Research Article

Short-Long Term Load Prediction Based on GM-BP Model

Fang Yuan,¹ Dong Li,¹ Wencui Li,¹ and Shenliang Wang²

¹State Grid Henan Information & Telecommunication Company, Zhengzhou 450052, Henan, China
²Nanjing Nari Information & Communication Technology Co., LTD., Nanjing 211106, Jiangsu, China

Correspondence should be addressed to Fang Yuan; liwenxue884813@outlook.com

Received 21 June 2022; Revised 26 August 2022; Accepted 29 August 2022; Published 10 September 2022

Abstract

With the continuous development of the smart grid, rational power planning can greatly increase economic benefits. The power load is affected by temperature, humidity, and other factors, which presents typical nonlinear and random characteristics. Therefore, a combined prediction model based on grey model GM (1, N) and BP neural networks is proposed, where GM is improved by weighting the original data, selecting appropriate initial conditions and self-adaptive optimizing model parameters, while PSO algorithm is used to automatically optimize the global optimization. Finally, an experiment is carried out based on the actual power load data of a certain area, and the results show that the proposed method realizes higher accuracy in load prediction, and has a stronger generalization ability for small samples and low SNR conditions.

1. Introduction

With the reformation of electricity marketization, reasonable power planning can greatly increase social and economic benefits. Power load forecasting is from the perspective of the whole power supply chain, which solves some influential factors with a high correlation degree of electricity load through some mathematical formula and predicts the process of meeting certain requirements in the short-term or long-term. Accurate power load forecasting of the power system can more reasonably and efficiently plan the start-up and shutdown of generating units at the generation end, so as to maintain the continuous safe and stable operation of the power grid, greatly reduce the demand of energy storage spent at the generation side, and make the maintenance plan of generating units more reasonable [1, 2].

Higher prerequisites are advanced for the dependability of power supply, and the factors affecting the power load are gradually increasing, which leads to the increasing complexity of power grid construction. Therefore, it is necessary to comprehensively convert more factors into actual factors and add them into the prediction model, so as to provide a more accurate and reliable basis for power network planning.

However, due to the nonlinear and fluctuating characteristics of power load data, the traditional method is difficult to accurately predict the power load [3, 4]. With the development of intelligent algorithms, genetic algorithm [5] and artificial neural network [6] are introduced into the field of power load forecasting. Among them, compared with other prediction methods, the advantage of the grey theory is that it does not need too much sample data and there is no strict requirement for the distribution of samples. GM pays more attention to the real information. Therefore, it can make an accurate prediction of the system only by analyzing and processing the information that has been mastered. For the medium and long-term load forecasting with less information, the GM has achieved better results in the field of power load forecasting. However, when GM is applied to load prediction scenarios with multiple original data and a long-time scale, the prediction accuracy will deteriorate. Many scholars have optimized the grey prediction data by smoothing the original data. Tan et al. proposed a method of using exponential smoothing to process raw load data with large fluctuations [9]. Based on the grey prediction method of discrete Fourier change,
Fourier transform is applied to the original data, and the characteristic components of data of various frequencies are decomposed. The original data are reconstructed and combined into low-frequency component data and high-frequency component data. The grey prediction model is used to predict the low-frequency data, and the high-frequency data are processed separately. Finally, the comprehensive results are output, which can shield the interference of random factors to the data prediction results to a certain extent [10]. At present, GM (1, 1) has been widely used in various kinds of power load forecasting research, the main reason is that the model can be solved easily in the case of lack of data and uncertainty. The advantages of the GM (1, 1) model are that the algorithm is simple, the operation and test are very convenient, and the data requirements are not high, and it can predict the future development trend without too much historical data. However, there are also many problems about GM (1, 1) found in the actual prediction, which are as follows:

1. The discrete degree of original data is negatively correlated with the prediction accuracy, which means that the traditional GM (1, 1) is only suitable for load series with exponential growth type.
2. When the large amount of data processing is poor, the grey data processing is better. Therefore, the increase of historical data cannot improve the accuracy of load forecasting.
3. The defect of the traditional GM (1, 1) model is that it is difficult to adapt to the power load with many influencing factors, and the factors of this model are single, so it is difficult to track the future trend of power forecasting in real time.
4. The complexity of the grey forecasting model is also the key factor affecting the accuracy of load forecasting. For example, when dealing with medium and long-term forecasting, a simple forecasting model often leads to the phenomenon that the growth rate is too fast, so the forecast results in the next few years are often not accurate.

In terms of machine learning, power load forecasting methods and theories are constantly emerging [11, 12], such as the time series method, fuzzy theory, Regression analysis, Support vector regression machine, Bayes, and neural network, which provide good theoretical support for power load forecasting, but also have certain limitations. Among them, the time series method requires high accuracy of historical data, and the short-term load forecasting is not sensitive to weather factors. Regression analysis: the quantitative relationship between the observed variables is quantitatively described from the perspective of statistical average significance, but it is easily limited by the scale of load data. Support vector regression machine has a good generalization ability, but the training time is excessively long due to the optimization of penalty coefficient, $\epsilon$ of loss function and $\gamma$ of the kernel function, especially when the training sample set is large. In addition, with the emergence of massive load data, it is obviously necessary to find a new method to meet the requirements of load forecasting and analysis under mass data [13]. As one of the power load forecasting methods, BP neural network has a strong nonlinear mapping ability. However, it will carry out a round of training for each load input and output sequence to calculate the correction of weights and thresholds of each layer of the network. Obviously, when the amount of data is very large, the amount of calculation will become large, and the serial training time of a single machine may reach several hours or even more. Meanwhile, when the total load samples increase, the prediction accuracy of the BP neural network may decrease, that is, the problem of “over fitting.” In training, the network overemphasizes the accuracy of the overall training effect of samples, which results in weak generalization ability and poor prediction effect in actual load prediction.

Therefore, this paper proposes a combination model of Grey Model (GM) and BP neural network. It fully combines the strong prediction ability of GM under the condition of a small sample and poor information and the approximation ability of BP neural network to arbitrary nonlinear function, which can realize high-precision prediction of power load and improve its generalization ability. In addition, aiming at the limitations of traditional GM and BP neural network, an improved GM is proposed, and the global optimization of the BP neural network is carried out automatically by using the particle swarm optimization (PSO) algorithm.

2. Power Load Forecasting Based on GM

2.1. Modeling Basis

2.1.1. Characteristics of Power Load. Power load is different from other products in its continuity and nonmass storage. At the same time, some manpower and material resources are required for the construction of power generation, substation, and distribution network in the initial stage of construction, so it is very important to make the power load meet users' needs reasonably. In order to achieve this goal, it is necessary to clarify the characteristics of power load. Power load is not only affected by people's production, such as heavy industry, light industry, agriculture, and so on but also affected by people's life, such as lighting, heating, and household appliances with different power. At the same time, the power load is also affected by weather, economic development, and major holidays. Therefore, its characteristics are Variability, Periodicity, and Continuity:

2.1.2. Characteristics of GM. Unlike traditional models, grey modeling does not establish difference equations for sequence, but preprocesses sample data by accumulative generation, and then establishes corresponding differential equations for processed data. As the original historical data is affected by many external factors, the original sequence is always discrete, which is called the grey sequence or grey process. The modeling of such a grey sequence is a grey model. Differential equation modeling is mainly based on the following points:
(1) Grey system is a system that uses grey theory to study grey quantity. Grey quantity is a random quantity and can be changed in a certain range. Therefore, the research process of the grey system is a process of changing in a certain range.

(2) The original load data are all fluctuating and irregular. The grey system transforms it into a new series of numbers with exponential growth, which is in line with the growth form of the first-order differential equation.

(3) According to the specific requirements, the grey theory can adopt different generating methods to establish the corresponding residual GM (1, 1) model, so as to make the prediction result more accurate.

(4) For some complex systems, GM (1, n) modeling can be used to achieve high accuracy prediction.

(5) The predicted value of the model obtained directly cannot be directly applied, and can only have a reference value after the reverse operation. At the same time, the prediction model needs to be tested before prediction.

3. Power Load Forecasting Model Based on Improved GM (1, n)

Grey modeling realizes the modeling and analysis of grey-generated data, and the first-order linear differential equation can be constructed as follows:

$$\frac{dx^\prime}{dt} + \lambda x^\prime = u,$$

where $\lambda$ and $u$ are the parameters of the differential equation.

The original data vector is $x = [x_1, x_2, \ldots, x_k, \ldots, x_K]$, where $x_k$ is the value of the $k$-th dimension of the original data vector, and the dimension of the original data vector is $K$. The vector for the generation of the data obtained through accumulation is $x' = [x_1', x_2', \ldots, x_k', \ldots, x_K']$, among them, $x_k'$ is the value of the $k$-dimensional data of the generated data vector and it satisfies the following formula:

$$x_k' = \sum_{i=1}^{k} x_i.$$

The first-order differential equation is expressed as follows:

$$\frac{dx_k'}{dt} = \frac{x_k' - x_{k-1}'}{k + 1 - k} = x_{k+1},$$

where $x_k'$ is the mean value of $k$ and $k + 1$, and is substituted with formulas (1) into (2), thus

$$x_{k+1} = \frac{1}{2}a(x_k' + x_k), \quad k = 1, \ldots, k - 1.$$

Equations (1)–(4) have a total of $k - 1$, which can be described in the form of a matrix as follows:

$$Y_K = BA$$

$$Y_K = \begin{bmatrix} x_2 \\ x_3 \\ \vdots \\ x_K \end{bmatrix}, \quad A = \begin{bmatrix} a \\ u \end{bmatrix}, \quad B = \begin{bmatrix} 1/2(x'_1 + x'_2) \\ 1/2(x'_2 + x'_3) \\ \vdots \\ 1/2(x'_{K-1} + x'_K) \end{bmatrix}.$$  \hspace{1cm} (5)

Furthermore, the solution of the equation can be obtained as follows:

$$x_{ki} = (x_1 - \frac{\bar{u}}{a})e^{-\lambda k} + \frac{\bar{u}}{a}, \quad (k = 0, 1, \ldots, K - 1).$$  \hspace{1cm} (6)

It is reduced to the actual predicted value by reduction

$$x_{k+1} = x_k' - x_k = \left(1 - e^{-\lambda k}\right)(x_1 - \frac{\bar{u}}{a})e^{-\lambda k}.$$  \hspace{1cm} (7)

After solving $\lambda$ and $\bar{u}$, the general solution can be obtained as follows:

$$x_k' = \frac{\bar{u}}{a} + ce^{-\lambda k},$$  \hspace{1cm} (8)

where $c$ is the parameter of a general solution.

The traditional GM usually assumes that the initial conditions are

$$x_1' = x_1.$$  \hspace{1cm} (9)

Since the latest data usually contains more information associated with it, the initial improvement can be made as follows:

$$x_K' = x_K.$$  \hspace{1cm} (10)

The improved GM is obtained:

$$x_{k+1}' = \left(x_k - \frac{\bar{u}}{a}\right)e^{-\lambda(k - K + 1)} + \frac{\bar{u}}{a}.$$  \hspace{1cm} (11)

The traditional grey model takes $x_k'$ as the mean value of $k$ and $k + 1$, and the improved grey model takes $x_k'$ as follows:

$$x_{k+1}' = \lambda_k x_{k+1} + (1 - \lambda_k)x_k', \quad (k = 1, \ldots, K - 1),$$  \hspace{1cm} (12)

where $\lambda_k$ is the weighting factor to be optimized.

Finally, the $K - 1$ weighting factor $\lambda_1, \ldots, \lambda_{K-1}$ are taken as a variable, a genetic algorithm is used to optimize, and the optimization objective function is the minimum variance between the predicted result and the actual result.
where \( p_k \) is the actual load, \( \hat{p}_k \) is load forecasting.

4. Power Load Forecasting Based on GM-BP Model

4.1. BP Neural Network Model. Figure 1 is the structure diagram of a typical three-layer BP neural network. The layers are fully interconnected, and there is no interconnection between the same layer.

Among them: \( x_i \) represents the input of the \( j \)-th node in the input layer; \( o_k \) is the threshold of the \( k \)-th node in the output layer. \( M \) is the dimension of input signal; \( q \) is the total number of hidden layer nodes; \( L \) is the node dimension of the output layer; \( \psi \) is the excitation function of the output layer; the \( O_k \) represents the output of the \( k \)-th node.

Using the gradient descent algorithm in optimization, then using the iterative operation to solve the weights corresponding to the learning and memory problem, we add hidden nodes. The adjustable parameters of the optimization problem are increased, and the exact solution of the predicted load can be obtained. The weight and threshold of each layer are corrected in the direction of output to input. Combined with Figure 1, input signal transmission, weight, and threshold correction of the BP neural network can be calculated separately.

4.1.1. Forward Propagation of Input Signals. It can be known that the input \( net_i \) and output \( o_i \) of the \( i \)-th node in the hidden layer, and the input \( net_k \) and output \( o_k \) of the \( k \)-th node in the output layer are respectively shown in formulas (14) to (17):

\[
net_i = \sum_{j=1}^{M} w_{ij} x_j + \theta_i, \quad (14)
\]

\[
o_i = \phi(net_i) = \phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_j\right), \quad (15)
\]

\[
net_k = \sum_{i=1}^{q} w_{ki} \phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_j\right) + \alpha_k, \quad (16)
\]

\[
o_k = \psi(net_k) = \psi\left[\sum_{i=1}^{q} w_{ki} \phi\left(\sum_{j=1}^{M} w_{ij} x_j + \theta_j\right) + \alpha_k\right]. \quad (17)
\]

4.1.2. Back Propagation of Error Signal. As shown in formulas (18) to (21), the output of the adjusted network mapping can meet the expected value.

\[
\Delta w_{ki} = \eta \sum_{p=1}^{P} \sum_{k=1}^{L} (T^p_k - o^p_k) \psi'\left(net_k\right) o_i, \quad (18)
\]

\[
\Delta \alpha_k = \eta \sum_{p=1}^{P} \sum_{k=1}^{L} (T^p_k - o^p_k) \psi'\left(net_k\right), \quad (19)
\]

\[
\Delta w_{ij} = \eta \sum_{p=1}^{P} \sum_{k=1}^{L} (T^p_k - o^p_k) \psi'\left(net_k\right) w_{ki} \psi'\left(net_i\right)x_j, \quad (20)
\]

\[
\Delta \theta_i = \eta \sum_{p=1}^{P} \sum_{k=1}^{L} (T^p_k - o^p_k) \psi'\left(net_k\right) w_{ki} \psi'\left(net_i\right), \quad (21)
\]

where \( \Delta \theta_i \) is the threshold correction of the \( i \)-th node in the hidden layer. \( p \) is load sample index; \( \eta \) is the learning rate; \( P \) is the total number of training samples.

4.1.3. PSO Optimized BP. Since BP adopts the gradient descent method for network learning, there is the problem of initial value sensitivity. Improper selection of the initial value will lead to the convergence of the algorithm to the local extreme point, and the prediction performance of the model will decline. Therefore, PSO is used to optimize the initial value of BP globally to improve its performance.

The specific process is shown in Figure 2, which can be divided into the following five steps.

Step 1. PSO initialization. Set the number of population, particle position and speed, search space, search step size, value range, and the maximum iterations.

Step 2. The initial value of the BP neural network model is endowed with particles, and the initial velocity and position of the particles are randomly set in the parameter space.

Step 3. Calculate the fitness function value. The difference between \( y^* \) and squared \( f = (y - y^*)^2 \) was used as the fitness function in the current state and the corresponding \( P_g \) and \( P_i \).

Step 4. Update the position and velocity of particles according to formula (19) and equation (20) and calculate the updated fitness function value. If the fitness function value is greater than the result obtained in Step 3, the updated particle state will be the current state.

Step 5. Judge whether the termination condition is satisfied. If so, the particle at this time is the optimal initial value of the BP neural network; otherwise, return to Step 3

4.2. GM-BP Model. The schematic diagram of the proposed combined model is shown in Figure 3.
It can be seen that the whole algorithm mainly includes three parts, in which the data preprocessing realizes the function of data normalization and transforms the data from different sources and dimensions into the same standard range. The GM \((1, n)\) model mainly improves the prediction performance and generalization ability of the algorithm under the condition of small samples and poor information and outputs the prediction value and prediction residual. In addition, BP neural network takes the prediction output, residual, and preprocessed original data of GM \((1, n)\) model as the input. Then, it models the nonlinear information contained in the data and finally outputs the prediction results. Its core is to improve the adaptability of the algorithm to the nonlinear information that GM \((1, n)\) model cannot process.

5. Experiment and Results

5.1. Midlong Term Power Load Forecasting. Taking the power load data of a city from 2013 to 2021 as an example, the improved medium and long-term load forecasting method is used to predict the overall power load of the city. The accuracy of the model is verified by the actual annual power load data of the city from 2013 to 2021. The power load data from 2013 to 2021 are shown in Table 1. From 2013 to 2021, the overall annual power load of the city shows an increasing trend, which is mainly due to the improvement of residents’ living standards. As a result, the maximum annual load peak valley difference increases from 177.4 billion kWh in 2013 to 400.2 billion kWh in 2021.

For the annual power load, the influencing factors more than 0.9 are the household consumption index and regional annual GDP power consumption structure. After considering these factors, GM \((1, n)\) and GM-BP models were established respectively. Figure 4 shows the error curve of annual load forecasting results between GM-BP and traditional GM \((1, n)\).

Compared with the traditional grey forecasting method, the annual load graph curve predicted by the improved grey prediction model method is closer to the actual annual load graph curve, and the change of the predicted graph curve is more gentle. In addition, in the midlong term load forecasting, the relative error of traditional GM prediction is
more than 5%, and the maximum relative error is 12.5% in 2018, while the relative errors of the GM-BP model are less than 5%. Therefore, the GM-BP model has higher prediction accuracy than the traditional GM.

5.2. Short Term Power Load Forecasting. Short-term load forecasting predicts a 24-hour load in a day. Therefore, it is necessary to distinguish the attributes of forecast days in load forecasting. In this paper, the daily load curve from October 1, 2021 to October 5, 2021 is selected as the known data to predict the load curve of the next working day. The error curve of short-term load prediction results between GM-BP and GM (1, n) is shown in Figure 5.

When the traditional GM is applied to short-term load forecasting, the relative error of load forecasting in each period is controlled within 6%. In 1, 13, 14, 18, 20, and 24 periods, the relative error of prediction is more than 4%. In addition, when the improved GM is applied to load forecasting, the relative error of load forecasting is only 3.52% in 21 periods, and the relative errors in other periods are less than 3%. There are some periods when the GM deviates from the actual load significantly. For example, in the evening peak load period, the predicted daily load curve of the GM-BP model is more consistent with the actual load curve than the traditional GM, which can be seen that the GM-BP model also has higher accuracy in short-term load forecasting.

5.3. Prediction Performance of Combined Model

5.3.1. Sample Size. The data from the first 10 days to the first 60 days of October 2021 are used as the training samples to complete the model training. The power load data on October 5 are predicted. Figure 6 shows the 24-hour average prediction results. Meanwhile, for comparison, the prediction results obtained by using the improved GM (1, n) model and BP model under the same conditions are given.

Figure 6 shows that when the number of training samples increase, the prediction performance of the three methods has improved to varying degrees. Once the training samples reach 20 days, the prediction accuracy of the proposed combination model (95.7%) has stabilized, which is close to the case of 60 days of training samples (96.9%). When the performance of GM (1, n) model is near the optimal (60 days, 91.5%), the corresponding number of training samples is 30 days (90.8%). When the prediction performance of the BP neural network model is close to the optimal (60 days, 93.2%), the number of training samples required is 40 days (93.1%).
5.3.2. Signal to Noise Ratio. In practical engineering practice, there will inevitably be some useless information such as random error and interference in the collected data. To further verify the prediction of different methods in the presence of random error and interference, Gaussian white noise is adopted. Similarly, the data on October 5, 2021 is taken as the test data, and that of the first 60 days is used as the training sample, Gaussian white noise with a mean value of 0 and noise variance of $0.1 \sim 0.9$ are added to the training samples (the interval of noise variance is 0.1). Three methods are used to carry out the prediction test, and the results are shown in Figure 7.

It can be seen that the prediction performance of the three methods decreases in varying degrees when the noise variance increases. Among them, the performance of GM (1, $n$) model declines most obviously, and the prediction performance of the proposed combination model is more stable, reflecting the stronger noise robustness. As can be seen from Figure 7, under the condition of low SNR (noise variance greater than 0.4), the proposed combined model can still achieve a prediction accuracy of better than 85%, while the prediction accuracy of the other model is lower than 70%, which cannot meet the practical application requirements.

6. Conclusion

In this paper, a combined prediction model based on GM (1, $N$) and BP neural networks is proposed. The experimental results based on the actual power load data of a certain area show that the proposed model can understand higher expectation exactness than single GM (1, $n$) model and BP neural network model, and has more obvious advantages under the condition of small samples and low signal-to-noise ratio. In the future work, various factors affecting load such as GDP value, weather, policy, and other information will be deeply mined to find the law of the impact of various data on load, reduce the grey level in load prediction, and modify the prediction results.

Data Availability

The dataset can be accessed from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

[1] W. J. Zhang, J. Qin, F. Mei, J. Fu, B. Dai, and W. Yu, “Short-term power load forecasting using integrated methods based on long short-term memory,” Science China(Technological Sciences), vol. 63, no. 4, pp. 614–624, 2020.
[2] Z. Yuan and C. Yu, “Data integrity analysis of power engineering based on DBSCAN clustering,” Journal of Shenyang University of technology, vol. 41, no. 3, pp. 246–250, 2019.
[3] T. S. Li, W. Wang, and Z. Liu, “GELM-WNN forecasting method for power load considering Uncertain interval,” Computer Engineering, vol. 45, no. 7, pp. 315–320, 2019.
[4] X. Guan, Y. Tong, and Z. Gao, “Short term load forecasting based on EMD fractal theory,” Electronic design engineering, vol. 24, no. 1, pp. 184–186, 2016.
[5] Z. Yang, W. song, and K. Cao, “Short term load forecasting based on FARIMA model based on quantum genetic algorithm,” Sensors and Microsystems, vol. 38, no. 10, pp. 143–145, 2019.
[6] Y. Dang, T. Xu, and Y. Gao, “Short term load forecasting based on Haar wavelet and arimax model,” Journal of Shanghai University of technology, vol. 41, no. 1, pp. 64–70, 2019.
[7] J. Sun, X. Qian, and Y. Wang, “Improved grey load forecasting model based on,” Electrotechnics, vol. 19, pp. 28–31, 2019.
[8] Y. Yang and M. Su, “Analysis of power system short term load forecasting method,” China electric power enterprise management, vol. 12, pp. 92-93, 2018.
[9] W. Tao, J. Zhang, and H. Pan, “Spatial load forecasting based on bi level Bayesian classification,” Chinese Journal of electrical engineering, vol. 27, no. 7, pp. 13–17, 2007.
[10] F. Tan, Z. Zhang, C. Zhu, and J. Zhang, "An improved exponential smoothing load forecasting method," \textit{Power demand side management}, vol. 18, no. 06, pp. 22–26, 2016.

[11] Y. Cao and X. Yin, "Research on power system Harmonic Detection Technology based on BP neural network and all-phase fast Fourier transform," \textit{The modern electronic technology}, vol. 40, no. 01, pp. 125–128+133, 2017.

[12] W. Li, D. Sheng, and J. Chen, "Application of dual BP neural network combination model in real-time data prediction," \textit{Chinese Journal of electrical engineering}, vol. 27, no. 17, pp. 94–97, 2007.

[13] P. Liu, Y. Huang, and W. Chen, \textit{Hadoop-open a Shortcut to Cloud Computing} Electrical Industry Press, Beijing, 2011.