International Journal of Electric Power and Energy Studies
ISSN: 3006-2004 (Print), ISSN: 3006-0826 (Online) | Volume 3, Number 1, Year 2024
DOI: https://doi.org/10.62051/ijepes.v3n1.05
Journal homepage: https://wepub.org/index.php/IJEPES/index

Load Forecasting Method based on Long Short-term Memory Network and Business Expansion Installation

Libo Cui, Jing Zhao, Chunwei Guan

State Grid Shandong Electric Power Company Qingdao Power Supply Company, Qingdao, 250022, China

ABSTRACT

In view of the problem that the traditional monthly load forecasting method lacks considering the internal load influencing factors, this paper puts forward the monthly load forecasting method considering the industry expansion. On the basis of the similar load development trend, the benchmark value of the planning annual forecast is determined. Then, the basic data of uncertainty and selection factors were collected, and the correlation model affecting the bias of load size factors and the percentage of load deviation was established based on the long-and short-term memory network, and then the benchmark value of monthly load prediction value was corrected. The comparison of the forecast results of considering the actual expansion, the uncorrected expansion and the industry expansion shows that the actual expansion has an important impact on the monthly load and can effectively improve the prediction accuracy.

KEYWORDS

Monthly Load Forecast; Business Expansion and Installation; Long-term and Short-term Memory Network.

1. INTRODUCTION

Load prediction is the key element of distribution network planning. In traditional planning, load prediction includes power elasticity coefficient method, single consumption method, departmental prediction method, per electricity quantity method, average growth rate method, linear growth trend method, exponential curve growth trend method, etc[1]. Under the background of the gradual opening of the power market, the traditional planning load forecast has been difficult to meet the demand of the new power system. On the one hand, it is necessary to further improve the accuracy and accuracy of load forecasting, provide data support for the planning and construction of the incremental market, improve the fine management; and the influence of new uncertainty factors in the construction of new power system, which is difficult to ignore the influence, especially the sensitivity of industrial users to the electricity price, which directly affects the total load change, but the specific mechanism is relatively complex and difficult and lacks sufficient data, so in the form of correction parameters[2-3].

In view of the shortcomings in the existing technology, this paper provides an active planning method of the power distribution network based on the load prediction correction. Compared with the traditional power grid load prediction, the quantitative correction of the influence of uncertain factors is added. Considering the impact of load expansion demand and peak and valley price difference, the benchmark value of planning annual forecast is determined on the basis of similar load development trend. Then, the basic data of uncertainty and selection factors were collected, and the correlation model of the deviation of load size was established based on LSTM, and the benchmark value of load prediction value was corrected. The example shows that the prediction method can effectively improve the accuracy of monthly load prediction.

2. ANALYSIS OF THE IMPACT OF INDUSTRY EXPANSION ON THE MONTHLY LOAD

Capacity is an important factor restricting customer demand for electricity. Business expansion capacity is the base for an impact on future load. However, after the business expansion, customers need to carry out the debugging of electrical equipment and other aspects, and the electricity consumption needs a certain time to achieve the stability. Therefore, the impact of the business expansion capacity on the load is gradual, and the stable cycle after the business expansion and the monthly impact ratio should be studied.

2.1. Analysis of Growth Curve of Power Expansion

Business expansion includes new installation, capacity increase, suspension, suspension and recovery, capacity reduction, capacity reduction and recovery, and account cancellation. Here to take the new installation business as an example for analysis. As the electricity consumption of customers increases rapidly after the newly installed business, and then tends to stabilize, showing a growth trend. Therefore, the growth curve Logistic curve model can be used to fit and analyze the electricity consumption of customers.

2.2. Extraction of the Growth Cycle of Industry and Power Expansion

The electricity growth curve obtained from the fitting is the power expansion of typical customers. In order to obtain the data representing the growth of electricity consumption after newly installed business in the industry, this paper uses k-mean clustering to group the power expansion of typical customers as the growth of electricity expansion in the industry.

K-mean clustering is a division-based clustering method, takes the distance between the data object and the cluster center as the basis. The specific steps of the algorithm are as follows:

Step 1: Arbitrarily select k data points from n data as the initial cluster center $\boldsymbol{M}_i\left(i=1,2,\cdots,k\right)$.

Step 2: Calculates the distance between each data point $\boldsymbol{X}_j\left(j=1,2,\cdots,n\right)$ and the cluster center $D_j=\left|\boldsymbol{M}_i-\boldsymbol{X}_j\right|$, and divides the data according to the minimum distance to form the cluster $\boldsymbol{C}_i\left(i=1,2,\cdots,k\right)$.

Step 3: Takes the mean of the data in each class cluster as the center of the updated class cluster.

$$M_i=\sum_{x_j\in C_i}X_j/N_i \tag{1}$$

Where N_i is the number of data contained in the class cluster $\boldsymbol{C}_i$.

Step 4: Repeats steps 2 and 3 until the cluster center no longer changes.

After obtaining the growth situation of the industry, we continue to fit with the growth curve, and the model with high fitting degree is selected for analysis. Since there will still be some fluctuation even after the electricity is stable, this paper analyzes the change of the electricity in the sample during the industrial expansion, and finds that the power consumption basically fluctuates within 7%, so the change of electricity consumption is not more than 7% as the stable judgment criterion. Then divide the stable electricity from the electricity of each month to get the growth proportion of the monthly electricity, and the proportion will be taken as the monthly impact proportion of the industry expansion capacity on the load.

In this paper, we select the newly installed customers with large electricity consumption in the machinery industry, without considering the new installed customers affected by the Spring Festival or policies, and finally select 10 customers. After fitting the power consumption of 10 customers with the growth curve, the k-mean clustering of the customers. The data obtained from the clustering were then fitted with a growth curve, and the results are shown in Figure 1. Further analysis of the curve shows that the stability period of the new installed business in the industry is 4 months, and the monthly impact ratio is 10%, 46%, 79% and 100% respectively.

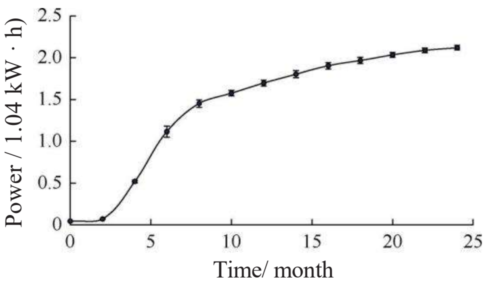

Fig. 1 Growth curve of certain industry

2.3. Actual Business Expansion and Increment Extraction

From the above analysis, it can be seen that the influence of the capacity of business expansion on the power load is delayed, and there is no direct connection between the capacity of business expansion capacity and the load of the month, so this paper considers to restore the increment of business expansion, which has an actual impact on the monthly load.

If the load stability month is n months, the application capacity of new installed business in month k is U_k, and the monthly load impact ratio after this business is $a1, \ldots, an$, then the application capacity of new installed business in month k in month $j(j \geq k)$is:

$$C_j = \begin{cases} U_k a_1 & j = k \\ U_k (a_{j-k+1} - a_{j-k}) & k < j < k+n-1 \end{cases} \tag{2}$$

The expansion business of different industries accumulates the expansion increment of the month, that is, the expansion increment that has an actual impact on the monthly load.

3. LOAD FORECASTING MODEL

Step 1: Obtain the substation capacity, time of use electricity price, industry type, daily electricity load, distribution transformer feeder and user access relationship, GDP annual growth rate of the regional power grid within a certain period of time, and form a sample dataset. Obtain relevant load indicators for different industries through k-means clustering, including monthly maximum load, monthly minimum load, monthly average load, monthly average daily load rate, monthly maximum peak valley difference, and monthly average daily peak valley difference rate. By fitting the growth curve and k-means clustering, the load growth cycle of different types of business expansion can be obtained, and the actual business expansion increment can be calculated by combining the monthly reported installation capacity. Construct feature vectors based on the sample set and processed load data: $X(t) = [x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t), x_8(t), x_9(t)]$.

Among them, $x_1(t), x_2(t), x_3(t), x_4(t), x_5(t), x_6(t), x_7(t), x_8(t), x_9(t)$ represents: peak valley electricity price difference, industry type, average load of the year before the forecast month, average load of the two years before the forecast month, average load of the month before the forecast month, average load of the two months before the forecast month, average daily load rate of the month, maximum peak valley difference of the month, and business expansion increment.

Step 2: Calculate the distance between each feature matrix and determine the similarity between factors related to monthly load changes in different industries on different dates based on the distance between feature matrices; Then, based on the judgment results of dynamic sorting based on similarity, obtain sample values with similar changes, and use their monthly maximum load value as the benchmark value $T'_{\text{base},\ t}$.

The feature distance is implemented using the grey correlation coefficient method, and the predicted feature vector for the day is $X = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, ..., x_m]$,m is the number of eigenvector components. The feature vector of the nth historical month's factors is $X(n) = [x_1(n), x_2(n), x_3(n), x_4(n), x_5(n), x_6(n), x_7(n), x_8(n), ..., x_m(n)]$ .Normalize the i-th feature vector component for the nth historical month as follows:

$$x_i'(n) = [x_i(n) - x_{i\min}(n)]/[x_{i\max}(n) - x_{i\min}(n)] \tag{3}$$

Where $x_i(n)$ represents the i-th eigenvector component, and $x_{i\max}(n)$ 、 $x_{i\min}(n)$ are the maximum and minimum values of the i-th eigenvector.

Using the grey correlation coefficient method, predict the correlation coefficient $\varepsilon_i(n)$ between the i-th historical feature vector of the month and the nth month.

$$\varepsilon_i(n) = \frac{\min_n \min_i |x_i(n) - x_i'(n)| + \rho \max_n \max_i |x_i(n) - x_i'(n)|}{(|x_i(n) - x_i'(n)|) + \rho \max_n \max_i |x_i(n) - x_i'(n)|} \tag{4}$$

The grey correlation R_n formula for the forecast month and n month is:

$$R_n = \frac{1}{m} \sum_{i=1}^{m} \varepsilon_i(n) \tag{5}$$

Step 3: Build a training dataset for the modified model based on the data sample set, define the actual load value of year t as T_{nt} , and determine the load value of similar time periods as T'_{nt} . Therefore, the monthly load value deviation is $\Delta T_{nt} = T_{nt} - T'_{nt}$; The deviation between the feature matrix $X(i)_{nt}$ of the i-th related factor in the nth year t month and the feature matrix $X(i)'_{nt}$ of similar time period related factors is $\Delta X_{nt} = X(i)_{nt} - X(i)'_{nt}$; Modeling the related factor deviation ΔX_{nt} and monthly load deviation ΔT_{nt} using the LSTM algorithm, with the related factor deviation ΔX_{nt} as the input and the monthly load deviation ΔT_{nt} as the output of the model. The monthly load deviation prediction model is obtained by training the sample set;

Step 4: Using the feature matrix $X(i)_{nt}$ of the relevant factors in the previous period of the predicted month and the determined feature matrix $X(i)_{nt}^{'}$ of the relevant factors in similar periods as inputs to the monthly load deviation prediction model, obtain the monthly load deviation ΔT_{nt} output by the monthly load deviation prediction model;

Step 5: Predict the monthly load according to the following equation:

$$T_{f,\ t} = \Delta T_t + T_{base,\ t}^{'} \tag{6}$$

In the formula, $T_{f,\ t}$ is the predicted monthly load value, ΔT_t is the monthly load deviation obtained in step 4, and $T_{base,\ t}^{'}$ is the monthly load benchmark value obtained in step 2.

4. SIMULATION EXAMPLES

Table 1. Forecasting results

month	actual value	Consider the actual increase in business expansion/(10^7kWh)		Not considering business expansion and installation/(10^7kWh)	
		PV	relative error/%	PV	relative error%
1.	16.37	16.70	0.32	18.81	2.43
2.	7.90	8.65	0.75	10.62	2.72
3.	14.79	15.32	0.52	18.17	3.37
4.	18.12	17.65	0.47	21.25	3.12
5.	17.20	16.82	0.38	19.90	2.70
6.	17.26	17.33	0.07	19.27	2.00
7.	17.25	16.78	0.48	18.66	1.41
8.	17.35	17.24	0.11	19.96	2.61
9.	16.94	16.60	0.34	19.84	2.91
10.	15.18	15.37	0.19	17.22	2.04
11.	16.60	16.13	0.47	19.15	2.55
12.	16.70	16.68	0.02	20.16	3.46
Average relative error/%			0.34		2.61

This article selects the monthly electricity consumption and expanded installed capacity data of the mechanical manufacturing industry in Shandong Province from 2017 to 2022 for simulation testing. The data from 2017 to 2011 were used as training samples to establish a prediction model; The data for 2022 is used as a prediction sample to test the prediction performance.

To verify the effectiveness of the method proposed in this article, a comparative analysis will be conducted on three scenarios: actual business expansion, consideration of uncorrected business expansion, and non consideration of business expansion and installation. Among them, considering uncorrected business expansion refers to the impact of various business expansion and installation services on load without lag, with a stable period of one month, and the calculated business expansion increment.

Combining the expansion of the growth cycle of the mechanical manufacturing industry, the results calculated using long short-term memory networks are shown in Table 1. Without considering the expansion and installation of the industry, there is a significant deviation in the forecast for January and February. This is due to the impact of the Spring Festival, and it is difficult to determine the change pattern based solely on historical electricity data. When considering business expansion and installation, the actual business expansion increment is adopted, which greatly improves the accuracy of electricity forecasting in January and February. This is because most customers suspend their business during the Spring Festival period, and the stable period of business suspension is one month. Therefore, the actual business expansion increment during the Spring Festival period has a corrective effect on the predicted values.

5. SUMMARY

This paper uses LSTM to predict the monthly load of power system, and studies the stable period and monthly influence ratio of power expansion. After the expansion of business, the impact of business expansion capacity on the load is gradually, and there is a lag. This analysis can be quantified by growth curve fitting and k-mean clustering. The load forecasting considering the actual business expansion increment and the peak-valley price difference provides more useful information for the monthly load forecasting compared to the forecasting method considering the uncorrected industry expansion increment and only the historical load, thus making the forecast more accurate. Through the extraction of the growth cycle of industry expansion, the industry expansion is optimized, the problem of "reporting large and using small" caused by the lag of actual industry expansion is solved, the transformer capacity is reasonably planned, and the economic benefit of power grid investment is improved.

REFERENCES

[1] Biswal B ,Deb S ,Datta S , et al.Review on smart grid load forecasting for smart energy management using machine learning and deep learning techniques[J].Energy Reports,2024,123654-3670.

[2] Jiang Mengyang, Cheng Haozhong, Wu Zhen, et al. Monthly Load Forecasting Method Using Vector Machines Considering Business Expansion and Installation [J]. Proceedings of the CSU-EPSA, 2017, 29 (07): 1-6.

[3] Zhou C ,Guo D ,Hu H , et al.TCN Multi-time-scale Transformation and Temporal-Attention Neural Network for Monthly Electricity Consumption Forecasting[J].Recent Advances in Electrical & Electronic Engineering,2023, 16(8):872-883.

[4] Sing C L ,Zhenyao M ,Ting W , et al.Load forecasting based on deep neural network and historical data augmentation[J].IET Generation, Transmission & Distribution,2020,14(24):5927-5934.