Peak Load Forecasting for Electrical Bus based on Limited Historical Data under Complex Meteorological Condition

Heyan Zhu¹, Changyong Yu¹, Qingkui He²,*, Kun Song¹, Jingbo Liu¹, Yanlu Xu¹, Tianqi Lu¹
¹State Grid Liaoning Electric Power Company Limited Economic Research Institute, Shenyang, China
²School of Northeast Electric Power University, Jilin, China
*Corresponding author e-mail: 2605055605@qq.com

Abstract. The historical data of peak load for electrical bus are limited and fluctuated violently. The fluctuation ability of peak load for electrical bus is nonlinear and randomly. So its prediction accuracy is low. In order to improve the accuracy of peak load forecasting for electrical bus, a peak load forecasting for electrical bus method based on limited historical data under complex weather conditions is proposed. Firstly, the influence of natural meteorology, society and other factors on peak load fluctuation for electrical bus is analysed; Secondly, based on the reduction of redundancy between features of potential feature set, the feature importance ranking is obtained by conditional mutual information (CMI). Then, according to the improved particle swarm optimization extreme learning machine suitable for small sample training, the forward feature selection is performed to determine the optimal feature subset. Finally, based on the optimal feature subset, the optimal peak load forecasting model is established.

1. Introduction
Peak load prediction of electrical bus is an important basis to ensure reliable and stable operation of power system. However, the peak load of electrical bus is affected by many factors, which is easy to fluctuate violently and difficult to predict. Therefore, it is very important to realize high precision peak load prediction for electrical bus.

In the process of bus load prediction modeling, it is necessary to consider the influence of many factors on bus load. Literature [1] studied the influence of environmental factors on bus load and the relationship between influencing factors and bus load through the quick attribute reduction method. It found out the factors that have a great influence on bus load, and improved the accuracy of bus load prediction; in the literature [2], abnormal data in the data were found and adjusted by establishing the eigenvector matrix. According to the load under test and the similar set, the similar day is selected to realize the comprehensive prediction of bus load; in the literature [3], the power supply and load components of the active network are respectively studied, and the bus load prediction is optimized. In above literatures, the influence of various factors such as natural meteorology and society on bus load is not fully analyzed, and in the bus load prediction, there is no feature selection when considering many factors.
Most of the existing research on peak load forecasting focuses on the influencing factors and forecasting methods of peak load in urban power grid. In literature [4], the internal causes of urban load fluctuation are analyzed, and the influence of multi-source data on load fluctuation is considered. Pearson correlation coefficient and mutual information were used to calculate the correlation between multi-source data and load fluctuation. Natural, social and other characteristics strongly correlated with load fluctuation were selected. Support vector regression model was applied to predict peak load; in literature [5], a peak load prediction model of Cartesian Genetic Programming based on the evolution of artificial neural network was proposed and peak load is predicted by optimal characteristics. Although existing studies have improved the peak load prediction accuracy to some extent, targeted analysis has not been carried out on the small sample problem with limited historical data in peak load prediction. Extreme Learning Machine [6, 7, 8, 9] (ELM) requires fewer training samples, which can solve the problem of peak load prediction of electrical bus with the characteristics of nonlinear fluctuation, and the prediction accuracy is high.

Aiming at the problems in the above studies, in order to analyze the influence of complex meteorological data on peak load for electrical bus and solve the shortage of limited historical data of peak load and low accuracy of load prediction, a method of peak load for electrical bus prediction based on finite historical data is proposed under complex meteorological conditions. Firstly, the influence of natural meteorological factors and social factors on the fluctuation of peak load for electrical bus is analyzed; then, according to different influence factors on peak load for electrical bus, Conditional Mutual Information was used to analyze the correlation of factors affecting peak load for electrical bus. On this basis, the Improved Particle Swarm Optimization Extreme Learning Machine (IPSO-ELM) suitable for small sample training was adopted as the peak load predictor of electrical bus, and the IPSO-ELM prediction accuracy was taken as the decision variable to determine the optimal feature subset; Finally, according to the optimal feature subset, the optimal model for peak load prediction of different bus is established.

2. Peak load data analysis for electrical bus

Figure 1 shows the actual values of peak load and daily load of two bus in a city in northeast China in August 2018. As can be seen from figure 1, the maximum and minimum values of bus 1 peak load for bus 1 in one month are 160.6MW and 60.5MW respectively; the maximum and minimum values of peak load for bus 2 in one month are 170.7MW and 112.2MW respectively. Compared with city power load, bus load fluctuates greatly. Moreover, the peak load for electrical bus only accumulates a set of data every day, and the historical load data is limited, so the model training is more difficult.

Figure 1. Daily load and peak load for electric bus.
3. Feature selection based on conditional mutual information

3.1. Construction of original feature set for peak load prediction of electrical bus

Natural meteorological factors (longitude, latitude, temperature, air pressure, humidity, wind direction, wind speed, etc.) and social factors (dates, holidays, etc.) will cause the peak load of bus to fluctuate. According to the analysis of complex meteorological factors and relevant literature research, the original feature set constructed is shown in table 1.

| Number | Type                        | Characteristics                                                                 |
|--------|-----------------------------|---------------------------------------------------------------------------------|
| 1      | Meteorological             | $F_T, F_{T_{ave}}, F_{T_{(max,d-1)}}, F_{T_{(ave,d-1)}}, F_d, F_{H}, F_{W}, F_{W1}$ |
| 2      | The geographical position  | $F_{G1}, F_{G2}$                                                               |
| 3      | Data                       | $F_{d1}, \ldots, F_{d7}, F_{j1}, F_{j2}, F_{h1}, F_{h2}$                       |
| 4      | Load                       | $F_{L_{(max,d-1)}}, F_{L_{(max,d-2)}}, \ldots, F_{L_{(max,d-30)}}, F_{L_{(t-15)}}, F_{L_{(t-30)}}, \ldots, F_{L_{(t-120)}}$ |

Note:
(1) $F_T$ and $F_{T_{ave}}$ respectively represent the peak and mean values of temperature on the forecast day; $F_{T_{(max,d-1)}}$ represents the peak temperature of the previous day; $F_{T_{(ave,d-1)}}$ represents the average temperature of the day before the forecast; $F_d$, $F_{H}$, $F_{W}$, $F_{W1}$ represent air pressure, humidity, wind direction and wind speed at the predicted time respectively;
(2) $F_{G1}$, $F_{G2}$ respectively represent the longitude and latitude of the bus to be predicted;
(3) $F_{d1}$ to $F_{d7}$ represent the date within the week; $F_{j1}$ and $F_{j2}$ represents working day and non-working day; $F_{h1}$ and $F_{h2}$ represent holidays and normal days respectively;
(4) $F_{L_{(max,d-1)}}$ represents the peak bus load predicted the day before, $F_{L_{(max,d-2)}}$ represents the peak load forecast for the previous two days, and so on; $F_{L_{(t-15)}}$ represents the bus load 15 minutes before the peak load of the day before the prediction date, $F_{L_{(t-30)}}$ represents the bus load 30 minutes before the peak load of the day before the prediction date, and so on.

3.2. Conditional mutual information

In the process of peak load prediction for electrical bus, $D$ is set as the original feature set including natural meteorology, society and other factors; $Q$ represents bus peak load value; $Z$ is the selected features. Under $Z$ condition, the conditional mutual information between $D$ and $Q$ is:

$$ F(D,Q | Z) = \sum_{d \in D} \sum_{q \in Q} \sum_{z \in Z} P(d,q,z) \log \frac{P(d,q | z)}{P(d | z)P(q | z)} $$

In formula (1), $P(d | z)$, $P(q | z)$ is probability density functions of $D$ and $Q$ respectively under $Z$ condition, $P(d,q | z)$ is the joint probability density function of $D$ and $Q$ under $Z$ condition; $P(d,q,z)$ is the joint probability density function of $D$, $Q$ and $Z$. 


4. The principle of ultimate learning machine based on improved particle swarm optimization

4.1. Extreme learning machine

\{ (x_i, t_i) \}_{i=1}^N \) Represents N samples, the input data is \( x_i = [x_{i1}, x_{i2}, \cdots, x_{in}]^T \in R^n \), and the target output value is \( t_i = [t_{i1}, t_{i2}, \cdots, t_{im}]^T \in R^m \). Then the ELM network model of single hidden layer neural network (number of hidden layer nodes is L) can be expressed as

\[
o_j = \sum_{i=1}^{L} \beta_j g(\omega_i \cdot x_j + b_j) \quad j = 1, 2, \cdots, N
\]  

(2)

In formula (2), \( o_j \) represents the network output value; \( g \) represents the activation function; \( \omega_i \) is the input weight; \( \beta_j \) is the output weight; \( b_j \) is the bias of the hidden layer unit.

When there is no error, the activation function is infinitely close to any N samples, namely

\[
\sum_{i=1}^{N} \left\| o_i - t_i \right\| = 0
\]  

(3)

According to (3), it can be obtained

\[
\sum_{i=1}^{L} \beta_j g(\omega_i \cdot x_j + b_j) = t_j \quad j = 1, 2, \cdots, N
\]  

(4)

The matrix form of N equations in equation (4) is

\[
H \beta = T
\]  

(5)

\( H \) Represents output of node of hidden layer; \( T \) is the expected output.

Through obtaining \( \hat{\beta} \), \( \hat{\omega} \) and \( \hat{b} \) to achieve ELM training

\[
\left\| H(\hat{\omega}, \hat{b}) \beta_1 - T \right\| = \min_{\omega, b, \beta} \left\| H(\omega, b) \beta - T \right\| \quad i = 1, 2, \cdots, L
\]  

(6)

The minimum loss function equivalent to (6) is

\[
E = \sum_{j=1}^{N} \left( \sum_{i=1}^{L} \beta_j g(\omega_i \cdot x_j + b_j) - t_j \right)^2 \quad j = 1, 2, \cdots, N
\]  

(7)

In ELM, once \( \omega \) and \( b \) are randomly determined, a unique \( H \) can be obtained. Accordingly, the ELM structure is determined.

4.2. ELM parameter optimization based on improved particle swarm

Taking the original feature set as an example to construct the optimal process of peak load prediction model for electrical bus, the process of optimizing ELM through IPSO is as follows:
(1) The ELM model of peak load prediction for electrical bus is constructed based on the original feature set (the feature set is shown in table 1), and \( \omega \), \( b \) of ELM is randomly generate; 

(2) The sample \( X \) of the original feature set information needed for peak load prediction for electrical bus is determined, and the ideal accuracy or number of iterations is the end condition; 

(3) According to the sample data, normalization process is carried out; 

(4) Fitness value was obtained according to MAPE of the peak load prediction of ELM for electrical bus, and the optimal fitness value of the current individual and group was determined. 

(5) On the basis of traditional PSO, particle velocity and position are updated according to formulas (11) and (12). 

(6) Firstly, the fitness value of the current particle is calculated; then compared with the historical optimal value, if it is better, the optimal solution of the particle is updated. Otherwise, the optimal fitness value of individuals is maintained. 

(7) If the fitness value of the current particle is better than the group optimal solution, the optimal solution is updated. Otherwise, the group optimal solution is maintained. 

(8) If the end condition is not reached, return 3). Otherwise, the optimal solution of \( \omega \) and \( b \) are substituted into ELM to build the optimal peak load prediction model for electrical bus. 

In the subsequent feature selection process, the corresponding optimal ELM predictor was constructed according to feature sets of different dimensions, and the above methods were adopted. 

5. Experiment and analysis 
Peak load for electrical bus and meteorological information data of a city in northeast China in 2018 were applied in the study. The data of January, April, August and October were used as verification set, the data of July were used as test set, and the remaining 7 months were used as training set to carry out targeted feature selection for different bus. In order to prove the advanced of the new method, IPSO-ELM was compared with ELM and BPNN. Mean Absolute percentage Error (MAPE) and Root Mean Square Error (RMSE) were used to evaluate the prediction effect of the model. 

5.1. Feature selection analysis based on CMI 
Figure 2 shows the feature importance degree of the original features to the peak load for bus 1 and bus 2 obtained through CMI respectively. Correlation between peak load of different bus and features is different. 

![Figure 2. Feature importance analysis.](image-url)
Figure 3. Selection of optimal feature set (bus 1).

Figure 3 shows the results of feature dimension and MAPE values for feature selection using three methods. As can be seen from figure 3, when IPSO-ELM, BPNN and ELM are respectively used as predictors and feature subset dimensions are 25, 28 and 34, respectively, the error of peak load prediction for electrical bus is the smallest. According to figure 3, when BPNN, ELM and IPSO-ELM are respectively used as predictors according to the optimal feature set, the MAPE value of peak load prediction of electrical bus 1 is 4.54%, 3.75% and 3.04%, respectively. Among the three predictors, the MAPE value of peak load prediction of IPSO-ELM is the smallest, which shows the advantages of high prediction accuracy of IPSO-ELM. Similarly, the optimal feature subset of bus 2 can be determined.

5.2. Analysis of peak load prediction results for electrical bus

In order to prove the advantages of high accuracy of the proposed peak load prediction method for electrical bus, the peak load prediction results of bus 1 and bus 2 in July 2018 are listed in a city of northeast China. In figure 4, the targeted models of two bus and the prediction results of peak load for electrical bus using the original feature set are presented. Table 2 corresponds to the peak load prediction error in figure 4. Under the optimal model, the MAPE of peak load prediction for bus is 3.04% and 2.98% respectively. Under the model of the original feature set, the MAPE of peak load prediction for bus is 3.89% and 4.01% respectively. It can be seen from the comparison that the optimal model has higher prediction accuracy.

In figure 5, IPSO-ELM, ELM and BPNN are respectively used as peak load predictors to predict the peak load for electrical bus 1 in July in a city of northeast China in 2018. Table 3 corresponds to the prediction errors of the three peak bus load prediction methods in figure 5. As can be seen from table 3, MAPE of peak load predicted for electrical bus by IPSO-ELM is the smallest. Therefore, the proposed method has higher prediction accuracy.
Figure 4. Peak load prediction

Figure 5. Three methods of peak load prediction

Table 2. Error of peak load prediction

| Bus  | IPSO-ELM | Original model |
|------|----------|----------------|
|      | MAPE (%) | RMSE(MW)       |
|      | MAPE (%) | RMSE(MW)       |
| Bus 1| 3.04     | 9.45           |
| Bus 2| 2.98     | 9.12           |

Table 3. Prediction errors of the three methods

| Methods | MAPE (%) | RMS(MW) |
|---------|----------|---------|
| BPNN    | 4.23     | 12.83   |
| ELM     | 3.93     | 11.24   |
| IPSO-ELM| 3.04     | 9.45    |
6. Conclusion
In order to overcome the problem that bus peak load is affected by many factors and has few sample data, a new method of bus peak load prediction based on limited historical data is proposed in this paper. The new method has the following advantages:

1. Based on the CMI value, feature selection of peak load prediction for electrical bus was carried out to reduce the influence of feature redundancy on prediction accuracy during peak load prediction.
2. The optimal prediction model for different bus is constructed to improve the prediction accuracy of different bus.

Acknowledgments
In the process of writing my thesis, State Grid Liaoning Electric Power Company Limited Economic Research Institute and the surrounding students have provided me with valuable help. First of all, I sincerely thank State Grid Liaoning Electric Power Company Limited Economic Research Institute for providing experimental data in the project of accurate prediction of power grid load based on meteorological data under the trend of diversified electricity consumption (SGTYHT/18-JS-206); Then, I would like to sincerely thank my classmates for their wise advice.

References
[1] Long Danli, Li Jinghua, Wei Hua. Solution of Multi Environmental Factor-Influenced Bus Load Forecasting by Rough Set Method, J. The grid technology, 2013, 37 (05): 1335 - 1340.
[2] Sun Qian, Yao Jiangang, Zhao Jun, etc. Short-term Bus Load Integrated Forecasting Based on Selecting Optimal Intersection Similar Days, J. Chinese journal of electrical engineering, 2013, 33 (04): 126 - 134+17.
[3] Zhu Zelei, Cheng Xin, Yang Guizhong, etc. Collaborative Scheduling Optimization Method for Provincial Grid and Regional Grid Based on Refined Busload Forecast, J. Chinese journal of electrical engineering, 2017, 37 (03): 665 - 676.
[4] Pan Z, Min J. Peak load forecasting based on multi-source data and day-to-day topological network, J. Iet Generation Transmission & Distribution, 2018, 12 (6): 1374 - 1381.
[5] Khan G M, Arshad R. Electricity Peak Load Forecasting using CGP based Neuro Evolutionary Techniques, J. International Journal of Computational Intelligence Systems, 2016, 9 (2): 376 - 395.
[6] Peng Xiangang, Zheng Wei Qin, Lin lixiang, etc. Short-term Load Forecasting Method Based on Outlier Robust Extreme Learning Machine Considering Adaptive Load Detection and Repair, J. Chinese journal of electrical engineering, 2016, 36 (23): 6409 - 6417+6606.
[7] Lu Jinling, Sheng Feifei, Zhao Hongshan. Fault diagnosis method of wind turbine main bearing based on extreme learning machine, J. Renewable energy, 2016, 34 (11): 1588 - 1594.
[8] Hu K, Zhou Z, Weng L, etc. An Optimization Strategy for Weighted Extreme Learning Machine based on PSO, J. International Journal of Pattern Recognition & Artificial Intelligence, 2016, 31 (01): 1751001-.
[9] Naz A, Javed M U, Javaid N, etc. Short-Term Electric Load and Price Forecasting Using Enhanced Extreme Learning Machine Optimization in Smart Grids, J. Energies, 2019, 12 (5).