Power System Peak Regulation Demand Forecasting Based on LSTM Neural Network

Yi Lin[1,*], Wei Wu[1,**], Wei Lin[1***], Azhu Lin[2,****]

[1]State Grid Fujian Economic Research Institute, Fuzhou, China
[2]College of Electrical Engineering and Automation, Fuzhou University, Fuzhou, China

Abstract—In the context of a high proportion of renewable energy integrated to the power grid, the net load may has significant fluctuations, and it is necessary to quantify peak regulation demand of power system. This paper stablishes a peak regulation demand prediction model based on long short-term memory (LSTM) method by training historical data. The typical data such as load and renewable energy output are selected as the input vector, and correlation coefficients are used to process and simplify the input vector. The historical prediction errors are used to set margins for peak regulation demand prediction. The case study shows that the proposed model can effectively predict the peak regulation demand of the power system.

1 Introduction

With the large-scale integration of renewable energy into the grid, the demand for peak regulation in the power grid is increasing [1]. Accurate prediction of peak regulation demand is beneficial for the safe and stable operation of the power system. The randomness and uncertainty of the output of renewable energy sources such as wind power and photovoltaic greatly increase the difficulty of determining peak regulation demand.

Currently, research on measuring the impact of renewable energy integration on peak regulation demand mainly includes: calculating the grid peak regulation demand caused by new energy integration using scenario reduction method after sampling the output of renewable energy led by wind power [2]; Reference [3] extracted typical scenarios of wind power anti peak shaving, used back propagation (BP) neural network to calculate the fitness of evaluation indicators for each typical scenario, and further data mining models are used to obtain the peak shaving capacity caused by wind power. There is little literature on the methods for determining the overall peak regulation demand of the power system, and more research focuses on optimizing resource allocation within a framework for determining the peak regulation demand of the power grid. Literature [4] focuses on the study of the peak regulation demand pressure that can be alleviated by introducing energy storage into the power grid with a high proportion of renewable energy. In the calculation example, a scenario generation method based on Quantile regression analysis and Gaussian Mixture model clustering is proposed to describe the peak regulation demand of the power system; The research objective of reference [5] is to optimize the real-time response strategy of power system with renewable energy. The random dynamic changes in renewable energy output are considered. The peak regulation demand of the system is described as a Gaussian Markov process.

The application of deep learning methods in the power grid can solve complex nonlinear relationships through model training to achieve the prediction results [6,7]. As a mature learning algorithm, long short-term memory (LSTM) has great application value in data prediction. Especially in the modeling and prediction of time series. It can overcome the gradient explosion problem in the training process of traditional recurrent neural network (RNN) methods. Reference [8] uses LSTM method to achieve load prediction in the power system, and added a sequence to sequence (Seq2Seq) structure in the application process to improve the final prediction accuracy. Reference [9] first discusses the optimal values of hyperparameters for LSTM networks, and sets up numerical examples to compare their advantages and disadvantages with traditional machine learning algorithms. The findings indicate that the utilization of LSTM techniques can significantly decrease the system's prediction errors. In line with this, Reference [10] successfully employed the LSTM method alongside real-time network data to attain accurate wind speed predictions.

This paper applies LSTM technology to the field of peak regulation demand prediction in power systems, and proposes a input features selection method which can effectively prove peak regulation demand prediction accuracy. Simulation studies are carried out to verify the proposed method.

2 Basic principles of LSTM prediction

LSTM is a widely used recurrent neural network in deep learning, which can overcome the problem of vanishing

*liny-02@163.com; **wuw_fjjy@163.com; ***linwei@tju.edu.cn; ****675250747@qq.com

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

gradients by "backpropagating through time" during training. LSTM networks are capable of handling complex sequence problems in machine learning by constructing large-scale recurrent neural networks and achieving good results [11].

LSTM primarily implements data discarding and transmission through four control gates within a single recurrent structure. Moreover, the prediction performance of time-series data is strongly correlated with the network depth, learning rate, and temporal correlation of the input data. The overall framework of LSTM can be represented as follows:

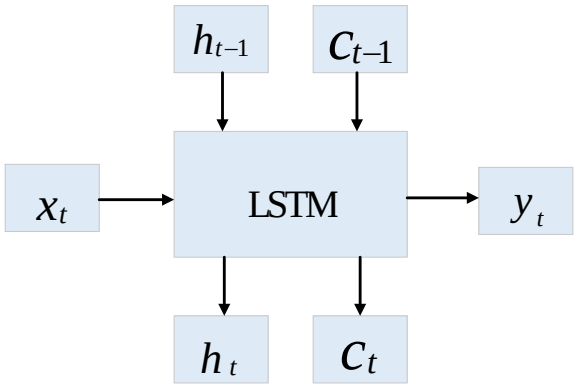


Figure 1 Overall framework of LSTM.

In Figure 1, there are three inputs, namely h_{t-1} and c_{t-1}. After being processed by the central LSTM module, the outputs are y_t, h_t and c_t. Here, x_t represents the input of the current round, h_{t-1} represents the state output of the last round and c_{t-1} represents the global information carrier of the last round. Furthermore, y_t rep x_t resents the output of the current round, h_t represents the state output of the current round, and c_t represents the global information carrier of the current round. Additionally, the internal structure of the LSTM module is illustrated in the following diagram.

Furthermore, the internal structure of the LSTM module is drawn as shown in Figure 2:

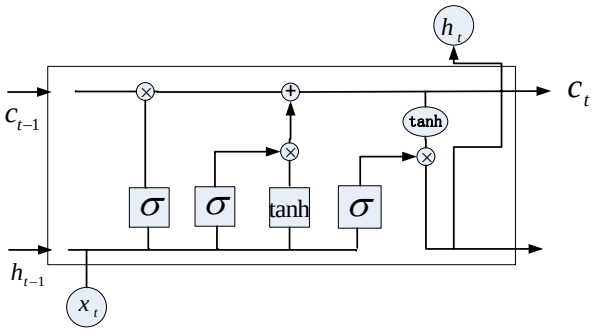


Figure 2 Internal Structure of the LSTM Neural Network Neuron

Among them, σ represents the sigmoid activation function, tanh represents the tanh activation function. By using these functions and the joint weight matrix, a series of calculations are performed to selectively remember, learn and update the input quantities. Specifically, the "gate" structure is used to implement the transmission of information.

An activation unit has three types of gates:

The forget gate has the responsibility to determine which information should be disregarded. For each number in the cell state, this gate reads and produces a value ranging from 0 to 1. A value of 1 indicates that the information is completely retained, while a value of 0 indicates that it is completely discarded. By selectively forgetting unnecessary information, the model can remember long-term dependencies that are crucial for many Natural Language Processing (NLP) tasks like language translation and sentiment analysis.

The decision of which input values to use to update the memory state rests with the input gate. A sigmoid layer is employed to decide which information requires updating, while a tanh layer is utilized to generate potential updates. The cell state is then updated by combining these two values. The input gate is critical for the model to be able to selectively update and forget information as needed, allowing for accurate predictions and long-term memory storage

The output values are determined by the output gate according to the input and memory state. A tanh layer is used to process the cell state prior to output, while a sigmoid layer is used to decide which sections of the cell's information to output. By selectively choosing which cell information to output, the model can make accurate predictions and maintain long-term memory. The output gate is a crucial component for the model to control the output and ensure proper functioning

Each gate's weight within an activation unit is obtained through training. In addition to variable prediction via feature learning, LSTM can further learn the complex internal patterns of time series data, which improves prediction accuracy.

3 Data processing and model establishment

3.1 Determining input and output variables

3.1.1 Selection of input parameters

According to the clearing process of the auxiliary service market for peak shaving and the speculation on the main factors affecting peak shaving, the input variables are set as follows: maximum daily load $X1$, minimum daily load $X2$, maximum daily wind power output $X3$, minimum daily wind power output $X4$, maximum solar voltage output $X5$, Pearson correlation coefficient $X6$ between the daily load output sequence and the time series of the total wind power photovoltaic output of the day. Except for the correlation coefficient, the unit of other characteristic quantities is taken as MW.

3.1.2 Definition of output

The dependent variable, that is, the predicted output result, is taken as the daily peak regulation demand. Y is taken as the daily peak shaving demand. The data set is processed, the daily peak regulation demand capacity of the system is calculated, and the daily peak regulation demand is jointly determined according to the maximum value of peak valley difference of daily load, rotating reserve and peak shaving capacity caused by wind power and photovoltaic. Among them, the rotating reserve is taken as 7.5% of the maximum load. The demand of all operation days in the dataset is calculated as follows:

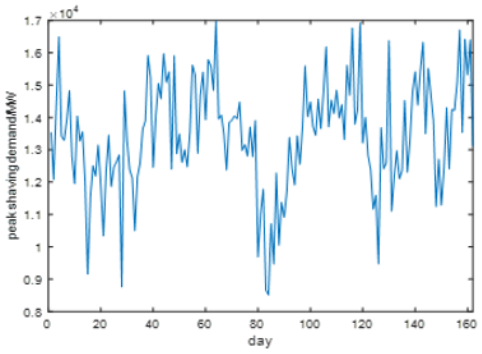


Figure 3 System peak shaving demand capacity.

The historical data of system peak shaving demand capacity is shown in Figure 3, the data relationship between peak shaving demand and input variables is explored through the above LSTM prediction model. Although the specific data fitting expression between input and output cannot be obtained by machine learning, the output value of peak regulation demand in the future can be predicted directly through the model training of the algorithm.

3.2 Data analysis

The correlation between the input and output variables is shown in Figure 4, the correlation between the variables can be analyzed to determine the most important input variables. The input feature vector can be reduced by removing features that have little correlation with the predicted variable. The input features with strong correlation also need to be reduced to independent variables.

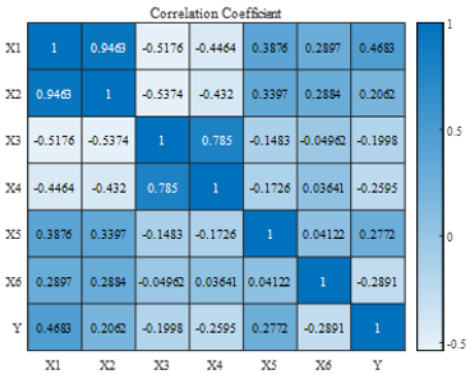


Figure 4 Correlation analysis between input and output variables.

Before predicting, preliminary analysis of the correlation between variables is made. Firstly, when the absolute value of the correlation coefficient between the feature quantity and the peak regulation demand quantity is greater than 0.8, the feature quantity will be finally retained. However, there is no such case in this scenario. Secondly, the feature quantities whose absolute value of correlation coefficient with the peak regulation demand quantity is greater than 0.2 are retained, which are X_1, X_2, X_4, X_5, X_6, respectively. Finally, the feature quantities with an absolute value of correlation coefficient greater than 0.6 between two feature quantities and a smaller absolute value of correlation coefficient with Y are deleted, so the remaining feature quantities for prediction *are* X_1, X_4, X_5, X_6.

3.3 Creating machine learning models

3.3.1 Parameter

In TensorFlow environment, keras is called to build a neural network to realize LSTM prediction. The input and output are as described above.

3.3.2 Explanation of parameter settings:

Optimization Algorithm: Adam algorithm is used, which can use the best initial values for each configuration parameter.

Learning rate: The learning rate controls the amount of weight updates based on estimated gradients at the end of each batch, and can greatly affect the trade-off between the speed and performance of the learning of the model. In this work, a learning rate of 0.1 is used.

Batch size: The batch size is the number of samples between updates of the model weights. In this scenario, a value of 32 is utilized as the default.

Number of iterations: A value of 200 is used in this work, as it has been found to provide good convergence characteristics after 200 iterations.

Activation function: The activation function is a weighted activation that is transferred between neurons, and is usually fixed by the framework and proportion of the input or output layer. In this work, the relu function is used.

Apart from the aforementioned feature quantities, there are additional parameters that can be optimized, such as the optimization of LSTM architecture, memory units, determination of the optimal number of hidden layers, and weight initialization. However, in this work, a single-layer LSTM architecture is used, which is simple and applied for peak regulation demand forecasting.

3.3.3 Splitting the training and testing datasets:

The sample data is sorted chronologically and 80% of the data from each season is used as the training set, while the remaining 20% is used as the testing set. Specifically，130 out of the 162 data sets are used as the training samples, and the remaining 32 are used as the testing samples.

3.3.4 Model selection and training:

We select a single-layer LSTM model and utilize the training set to train the model.

3.3.5 Prediction of testing data:

The previously developed model is utilized directly to forecast the testing dataset. The corresponding 6 features of the operating day are inputted, and the corresponding prediction results are obtained.

3.3.6 Evaluation method

By comparing the actual peak regulation demand of the test set with the peak regulation demand prediction results obtained by using the LSTM prediction model, the accuracy of the model prediction can be evaluated. In this paper, the sum of squared residuals (SSE), mean absolute percentage error (MAPE) and root mean square error (RMSE) are used to measure the prediction accuracy [11-12]. The value range of the three error indicators is $[0,+\infty)$, when the predicted value and the real value are completely consistent, the value is equal to 0; The greater the error, the greater the value.

4 Case analysis

4.1 Comparison of test set results

As previously stated，20% of the data is allocated to the test set. The comparison between the LSTM prediction results of 32 operation days and the real peak regulation demand of these operation days is shown in Fig.5.

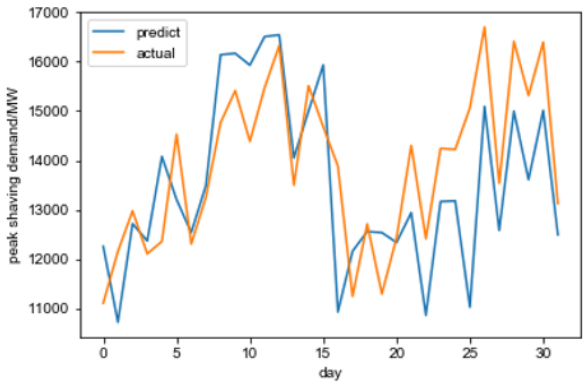


Figure 5 Comparison of predicted results and actual values of test set

The overall accuracy of peak regulation prediction of LSTM in scenario 1 is about 92%. The model exhibits a better predictive performance, as its RMSE value is 1382.57 and its deviation from the average value of the real peak regulation demand capacity (13417.26 MW) on the operation day is approximately 10.3%.

4.2 Multi scenario analysis of input feature selection

In order to study the impact of different types and quantities of feature inputs on the final peak regulation prediction results, the following four scenarios and series indicators are set for comparative analysis of peak regulation demand prediction errors.

Table 1 Peak regulation demand prediction error statistics of different characteristic quantities

	Input variables	MAPE	RMSE	SSE
Scenario 1	X_1, X_4, X_5, X_6	0.0806	1382.57	6.12E+07
Scenario 2	$X_1 \sim X_6$	0.0813	1448.53	6.71E+07
Scenario 3	X_5, X_6, X_7	0.1262	2142.38	1.47E+08
Scenario 4	$X_1 \sim X_8$	0.1114	2036.67	1.33E+08

Fig.5 shows the prediction results of scenario 1. Table 1 shows the statistical results of peak regulation demand prediction error with different characteristic quantities. At this time, there are no net load input variables and the correlation filtering of features has been completed. In addition, two input variables X_7 and X_8 are added in scenario 3 and scenario 4 to represent the maximum and minimum net load respectively. The setting scenarios of the results before and after the net load input participating in the peak regulation demand prediction is compared and analyzed. Under this parameter setting, the peak regulation demand prediction value without adding the net load information is more accurate.

4.3 Setting of peak shaving error margin

In order to further consider the impact of prediction error, the probability and statistics method of kernel density function is adopted, and the historical prediction data and real data are used to convert the single prediction value into the prediction interval. The specific measures are as follows:

Use daily data as the time scale for statistical analysis and collect historical data for peak regulation demand in the previous year, denoted as dataset A_1. We use the LSTM method to perform peak regulation demand prediction for one year and denote the predicted values as dataset A_2. We subtract the predicted peak regulation demand values of dataset A_1 from the actual values of dataset A_2 on the corresponding days to obtain the difference, which is denoted as dataset B. We use the maximum daily load as the reference value to normalize all elements in dataset A_1, A_2, B. Finally, we obtain the empirical distribution of the standardized prediction error based on the LSTM method, which serves as a measure of uncertainty.

The probability density function is fitted to the unitary data series, and the nonparametric kernel density estimation is used for fitting calculation. After obtaining

the probability density function of peak regulation demand error, determine the confidence level that the deviation value needs to meet, and obtain the upper and lower limits of the peak regulation demand prediction interval under the condition of meeting the confidence level, as shown in the following formula:

$$\left[Y_* + G_{(\frac{\alpha}{2})}, Y_* + G_{(1-\frac{\alpha}{2})} \right] \quad (1)$$

where Y_* is the unit value calculated based on the maximum load of the day in the data set of the initial forecast of annual peak regulation demand using the LSTM method; $G(\alpha/2)$ and $G(1-\alpha/2)$ are the inverse functions of the distribution function respectively representing the sum of the confidence degree, which meet the following constrains:

$$P\left\{ x \le G(\alpha/2) \right\} = \alpha/2 \quad (2)$$

$$P\left\{ x \le G(1-\alpha/2) \right\} = 1-\alpha/2 \quad (3)$$

In the calculation, the probability density function is obtained by integrating the corresponding confidence degree with the complex Simpson formula.

Finally, the deviation between the predicted value and the actual value of the peak regulation demand is adopted, and the peak regulation demand after considering the margin of peak regulation demand prediction error is updated as shown in Fig.6.

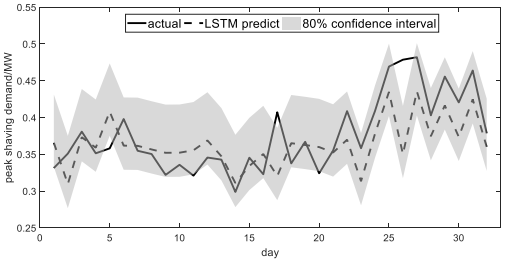


Figure 6 Results after considering the capacity margin of peak regulation demand prediction.

5 Summary

This paper uses the data-driven method to obtain the historical peak regulation demand and a series of load and wind power output data, and uses the LSTM method to train the historical peak regulation demand data to obtain the model that can be used to predict the peak regulation demand of power system. The sensitivity analysis of the input variables verify the necessity of using correlation coefficient to select variables. Considering the error of peak regulation prediction, a margin of peak regulation demand capacity considering confidence is set for it. And the 80% confidence interval of peak shaving capacity can effectively envelope the actual peak shaving demand.

Acknowledge

This research is funded by the Science and Technology Project of State Grid Fujian Electric Power Co., LTD., China (No. 52130N22000C).

References

1. Olawumi T O,Chan D W M. A scientometric review of global research on sustainability and sustainable development[J]. Journal of Cleaner Production, 2018(183): 23 1-250.
2. Wang Wei, Xu Jing, Zhao Xiang, et al. Peak shaving operation status analysis of coal-fired power units in China [J]. Southern Energy Construction, 2017, 4(01): 18-24.
3. Zhenpeng H, Zixuan W, Jianlin L, et al. A novel capacity demand analysis method of energy storage system for peak shaving based on data-driven[J]. Journal of Energy Storage, 2021, 39:102617.
4. Analysis of energy storage demand for peak shaving and frequency regulation of power systems with high penetration of renewable energy
5. Hierarchical scheduling learning optimisation of two-area active distribution system considering peak shaving demand of power grid
6. A. Gasparin, S. Lukovic, C. Alippi, Deep learning for time series forecasting: The electric load case, CAAI Trans. Intell. Technol. 7 (1) (2022) 1–25.
7. E. Tuna and A. Soysal, "LSTM and GRU based Traffic Prediction Using Live Network Data," 2021 29th Signal Processing and Communications Applications Conference (SIU), Istanbul, Turkey, 2021, pp. 1-4
8. An improved LSTM-Seq2Seq-based forecasting method for electricity load
9. Torres, J.F., Martínez-Álvarez, F. & Troncoso, A. A deep LSTM network for the Spanish electricity consumption forecasting. Neural Comput & Applic 34, 10533–10545 (2022).
10. Wind Speed Forecast using LSTM and Bi-LSTM Algorithms over Gabal El-Zayt Wind Farm
11. Shi X , Chen Z , Wang H ,et al.Convolutional LSTM Network: A Machine Learning Approach for Precipitation Nowcasting[J].MIT Press, 2015.
12. Li M , Chen G , Lin Z ,et al. Combination Forecasting Model Based on Drift[C]//International Joint Conference on Computational Sciences & Optimization.IEEE Computer Society, 2009.