Energy 262 (2023) 125439

Contents lists available at ScienceDirect

Energy

journal homepage: www.elsevier.com/locate/energy

Prediction of electricity consumption based on GM(1,N_r) model in Jiangsu province, China

Xiaoyi Du [a,*], Dongdong Wu [b], Yabo Yan [a]

[a] School of Business, Jiangnan University, Wuxi, Jiangsu Province, 214122, PR China
[b] College of Tourism and Service Management, Nankai University, Tianjin, 300350, PR China

ARTICLE INFO

Keywords:
Electricity consumptions forecasting
Grey forecasting model
GM(1,N_r) model
Grey incidence analysis
Particle swarm optimization algorithm

ABSTRACT

In this paper, GM(1,N_r) model is established to improve the traditional GM(1,N) model from three aspects: (1) transforming the original sequence to satisfy the modeling conditions with particle swarm optimization algorithm; (2) introducing grey incidence analysis to obtain the grey incidence ranking and carrying out stepwise test for significant variables to determine the number of variables; and (3) predicting the related factor sequence through the improved GM(1,1) model. Empirical analysis shows that the proposed GM(1,N_r) model has remarkable good prediction performance compared with the traditional grey forecasting model. It is also demonstrated that the extraction of influencing factors can significantly improve the prediction effectiveness, especially when pursuing the best fitting effect on small sample data. The findings indicate that the electricity consumptions of Jiangsu Province in the next several years will be at a high level and keep rising, with a predicted value of 9712.48 billion kilowatt-hours in 2030. The findings can help the government and energy related institutions to develop management policies on energy demand, and the proposed model can also be extended for the application in other regions.

1. Introduction

Energy plays an important role in promoting social and economic development but it is not easy to store [1,2]. Oversupply of energy will lead to the increase of operating costs, while short supply will have a negative impact on both economic development and social livelihood. Therefore, accurate assessment of energy consumptions and balance economic and social benefits with a foresight perspective are the key issues for energy productions and consumptions [3,4]. Among all kinds of energy, the electricity is the secondary energy obtained after the conversion from primary energy processing. Energy as one of the basic resources and electricity industry as the basic industry have a close relationship with the national economy and peoples' livelihood. Hence, it is of great significance to make the development of the specific electricity consumption clear, especially in the adjustment period of energy consumption structure.

Abundant studies have been carried out by scholars on energy forecasting, especially electricity consumption forecasting [5–8]. Existing studies mainly focus on the identification of the factors affecting energy consumptions and the prediction of energy consumptions. For the research stream of the identification of influence factors, scholars believed that energy consumption is closely related to national economy and peoples' livelihood, and of course electricity is not an exception [2,9]. Ali et al. found that higher household income level is related to the more consumption of clean energy resources [10]. Jian et al. believed that education, security, and social globalization have a negative impact on energy consumption and long-term CO_2 emissions [11]. Murshed et al. found that economic growth and household consumption expenditure show a positive effect on the primary energy and electricity demand [12]. Evidence from the work of Liddle et al. showed that population, age structure, household size, urbanization, and population density affect the way of carbon emissions and energy consumptions [13]. The study of Wang et al. showed that population, urbanization level, living standard, energy consumption structure, and energy use efficiency are the direct factors affecting energy demand changes, among which economic development level is the core factor [14]. Yang et al. concluded that the factors affecting regional energy efficiency include socio-economic factors (such as energy consumption composition, industrial structure, per capita GDP, etc.) and environmental factors (such as geographical location, topographical conditions, weather, sunshine duration, etc.) [15]. Meng et al. constructed an index system of influencing factors of energy efficiency in China, which contains

* Corresponding author.
E-mail addresses: dxy_jnu@163.com (X. Du), dwu@mail.nankai.edu.cn (D. Wu), 476182176@qq.com (Y. Yan).

https://doi.org/10.1016/j.energy.2022.125439
Received 11 January 2022; Received in revised form 27 August 2022; Accepted 10 September 2022
Available online 14 September 2022

industrial structure, technological progress, energy consumption structure, economic development level, marketization degree and openness degree [16]. Meangbua et al. showed that socio-economic and demographic factors can influence energy changes, and temperature and education are considered to be key driving factors [17]. Besides, population, technological progress, secondary industry, fixed asset investment, and openness may promote the increase of energy consumption, of which population is the most important factor affecting energy consumption [18].

For the research stream of electricity consumption prediction, several models are commonly used, such as multiple linear regression method [19–22], time series analysis method [23], intelligent algorithm [24] and grey system method [25]. Kim et al. used SARIMA, SARIMA + GARCH, Holt-Winters method and ARIMA with Fourier transformation models to forecast the electricity consumption of an industrial manufacturing building [26]. Tang et al. established binary nonlinear fitting regression and support vector regression models to predict the electricity consumption of urban rail transit [27]. Zolfaghari and Golabi offered a hybrid model which combines adaptive wavelet transform, long short-term memory (LSTM) and random forest algorithm to predict the EP in hydroelectric power plant [28]. Considering the influence of temperature, De et al. predicted daily electricity consumption in Italy through statistical modeling [29]. To simulate the effects of seasons and trends, Hamzacebi et al. developed four different artificial neural network models and chose the better one to predict monthly electricity consumption in Turkey [30]. Lin et al. proposed a short-term time--phased electricity consumption prediction model based on LSTM with an attention mechanism [31]. Ercan et al. proposed a model that uses artificial neural networks as a machine learning method to predicts the electricity consumption levels in dwellings as lower consumption and higher consumption classes [32]. Wei et al. constructed a comprehensive time series prediction model based on component data [24]. Rana et al. studied the feature selection of a neural network-based method for calculating prediction intervals, and then evaluated the half-hour electricity demand data in Australia and the UK [33]. Chen et al. proposed a novel data-driven framework to predict the annual household electricity consumption using ensemble learning technique [34].

The grey forecasting model proposed by Prof. Deng Julong shows its advantages that the model only needs 4 data at least and the calculation is relatively simple and quick. Therefore, grey forecasting model is gradually adopted by scholars [35]. According to different time range, forecasting can be divided into short-term, medium-term, and long-term forecast. Grey forecasting is more suitable for short-term and medium-term forecast. Due to the defects of the traditional model in terms of the background value and initial value, scholars have proposed several improvement methods to overcome those problems and achieved remarkable achievements [36–38]. Traditional GM(1,1) model has been modified by using the Markov chain method to forecast the coal consumptions [39]. Feng et al. adopted GM(1,1) model to forecast the total energy, coal energy and clean energy consumptions in China [40]. Genetic-algorithm-based remnant GM(1,1) model are proposed and demonstrated to be superior than any other GM(1,1) variants [41]. Wang and Song proposed a new nonlinear dynamic grey model, namely NMGM(1,1,alpha), to predict the oil consumptions in China accurately [42]. Besides, a new residual GM(1,1) model using neural networks has been proposed [43]. A new seasonal cycle GM(1,1) has been proposed to forecast the railway passenger volume [44].

Based on data conversion of original data series and the optimization of background values, Li and Zhang proposed the TBGM(1,1) model to predict the total energy consumption of Shanghai, China [45]. GM(1,1) model, NGBM(1,1) model and grey Verhulst model have been compared for theoretical derivation and scientific verification, and then those models were adopted to predict the growth trend of renewable energy consumption in China [46]. Hamzacebi et al. proposed a seasonal GM(1,1) model to solve the seasonal problem, and then predicted the monthly electricity demand in Turkey [47]. Applying the upper and lower limit sequences to construct the GM(1,1) model after residual correction, Hu et al. constructed a nonlinear interval grey prediction model [48]. Tsai et al. used the traditional GM(1,1) model and nonlinear grey Bernoulli model to carry out theoretical derivation and verification [49]. Xu et al. proposed a new grey model, namely IRGM(1,1) model, with optimal time response function, and was adopted to forecast electricity consumption [50]. Wu et al. established a multivariate grey forecasting model considering the total population to predict the electricity consumption [51].

There is no doubt that the multivariate linear regression method, intelligent algorithms, and time series analysis method require a large quantity of data. In addition, the time cost for calculation of intelligent algorithms is high. Instead, the grey forecasting model has better adaptability for the small data problem, which shows its advantages in modeling forecasting model when sufficient data is hard to come by. Most of the existing studies adopt grey univariate prediction models, however those kinds of models have the defects that only contain the system characteristic data sequence and ignore the influence of the system. Grey multivariable model GM(1,N) can be used to overcome the shortcomings by considering the influence of factors related to the sequence of system characteristic from a systematic perspective. Previous studies also confirmed that the prediction effect of GM(1,N) model is usually better than GM(1,1) model.

However, there are some structural defects for the traditional GM(1,N) model. Firstly, the sequences for GM(1,N) model need to satisfy both the non-negative and quasi-smooth conditions, which hinders the widely application of GM(1,N) model. In this paper, two constants are introduced to transform the original sequence to satisfy the basic modeling conditions and the particle swarm optimization (PSO) algorithm is used to determine the optimal value. Secondly, as the existence of some unnecessary factors, the prediction accuracy of the traditional GM(1,N) model may not be high. In this paper, grey incidence ranking, obtained by grey incidence analysis, is subsequently used for the stepwise test of significant variables, so as to determine the best "N" value in the traditional GM(1,N) model. Thirdly, the GM(1,1) model is requisite when constructing the traditional GM(1,N) model, but the deficiencies of the GM(1,1) itself may lead to large prediction errors. Therefore, the improved GM(1,1) model, metabolism GM(1,1) model will be introduced to the traditional GM(1,N) model to improve the fitting precision.

The remainder of this paper is structured as follows. Section 2 illustrates the construction of the proposed GM $(1,N_r)$ model and also the way of prediction effect verification. Section 3 carries out empirical analysis and forecasts the electricity consumptions in Jiangsu Province. Section 4 provides the conclusion and policy suggestions.

2. Methodology

2.1. The construction of GM $(1,N_r)$ model

This paper improves the traditional GM(1,N) model from three aspects: (1) transforming the original sequence to satisfy the modeling conditions with particle swarm optimization algorithm; (2) introducing grey incidence analysis to obtain the grey incidence ranking and carry out stepwise test for significant variables to determine the number of variables; and (3) predicting the related factor sequence through metabolism GM(1,1) model. We name the improved GM(1,N) model as GM$(1,N_r)$ model. Fig. 1 shows the predicting procedures.

Definition 1. To obtain the original data points, we assume the non-negative original sequences as$X_i^{(0)} = (x_i^{(0)}(1), x_i^{(0)}(2), ..., x_i^{(0)}(n))$, $i = 1, 2, ..., N_r$, where $X_1^{(0)}$denotes the data of the system behavior variable. The driving variables or relative factors are denoted as $X_2^{(0)}, X_3^{(0)}, ..., X_{N_r}^{(0)}$, the 1-AGO sequence of $X_i^{(0)}$is represented by $X_i^{(1)} = (X_i^{(1)}(1), X_i^{(1)}(2), ..., X_i^{(1)}(n))$, where $X_i^{(1)}(k) = \sum_{i=1}^{k} X_i^{(0)}(i), k = 1, 2, ..., n$ and the background

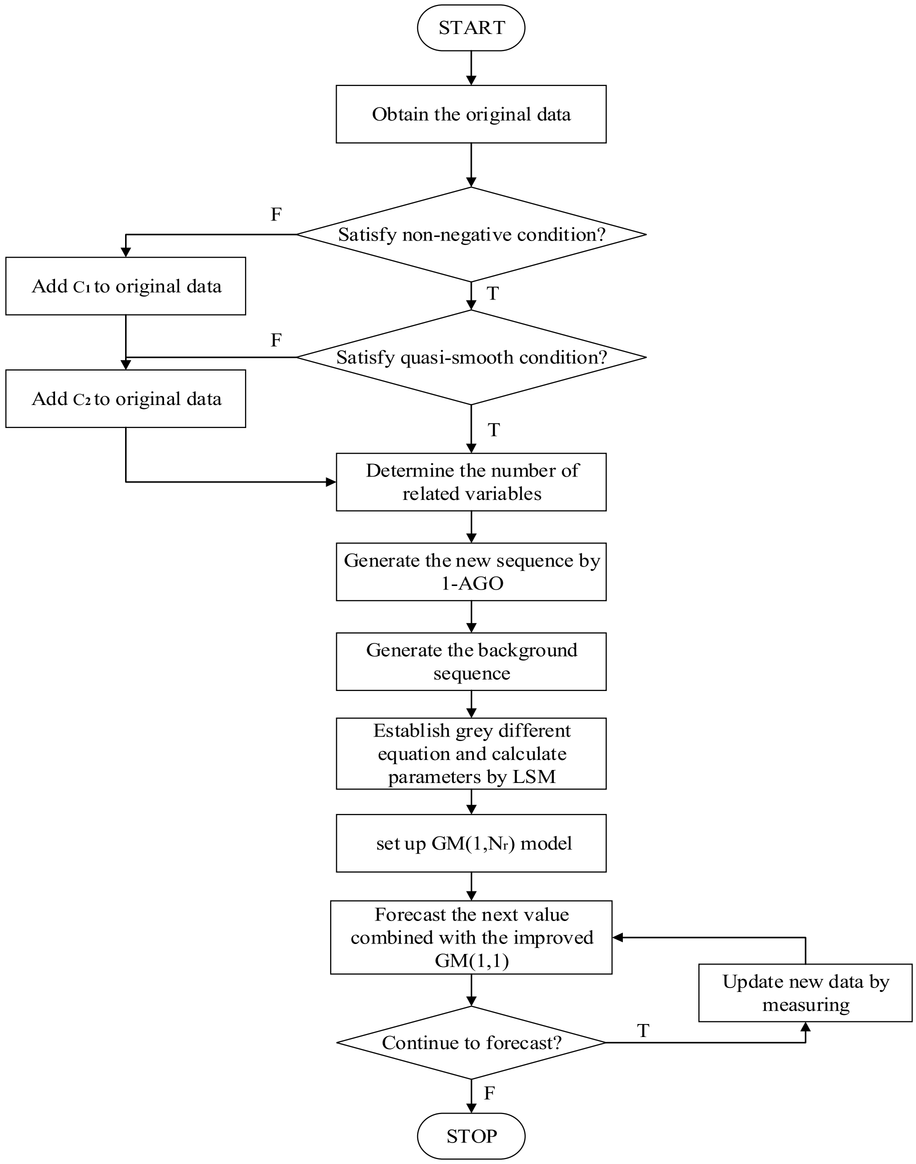

Fig. 1. The procedure of predicting progress using GM(1,N_r) model.

value is $z_1^{(1)}$. Set up GM(1,N_r) model as $x_1^{(0)}(k) + az_1^{(1)}(k) = \sum_{i=2}^{N_r} b_i x_i^{(1)}(k)$, and its whiten differential equation is $\frac{dx_1^{(1)}(t)}{dt} + az_1^{(1)}(t) = \sum_{i=2}^{N_r} b_i x_i^{(1)}(t)$.

The traditional grey multivariable model requires the sequence to satisfy non-negative and quasi-smooth conditions [52], so GM(1,N) model cannot be established if the sequence does not satisfy the conditions. In this paper, the initial state of the original sequence is transformed by introducing two constants to satisfy the conditions and the applicability of the grey multivariable model can be expanded.

Theorem 1. *Assume that the initial state of the original sequence is $X_{iraw}^{(0)}$, where $X_{1raw}^{(0)}$ denotes the data of the system behavior and $X_{iraw}^{(0)}(i=2,3,\dots,N)$ denotes the relative factors, then the original sequence transformation is expressed by*

$$X_i^{(0)} = \left(x_i^{(0)}(1), x_i^{(0)}(2), \dots, x_i^{(0)}(n)\right) \geq 0, n \geq 4 \tag{1}$$

$$X_i^{(0)} = X_{iraw}^{(0)} + c_1 + c_2, i = 1, 2, \dots, N_r$$

where c_1 and c_2 are named as the first and second non-negative additive factors, which meet the following conditions respectively:

$$c_1 = \begin{cases} 0 & if: \forall k \in (1,2,\dots,n), x_{iraw}(k) \geq 0 \\ \max_k\{|x_{iraw}(k)|\} & if: \exists k \in (1,2,\dots,n), x_{iraw}(k) < 0, i = 1,2,\dots,N_r \end{cases} \tag{2}$$

$$c_2 = \begin{cases} 0 \quad if: \text{Equation (12) issatisfied} \\ \begin{cases} c_2 \geq 0 \\ c_2 \geq \max\left\{\dfrac{2(x_{iraw}(k)+c_1) - \sum_{i=1}^{k-1}(x_{iraw}(i)+c_1)}{k-3}\right\} \end{cases} \quad if: \text{Equation (12) is unsatisfied}, \; i = 1, 2, ..., N_r \end{cases} \tag{3}$$

Theorem 2. *For any data sequence, there always exists two non-negative additive factors c_1 and c_2, which can be added to this sequence to satisfy the non-negative and the quasi-smooth conditions* [53].

Proof. *Considering a prediction problem, the initial state of the original sequence is expressed as*:

$$X_{1raw}^{(0)} = \left(x_{1raw}^{(0)}(1), x_{1raw}^{(0)}(2), ..., x_{1raw}^{(0)}(n)\right), n \geq 4$$

If there exists a negative number in the above sequence, a non-negative number c_1 is easy to find using Equation (5). Through the transformation of Equation (4), the initial state of the original sequence can be changed into a non-negative sequence.

$$X_{1T}^{(0)} = \left(x_{1T}^{(0)}(1), x_{1T}^{(0)}(2), ..., x_{1T}^{(0)}(n)\right) = \left(x_{1raw}^{(0)}(1)+c_1, x_{1raw}^{(0)}(2)+c_1, ..., x_{1raw}^{(0)}(n)+c_1\right)$$
$$x_{1raw}^{(0)}(k)+c_1 \geq 0, k=1,2,...,n \tag{4}$$

$$c_1 = \max\left\{\left|x_{1raw}^{(0)}(1)\right|, \left|x_{1raw}^{(0)}(2)\right|, ..., \left|x_{1raw}^{(0)}(n)\right|\right\} \tag{5}$$

By Equation (4), the new sequence $X_{1T}^{(0)}$ satisfies the nonnegative condition, and then generate its 1-AGO sequence as below:

$$X_{1T}^{(1)} = \left(x_{1T}^{(1)}(1), x_{1T}^{(1)}(2), ..., x_{1T}^{(1)}(n)\right) \tag{6}$$

$$x_{1T}^{(1)}(k) = \sum_{i=1}^{k} x_{1T}^{(0)}(i), k = 1, 2, ..., n \tag{7}$$

Grey forecasting modeling requires $X_{1T}^{(0)}$ and $X_{1T}^{(1)}$ to satisfy the quasi-smooth condition. If the requirement is not satisfied, a non-negative number c_2 is added to the sequence through Equation (9), then the final state of the original sequence is presented as Equation (8):

$$X_1^{(0)} = \left(x_1^{(0)}(1), x_1^{(0)}(2), ..., x_1^{(0)}(n)\right) \tag{8}$$

$$x_1^{(0)}(k) = x_{1T}^{(0)}(k) + c_2, k = 1, 2, ..., n; \; c_2 \geq 0 \tag{9}$$

$$X_1^{(1)} = \left(x_1^{(1)}(1), x_1^{(1)}(2), ..., x_1^{(1)}(n)\right) \tag{10}$$

$$x_1^{(1)}(k) = \sum_{i=1}^{k} x_1^{(0)}(i), k = 1, 2, ..., n \tag{11}$$

From Equation (9), it is obvious that $x_1^{(0)}(k) \geq x_{1T}^{(0)}(k), k = 1, 2, ..., n$. Therefore, we can obtain $x_1^{(0)}(k) \geq 0, k = 1, 2, ..., n$, which satisfies non-negative conditions. Then, the necessary condition of c_2 are deduced according to the quasi-smooth condition:

$$0 \leq \rho(k) = \frac{x_1^{(0)}(k)}{x_1^{(1)}(k-1)} \leq 0.5, k = 4, 5, ..., n \tag{12}$$

Replace Equation (11) into Equation (12), then we can get

$$0 \leq \frac{x_{1T}^{(0)}(k) + c_2}{\sum_{i=1}^{k-1}\left(x_{1T}^{(0)}(i) + c_2\right)} \leq \frac{1}{2}$$

$$0 \leq \frac{x_{1T}^{(0)}(k) + c_2}{\sum_{i=1}^{k-1}\left(x_{1T}^{(0)}(i)\right) + (k-1)c_2} \leq \frac{1}{2}$$

$$0 \leq \frac{x_{1T}^{(0)}(k) + c_2}{x_{1T}^{(1)}(k) + (k-1)c_2} \leq \frac{1}{2}$$

Thus,

$$\begin{cases} c_2 \geq 0 \\ c_2 \geq \max\left\{\dfrac{2x_{1T}^{(0)}(k) - x_{1T}^{(1)}(k-1)}{k-3}\right\} \end{cases}, k = 4, 5, ..., n \tag{13}$$

Therefore, any data sequence can be transformed and then be used to establish the grey forecasting model by selecting c_2 to satisfy Equation (13).

Although we can get a necessary condition of c_2 by Equation (13), the variation of its specific value may result in different prediction errors. To minimize the prediction errors, we adopt the nonlinear programming method to determine the best value of c_2.

Letting minimizing the average relative errors of the model as the objective, the known conditions such as model parameters and time response are modelled as the constraints. Then, the nonlinear programming problem can be constructed as follows:

$$\min_{c_2} avg(e(k)) = \frac{1}{n-1}\sum_{k=2}^{n}\frac{\left|\widehat{x}_1^{(0)}(k) - x_1^{(0)}(k)\right|}{x_1^{(0)}(k)}, i = 2, 3, ..., N$$
$$s.t. \begin{cases} x_1^{(0)}\left(k\right) = -az_1^{(1)}\left(k\right) + \sum_{i=2}^{N} b_i\left(x_i^{(1)}\left(k\right)\right) \\ \text{Equation (2)} \\ \text{Equation (3)} \end{cases} \tag{14}$$

The advantage of PSO algorithm is that it can find the global optimal solution more easily. Practice has proved that compared with other traditional optimization algorithms, the solution speed of PSO algorithm is faster with fewer parameter adjustments, and it has a great advantage in the global search for the optimal solution. Therefore, PSO algorithm is used to solve Equation (14) and the optimal value can be obtained.

Grey incidence analysis is generally used to determine the number of variables related to the system variables, but its defect lies in that the inclusion of certain influencing factors may also lead to the accuracy reduction of grey forecasting model. Therefore, a new technique was proposed in this paper to determine the optimal number of related variables, which grey incidence degree is combined with fitting and prediction results.

Specifically, the procedures are summarized as follows. Firstly, obtain the grey incidence degree between the system characteristic sequence and the related factor sequences using grey incidence analysis, and then grey incidence ranking can be obtained. Secondly, stepwise test is carried out. According to the grey incidence ranking from high to low, we can select one, two, …, N numbers of related factors each time to establish GM(1,N) model, and then the fitting and prediction effect can be calculated. Finally, the optimal "N" value, called "N_r", is determined, which derives from the model with best performance.

Theorem 3. *X_1 and $X_i (i = 2,3,...,N)$ are the indicators as assumed before, the grey incidence degree between the system characteristic sequence and the related factor sequences is calculated as*

$$R_i = \frac{1}{n}\sum_{t=1}^{n}\xi_i(t) \tag{15}$$

where $\xi_i(t) = \frac{\min_i\min_t|x_0(t)-x_i(t)|+\rho\max_i\max_t|x_0(t)-x_i(t)|}{|x_0(t)-x_i(t)|+\rho\max_i\max_t|x_0(t)-x_i(t)|}$, is called grey incidence coefficient, $0 < \rho < 1$. In general, $\rho = 0.5$.

Theorem 4. *$X_i^{(0)}, X_i^{(1)} (i = 2,3,...,N)$, $Z_1^{(1)}$ are the indicators as assumed before, the matrix B, Y_{N_r} and the parameter matrix P_{N_r} can be expressed as*

$$B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & \cdots & x_{N_r}^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & \cdots & x_{N_r}^{(1)}(3) \\ \vdots & \vdots & & \vdots \\ -z_1^{(1)}(n) & x_2^{(1)}(n) & \cdots & x_{N_r}^{(1)}(n) \end{bmatrix}, Y_{N_r} = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(n) \end{bmatrix}, P_{N_r} = \begin{bmatrix} a \\ b_2 \\ \vdots \\ b_{N_r} \end{bmatrix}$$

Then, the model parameters are estimated as follows:

$$\begin{cases} P_{N_r} = B^{-1}Y_{N_r}, |B| \neq 0 \ if \ n = N_r + 1 \\ P_{N_r} = \left(B^TB\right)^{-1}B^TY_{N_r} \ if \ n > N_r + 1 \\ P_{N_r} = B^T\left(BB^T\right)^{-1}Y_{N_r} \ if \ n < N_r + 1 \end{cases} \tag{16}$$

Theorem 5. *The matrix B, Y_{N_r} and P_{N_r} are the indicators as assumed before, the predicted values of $X_2^{(0)}, X_3^{(0)}, ..., X_N^{(0)}$ are obtained from the metabolism GM(1,1) model [54], which is based on the principle of new information priority. Then, the fitted and predicted values can be deduced by*

$$x_1^{(0)}(k) = -ax_1^{(1)}(k-1) + \beta_2x_2^{(1)}(k) + \cdots + \beta_{N_r}x_{N_r}^{(1)}(k), k = 2,3,...,n \tag{17}$$

where $\beta_i = \frac{b_i}{1+0.5a}$, $\alpha = \frac{a}{1+0.5a}$

Proof. According to $z_1^{(1)}(k) = 0.5x_1^{(1)}(k) + 0.5x_1^{(1)}(k-1)$, $x_1^{(1)}(k) = x_1^{(1)}(k-1) + x_1^{(0)}(k)$, the following formulas can be obtained:

$$z_1^{(1)}(k) = x_1^{(1)}(k-1) + 0.5x_1^{(0)}(k)$$

$$x_1^{(0)}(k) + a\left(x_1^{(1)}(k-1) + 0.5x_1^{(0)}(k)\right) = \sum_{i=2}^{N} b_ix_i^{(1)}(k)$$

$$(1+0.5a)x_1^{(0)}(k) = \sum_{i=2}^{N} b_ix_i^{(1)}(k) - ax_1^{(1)}(k-1)$$

$$x_1^{(0)}(k) = \sum_{i=2}^{N}\frac{b_i}{1+0.5a}x_i^{(1)}(k) - \frac{a}{1+0.5a}x_1^{(1)}(k-1)$$

Suppose that $\beta_i = \frac{b_i}{1+0.5a}, i = 2,3,...,N_r$; $\alpha = \frac{a}{1+0.5a}$, then:

$$x_1^{(0)}(k) = -ax_1^{(1)}(k-1) + \beta_2x_2^{(1)}(k) + \cdots + \beta_{N_r}x_{N_r}^{(1)}(k), k = 2,3,...,n$$

Based on the above analysis, the prediction procedure of GM(1,N_r) model is summarized as follows:

Step1: Check whether the original sequences satisfy both the non-negative and quasi-smooth conditions. If not, transforming the original sequences by Equation (1) to add the non-negative addictive factors, and then use particle swarm optimization algorithm to obtain the parameter value. Instead, go to the second step.

Step2: Grey incidence analysis is carried out by Equation (15) and the number of variables is determined by the grey incidence ranking of the sequence and the result of fitting and prediction. Then, "N_r" is determined to construct the GM(1,N_r) model.

Step3: Solving the parameter matrix P_{N_r} of the model by Equation (16) and calculating the fitted and predicted values for system characteristic data by Equation (17).

2.2. *The verification of the proposed model*

We utilize seven evaluation ways to measure the predictive performance of the alternative models [55,56], including absolute percentage error (APE), mean absolute percentage error (MAPE), fitting degree (FD), mean absolute error (MPE), root mean square error (RMSE), directional statistics (DS), and Diebold-Mariano (DM) test. The formulations are described as follows:

$$APE = \frac{|\widehat{x}(k) - x(k)|}{x(k)} \times 100\% \tag{18}$$

$$MAPE = \frac{1}{n}\sum_{k=1}^{n}\frac{|\widehat{x}(k) - x(k)|}{x(k)} \times 100\% \tag{19}$$

$$FD = 100 - MAPE \tag{20}$$

$$MAE = \frac{1}{n}\sum_{k=1}^{n}|\widehat{x}(k) - x(k)| \times 100\% \tag{21}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{k=1}^{n}(|\widehat{x}(k) - x(k)|)^2} \tag{22}$$

$$DS = \frac{1}{n}\sum\nolimits_{k=1}^{n}F_k \times 100 \tag{23}$$

$$DM = \frac{\overline{d}}{\sqrt{Var(\overline{d})}} \tag{24}$$

where $\widehat{x}(k)$ and $x(k)$ represent the forecasted value and the original value, and n represents the total number of periods. For Equation (23), $F_k = 1$ if $(x(k+1) - x(k)) \times (\widehat{x}(k+1) - x(k)) \geq 0$, otherwise $F_k = 0$. For Equation (24), $\overline{d} = \frac{1}{n}\sum_{k=1}^{n}d_k$, $d_k = (x(k) - f_{1,k})^2 - (x(k) - f_{2,k})^2$; $Var(\overline{d}) = \frac{1}{n}(\gamma_0 + 2\sum_{i=1}^{n}\gamma_i), \gamma_i = Cov(d_k - d_{k-i})$. $f_{1,k}$ and $f_{2,k}$ represent the predicted value obtained by the first and second model, respectively.

3. Empirical analysis and results

3.1. *Data and descriptive statistic*

In 2020, the total GDP in Jiangsu Province is close to 10.28 trillion yuan, ranking the second in China. Electricity consumptions can reflect economic development, so the case analysis in Jiangsu Province can not only provide decision-making basis for resource planning and management policies, but also provide reference for economy-related governance. The influence factors of electricity consumptions (x_0, billion kilowatt hours) are generally several representative indicators, which are selected from economic factors and other factors related to the electric power development.

Due to the availability of limited data, we selected nine factors including GDP of the second industry (x_1, billion yuan), total number of employees in the secondary industry (x_2, million people), major operating revenue of industrial enterprises above designated size (x_3, billion

Table 1
Electricity consumptions and the main influencing factors.

Year	x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9
2010	3864.37	21861.48	19.99	91077.41	17121.03	22273.00	64.58	103.89	53.03	60.337
2011	4281.62	25239.22	20.18	107030.09	20699.22	25570.00	69.92	172.13	58.15	66.275
2012	4580.90	27158.78	20.32	119286.78	23309.75	28808.00	75.32	220.05	58.20	68.566
2013	4956.62	29149.42	20.41	133605.91	26752.50	31585.00	82.29	201.65	64.01	77.382
2014	5012.54	31048.84	20.45	141955.99	30174.27	34346.00	85.99	163.73	64.29	78.63
2015	5114.70	33371.77	20.34	147074.45	33931.69	37173.00	95.29	76.15	68.70	81.183
2016	5458.95	35041.53	20.27	156591.04	38269.57	40152.00	101.48	89.12	73.60	88.862
2017	5807.89	39124.11	20.12	148996.61	42700.49	43622.00	114.57	58.93	76.21	102.186
2018	6128.27	42129.37	19.93	127777.89	46936.47	47200.00	126.96	32.34	79.01	102.879
2019	6264.36	44270.51	19.61	114089.17	50852.05	51056.00	132.88	59.01	81.76	110.145
2020	6373.71	44226.43	19.43	120699.39	53955.83	53102.00	141.46	37.43	86.89	108.284

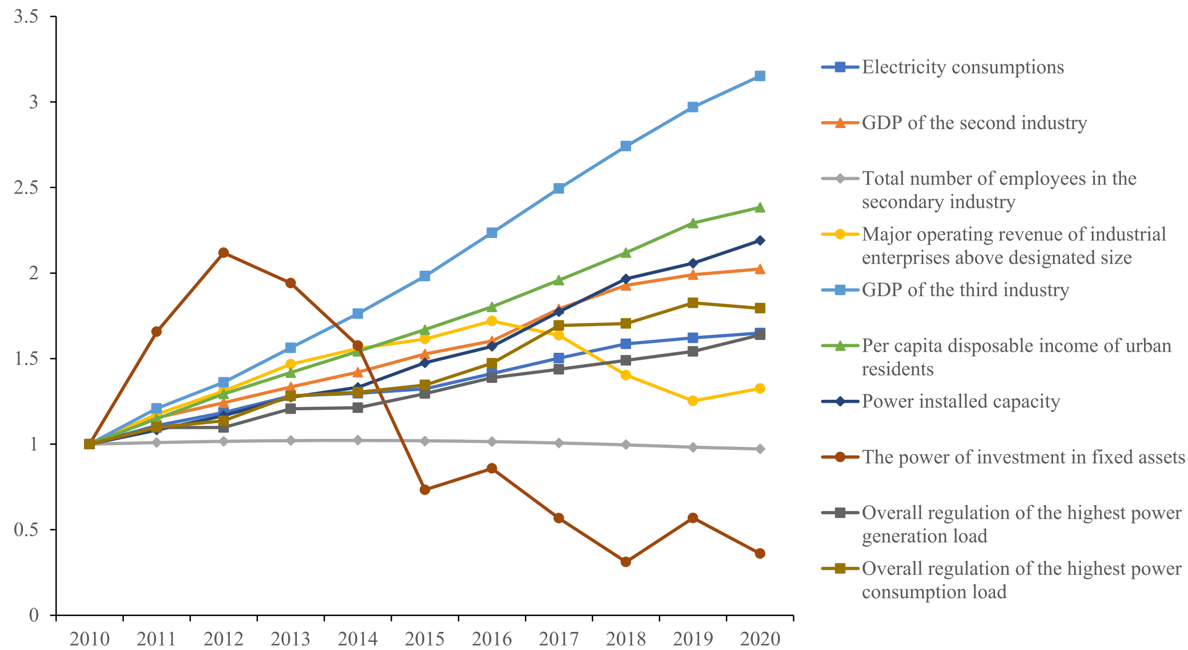

Fig. 2. Variation tendency of electricity consumptions and main influencing factors.

Table 2
Grey incidence degree between electricity consumption and its relevant factors.

Influencing factors	Grey incidence degree	Ranking
x_8	0.93026	1
x_9	0.77391	2
x_6	0.70592	3
x_1	0.55855	4
x_5	0.54558	5
x_4	0.53957	6
x_3	0.51798	7
x_2	0.50542	8
x_7	0.50374	9

yuan), GDP of the third industry (x_4, billion yuan), per capita disposable income of urban residents (x_5, yuan), power installed capacity (x_6, million kilowatt hours), the power of investment in fixed assets (x_7, billion yuan), overall regulation of the highest power generation load (x_8, million kilowatt hours), overall regulation of the highest power consumption load (x_9, million kilowatt hours).

The statistical data of electricity consumptions and main influencing factors in Jiangsu Province from 2010 to 2020 are provided in Table 1. The data are collected from Jiangsu Provincial Bureau of Statistics (available at: http://tj.jiangsu.gov.cn/2021/index.html).

The data in Table 1 are initialized and their initial values are depicted as shown in Fig. 2. We can see the data trend of related factors and system behavior sequences. Except for three factors (i.e., the investment in fixed assets of power supply, major operating revenue of industrial enterprises above designated size, and total number of employees in the secondary industry), the other factors' sequences show an upward trend basically. The data characteristics are consistent with system behavior sequences, which can be used for modeling and further analysis.

Through the calculation steps of grey incidence analysis, the influencing factors of electricity consumptions are determined, and the results can be obtained in Table 2. The result shows that the biggest factor impacting the electricity consumptions in Jiangsu Province is the overall regulation of the highest power generation load, followed by the overall regulation of the highest power consumption load, GDP of the second industry, per capita disposable income of urban residents, GDP of the third industry, major operating revenue of industrial enterprises above designated size, total number of employees in the secondary industry, and the last is the power of investment in fixed assets.

3.2. Identifying the number of related factors

Firstly, check whether the experimental data satisfies both the non-negative and the quasi-smooth conditions. The result shows that the data can be applied directly to establish the grey forecasting model. The grey incidence degree between electricity consumptions and its relevant factors can be given through grey incidence rankings as shown in

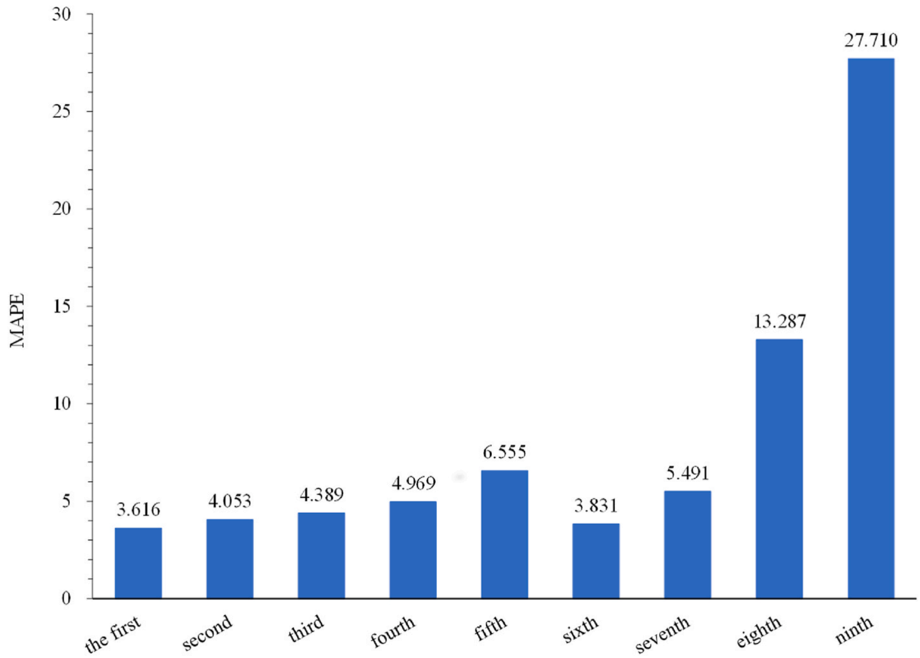

Fig. 3. Stepwise test result.

Table 3
Fitted values of electricity consumptions by different models.

	Year	Actual value	GM(1,N_r)		Traditional GM(1,N)		Traditional GM(1,1)		Metabolism GM(1,1)		FDGM		Holt ES	
			Fitted value	APE (%)	Fitted value	APE (%)	Fitted value	APE (%)	Fitted value	APE (%)	Fitted value	APE (%)	Fitted value	APE (%)
train set	2010	3864.37												
	2011	4281.62	3881.04	9.4	4281.62	0.094	4343.91	1.5	4384.89	2.4	4345.27	0.015		
	2012	4580.90	4964.75	8.4	4580.90	0.084	4556.71	0.5	4578.13	0.1	4558.05	0.005	3864.37	15.6
	2013	4956.62	5075.36	2.4	4956.62	0.024	4779.94	3.6	4779.88	3.6	4781.25	0.035	4202.34	15.2
	2014	5012.54	4963.29	1.0	5012.54	0.010	5014.11	0.0	4990.53	0.4	5015.38	0.001	4512.35	10.0
	2015	5114.70	5263.99	2.9	5114.70	0.029	5259.74	2.8	5210.46	1.9	5260.97	0.029	4875.31	4.7
	2016	5458.95	5630.17	3.1	5458.95	0.031	5517.41	1.1	5440.08	0.3	5518.59	0.011	4990.10	8.6
	2017	5807.89	5828.21	0.3	5807.89	0.003	5787.70	0.3	5678.73	2.2	5788.82	0.003	5092.17	12.3
	2018	6128.27	6041.72	1.4	6128.27	0.014	6071.23	0.9	5928.02	3.3	6072.28	0.009	5390.28	12.0
	MAPE (%)			3.616		3.616		1.345		1.77		1.347		11.21
test set	2019	6264.36	6353.01	1.4	6353.01	0.014	6368.66	1.7	6497.88	3.7	6369.63	0.017	5731.53	8.5
	2020	6373.71	6585.38	3.3	6644.19	0.042	6680.65	4.8	6534.56	2.5	6681.53	0.048	6056.30	6.7
	MAPE (%)			2.368		2.829		3.24		3.131		3.255		6.743

Table 4
Range of MAPE [58].

MAPE(%)	Prediction performance
<10	Excellence
10–20	Good
20–50	Reasonable
>50	Incorrect

Table 2.

The result shows that grey incidence degree between the related factors and the electricity consumptions were greater than 0.5. Therefore, the related factors will have a significant impact on the forecasting of electricity consumptions. Besides, according to the rankings, we select the first, second, …, fourth variable (from high to low) to establish the GM(1,N_r) model and calculate the MAPE of fitted values as shown in Fig. 3. When we choose the first variable, the MAPE of fitted values is the smallest (3.616%). Besides, the value tends to increase if more factors

Table 5
Comparisons between GM(1,N_r) model and other models.

	GM(1,N_r)	Traditional GM(1,N)	Traditional GM(1,1)	Metabolism GM(1,1)	FDGM	Holt ES
FD	97.632	97.171	96.76	96.87	96.745	93.257
MAE	150.163	359.133	205.619	197.185	206.547	425.119
MAPE	2.368	2.829	3.24	3.13	3.255	6.743
RMSE	162.273	284.639	229.228	200.505	230.041	438.552
DS	100	100	100	100	100	0
ranking	1	2	4	3	5	6

are involved. So, the number of relevant variables of GM(1,N_r) model is determined as 1 and the parameter N_r is set as 2.

3.3. Forecasting electricity consumptions in Jiangsu Province

In this paper, the experiment of the proposed GM(1,N_r) model is carried out by MATLAB R2017b. Due to the availability of limited data, the dataset is divided into training set (from 2010 to 2018) and test set (from 2019 to 2020). At the same time, we establish the traditional GM (1,N) model, the traditional GM(1,1) model, metabolism GM(1,1) model, fractional discrete grey model (FDGM) [57], and exponential smoothing method (Holt ES) for comparative analysis. The results of the above six models are presented in Table 3. The absolute percentage error and mean absolute percentage error of the fitted values for each forecasting model are calculated according to Equations (18) and (19).

Model fitting effect is measured according to the prediction accuracy standard of MAPE as shown in Table 4. The MAPE of six models are all below 10%, and the model prediction grade is excellent. As for the value of FD, grey forecasting model is above 95% at high level, while Holt ES model is below 95%. These evidences indicate that grey forecasting model has better fitting performance for small sample data.

For model evaluation, the prediction effect of the model is more important than the fitting degree, because the prediction effect can better reflect the stability and reliability of the model's prediction ability. For the test set, the performance of the proposed GM(1,N_r) is significantly better than other models, and the prediction of MAPE is 2.368%, which is at the lowest level among all the models, such as traditional GM(1,N) model (2.829%), traditional GM(1,1) model (3.24%), metabolism GM (1,1) model (3.13%), FDGM model (3.255%), and Holt ES model (6.743%) as shown in Table 5. Compared with its own fitting effect, the accuracy of the proposed model is improved. However, traditional GM(1,1) model, metabolism GM(1,N) model, and FDGM model decrease in different degrees. FDGM model is unable to show its advantage, which means this model cannot be applied in this forecasting case. With regard to other indexes, such as MAE and RMSE, the values of GM(1,N_r) areae obviously the lowest, and its FD is the highest. Besides, DS and DM test are also provided for further analysis. These specific values can be seen in Tables 5 and 6 and their comparisons are depicted in Fig. 4. DS result shows that the directional forecasting accuracy of these models(100) is good except Holt ES model(0). And We can deduce that our model has a significant probability to perform well based on DM test.

Fig. 5 shows the line graph of the actual sequence, GM(1,N_r) model, traditional GM(1,N) model, traditional GM(1,1) model, metabolism GM (1,1) model, FDGM model and Holt ES model. We can find that the fitting and prediction sequences of GM(1,N_r) model, traditional GM(1, N) model, traditional GM(1,1) model, and metabolism GM(1,1) model are relatively close to the actual data sequence with similar morphology, yet the FDGM model and Holt ES model show obvious deviation from the original sequence. In addition, the performance of metabolism GM (1,N) is better than that of traditional GM(1,1) model in predicting stage, which proves the rationality and effectiveness of this optimization to a certain extent.

In general, GM(1,N_r) performs best of all benchmark models. The prediction effect of GM(1,N_r) model is better than that of the traditional GM(1,N), which indicates that the improvement of the traditional GM(1, N) model is effective, and also demonstrates that the metabolism GM (1,1) model can improve the model accuracy. Holt ES model is also unsatisfactory in terms of prediction effect, because its MAPE is the largest among all the benchmark models, indicating its weak ability to reduce uncertainty factors. Univariate grey forecasting models (i.e., traditional GM(1,1) model, metabolism GM(1,1) model) are better than other models in terms of fitting effect. Though the prediction accuracy is slightly lower, it is still higher than the FDGM model and the Holt ES model, showing the superiority of grey forecasting model. Because of the lack of information extracted from relevant factors, it is slightly inferior to the multi-variable grey forecasting model.

Based on above analysis, it can be preliminarily judged that the proposed GM(1,N_r) model has the best adaptability to predict electricity consumptions in Jiangsu Province. It is also demonstrated that the extraction of influencing factors can significantly improve the prediction effectiveness, especially when pursuing the best fitting effect on small sample data. Meanwhile, it is essentials to determine the number of influencing factors by combining the grey incidence analysis and checking the fitted effects.

After verifying the rationality of the proposed model in this paper, the electricity consumptions of Jiangsu Province can be predicted. Firstly, the main relevant factor of the next 10 years, i.e., overall regulation of the highest power generation load, need to be obtained. Due to the lack of actual data at present, this paper adopts metabolism GM(1,1) model to infer the appropriate value. In order to ensure that the grey forecasting model is suitable for prediction, the grey exponential rate is checked before modeling. The results show that the fitted value of MAPE is 1.252%, indicating an excellent accuracy. Therefore, the prediction could be made based on the metabolism GM(1,1) model, and the prediction results are shown in Table 7.

Based on the prediction result, the GM(1,N_r) model is constructed, and the predicted values from 2021 to 2030 are calculated as shown in Fig. 6. The analysis shows that the electricity consumptions of Jiangsu Province in the next 10 years is still at a high level and will keep rising. Specifically, it will achieve 6871.47 billion kilowatt-hours (KWHs) in 2021, 7129.17 billion KWHs in 2022, 7413.49 billion KWHs in 2023 and so on. The specific values can provide reference for government departments to make economic policy related decisions.

Table 6
The Diebold-Mariano test results of different models.

	Traditional GM(1,N)	Traditional GM(1,1)	Metabolism GM(1,1)	FDGM	Holt ES
GM(1,N_r)	−1.054 (0.000)*	0.703 (0.000)*	0.791 (0.000)*	0.711 (0.000)*	−1.071 (0.000)*
Traditional GM(1,N)		0.893 (0.000)*	1.147 (0.000)*	0.896 (0.000)*	−1.071 (0.000)*
Traditional GM(1,1)			−0.436 (0.000)*	1.068 (0.000)*	−1.071 (0.000)*
Metabolism GM(1,1)				0.476 (0.000)*	−1.071 (0.000)*
FDGM					−1.071 (0.000)*

Note: "*" indicates significant level at 1%.

4. Conclusion and policy implications

4.1. Conclusion

In this paper, a novel grey forecasting model i.e., GM(1,N_r) model is constructed by combining particle swarm optimization algorithm, grey incidence analysis and the metabolism GM(1,1). Besides, the proposed model is compared with traditional GM(1,N) model, traditional GM(1,1) model, metabolism GM(1,1) model, FDGM model and Holt ES model to demonstrate its superiority. The following findings can be drawn:

(1) The proposed GM(1,N_r) model is more effective than the other two grey forecasting models, which indicates that the proposed model can predict the electricity consumption in Jiangsu Province in a more accurate way. To overcome the defects of the traditional grey forecasting model, we develop three improvements to the traditional GM(1,N) model: Firstly, introducing two

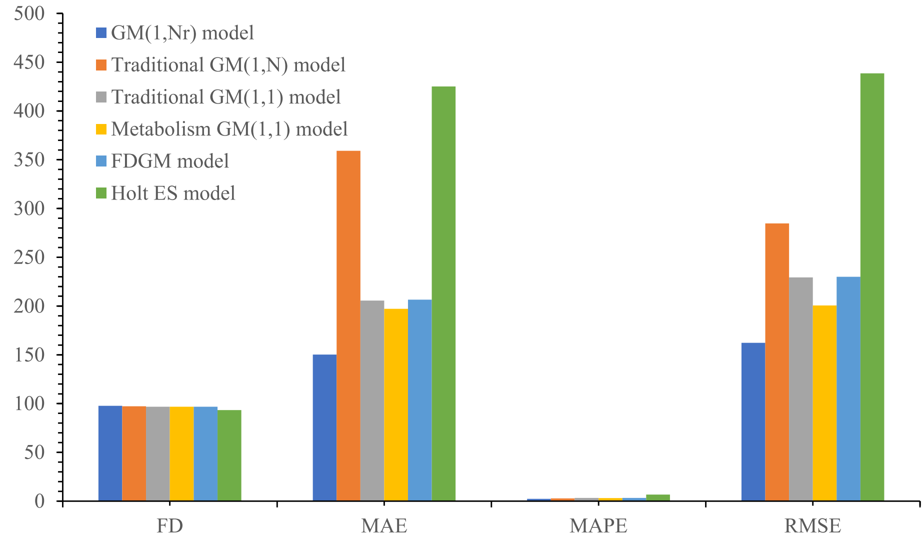

Fig. 4. Comparison between the GM(1,N_r) model and the other models.

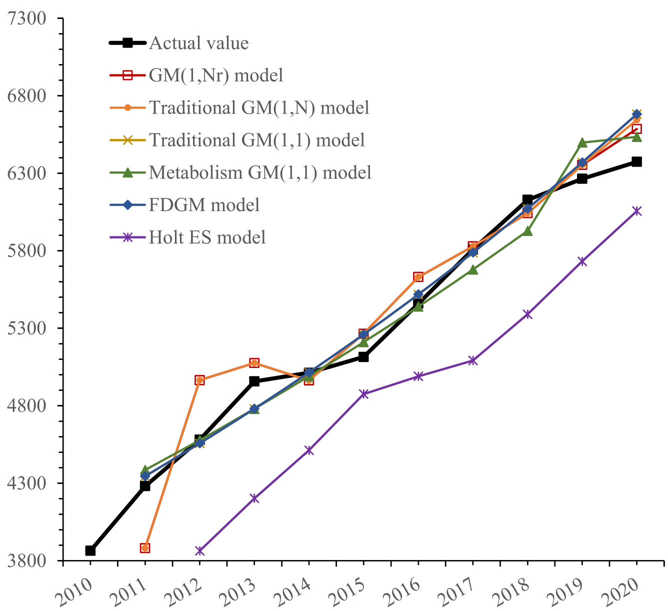

Fig. 5. Comparisons of the fitted value of six models with the actual value.

constants to satisfy the basic modeling conditions and using particle swarm optimization to find the parameter value; Secondly, the number of related factors is determined by grey incidence analysis and stepwise test of the fitting and prediction errors; Thirdly, the metabolism GM(1,1) is used to improve the prediction accuracy of GM(1,N_r) model. The proposed model is applied to the prediction case of electricity consumptions in Jiangsu Province, and also compared with the fitted results of traditional grey forecasting models and time series models. The results show that the fitting accuracy of GM(1,N_r) model is significantly higher than that of the other five models.

(2) By using the grey incidence analysis, it is found that the following factors, such as GDP of the second industry, total number of employees in the secondary industry, major operating revenue of industrial enterprises above designated size, GDP of the third industry, per capita disposable income of urban residents, power installed capacity, the power of investment in fixed assets, and overall regulation of the highest power generation load have a significant impact on the electricity consumptions. Therefore, electricity consumptions in Jiangsu Province is attributed to the level of economic development, internationalization, living standards of residents, and relevant investment in electricity, which coincides with the conclusion of existing studies. Based on the GM(1,N_r) model, the amount of electricity consumption is predicted. The results show that the value will continue to increase in the future.

4.2. Policy implications

Under the background of power consumption increasing with economy, we can solve the contradiction between power supply and demand from both sides. On the supply side, China mainly relies on thermal power generation at present. According to the latest data from the National Bureau of Statistics, the cumulative thermal power generation has reached 5.28 trillion KWHs in 2020, accounting for 71.19% of the total power generation. In recent years, the trend of thermal power generation has been gradually decreasing. And thermal power mainly depends on coal, which will produce a large amount of CO_2, and then bring a series of environmental problems [59].

Thus, China needs to explore a low carbon development mode when using coal or other clean or new energy to further reduce coal-fired power and achieve the goal of the proposed “carbon dioxide peaking” and “carbon neutrality” [60]. On the demand side, due to the particularity of electricity, i.e., electricity demand is concentrated, it is also crucial to reduce peak load and achieve balanced electricity consumptions. From the perspective of government and enterprises, more technical measures should be explored to optimize the way of electricity consumptions, such as load monitoring, Ultra High Vacuum transmission, energy storage, and other load management technologies. From the perspective of publics, it is necessary to raise awareness of environmental protection, cultivate the atmosphere of saving electricity in

Table 7
Predicted value of the main relevant factor.

Year	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030
Predicted value	8986.34	9323.36	9695.19	10084.74	10474.42	10891.61	11319.06	11761.70	12224.63	12701.75

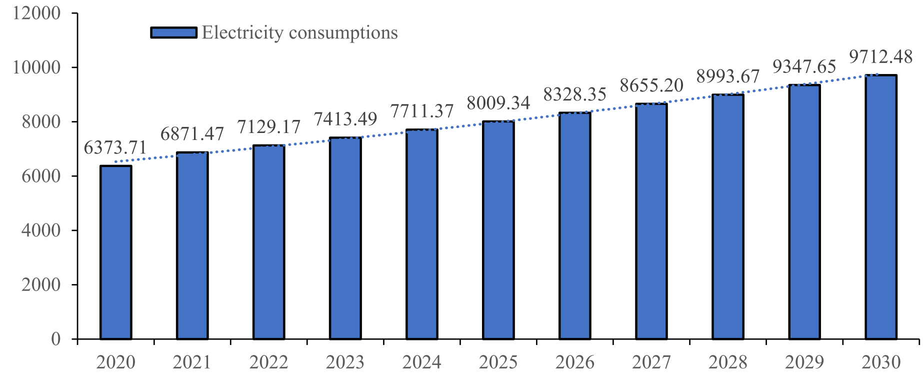

Fig. 6. Predicted value of electricity consumptions from 2021 to 2030.

the whole society.

4.3. Limitation and future directions

The GM(1,N_r) model proposed in this paper is an effective tool to analyze and predict electricity consumptions in Jiangsu province, but it should be noted that this paper remains some limitations. For example, in the selection of related factors, qualitative factors are not considered. In fact, some qualitative factors may also have a strong impact on electricity consumptions, such as electricity price, government financial support and so on. In the future, the impact factors could be further quantified to improve the forecasting prediction of electricity consumptions. In addition, the proposed GM(1,N_r) model can be further optimized, such as utilizing background value optimization to reduce the impact of the lack of the model itself on prediction effectiveness.

Credit author statement

Xiaoyi Du: Methodology, Writing – original draft & Editing, Software; **Dongdong Wu**: Visualization, Writing - Review; **Yabo Yan**: Writing - Review, Resources;

Declaration of competing interest

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.

Data availability

The data that has been used is confidential.

Acknowledgments

The authors would like to thank editors and anonymous reviewers for their helpful comments and suggestions.

References

[1] Pata UK, Caglar AE. Investigating the EKC hypothesis with renewable energy consumption, human capital, globalization and trade openness for China: evidence from augmented ARDL approach with a structural break. Energy 2021;216:119220.
[2] Zhao W, Zhong R, Sohail S, Majeed MT, Ullah S. Geopolitical risks, energy consumption, and CO_2 emissions in BRICS: an asymmetric analysis. Environ Sci Pollut Control Ser 2021;28(29):39668–79.
[3] Ghalehkhondabi I, Ardjmand E, Weckman GR, Young WA. An overview of energy demand forecasting methods published in 2005-2015. Energy Systems 2017;8(2):411–47.
[4] An N, Zhao W, Wang J, Shang D, Zhao E. Using multi-output feedforward neural network with empirical mode decomposition based signal filtering for electricity demand forecasting. Energy 2013;49(1):279–88.
[5] Hernandez L, Baladron C, Aguiar JM, Carro B, Sanchez-Esguevillas AJ, Lloret J, Massana J. A survey on electric power demand forecasting: future trends in smart grids, microgrids and smart buildings. IEEE Commun. Surveys Tutorials 2014;16(3):1460–95.
[6] Lopez-Martin M, Sanchez-Esguevillas A, Hernandez-Callejo L, Arribas JI, Carro B. Novel data-driven models applied to short-term electric load forecasting. Appl Sci 2021;11(12):5708.
[7] Lopez-Martin M, Sanchez-Esguevillas A, Hernandez-Callejo L, Arribas JI, Carro B. Additive ensemble neural network with constrained weighted quantile loss for probabilistic electric-load forecasting. Sensors 2021;21(9):2979.
[8] Mariano-Hernández D, Hernández-Callejo L, Solís M, Zorita-Lamadrid A, Duque-Perez O, Gonzalez-Morales L, Santos-García F. A data-driven forecasting strategy to predict continuous hourly energy demand in smart buildings. Appl Sci 2021;11(17):7886.
[9] Xu G, Schwarz P, Yang H. Adjusting energy consumption structure to achieve China's CO_2 emissions peak. Renew Sustain Energy Rev 2020;122:109737.
[10] Ali HS, Nathaniel SP, Uzuner G, Bekun FV, Sarkodie SA. Trivariate modelling of the nexus between electricity consumption, urbanization and economic growth in Nigeria: fresh insights from Maki Cointegration and causality tests. Heliyon 2020;6(2):e03400.
[11] Jian L, Sohail MT, Ullah S, Majeed MT. Examining the role of non-economic factors in energy consumption and CO_2 emissions in China: policy options for the green economy. Environ Sci Pollut Control Ser 2021;28(47):67667–76.
[12] Murshed M, Alam M. Estimating the macroeconomic determinants of total, renewable, and non-renewable energy demands in Bangladesh: the role of technological innovations. Environ Sci Pollut Control Ser 2021;28(23):30176–96.
[13] Liddle B. Impact of population, age structure, and urbanization on carbon emissions/energy consumption: evidence from macro-level, cross-country analyses. Popul Environ 2014;35(3):286–304.
[14] Wang JM, Kang JJ. Analysis on influencing factors of energy demand based on interpretative structural modeling. Electr power 2017;50(9):31–6 (in Chinese).
[15] Yang HL, Shi D, Xiao J. Influence of environmental factors on energy efficiency: analysis on regional theoretic energy saving potential and real energy saving potential. China Indus. Econ. 2009;253(4):73–84 [in Chinese)].
[16] Meng FS, Li MY. Research on evaluation of energy efficiency and influence factors in China based on optimal combination weights. Oper Res Manag Sci 2013;22(6):153–60 [in Chinese)].
[17] Meaugbua O, Dhakal S, Kuwornu JK. Factors influencing energy requirements and CO_2 emissions of households in Thailand: a panel data analysis. Energy Pol 2019;129:521–31.
[18] Jiang L, Ji M. A study on influencing factors of energy consumption in Shanghai based on STRIPAT model. Shangai Environ. Sci. 2011;30(6):240–4 [in Chinese)].
[19] Lee GC. Regression-based methods for daily peak load forecasting in South Korea. Sustainability 2022;14(7):3984.
[20] Chapagain K, Kittipiyakul S, Kulthanavit P. Short-term electricity demand forecasting: impact analysis of temperature for Thailand. Energies 2020;13(10):2498.
[21] Chapagain K, Kittipiyakul S. Performance analysis of short-term electricity demand with atmospheric variables. Energies 2018;11(4):818.
[22] Toktarova A, Gruber L, Hlusiak M, Bogdanov D, Breyer C. Long term load projection in high resolution for all countries globally. Int J Electr Power Energy Syst 2019;111:160–81.
[23] Velasquez CE, Zocatelli M, Estanislau FB, Castro VF. Analysis of time series models for Brazilian electricity demand forecasting. Energy 2022;247:123483.
[24] Tajeuna EG, Bouguessa M, Wang SR. Mining customers' changeable electricity consumption for effective load forecasting. ACM Transactions on Intelligent Systems and Technology 2021;12(4):1–26.
[25] Wang J, Du P, Lu H, Yang W, Niu T. An improved grey model optimized by multi-objective ant lion optimization algorithm for annual electricity consumption forecasting. Appl Soft Comput 2018;72:321–37.
[26] Kim M, Kim J. Forecasts of electricity consumption in an industry building. Korean J Appl statistics 2018;31(2):189–204.
[27] Tang Z, Yin H, Yang C, Yu J, Guo H. Predicting the electricity consumption of urban rail transit based on binary nonlinear fitting regression and support vector regression. Sustain Cities Soc 2021;66:125529.

[28] Zolfaghari M, Golabi MR. Modeling and predicting the electricity production in hydropower using conjunction of wavelet transform, long short-term memory and random forest models. Renew Energy 2021;170:1367–81.
[29] De FM, Alessandri A, Ruti PM. Electricity demand forecasting over Italy: potential benefits using numerical weather prediction models. Elec Power Syst Res 2013; 104:71–9.
[30] Hamzacebi C, Avni EH, Cakmak R. Forecasting of Turkey's monthly electricity demand by seasonal artificial neural network. Neural Comput Appl 2019;31(7): 2217–31.
[31] Lin ZF, Cheng LL, Huang GH. Electricity consumption prediction based on LSTM with attention mechanism. IEEE Trans. Elec. Electronic Eng. 2020;15(4):556–62.
[32] Ercan U, Irmak S, Cevik KK, Canbazoglu E. Estimating electricity consumption levels in Dwellings using artificial neural networks. Sosyoekonomi 2021;28(46): 173–86.
[33] Rana M, Koprinska I, Khosravi A. Feature selection for interval forecasting of electricity demand time series data. In: Artificial neural networks. Cham: Springer; 2015. p. 445–62.
[34] Chen KL, Jiang JC, Zheng FD, Chen KJ. A novel data-driven approach for residential electricity consumption prediction based on ensemble learning. Energy 2018;150:49–60.
[35] Deng JL. Grey systems *theory*. Huazhong University of Science and Technology Press; 2002.
[36] Pi D, Liu J, Qin X. A grey prediction approach to forecasting energy demand in China. Energy Sources, Part A Recovery, Util Environ Eff 2010;32(16):1517–28.
[37] Yousur MU, Al-Bahadly I, Avci E. A modified GM(1,1) model to accurately predict wind speed. Sustain Energy Technol Assessments 2021;43:100905.
[38] Yuan C, Zhu Y, Chen D, Liu S, Fang Z. Using the GM(1,1) model cluster to forecast global oil consumption. Grey Syst Theor Appl 2017;7(2):286–96.
[39] Jia ZQ, Zhou ZF, Zhang HJ, Li B, Zhang YX. Forecast of coal consumption in Gansu Province based on Grey-Markov chain model. Energy 2020;199:117444.
[40] Feng SJ, Ma YD, Song ZL, Ying J. Forecasting the energy consumption of China by the grey prediction model. Energy Sources B Energy Econ Plann 2012;7(4):376–89.
[41] Hu YC. A genetic-algorithm-based remnant grey prediction model for energy demand forecasting. PLoS One 2017;12(10):e0185478.
[42] Wang Q, Song X. Forecasting China's oil consumption: a comparison of novel nonlinear-dynamic grey model (GM), linear GM, nonlinear GM and metabolism GM. Energy 2019;183:160–71.
[43] Hu YC. Energy demand forecasting using a novel remnant GM(1,1) model. Soft Comput 2020;24(18):13903–12.
[44] Wang H, Wang Y, Wu D. A new seasonal cycle GM(1,1) model and its application in railway passenger volume forecasting. Grey Syst Theor Appl 2022;12(2):293–317.
[45] Li K, Zhang T. A novel grey forecasting model and its application in forecasting the energy consumption in Shanghai. Energy Systems 2021;12(2):357–72.
[46] Tsai SB, Xue Y, Zhang J, Chen Q, Liu Y, Zhou J, Dong W. Models for forecasting growth trends in renewable energy. Renew Sustain Energy Rev 2017;77:1169–78.
[47] Hamzaçebi C. Primary energy sources planning based on demand forecasting: the case of Turkey. J Energy South Afr 2016;27(1):1–10.
[48] Hu YC, Chiu YJ, Yu CY, Tsai JF. Integrating nonlinear interval regression analysis with a remnant grey prediction model for energy demand forecasting. Appl Artif Intell 2021;35(15):1490–507.
[49] Tsai SB. Using grey models for forecasting China's growth trends in renewable energy consumption. Clean Technol Environ Policy 2016;18(2):563–71.
[50] Xu N, Dang Y, Gong Y. Novel grey prediction model with nonlinear optimized time response method for forecasting of electricity consumption in China. Energy 2017; 118:473–80.
[51] Wu L, Gao X, Xiao Y, Yang Y, Chen X. Using a novel multi-variable grey model to forecast the electricity consumption of Shandong Province in China. Energy 2018; 157:327–35.
[52] Liu SF. Grey system Theory and its application. Science Press; 2010.
[53] Guo H, Xiao X, Forrest J. A research on a comprehensive adaptive grey prediction model CAGM(1,N). Appl Math Comput 2013;225:216–27.
[54] Dang YG, Liu SF, Liu B. GM model with xˆ(1)(n) as initial condition. Chinese J Manag. Sci. 2005;13(1):132–5 [in Chinese)].
[55] Xu P, Aamir M, Shabri A, Ishaq M, Aslam A, Li L. A new approach for reconstruction of IMFs of decomposition and ensemble model for forecasting crude oil prices. Math Probl Eng 2020;2020:1325071.
[56] Gao W, Aamir M, Shabri AB, Dewan R, Aslam A. Forecasting crude oil price using Kalman filter based on the reconstruction of modes of decomposition ensemble model. IEEE Access 2019;7:149908–25.
[57] Wu L, Liu S, Yao L. Discrete grey model based on fractional order accumulate. Syst. Eng. Theory Practice 2014;34(7):1822–7 [in Chinese)].
[58] Hu YC. Electricity consumption prediction using a neural-network-based grey forecasting approach. J Oper Res Soc 2017;68(10):1259–64.
[59] Wu D, Wang Y, Qian W. Efficiency evaluation and dynamic evolution of China's regional green economy: a method based on the Super-PEBM model and DEA window analysis. J Clean Prod 2020;264:121630.
[60] Wu D, Li H, Wang Y. Measuring sustainability and competitiveness of tourism destinations with data envelopment analysis. J Sustain Tourism 2022. https://doi.org/10.1080/09669582.2022.2042699.