Incremental Distribution Network Forecasting for Different Industries Based on Long and Short Term Memory Network

Yongshang Ji, Wensheng Li, Shuhuai An, Zhi Yin, Jifu Qiu, Bo Peng

Qingdao Power Supply Company of State Grid Shandong Electric Power Company, Qingdao, 266002, China

Abstract. In order to improve the quantitative evaluation and prediction ability of power grid investment benefit and solve the optimization problem of investment rhythm caused by the increase of user load in the new park, an incremental distribution network load forecasting method based on long-term and short-term memory network is proposed. Considering multiple influencing factors of investment decision-making, grey correlation analysis is carried out on the factors affecting saturated load, so as to quantitatively determine the impact Degree and size. Then, the influencing factors are used as independent variables, and the demand of electric power or electricity is used as the dependent variables to establish the prediction model to realize the medium and long-term load high-precision forecasting under the condition of small sample and high uncertainty.

Keywords: Medium-and long-term load forecasting; deep learnining; industrial classification; incremental distribution network.

1. Introduction

Medium and long-term power load forecasting has important guiding significance for power grid planning scheme and investment and construction progress. At present, time series nonlinear trend forecasting [1], comprehensive forecasting model with weighted combination of various forecasting models [2], load forecasting model considering urban load fluctuation and economic impact [3-4], intelligent forecasting algorithm based on neural network [5] have been proposed at home and abroad, but the above methods are extrapolated by intelligent analysis method on the basis of historical load, At the same time, multiple factors are considered to improve the accuracy of prediction. For newly developed urbanized new areas or new parks, due to the lack of historical sample data, the traditional medium and long-term load forecasting algorithm is difficult to meet this kind of incremental load forecasting.

In view of the characteristics of few historical samples, literature [7] uses Monte Carlo simulation analysis method to deal with the uncertainty of data, and literature [8] proposes a multi factor grey neural network combined prediction model, which can simultaneously reflect the characteristics of nonlinear fluctuation and growth trend under the influence of multiple factors, Literature [9] uses the combined load forecasting model based on support vector machine and Markov chain to deal with the problem of multi factor influence. Literature [10] proposes a forecasting method based on gm-grnn to weaken the randomness and increase the regularity in the original data. Literature [11] uses the associated fuzzy neural network to improve the forecasting accuracy.

The load forecasting algorithm based on artificial intelligence has been well applied in the overall load forecasting of power system. For single industry users, this paper proposes an incremental distribution network prediction method based on long-term and short-term memory network. Aiming at the problem of insufficient sample number of related factors for load forecasting of newly-built industrial users, through the analysis of the growth law of user load in different industries and different formats, the characteristic labels including industry, operation years, annual power consumption, annual peak load and so on are extracted to train the forecasting model. Input the industry attributes, operating years and other information of new industrial users into the trained neural network model to realize the medium and long-term load forecasting. Control the investment rhythm according to the predicted load level under different time scales in the current year, the next 5 years and the next 10 years.

2. Load forecasting information model of different business types

The power load of different business types can reflect the current industrial energy consumption level and the fluctuation of business types. The load characteristic indicators are the user's annual power consumption and annual peak load. Regional medium and long-term load

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

forecasting is affected by economic factors, income level and consumption concept, industrial structure adjustment, power supply capacity, climate and temperature, policy factors, power demand side management, etc.. Among them, economic factors directly affect the power demand of users in different industries, and the income level and consumption concept affect the living power consumption level and the power consumption level of general industry and commerce. The adjustment of industrial structure affects the power consumption of heavy industry, the power supply capacity affects the annual maximum load, the temperature and climate affect the temperature sensitive load in the power load, and the policy factors and power demand side management affect the power consumption structure.

The influence degree of different influencing factors on different industries is different, and the law of power consumption is different, but the growth of industrial output value has a strong correlation with the growth of industrial user load, and on the time scale of year,. According to the load characteristics of users in different industries and specific industrial users, the factors affecting the medium and long-term load include economic factors, industry type and user operation years. See Table 1 for details.

Tab.1 Information model

Number	Characteristic quantity
1.	Industry type
2.	Operating years
3.	GDP
4.	Industrial output value
5.	Investment scale
6.	Annual power consumption
7.	Annual peak load

According to the classification standard of power consumption in national economy, the types of industrial industries are preliminarily divided, including mining, timber and bamboo mining and transportation; Tap water production and supply; Food, beverage and tobacco manufacturing; Textile industry; Papermaking and paper products industry; Production and supply of electricity, steam and hot water; Petroleum processing industry; Coking, gas and coal products industry; Machinery manufacturing; Electronic equipment manufacturing; Metal products industry; Building materials processing industry, etc. In order to improve the accuracy of prediction, this paper further subdivides the types of industrial industries, including cement industry, machinery manufacturing industry, metal smelting industry, washing machine manufacturing industry, automobile parts processing, garment manufacturing, ceramic manufacturing and other industries.

In order to master the power consumption law of different industries and improve the accuracy of medium and long-term load forecasting in different industries, it is necessary to analyze the power consumption characteristics of different industries and further classify the loads according to industries.

In this paper, fuzzy c-means FCM (fuzzy c-means) algorithm is used to classify various industries. FCM algorithm can divide a sample set without category mark into several subsets according to certain criteria and laws when the boundary between things is not clear, so that similar samples can be classified into one class and dissimilar samples can be classified into different classes. The correlation matrix R obtained from the above grey correlation analysis is studied as the eigenvector matrix R = [R1, R2,..., RN] of FCM algorithm, and the eigenvectors RJ (J = 1, 2,..., n) of N industries are divided into C categories.

Good clustering results should meet the small gap between the same type of industries and the large gap between different types of industries. Each industry should have a large membership degree for its classification. This paper introduces three indicators to reflect the advantages and disadvantages of clustering effect from different aspects: I1 describes the distance between all industry feature vectors and their clustering centers under the same classification. The smaller the indicator, the better the classification effect; I2 describes the differences between different classifications. The larger the index, the more obvious the differences between different classifications; I3 describes the average membership of all industries. The larger the index, the clearer the classification. The calculation formula is:

$$\begin{cases} I_1 = \sqrt{\dfrac{1}{c}\sum_{i=1}^{c}\sum_{j}^{x_j \in i} d_{ij}^2} \\ I_2 = \dfrac{1}{(c-1)c}\sum_{i \neq j} d(\boldsymbol{C}_i, \boldsymbol{C}_j) \\ I_3 = \dfrac{1}{n}\sum_{j=1}^{n} U_{j,\max} \end{cases} \tag{1}$$

Where: $d(C_i,C_j)$ is the Euclidean distance between the cluster centers of classification i and classification j; $U_{j,Max}$ is the maximum value of the membership degree of industry j eigenvector.

Select different category numbers c, use FCM algorithm for clustering calculation, compare the values of three indicators under each category number, and select the best category number.

3. Medium and long term load forecasting model based on LSTM

It is difficult to establish an accurate data model to fit between the multiple factors affecting the medium and long-term load forecasting of industry users and the forecasting results. Therefore, a neural network forecasting method with strong adaptability is considered to comprehensively consider a variety of factors, so as to improve the accuracy of forecasting. In order to make up for the lack of historical data of the industrial park, the medium and long-term load influencing factors of various industries are analyzed as the input of the prediction model.

3.1 LSTM network structure

The hidden layer of LSTM is composed of triple gate structure. In the process of forward calculation, each memory unit is composed of input gate, forgetting gate and output gate, and each gate judges whether to transfer the operation result downward according to the activation function.

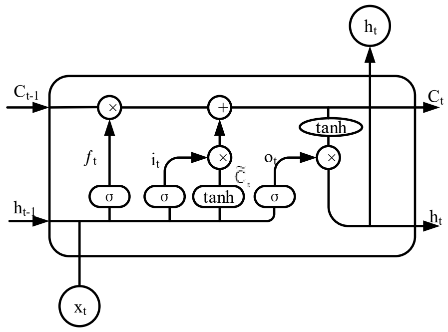

Fig. 1 Topology of LSTM

The transmission of LSTM signal mainly experiences the selection, update and output of signal. Judge whether all or part of the information learned at the previous time passes or not according to the output of the previous time and the input of the current time. At the same time, the output is mapped to the hidden layer to generate new candidate values that may be added to the unit status, and the two parts of values are combined for updating. After an initial output value is obtained, it is scaled and multiplied by the output value to obtain the final output value. When LSTM is used for long-term prediction, the new prediction value can replace the old prediction value in turn for the prediction of the next stage, so as to realize circular prediction.

1) Forget the door floor. The value between 1-0 and 1-0 of the cell is calculated as follows:

$$\mathrm{f_t} = \sigma(\mathrm{W_{xf}x_t} + \mathrm{W_{hf}h_{t-1}} + \mathrm{b_f}) \quad (2)$$

2) Enter the door level. Control the proportion of input, take the input x_t of the current layer and the output h_{t-1} of the hidden unit at the previous time as the input, and the output result i_t as the information to be updated. The formula is as follows:

$$i_t = \sigma(W_{xi}x_i + W_{hi}h_{t-1} + b_i) \quad (3)$$

Then update the cell state. The new cell state C_t is determined by the output results of the old state, forget gate and input gate. The formula is as follows:

$$\tilde{C}_t = \tanh\,(W_{xc}x_t + W_{hc}h_{t-1} + b_c) \quad (4)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (5)$$

3) Output gate layer. First, run a sigmoid layer to determine which part of the cell state will be output. Then, the cell state is processed through a tanh and multiplied by the output of sigmoid gate as the output.

$$o_t = \sigma(W_{xo}x_t + W_{ho}h_{t-1} + b_o) \quad (6)$$

$$h_t = o_t * \tanh\,(C_t) \quad (7)$$

Where and tanh are activation functions; W_x represents the weight from the input layer to the hidden layer; W_h represents the weight of recursive connection; b represents the offset term.

3.2 Multi dimensional and multi-step prediction of annual power consumption and annual maximum load of industrial users

Based on the network characteristics of LSTM and the incremental distribution network based on long-term and short-term memory network, this paper builds a three-layer single-cell neural network model. The input variables are the data in the established information model, and the output variables are annual power consumption and maximum load.

According to formula (6) and formula (7), the following training is carried out. The specific steps are as follows:

Initialization of data samples. The method of min max is adopted to normalize the data to unify the data between [0,1] and eliminate the difference caused by dimension. The calculation formula is as follows:

$$I = \frac{i - i_{\min}}{i_{\max} - i_{\min}} \quad (8)$$

(2) Preparation of input variables. The data of load forecasting information model is obtained through data preprocessing, and the associated information model sample data is obtained through K-means clustering as the data input of LSTM.

(3) Determination of the number of hidden layer nodes. In general, the number of nodes in the hidden layer is determined according to the following formula:

$$N = \sqrt{n+m} + a \quad (9)$$

Where, N represents the number of hidden layer nodes, n represents the number of input nodes, m represents the number of output nodes, and a is an integer between 1 and 10.

Select the most appropriate number of hidden layer nodes according to the minimum root mean square error. The root mean square error formula is as follows:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(Y_0 - Y_p)^2}{s}} \quad (10)$$

The root mean square logarithm error is:

$$RMSLE = \sqrt{\frac{1}{s}\sum_{i=1}^{n}(\lg(Y_0+1)^2 - \lg(Y_p+1)^2)} \quad (11)$$

The determination coefficient is:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(Y_0 - Y_p)^2}{\sum_{i=1}^{n}(Y_0 - Y_{mean})^2} \quad (12)$$

Where, s is the number of samples, Y_0 is the original value, Y_p is the predicted value and Y_{mean} is the sample mean value.

(4) Initialization of network parameters. The LSTM network is implemented using the keras framework. Keras can call tensor flow and adjust the maximum number of iterations, learning rate, activation function and other parameters.

(5) Training of network model. BPTT training algorithm is used for gradient detection, and SGD algorithm is used to update the weight until the requirement of minimum error is met.

In the experiment, the characteristic information in the training sample set is used as the input of the network, the network model is trained offline, the parameters are initialized, the learning step is set to 0.01, the input dimension is 16, the number of nodes in the hidden layer is 32, and the number of nodes in the output layer is 7, that is, the structure design of the network is 16-32-7, all elements of the weight parameter matrix are initialized to (-1,1), and the error limit error is error_Gate set to 1×10^{-4}。Figure 2 shows the variation curve of the average error of LSTM model with the number of iterations in the training process.

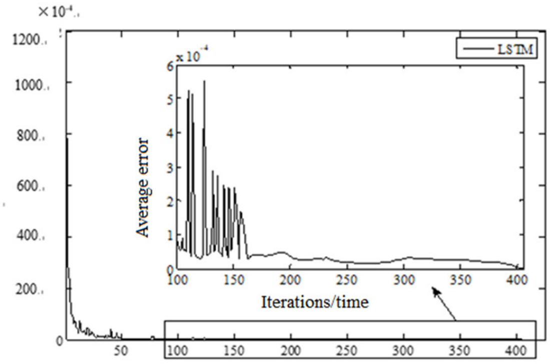

Fig. 3 Average error curve

(6) Error standard. In order to verify the superiority of the proposed prediction model, the absolute value of the maximum relative error and the absolute value of the average relative error are used as the error standards, as shown below:

$$E_m=\frac{\max\left|\hat{y}_t-\hat{y}_{st}\right|}{\hat{y}_{st}}\times100\% \tag{12}$$

$$E_a=\frac{1}{n}\sum_{1}^{n}\frac{\left|\hat{y}_t-\hat{y}_{st}\right|}{\hat{y}_{st}}\times100\% \tag{13}$$

Where, $\hat{y}_t$ is the prediction result of the target year; $\hat{y}_{st}$ is the actual value of the target year; t is the target year, $t=1\cdots n$.

4. Case analysis

In order to verify the effectiveness of the method proposed in this paper, the actual annual power consumption and maximum load data of users in cement industry, machinery manufacturing industry, metal smelting industry, washing machine manufacturing industry, automobile parts processing, garment manufacturing, ceramic manufacturing and other industries in a province from 2009 to 2018 are collected to test the proposed prediction algorithm. The basic information of an industrial user in Qingdao is used for prediction.

Industry classification, operating years, regional GDP, industrial output value and investment scale are taken as model inputs, and annual power consumption and annual peak load are taken as model outputs.

Tab. 2 Binary code of industry classification

Industry classification	Corresponding code
Cement industry	(1,0,0,0,0,0,0,0)
Machinery manufacturing industry	(0,1,0,0,0,0,0,0)
Metal smelting industry	(0,0,1,0,0,0,0,0)
Washing machine manufacturing industry	(0,0,0,1,0,0,0,0)
Automobile parts processing	(0,0,0,0,1,0,0,0)
Garment manufacturing	(0,0,0,0,0,1,0,0)
Ceramic manufacturing	(0,0,0,0,0,0,1,0)
Other industries	(0,0,0,0,0,0,0,1)

Tab. 3 Power consumption of some industries in recent three years

Industry classification	Building material processing industry	Machinery manufacturing
Operating life (year)	12	4
Electricity consumption in 2018 /kWh	62000000	58035
Maximum load in 2018 /kw	17000	640
2017 electricity consumption/ kWh	58500000	54750
Maximum load in 2017 /kw	16000	600
Electricity consumption in 2016 /kWh	5518800	51000
Maximum load in 2016 /kw	15094	540

Tab. 4 Historical value and growth rate of load influencing factors

Year	GDP of the whole society / 100 million yuan		Fixed investment / yuan		Industrial output value / yuan	
	Actual value	Growth rate%	Actual value	Growth rate%	Actual value	Growth rate%
2013	7 627.52	11.98	5 006.32	47.07	13 384.25	2.62
2014	9 357.64	22.68	6 511.42	30.06	17 107.19	27.82
2015	11 480.32	22.68	7 510.67	15.35	21 528.34	25.84
2016	13 110.87	14.20	8 871.31	18.12	24 194.13	12.38
2017	14 689.94	12.04	10 121.21	14.09	27 283.28	12.77
2018	16 002.98	8.94	11 654.09	15.15	29 686.33	8.81

Input the sample data according to the prediction model, predict the user's power consumption from 2019 to 2020, and select the linear regression model for comparison. The results are shown in Table 5.

Tab.5 Electricity consumption forecast

Year	Actual value / kW·h	linear regression model/kW·h		Prediction model in this paper/kW·h	
		Predictive value	Relative error%	Predictive value	Relative error%
2019	85113	88987	4.55	84828	-0.33
2020	86160	94266	9.41	86605	0.52
E_m	—	9.41		0.52	
E_a	—	6.98		0.43	

The absolute value of average relative error is 0.43%, and the absolute value of maximum relative error is 0.52%. The prediction accuracy is much higher than that of linear regression model, which can provide reference for medium and long-term load forecasting of new users.

5. Conclusion

In this paper, an incremental distribution network load forecasting method based on long-term and short-term memory network is proposed. Considering multiple influencing factors affecting investment decision, the factors affecting saturated load are analyzed by grey correlation degree, and the influence degree and size are determined by specific quantification. The forecasting model is established with the influencing factors as independent variables and power or electricity demand as dependent variables High precision medium and long-term load forecasting under high uncertainty can be used as a forecasting method of medium and long-term load in the future.

Acknowledgements

Fund Project: Science and technology project of State Grid Shandong electric power company (5206021800D8).

References

1. ZOU Pinjing, YAO Jiangang,KONG Weihui,et al.Mid-long Term Power Load Forecasting Based on Multivariable Time Series Inversion Self-memory Model[J]. Proceedings of the CSU-EPSA, 2017,29(10):98-105.
2. YUAN Tie-jiang, YUAN Jian-dang,CHAO Qin,et al.Study on the comprehensive model of mid-long term load forecasting[J]. Power System Protection and Control ,2012,40(14):143-146.
3. LI Yiyan, YAN Zheng, FENG Donghan.Mid/long-term load forecasting model considering urbanization[J].Electric Power Automation Equipment, 2016, 36(04):54-61.
4. Zhang Guanying,Xian Yiming,Ge Leijiao,et al. Medium and long-term load forecasting model based on Verhulst-SVM under new normal economy[J]. Electrical Measurement & Instrumentation,2019, 56(1):102-107.
5. HI Jiaqi,ZHANG Jianhua.Load Forecasting Based on Multi-model by Stacking Ensemble Learning[J]. Proceedings of the CSEE,2019,39(14):4032-4042.
6. Liu Hengyong,Liu Yongli,Deng Shicong,et al.A power load identification method based on LSTM model [J]. Electrical Measurement & Instrumentation, 2019, 56(23):62-69.
7. ZHENG Zhi-jie, LI Lei, ZHAO Lan-ming.Medium and long term load forecasting considering data uncertainty[J].POWER SYSTEM PROTECTION AND CONTROL,2011(07):130-133.
8. GU Xi-hua, XING Mian,NIU Dong-xiao,et al. Medium and long-term load forecasting based on multifactor-influenced gray artificial neural network optimal combination algorithm in power system[J]. JOURNAL OF NORTH CHINA ELECTRIC POWER UNIVERSITY,2007,34(6):32-36.
9. Chen jianyong, SU yihao.A Forecasting Model of Medium/Long Term Power Load in Combination of the Support Vector Machine and Markov Chain Algorithms[J].Southern Power System Technology, 2012,6(1):54-57.
10. WU yaohua.Long term load forecasting based on GM-GRNN in power system[J]. RELAY, 2007, 35(6):45-53.
11. ZHAO Zhipu,GAO Chao, SHEN Yanxia,et al.A Method for Load Forecasting Based on Correlated Fuzzy Neural Network and Improved Artificial Bee Colony Algorithm.Electric Power,2018,51(2):54-60.