The Method of Short Term Load Forecasting for Micro Grid Using Limit Learning

Peng Ji¹, Deren Zhao¹, Fei Xia¹, Zongze Xia¹, Xiaobo Huang¹, Peixian Cong¹, Zhixiong Yang²,*, Zhuo Di¹, Li Song¹
¹State Grid Liaoyang Electric Power Supply Company, Liaoyang 111000, China
²Kunming Institute of Physics, Kunming 650223, China
*xiongmaoer39@126.com

Abstract. Considering the cost constraint and the uncertainty of power consumption, a short-term load forecasting method for microgrid based on kernel function extreme learning machine is proposed. The use of kernel extreme learning machine and heuristic genetic algorithm and time of training samples, the offline optimization of the parameters of prediction model and online load forecasting including periodic update of model parameters; through to ensure the timeliness of algorithm of optimal parameters, while reducing the computational complexity of online prediction system and historical data storage. Through short-term load forecasting for different capacity and type of user side microgrids, the accuracy of prediction results, the effect of parameter cycle update, the impact of prediction results on economic operation and the computational efficiency of prediction methods are analysed.

1. Introduction
Microgrid can promote the access of distributed clean energy, reduce environmental pollution and reduce power transmission loss, and improve the reliability of power supply by [1]. Short term load forecasting is an important part of microgrid energy management system. It is the basis for optimal operation of micro controlled sources such as micro gas turbines, diesel engines and energy storage. [2] Will directly affect the operation strategy [3] and power transactions of microgrid. The related research shows that the higher load forecasting error of the microgrid will lead to a significant increase in the operating cost by [4].

At present, the research results of short-term load forecasting for microgrid are relatively limited, and the computational complexity is high. [5]For small capacity to individual or collective user as the owner of the micro grid, in addition to ensure higher forecasting accuracy of short-term load givers should also reduce the computational load forecasting method of complexity, it is easy to realize [6] in the embedded terminal device.

Based on the above reasons, this paper based on the analysis of the micro grid load forecasting based on features using machine learning algorithm based on kernel function limit, includes the establishment of off-line parameter optimization and online load forecasting of micro grid short-term load forecasting model, and through the micro grid examples of various types and capacity of forecast accuracy, stability and computational efficiency of periodic update.
2. Kernel function limit learning machine

In this paper, ELM_k is chosen as the load forecasting algorithm. First, the neural network construction mechanism of the basic ELM algorithm is briefly explained, and its neural network function can be expressed as:

\[ f(x) = h(x)\beta \]  \hspace{1cm} (1)

ELM ensures the accuracy of regression prediction by minimizing the output error, i.e.

\[ \lim_{L \to \infty} \left\| f(x) - f_0(x) \right\| = \lim_{L \to \infty} \left\| \sum_{i=1}^{L} \beta_i h_i(x) - f_0(x) \right\| = 0 \]  \hspace{1cm} (2)

In the formula, \( L \) is the number of \( h(x) \) in the hidden layer neuron; \( F_0(x) \) is the expected function of the target value.

At the same time, the ELM algorithm ensures the generalization ability of neural networks by minimizing the output weight \( \beta \). Usually the \( \beta \) is obtained by the least square solution, and the calculation method is:

\[ \beta = H^\dagger O = H^\dagger (HH^\dagger)^{-1} O = H^\dagger \left( \frac{1}{C} + HH^\dagger \right)^{-1} O \]  \hspace{1cm} (3)

In the formula, \( H \) neural network hidden layer matrix; generalized inverse matrix \( \hat{H} \dagger \) is the predicted target value vector \( H \). According to the ridge regression theory, by increasing the normal number \( 1/C \), the results will be more stable and have better generalization ability.

For the ELM_k algorithm, a better regression prediction accuracy is obtained by introducing a kernel function.

\[ f(x) = h(x)H^\dagger \left( \frac{1}{C} + HH^\dagger \right)^{-1} O \]

\[ = \begin{bmatrix} K(x, x_1) \\ \vdots \\ K(x, x_N) \end{bmatrix}^\dagger \left( \frac{1}{C} + \Omega_{ELM} \right)^{-1} O \]  \hspace{1cm} (4)

\[ \Omega_{ELM(i,j)} = \exp(-\gamma \| x_i - x_j \|^2) \]

In the formula, \( \Omega_{ELM} \) for the selected kernel function, usually from the Gauss kernel function; \( N \) is the input dimension.

3. Microgrid load forecasting model

3.1. Demand and model construction of micro grid load forecasting

Compared with the large power grid operation conditions, micro short-term load forecast has the following characteristics: ① the micro grid cannot forecast the load with high performance servers, it is not appropriate at each load forecasting when the parameters were optimized; ② each time with large differences in dielectric properties, in order to improve the accuracy of load forecasting, need to build the load forecasting model time; ③ the online load forecasting may be completed by embedded system, and the prediction model has great limitations in computing complexity and storage space.
Considering the above characteristics of the micro grid load forecasting, based on the time of training samples, and select the ELM_ k genetic algorithm (Genetic Algorithm, GA) combination, build offline parameter optimization and online load forecasting of micro grid short-term load forecasting model, as shown in figure 1.

Figure 1. Short term load forecasting model of microgrid

3.2. Incomplete restoration of data
The historical data of the microgrid load may have the following problems:

(1) The data of the load sampling is incomplete, that is, the lack of a historical sample at a certain time, or the power sampling value of the time is empty.

(2) In the load history sampling data, there is a special event of blackouts. Its characteristic is that the load sampled over 2 h is 0, which is not the application category of short-term load forecasting.

(3) The history of sampled data included in the special holidays load sampling data, especially for special holidays to leave, have effect load characteristics on working days and holidays. Because this is not to short term load forecasting for special holidays as a key to the original load off in the sampling days load data processing. For the problems existing in the historical data, the experiments are filled or covered with the average value of the load data in the same period. The calculation method is as follows:

\[
x_i = \frac{x_{i-W} + x_{i-W}}{2} \quad W = 7 \times 24
\]

In the formula, \(x_i\) is the \(i\) load power sampling value; \(W\) is the number of theoretical samples for a week.

3.3. Load attributes
For the \(i\) load sample value \(x_i\), used in the study, the sampling time is \(t_i\) weeks, days before the average load information of \(w_i\) \(dai\), former \(dl_i\), weeks before the load lag load \(wl_i\) as its load attribute. Load data sampling interval is 1h, the sampling time \(t_i\) is 1–24 integer; for weeks \(w_i\), a value of 1–7 for the day before the integer; the average load calculation method for \(dai\).

\[
da_i = \frac{1}{24} \sum_{k=1}^{24} x_{i-k} \quad i = 25, \cdots, G
\]
In the form, \( G \) is the total number of sample samples. The day before the week before the load and lag load, calculated as follows:

\[
\begin{align*}
  dl_i &= x_{i-24} & i = 25, \cdots, G \\
  w_l_i &= x_{i-168} & i = 169, \cdots, G
\end{align*}
\]  

(7)

For ease of presentation, the \( A_i \) refers to all the load attributes of the \( i \) sample, that is,

\[ A_i = \{ t_i, w_l_i, da_i, dl_i, w_l_i \} \]

(8)

3.4. Off-line model construction and parameter optimization
Due to the big difference in the load characteristics between the working day and the holiday, the short-term load forecasting model is set up respectively for the two. For the selection of the validation samples, the parameter optimization of the current \( T \) time load forecasting model is set up, and the corresponding validation sample is.

\[ V_T = \{ x_i \mid A_i \}; \quad t_i = T \]

(9)

In the formula, \( V_T \) is the validation sample of the \( T \) time prediction model; \( x_i \) is the load. The predicted target value; \( A_i \) is the load attribute value of the sample.

The sample of the precursor of the \( T \) time, the sample of the day period and the interval of the weekly interval are used as the time-sharing training samples of the \( T \) time load forecasting model.

\[ M_T = \{ x_i \mid A_i \}; \quad t_i = \{ T - 1, T - 2, \\
T - 22, \cdots, T - 26, T - 166, \cdots, T - 170 \} \]

(10)

In the formula, \( M_T \) is a training sample for the \( T \) time prediction model.

Based on the ELM_\_k algorithm, the load forecasting model \( F_T^g \) is established by using the sample \( M_T \) and parameters \( C_T^g \) and \( \gamma_T^g \). The load attributes of the sample \( V_T \) are then predicted and verified.

\[
\begin{align*}
  F_T^g &= ELM_\_k(C_T^g, \gamma_T^g, M_T) \\
  \hat{x}_T^g &= F_T^g(A_i) \quad A_i \subset V_T
\end{align*}
\]

(11)

4. Experiment and result analysis
4.1. Microgrid load sample
For the short-term load forecasting of microgrid, the four month load of 5~7 microgrids with different capacities is selected as the original sampling data. Among them, D1 and D2 are mainly residential property loads, and R1 and R2 are mainly Hotel loads. The sampling time interval is 1h, the original sampling properties include sampling time and load power.
4.2. Microgrid load forecasting accuracy
For evaluating the accuracy of load forecasting, we choose MAPE, Maximum Error (ME) and Mean Square Error (MSE) as evaluation criteria.

\[ ME_i = \max |y_i(p) - \hat{y}_i(p)|; p = 1, \ldots, 10 \]  \hspace{1cm} (12)

\[ MSE = \frac{1}{S} \sum_{t=1}^{S} (y_t - \hat{y}_t)^2 \]  \hspace{1cm} (13)

In the formula, \( ME_i \) is the maximum error of \( p \) time load forecasting \( i \) moment; \( y_i \) goes \( I \) for predicting the output of the neural network to predict the target value of \( y_i \) value. In the first weeks of July, the error of load forecasting for each micro grid is shown as shown in Table 1, and the maximum errors at each moment are shown in Figure 2 to figure 5.

![Figure 2. Load forecasting results of micro-grid D1](image)

![Figure 3. Load forecasting results of micro-grid D2](image)
Figure 4. Load forecasting results of micro-grid R1

Figure 5. Load forecasting results of micro-grid R2

Table 1. Error prediction error of microgrid

| Micro grid | Average load/kW | MAPE (%) | MSE ($\times 10^{-2}$) |
|------------|-----------------|----------|------------------------|
| D1         | 140.85          | 10.03    | 7.62                   |
| D2         | 261.96          | 10.9     | 5.99                   |
| R1         | 1106.43         | 13.61    | 8.13                   |
| R2         | 1315.61         | 7.69     | 4.6                    |
4.3. Analysis of load forecasting results

In order to analyze the relationship between the load characteristics and the prediction accuracy, the load forecasting results of each micro grid are shown in Figure 2 to figure 5.

The results of the load forecasting of each micro grid are analysed.

1. The load forecasting errors of every microgrid in the holiday period between 25 and 72h (the area between the scribing lines) had a large load forecast. The main reason is that the load characteristics of holidays vary greatly, and the habits of electricity consumption are easy to mutate.

2. Microgrid load rise and decline along with the rapid change of power, it is also prone to have some prediction errors. Microgrid R1 is the most typical one, especially at the 31h load rising edge, which has the biggest prediction error of the prediction week.

3. The influence of the low load period of the microgrid on the prediction error. Although the MSE error of microgrid D2 is smaller than that of microgrid D1, the prediction error of D2 is at a low time, and the minimum load is only 116.4kW, which leads to the larger MAPE error of D2.

4. For D1 and R1 micro grid, the day before the similarity of load curve is low, it is easy to cause large error of load forecasting.

4.4. Performance analysis

The experiment machine uses the Core I3 380M processor and the 4G memory, and completes the related calculation in the Matlab 2012a environment. Taking the first week load forecast of July as an example, the single parameter optimization and single parameter optimization of each micro grid.

| Micro grid | Single parameter optimization time /s | Single load forecasting time /ms |
|------------|---------------------------------------|---------------------------------|
| D1         | 507.64                                | 0.508                           |
| D2         | 539.07                                | 0.530                           |
| R1         | 527.96                                | 0.565                           |
| R2         | 514.6                                 | 0.558                           |

Analysis table 2 shows that the short-term load forecasting method of the microgrid has high efficiency in two aspects of parameter optimization and load forecasting. In 10min, the optimization of the operation period of a microgrid can be completed, and the single load prediction time is less than 1ms.

5. Conclusion

The kernel extreme learning machine based on the proposed off-line parameter optimization and online load forecasting of micro grid short-term load forecasting methods include through example test and analysis, the following conclusions can be obtained: 1 on the various types of micro grid, the capacity, this method can achieve higher prediction accuracy; 2 by continuous micro grid load a multi week forecast, to verify the performance of stability in periodic updated prediction method; computational efficiency 3 the prediction method has high, easy to implement in embedded terminal device.

References

[1] Huang G B, Wang D H, Lan Y. Extreme learning machines: a survey [J]. International Journal of Machine Learning and Cybernetics, 2011, 2(2): 107 - 122.

[2] Huang G B, Zhou H, Ding X, et al. Extreme learning machine for regression and multiclass classification [J]. Systems, Man, and Cybernetics, Part B: IEEE Transactions on Cybernetics, 2012, 42 (2): 513 - 529.

[3] Cheng Song, Yan Jianwei, Zhao Dengfu, et al. Short-term load forecasting method based on ensemble improved extreme learning machine [J]. Journal of Xi’an Jiaotong University, 2009, 43 (2): 106 - 110.
[4] Mao Li, Wang Yuntao, Liu Xingyang, et al. Short-term power load forecasting method based on improved extreme learning machine [J]. Power System Protection and Control, 2012, 40 (20): 140 - 144.

[5] Toh K A. Deterministic neural classification [J]. Neural Computation, 2008, 20 (6): 1565 - 1595.

[6] Ghofrani M, Hassanzadeh M, Etezadi-Amoli M, et al. Smart meter based short-term load forecasting for residential customers [C]. North American Power Symposium (NAPS), 2011: 1 - 5.