td>
<td>9.735</td>
<td>5.354</td>
</tr>
<tr>
<td>Testing period</td>
<td>One year</td>
<td>19.811</td>
<td>8.303</td>
<td>11.890</td>
<td>10.785</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>35.970</td>
<td>7.566</td>
<td>18.982</td>
<td>11.900</td>
</tr>
<tr>
<td></td>
<td>One year</td>
<td>36.534</td>
<td>15.317</td>
<td>21.928</td>
<td>19.890</td>
</tr>
<tr>
<td></td>
<td>Two years</td>
<td>34.035</td>
<td>7.296</td>
<td>17.525</td>
<td>9.751</td>
</tr>
<tr>
<td></td>
<td>One year</td>
<td>36.534</td>
<td>15.317</td>
<td>21.928</td>
<td>19.890</td>
</tr>
</tbody>
</table>

Note. The optimum values of the above metrics are in bold type.

Figure 4: Continued.
Nevertheless, the SFBGM model is not flawless and has some limitations. For instance, compared with the machine learning method, the proposed model can hardly identify the mutation of data effectively. Although the limitations may block the extensions, the proposed model has been validated to have strong robustness, broad adaptability, and outstanding generalizability, which enriches the grey system theory.

4. Conclusion and Future Work

In view of the nonlinearity in the original sequences, a novel Bernoulli grey model, namely, the SFBGM model, is established based on the traditional NGBM (1, 1) model. Compared with the conventional NGBM (1, 1) model, the first-order accumulation is replaced by the $\gamma$-fractional order accumulation, and the time function term $(\dot{y}^\gamma + \lambda)$ is served as the grey action quantity. In the comparison of the traditional time series prediction method, two main improvements can be summarized as follows: (1) the self-adaptive parameters optimized by PSO can substantially strengthen the model’s flexibility and (2) the time function introduces the accumulative time effect into the formula, significantly improving the stability of the proposed model. As a result of the above, the predicting performances generated by the SFBGM model are more reliable.

For validation purposes, the novel model is utilized to predict the annual output of Photovoltaic modules and steel in China and Beijing, respectively. Subsequently, it can be discovered that the SFBGM model presents the best performance in comparison with the benchmarks. SX_he data sources are provided within the article.

Based on the aforementioned, the future work can be suggested: (1) the time function term $(\dot{y}^\gamma + \lambda)$ incorporated in the SFBGM model can be modified in the future for more substantial capacity to identify mutations; (2) this study focuses on the case with rare data, and the following study can investigate the application in the circumstances of sufficient data. In general, the SFBGM model is worth further research from both theoretical and practical perspectives.

Data Availability

The data sources are provided within the article.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

The study was supported by the National Natural Science Foundation of China (no. 71901191).

References


