Comparison of the Accuracy of Predicting Electricity Revenue by Predicting Electricity Quantity and Predicting Electricity Revenue by Historical Received Electricity Fees

Weipeng Qi*, Heng Yue, Siyang Xu, Xin Jin, Leilei Zhao

{Weipeng Qi*: feixue_7518@126.com}
{Heng Yue: yinhongling_123@163.com}
{Siyang Xu: hzgongdian1@163.com}
{Xin Jin: sxqiantanjiang@163.com}
{Leilei Zhao: yangxiaokun122@163.com}

State Grid Huitongjincai (Beijing) Information Technology CO., LTD Beijing ,100089, China

Abstract. Through in-depth research and analysis of historical data on regional electricity sales, we have successfully revealed the potential laws and unique characteristics of its development. Based on this research, we propose an innovative monthly electricity sales forecasting method. This method first predicts the quarterly electricity sales of the target month in the quarter, and then accurately predicts the electricity sales of the target month based on the proportion to the quarter and the predicted value of the quarterly electricity sales in the quarter. At the same time, we have considered the impact of the Spring Festival period and made revisions as needed. We used this new method to predict the monthly electricity sales of a city in Jiangsu from April 2020 to December 2022. The results show that our prediction has an average relative error of 2.35%, which is significantly improved compared to the previous method. This demonstrates the excellent performance of our proposed prediction method in terms of accuracy and reliability. This innovative method provides an effective and feasible new approach for monthly electricity sales forecasting, and provides useful references for research and practice in related fields.

Keywords: Data mining analysis; Winston method; Monthly electricity sales forecast.

1 Introduction

Electricity fee income is an important economic indicator in the operation of power supply enterprises. In China, unlike other commodities that adopt on-site equivalent transactions, electricity products are purchased before use. Therefore, the electricity fee income of power supply enterprises is not only affected by the amount of electricity used by users, but also by the full payment of electricity fees by users, which includes many random factors. At present, there are two main methods for predicting electricity revenue: one is to use time series models (including AR model, ARMA model, ARIMA model, etc.) for prediction[1]. Although these models are easy to operate, they have high data requirements; The second is to use machine learning methods for prediction. Although the model is relatively complex, it has high

MSIEID 2023, December 08-10, Guangzhou, People's Republic of China

DOI 10.4108/eai.8-12-2023.2344731

accuracy in organizing and fitting parameters, and has good performance in fitting any nonlinear trend[2].Based on this, this paper conducts research and analysis on the distribution patterns of historical data in regions, explores their development patterns and characteristics, and innovatively proposes a new method for monthly electricity sales prediction, providing an optional new tool for regional electricity forecasting work.

2 Data Analysis and Mining

The core of the LSTM network model is memory cells, which are responsible for storing and transmitting information. The memory cell consists of a linear unit and a nonlinear unit. The linear unit is a simple adder for summing the memory cell of the previous moment and the input from the current moment. The nonlinear unit is a sigmoid function that controls the flow of information.The input gate is used to control the input of the information. It consists of a sigmoid function and a point-multiplication operation. The sigmoid function is used to convert the input information into values between 0 and 1, and the point multiplication operation is used to multiply the input information with the output of the sigmoid function. The output of the input gate will be added to the memory cell.The forgetting gate is used to control the forgetting of information. It consists of a sigmoid function and a point-multiplication operation. The sigmoid function is used to convert the memory cell of the previous moment and the input of the current moment into values between 0 and 1, and the point multiplication operation is used to multiply the memory cell of the previous moment with the output of the sigmoid function. The output of the forgetting gate will be subtracted from the memory cell.The output gate is used to control the output of the information. It consists of a sigmoid function and a point-multiplication operation. The sigmoid function is used to convert the memory cell of the current moment and the input of the current moment into values between 0 and 1, and the point multiplication operation is used to multiply the memory cell of the current moment with the output of the sigmoid function. The output of the output gate will be used as the output of the current moment.The training method of LSTM network model is similar to the traditional RNN, using the back propagation algorithm. In the back-propagation algorithm, we need to compute the gradient of the loss function over the network parameters. However, the gating unit in the LSTM network model makes the calculation of the gradient more complicated. To solve this problem, we can adopt a method called "backpropagation weighting".The core idea of backpropagation weighting is to multiply the gradient of the gating unit by a weight for a greater contribution to the gradient. Specifically, we can multiply the output of the gating unit with the input of the gating unit to obtain a weight that can be multiplied by the gradient of the gating unit.

The core of load forecasting is to establish a mathematical model based on the historical data of the predicted object to express its development and changes, in order to obtain reasonable and reliable prediction results. Therefore, before establishing a mathematical model, it is necessary to conduct a comprehensive statistical analysis of the available historical data collected, and to study and explore the underlying laws of the development of historical data[3].The statistics, analysis, and mining of historical data refer to the organization and analysis of historical data in a certain region through the entire forecast. The electricity sales in each quarter of the year have seasonal periodic changes. Based on Figure 1, it can be observed that the electricity sales in a single quarter gradually increase with the progress of the year. In order to identify the relationship between the sales of electricity in each quarter over the years and the sales of electricity in each month of that quarter, the term "quarterly promotion" ζ is defined as the proportion of the sales of electricity in each month to the sales of electricity in that quarter, where ζ=the sales of electricity in a certain month/the sales of electricity in that quarter[4]. The results of historical data statistics are shown in Table 1. Research Table 1 found that except for January and February, the quarterly proportion ζ of all calendar years is very stable, with small fluctuations around a stable value. Taking October as an example, Figure 2 vividly illustrates the stability of the quarterly proportion ζ values for each month (excluding January and February). Due to the impact of the Spring Festival on electricity sales in January and February, as shown in Figure 3 and Figure 4, the fluctuation of quarterly proportion ζ values in these two months is significant; Therefore, when establishing mathematical models, factors affecting the Spring Festival should also be considered[5].

Table 1 Quarterly proportion of each month over the years(*indicates the month of the Spring Festival of that year).

Month	2015	2016	2017	2018	2019	2020	2021	2022
1	0.3592	*0.3082	0.3571	0.3425	*0.2913	0.3559*	*0.3157	0.3787
2	*0.2894	0.3227	*0.2667	*0.2874	0.3353	0.2645	0.3040	*0.2478
3	0.3514	0.3692	0.3762	0.3700	0.3734	0.3796	0.3802	0.3735
4	0.3243	0.3360	0.3187	0.2983	0.3378	0.3253	o.3330	0.3237
5	0.3330	0.3324	0.3314	0.3952	0.3283	0.3336	0.3294	0.3321
6	0.3427	0.3316	0.3498	0.3065	0.3338	0.3411	0.3377	0.3442
7	0.3460	0.3530	0.3432	0.3318	0.3393	0.3468	o.3334	0.3333
8	0.3461	0.3336	0.3384	0.3486	0.3469	0.3389	0.3560	0.3619
9	0.3079	0.3134	0.3184	0.3196	0.3138	0.3143	0.3090	0.3048
10	0.3188	0.3204	0.3421	0.3210	0.3225	0.3252	0.3209	0.3527
11	0.3264	0.3210	0.3255	0.3240	0.3277	0.3201	0.3203	0.3230
12	0.3548	0.3586	0.3504	0.3550	0.3498	0.3547	0.3588	0.3513

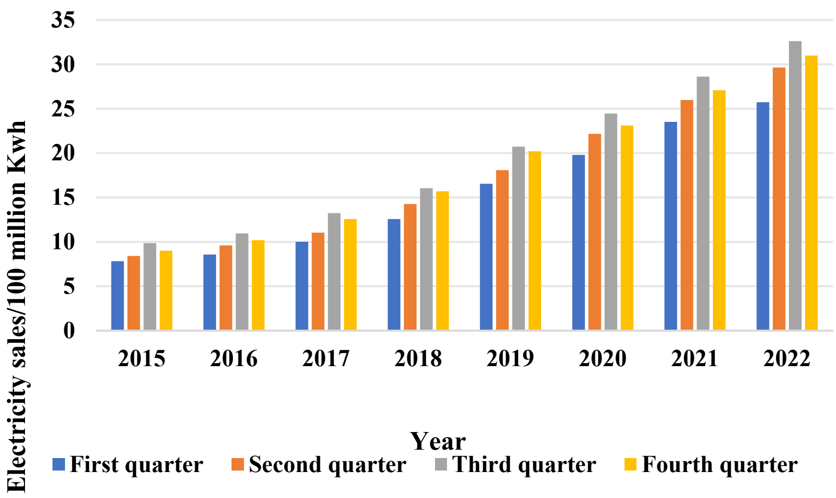

Fig. 1 Schematic diagram of the trend of electricity sales in a single quarter.

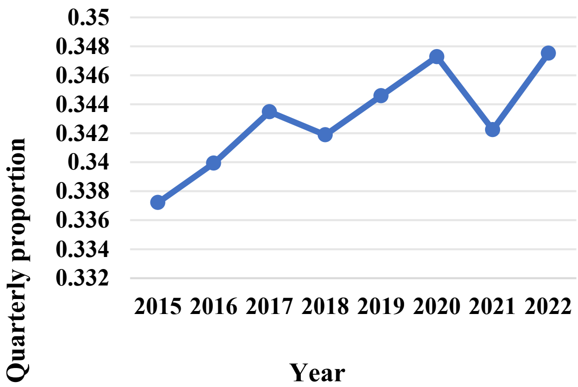

Fig. 2 Comparison Chart of Quarterly Proportion in October over the Years.

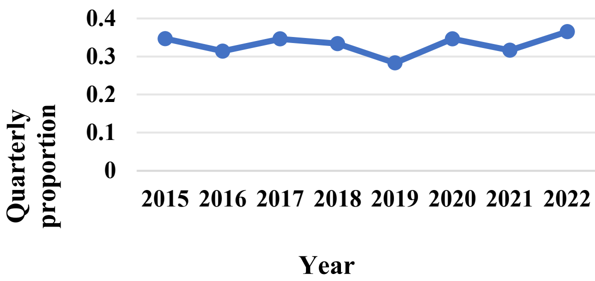

Fig. 3 Comparison chart of quarterly proportion in January over the years.

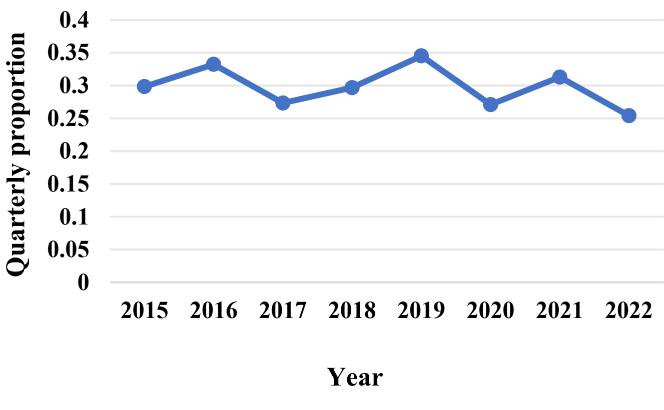

Fig. 4 Comparison chart of quarterly proportion in February over the years.

3 Establish a new method for predicting monthly electricity sales

Establishing a correct load forecasting model is a crucial step in electricity forecasting. Sometimes, due to improper model selection, the prediction error is too large, and it is necessary to change the model. Based on the mining and analysis of regional data over the years, the following important conclusions have been drawn: the sales of electricity in each quarter over the years have both a gradual growth trend and seasonal periodic changes in time series; Except for January and February, the quarterly proportion ζ of all other calendar years is very stable, with slight fluctuations around a stable value. When establishing mathematical models, factors affecting the Spring Festival should be considered in January and February[6].In order to fully consider the characteristics of seasonal cyclical changes and gradual growth of electricity sales, and to reduce errors caused by random fluctuations in a single month due to short time, the Winston method is used to predict the quarterly electricity sales of a certain month.The Winston method is a seasonal prediction method that combines factor analysis of time series with linear trends, seasonal changes, and regular changes, and exponential smoothing method. This method has three smoothing equations, which perform exponential smoothing on long-term trend S_t, trend increment b_t, and seasonal variation F_t. Then, the three smoothing results are combined with a prediction formula for extrapolation and prediction.The three smoothing formulas are:

$$S_t = \alpha \frac{Y_t}{F_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}) \tag{1}$$

$$b_t = \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1} \tag{2}$$

$$F_t = \gamma \frac{Y_t}{S_t} + (1-\gamma)F_{t-L} \tag{3}$$

In the formula: L is the length of the seasonal cycle, L=4; α、β、γ are the smoothing coefficient, both between 0 and 1; $Y_t$is the electricity sold during period t; The smoothing coefficients α、β、γ generally rely on experience to determine the initial values, then use historical sequences to make predictions and test them against actual errors, and finally use optimization methods to find the optimal values. Based on equations (1), (2), and (3), by selecting historical data from the first 20 quarters, the long-term trend S_t, trend increment b_t, and seasonal variation F_t for period t can be calculated. According to equation (4), the quarterly electricity sales during the t+k period can be calculated, which predicts the quarterly electricity sales for the predicted month.

$$\mathrm{E(n,m)} = \frac{\sum_{i=1}^{5} a_i \xi(n-i,m)}{\sum_{i=1}^{5} a_i} \times \mathrm{Y} \tag{4}$$

In the formula: a_i is the weight coefficient; E (n, m) is the predicted month m electricity sales in the nth year; ξ(n, m) is the quarterly promotion in month m of the nth year[7].It is generally believed that the impact period of the Spring Festival is from 2 days before the Spring Festival to 7 days after the Spring Festival.

$$\lambda_n = X_1/X_0 \tag{5}$$

$$\mathrm{E(n,1)'} = (31 - \mathrm{D_1})\mathrm{X_0} + \mathrm{D_1X_1} \tag{6}$$

$$\mathrm{E(n,2)'} = (31 - \mathrm{D_2})\mathrm{X_0} + \mathrm{D_2X_1} \tag{7}$$

$$\mathrm{E(n,1)'} + \mathrm{E(n,2)'} = \mathrm{E(n,1)} + \mathrm{E(n,2)} \tag{8}$$

According to equations (5), (6), (7), and (8), the predicted values E (n, 1) and E (n, 2) of January and February electricity sales in the nth year are known, and λ_n ,D,D_1,D_2 are also known. Therefore, the predicted corrected values $\mathrm{E(n,1)'}$, $\mathrm{E(n,2)'}$ for January and February in the nth year can be obtained.Before selecting a prediction model, first make a scatter plot of the city's historical electricity sales from 2013 to 2022 to predict its development trend. The historical electricity sales data is shown in Table 2.

Table 2 Electricity sold in a certain place from 2013 to 2022 (one billion kWh).

Time	2013	2014	2015	2016
Selling electricity	62.01	80.8	96.03	115.71
Time	2017	2018	2019	2020
Selling electricity	136.57	150.63	495.37	204.72
Time	2021	2022		
Selling electricity	233.52	246.31		

4 Experimental Results and Analysis

In order to verify the practical application of the new model proposed in this paper, the monthly electricity sales of some cities in Jiangsu from April 2020 to December 2022 are estimated. Using the new model, the estimated minimum estimation error is 0.05, the estimated estimation error is 5.1, and the estimated estimation error is 2.35. Using gray prediction, the minimum prediction error is 0.09, the maximum prediction error is 24.1, and the maximum prediction error is 9.06; Using the neural network prediction method, the minimum prediction error is 0.28, the maximum prediction error is 20.54, and the average prediction error is 6.78. Figure 5 shows that the prediction of the new model is very satisfactory, the prediction accuracy is better and more stable, which proves the application of the model. This new method is a short-term forecasting method for monthly electricity sales. is a possible method [8].

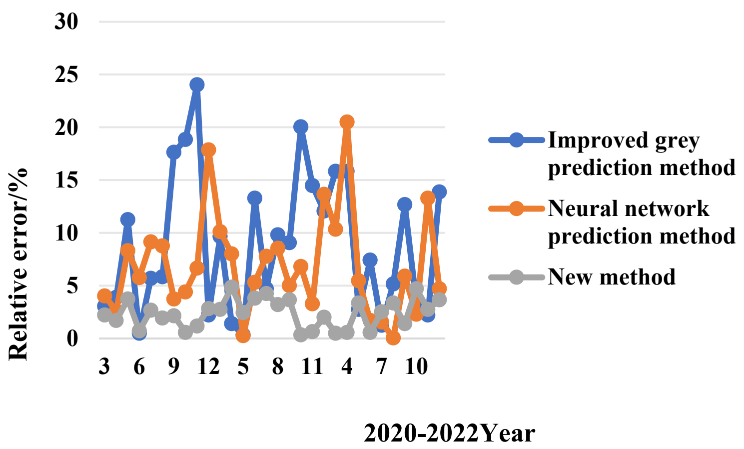

Fig. 5 Comparison of prediction accuracy of different prediction methods.

The reason why this new method has high accuracy and stability is due to the precise and in-depth mining of actual historical data in the region, which enables the new method to better reflect the stable factors of monthly electricity sales changes (quarterly proportion ζ), the increasing trend of monthly electricity sales changes, the periodicity of seasonal sudden changes, and the impact of the Spring Festival, thereby reducing the errors caused by the large random fluctuations in a single month due to short time[9-10].

According to the results of customer risk level classification of electricity bill recovery risk, it can be applied in actual business scenarios. It is suggested that labels and application scenarios can be designed for different positions such as meter reader and management. At each stage of the electricity bill recovery work, the meter reader uses the customer risk level label to screen customer groups and guide the electricity bill recovery work. First, in the meter reading stage, for high-risk, medium-risk customers, try to give priority to meter reading, face to face meter reading. According to the prompts of the mobile terminal, verify the contact information of the user, including the account number, account name, contact telephone number, mailing address, etc. Second, in the working stage of sending electricity and electricity charges, for high-risk and medium-risk customers, paper notice should be pasted and notify customers face to face as far as possible; for low-risk customers, electronic bill notification can be sent, including SMS bill, WeChat bill, email bill, etc. Third, in the working stage of sending fee notice, timely follow up the electricity payment of high-risk and medium-risk customers, and increase the frequency and intensity of payment. And according to the resource situation, take the telephone way to urge fees. For the management team, order the risk level, or increase the sending frequency of SMS messages for high-risk users, differentiate the charging content, and reduce the sending frequency of SMS messages for low-risk users.

5 Conclusion

Through in-depth research and analysis of the sales data of electricity in the region over the years, we have discovered some interesting patterns and characteristics, which are of great significance for understanding the development trends and patterns of the electricity market in the region. The following are the main conclusions of our research findings:However, in addition to the overall growth trend, we also found seasonal cyclical changes, where seasonal factors have an impact on electricity sales, leading to fluctuations within the quarter.The effectiveness of the new method: Our proposed method has been validated in monthly electricity sales forecasting. The analysis results of the example show that this method can better capture the stationary factors of monthly electricity sales changes (quarter to quarter ratio) ζ）、 The periodicity of growth trends, seasonal mutations, and the impact of the Spring Festival. By reducing the errors caused by short-term random fluctuations in a single month, this method improves the accuracy of monthly electricity sales prediction.Practical application value: We believe that our proposed new method for predicting monthly electricity sales is not only theoretically feasible, but also achieves good prediction accuracy in practical applications. This provides useful methodological support and necessary decision-making tools for power supply companies to strengthen power demand forecasting and power load management.

References

[1] Kim, H. G. . (2021). Effects of climate change on electricity consumption: a decomposition of industrial, residential, agricultural, and commercial sectors. Climate Change Economics (CCE), 12(44),96-102.

[2] Zhe, W. , Tianzhen, H. , Han, L. , & Piette, M. A. . (2021). Predicting city-scale daily electricity consumption using data-driven models - sciencedirect. Advances in Applied Energy, 2(14),54-58.
[3] Allee, K. D. , Baik, B. , & Roh, Y. . (2021). Detecting financial misreporting with real production activity: evidence from an electricity consumption analysis. Contemporary Accounting Research,96(7),32-36.
[4] Schinckus, C. , Nguyen, C. P. , & Chong, F. H. L. . (2022). Cryptocurrencies' hashrate and electricity consumption: evidence from mining activities. Studies in economics and finance,56(3), 39.
[5] Chen, X. , Song, C. , & Wang, T. . (2021). Analysis of electricity loss and electricity consumption law in low-voltage areas: a case study. Journal of Physics: Conference Series,184(7),366-369.
[6] Kesornsit, W. , Sirisathitkul, Y. , & Lin, C. J. . (2022). Hybrid machine learning model for electricity consumption prediction using random forest and artificial neural networks. Applied computational intelligence and soft computing,874(52),123-125.
[7] Kim, Y. T. , Kim, B. J. , & Kim, S. W. . (2022). Multi-level stacked regression for predicting electricity consumption of hot rolling mill. Expert Systems with Application(Sep.),85(9), 201.
[8] Yurtsever, O. . (2022). A practical methodology for predicting electricity consumption of urban residential buildings. Environmental Progress & Sustainable Energy,96(7),258-259.
[9] Zhou, X. , Lin, W. , Kumar, R. , Cui, P. , & Ma, Z. . (2022). A data-driven strategy using long short term memory models and reinforcement learning to predict building electricity consumption. Applied Energy, 306(14),652-658.
[10] Granell, R. , Axon, C. J. , Kolokotroni, M. , & Wallom, D. C. H. . (2021). Predicting electricity demand profiles of new supermarkets using machine learning. Energy and buildings(Mar.), 234(4),521-523.