Power load forecasting based on Bayesian neural network and particle swarm optimization

Zhengping Gao¹, Jing Shi¹, Hu Li¹, Chen Chen¹, Jian Tan¹ and Lixin Liu²

¹Economic and Technological Research Institute of State Grid Jiangsu Electric Power Company, Jiangsu 210008, China
²Beijing Tsingsoft Innovation Technology Co, Ltd, Beijing 100085, China

¹Corresponding author e-mail: intell_adal@163.com

Abstract. Accurate demand forecasting is very important for power system operation and planning. In order to forecast energy demand more accurately, a short-term power load forecasting model based on Bayesian neural network is proposed. Firstly, Pearson coefficient is used to analyze the influence of distributed energy on the load of the distribution network, and Bayesian neural network (BNN) is used to make predictions. Then, particle swarm optimization (PSO) is used to optimize the weights and thresholds of the neural network in order to obtain the best prediction effect. The experimental results show that the Pearson coefficient can quantitatively analyze the impact of distributed energy. The prediction accuracy of BNN-PSO based prediction model is significantly higher than that of other conventional methods, and the proposed prediction model has higher prediction accuracy.

1. Introduction

Load forecasting is to determine the load data at a certain time in the future according to the system operation characteristics, capacity increase decision-making, natural conditions and social impact and other factors, under the condition of meeting certain accuracy requirements, in which load refers to the power demand or power consumption. Load forecasting is an important part of power system economic dispatch and an important module of energy management system (EMS). Since load forecasting is based on the past and present of power load, the research object of load forecasting is uncertain event. In recent years, the power data collected by smart meters has the scale of big data, and shows the characteristics of data volume development from TB level to PB level, many structured types and fast processing speed. In the past, the research based on monthly electricity consumption has been upgraded to the research on daily electricity consumption behavior of consumers. It is more and more important to effectively analyze power big data and put forward targeted guidance in practical application scenarios [1-5].

Power load forecasting is the basic work of power grid control optimization and an important part of power system dispatching. At present, there are various forecasting methods for power load in the long, medium and short term, among which clustering technology has brought huge economic and social benefits to the marketing, pricing and user classification of the power industry in the field of medium and short-term load forecasting [1-3]. In reference [4-5], a user clustering algorithm based on adjustment potential index is proposed by comparing the user consumption data before and after peak valley price, which provides a certain reference for selecting users who voluntarily participate in demand response.
regulation. Reference [6] proposes a quadratic clustering algorithm, which improves the deficiencies in the load form similarity of the clustering algorithm based on Euclidean distance in the full-dimensional power system load curve. Due to the importance of power load forecasting, the research of power load forecasting has been the focus of many researchers. With the deepening of the research, more and more researches are on the combination of traditional forecasting methods and optimization methods or the combination of different prediction models or methods. In reference [7], the load forecasting model uses comprehensive meteorological factors as input data, and adopts Elman to establish short-term load forecasting model. The result shows that the method improves the prediction accuracy and is a practical and effective method. In reference [8], a cascaded neural network based on BP-RBF is proposed to predict power load. The no load factor is used as the input of BP network, and the output is the peak and valley load of the forecast day, which achieves good prediction effect. In reference [9], the medium and long-term power load is forecasted, and the grey prediction model is optimized by genetic algorithm. The results show that the improved model has high accuracy. In reference [10], the genetic algorithm is used to optimize BP neural network, and the prediction model of GA-BP neural network is established. The results show that the prediction accuracy is significantly improved after the improvement. These studies provide a theoretical basis for solving the power system load forecasting.

On the basis of cleaning the load data information in the area of distribution network, this paper proposes a load forecasting method of distribution network area considering the access of distributed generation. Firstly, the Pearson coefficient is used to analyze the influence of distributed energy on distribution network load, and the Bayesian neural network and particle swarm optimization algorithm are introduced. Finally, the load data of a distribution network area in Jiangsu Province are used to verify. The example shows that the Pearson coefficient effectively analyzes the influence factors of load change, and the algorithm in this paper can effectively analyze the influence factors of load change. The algorithm in this paper has a good application effect on the application scenarios of load forecasting in the distribution network area with distributed large-scale access.

2. Bayesian neural network

The performance of neural network is closely related to its network size, which is mainly determined by the weight and threshold of neural network. The neural network with small weight and threshold has better generalization ability. Bayesian neural network modifies the training performance function of the network by regularization, and limits the weights and thresholds, so as to improve the generalization ability of the neural network. In the traditional feedforward neural network, MSE formula is used as the training performance function:

\[ E_d = \frac{1}{N} \sum_{i=1}^{N} (e_i)^2 = \frac{1}{N} \sum_{i=1}^{N} (t_i - a_i)^2 \]  

(1)

Where \( N \) is the total number of samples; \( e_i \) is the error; \( t_i \) is the target output value; \( a_i \) is the neural network prediction output value.

On the basis of the original training performance function, Bayesian regularization method introduces the mean square deviation of network weights into the performance function to form the new training function shown in formula (2):

\[ msereg = \beta E_d + \alpha E_w \]  

(2)

\[ E_w = \frac{1}{N} \sum_{i=1}^{N} (W_i)^2 \]  

(3)
Where $msereg$ is the improved error function; $\alpha$ and $\beta$ are the regularization parameter; $E_w$ is the average value of the sum of squares of the network ownership values; $W_i$ is the network weight.

In the process of network training, the regularization parameters are adaptively adjusted by equation (4), so as to achieve the goal of optimal training.

$$\alpha = \frac{\gamma}{2E_w}, \quad \beta = \frac{N - \gamma}{2E_d}$$

(4)

Where $\gamma = N - 2 \text{atr}(H)^{-1}$; $H$ is the Hessian matrix of $msereg$.

The modified training performance function can reduce the occurrence of falling into local minimum or over-training during the training process, but it is difficult to obtain the optimal network parameters by the traditional weight threshold adjustment method.

3. Parameter optimization of prediction model

Although Bayesian neural network improves the generalization ability by modifying the error performance function, the traditional neural network parameter optimization method cannot obtain the optimal network parameters, and the basic particle swarm optimization algorithm has the premature convergence problem in the large-scale parameter optimization problem. Therefore, an improved particle swarm optimization algorithm is proposed to optimize the weights and thresholds of Bayesian neural networks.

PSO algorithm has been widely used due to its few parameters setting and fast convergence speed, but there are premature and local optimization problems in the optimization process.[14] Individual optimal extreme value and global optimal extreme value are generated in the process of particle updating. Individual optimal extreme value is the optimal solution in the process of individual optimization, and the global optimal extreme value is the optimal solution in the process of all particle optimization. The update formulas are as follows:

$$v_{id}^{k+1} = w \cdot v_{id}^k + c_1 \cdot r_1 \cdot (p_{id}^k - x_{id}^k) + c_2 \cdot r_2 \cdot (p_{gd} - x_{id}^k)$$

(5)

$$x_{id}^{k+1} = x_{id}^k + v_{id}^{k+1}$$

(6)

Where $r_1$ and $r_2$ are uniformly distributed random numbers between [0,1]; $c_1$ and $c_2$ are acceleration factors; $w$ is inertia weight; $v$ is flight speed of particles; $x$ represents particles; $p_{gd}$ and $p_{id}$ are global optimal extreme value and individual optimal extreme value respectively.

PSO is similar to most heuristic algorithms, such as genetic algorithm and ant colony algorithm. In the later stage of algorithm iteration, the population diversity will drop sharply and form "aggregation" phenomenon.[15], which leads to premature convergence.

4. Example analysis

The experimental data is from the power load of Baiyun District, Guangzhou Province. The prediction target is to forecast the load situation in the next hour, which belongs to the short-term prediction task, and the prediction target is the total load of the high-tech zone. Among them, the annual data of 2016 is the training data, and the first week of July 2019 is the test data. The accuracy indicators used in this paper include root mean square error (RMSE) and mean absolute error (MAE).
\[
\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |P_i - \overline{P}_i| \quad (7)
\]
\[
\text{RMSE} = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (P_i - \overline{P}_i)^2} \quad (8)
\]

Where \( n \) is the number of predicted samples, \( P_i \) and \( \overline{P}_i \) are the actual value and predicted value at time \( i \) respectively.

4.1. Quantitative analysis of distributed energy

Due to the high capacity of distributed generation in the selected area, the output of distributed generation has a certain impact on the load pattern of the distribution grid in the region. Taking the distributed wind power and photovoltaic power of the distribution grid in the jurisdiction as an example, the relationship between the load forecasting results and the photovoltaic output and wind power output is analyzed, and the Pearson correlation coefficient is used to analyze the correlation between the distributed generation output and the real value of load. In this paper, the simultaneous photovoltaic and wind power data in the jurisdiction area are input into the model as feature information. Table 2 shows the Pearson coefficient relationship between different levels of load and wind power and photovoltaic output in two scenarios. Further analysis shows that in the scenario where the access capacity of distributed generation in the selected area reaches a high proportion, the relevant factors affecting the output of distributed generation are closely related to the load change in the jurisdiction. In the jurisdiction with higher penetration rate of distributed generation, the greater the impact of distributed generation related information on load forecasting.

![Figure 1. Impact of wind power and photovoltaic output on load forecasting results.](image)

4.2. Analysis of prediction results

Using the distributed energy output of Section 4.1 as the feature input, the advantages and disadvantages of the prediction algorithm in this paper are analyzed. In order to compare the algorithm in this paper with other algorithms and reflect the optimization effect of this model algorithm, the comparison algorithms include linear model and SVM. It can be seen from Table 1 that the average absolute error
and root mean square error of the short-term load forecasting model in this paper are less than those of SVM and linear model forecasting model, which has higher prediction accuracy and stronger generalization ability, and the prediction result is closer to the actual value, and has better prediction effect.

Figure 2. Setting of input properties.

Table 1. Comparison of prediction results of several prediction models.

| Prediction model       | MAE  | RMSE |
|------------------------|------|------|
| Linear model           | 132.4| 128.3|
| SVM                    | 114.1| 118.5|
| The model in this paper | 76.6 | 78.11|

5. Conclusion

With the rapid development of smart grid and renewable energy, the access of high proportion of renewable power makes it more difficult to predict the level load of distribution network. Due to the volatility and randomness of distributed generation, the load at the level of distribution network is not the real power consumption form of users. Therefore, it is necessary to accurately predict the load change trend of distribution network with deep consideration of the output of distributed generation. Considering the characteristics of natural resources of light and wind, this paper proposes a power load forecasting method based on correlation coefficient analysis and Bayesian neural network for the distribution network area with high penetration of distributed generation, and analyzes the difference between load forecasting and traditional forecasting in the jurisdiction of distribution network.

Acknowledgments

This work was financially supported by Science and Technology Project (SGJSJY00GHJS1900035) of State Grid Jiangsu Electric Power Company.

References

[1] Wang Shuai. Short-term power combination forecasting method and evaluation of wind farms[D]. Hebei: North China Electric Power University, 2017.
[2] Fang Biwu, Liu Dichen, Wang Bo, et al. Short-term wind speed prediction based on wavelet transform and improved firefly algorithm to optimize LSSVM[J]. Power System Protection and Control, 2016, (8): 37-43.

[3] Feng Zhimin, Tian Li, Wu Daolin, et al. Research on IGSA optimization of short-term wind power prediction of LSSVM[J]. Renewable Energy, 2017, 35(11): 1699-1705.

[4] Le Wenhua, Wang Yingguan, Wang Jingcheng, et al. Traffic lane saturation prediction model based on improved TLBO-SVR[A]. Proceedings of the 2015 Annual Meeting of the National Metallurgical Automation Information Network[C]. 2015.

[5] Y. Xu, L. Wang, S.-y. Wang, M. Liu. An effective teaching-learning-based optimization algorithm for the flexible job-shop scheduling problem with fuzzy processing time[J]. Neuro computing, 2015, (148) : 260-268.

[6] Rao R V, Savsani V J, Vakharia D P. Teaching-Learning-Based Optimization: An optimization method for continuous non-linear large scale problems[J]. Information Sciences, 2012, 183(1): 1-15.

[7] Rao R V, Savsani V J, Vakharia D P. Teaching-learning-based optimization: A novel method for constrained mechanical design optimization problems[J]. Computer-Aided Design, 2011, 43(3): 303-315.

[8] S.C. Satapathy, A. Naik. Modified teaching-learning-based optimization algorithm for global numerical optimization: a comparative study[J]. Swarm Evol Comput, 2014, (16): 28-37.

[9] Z. Yang, K. Li, A. Foley, et al. A new self-learning tlbo algorithm for rbf neural modeling of batteries in electric vehicles[C]. 2014 IEEE Congress on Evolutionary Computation (CEC), IEEE, Beijing, 2014, pp. 2685–2691.

[10] J. Yan, K. Li, E.-W. Bai, et al. Hybrid probabilistic wind power forecasting using temporally local Gaussian process[J]. IEEE Trans. Sustain Energy, 2015, (99): 1-9.