

Electrical Load Forecasting Based on LSTM Neural Networks

Lei Guo^{1,2}, Linyu Wang^{1,2} and Hao Chen¹

¹State Grid (Suzhou) City Energy Research Institute, No. 9 Xuesen Road, Science & Technology Town, Huqiu District, Suzhou, China

²State Grid Energy Research Institute, No.18 Binhe Avenue, Future Science Town, Beiqijia Town, Changping District, Beijing, China

Abstract—The accurate prediction of electrical load is vital to the safety of power grid and the efficiency of energy. The development history of electrical load forecasting is introduced briefly in this paper. Then the relationship between electricity, environment and economy is also analyzed. Ljubljana's load data and environmental meteorological data of the previous time are applied to train LSTM networks to make prediction of the electrical load. The experimental results show that, in the case of abundant data with good quality, the LSTM network model can make quite acceptable short-term predictions of the power load based on previous energy consumption and environmental weather data.

Keywords—electrical load forecasting; time series; recursive neural network; long short-term memory

I. INTRODUCTION

Electricity is an important cornerstone of the modern civilization, which is closely related to industrial production, transportation and many other aspects. Compared with other resources such as coal and oil, electricity is more difficult to be stored. Its production, transportation and consumption are carried out at the same time. From the point of view of power production, excessive supply will cause huge waste of resources, but insufficient supply will also bring serious impact on people's daily life and cause economic losses. Therefore, it is important to accurately predict future electricity consumption, which can provide reference for future investment planning and construction of power grid.

From the point of view of electricity production, accurate forecast of electrical load can optimize and guide the production plan. At present, the growth rate of electricity production capacity exceeds the actual demand of local grid users, and many electrical facilities are idle and redundant for most of the year, which leads to the high cost of electrical enterprises. Therefore, we need more rational and efficient construction plan. If electrical load can be accurately predicted, it can guide the power investment planning, optimize the operation of electrical enterprises, reduce operation costs and improve profitability.

From the point of view of power transportation, correct forecast of electrical load in urban areas is an indispensable compass for the operation of power grid. Firstly, accurate

forecast of regional electrical load is needed to arrange the maintenance work of power grid scientifically and rationally. Secondly, forecast of electrical load can guide grid operation and allocation. In order to reduce the overall capacity demand of power grid construction, reduce operating costs, and avoid major accidents such as grid paralysis caused by the failure to meet peak load, the electrical load of multiple areas in different periods should be estimated in advance. Therefore, power grid enterprises have a strong demand for electrical load forecasting under their jurisdiction.

Based on the importance of electrical load forecasting, this paper first reviews the related research progress, analyses the key factors in power load forecasting, and classifies the power load forecasting according to the time period. Then principles of these methods and their application in load forecasting are introduced. With Ljubljana's load data and meteorological data, the short-term electrical load forecast based on long-term and short-term memory unit is carried out, and the results are analyzed and summarized.

II. RELATED WORK

A. Current Research Status

In recent years, electrical load forecasting technology has gradually changed from traditional methods to modern methods with artificial intelligence technology. The wide application of machine learning algorithms such as support vector machine and neural network greatly improves the accuracy of power prediction. Espinoza et al. [2] used least squares support vector machine (LS-SVM) to predict and evaluate the short-term electrical load under the influence of meteorological, date and other external factors, which has better accuracy and explanatory ability than other traditional methods. Luca Ghelardoni et al. [3] made a variety of improvement based on SVR and proposed SVP+SVB model with great performance. Wu et al. [4] proposed a novel model combining CNN with LSTM, which uses weather data and historical data to predict wind energy load with good results. Jia et al. [5] proposed a electrical information processing architecture based on convolutional neural network. Cheng et al. [6] used multi-layer neural network and LSTM to predict short-term electrical load and experimental results proved the superiority of deep neural network. Li et al. [7] used the principal component analysis

method to reduce the dimension of data and applied the support vector machine regression to predict short-term electrical load. The numerical experiments show that this integrated method possesses great effectiveness. AlRashidi et al. [8] used particle swarm optimization to predict the annual peak load of power system. The experiments using actual electrical data of Kuwaiti and Egyptian validated the effectiveness of this method. Parshall et al. [9] analyzed the data of power consumption and carbon dioxide emissions on a city scale, which revealed the relationship of power consumption and carbon emissions. Li et al. [10] used regression neural network and fruit fly optimization algorithm to forecast annual power load. The simulation experiments comparing this method with support vector machine and particle swarm optimization show that this method is effective in annual load forecasting. Xia et al. [11] used RBF neural network to predict short-term and long-term power load respectively. The results show that this method can obtain high accuracy and stability in electrical load forecasting. Qiu et al. [12] used the ensemble learning with deep belief network and support vector regression to predict electrical load and superiority of this method was demonstrated by experiments.

B. Important Factors of Electrical Load Forecasting

1) *Weather factor*: Weather factor is the most important factor in electrical load forecasting. The following factors are often considered: temperature, rainfall, wind speed and direction, cloudiness, humidity, and weather conditions (such as fog, breeze, haze) [13]. Temperature is the most important among them. When the temperature is too low or too high, people will consume more electricity to maintain the comfortable temperature. Therefore there is a strong correlation between power load and temperature. Precipitation can easily affect outdoor activities and some industrial processes. The increase of precipitation will decrease the power load of outdoor activities and increase the load of indoor activities. Cloudiness should be considered during the forecast of short-term power load, since it can increase the use of lighting. Humidity is also an important weather factor, which can describe the change of power load under some climatic conditions along with temperature.

2) *Social factors*: Social factors refer to the factors that affect load and can be used to further reduce power load forecasting errors, which include calendar information, holidays, population migration, social structure, and special festivals. Calendar information is the most important social factor. Human behavior is usually carried out by day, week, month, season and year, which attaches periodicity to power load patterns. Since people's daily lives consume electricity, the population migration will surely affect the load. Social structure is also necessary long-term load forecasting, because different social structures will lead to different power consumption and consumption patterns. When it comes to short-term electrical load forecast, special festivals must be considered to reduce prediction error.

3) *Economic factors*: Economic factors are negligible, the most important of which are GDP, oil price and heating fuel price. GDP represents value of a country's final products and the value of services produced in a given period of time, which can be used as an economic factor for medium and long term load forecasting. Since oil price can greatly affect industrial activities, it is vital for medium-term load forecasting. Load is usually negatively correlated with oil price, lagging behind oil price for several months.

C. Classification of Electrical Load Forecast

1) *Ultra short-term load forecasting*: Ultra short-term power load forecasting is used for short-term power generation optimization and real-time monitoring of power system, which is very important for frequency control and short-term power generation dispatch.

2) *Short-term power load forecasting*: Short-term load forecasting is the most commonly used type of load forecast, which provides time arrangement assistance for power generation facilities. Because of the long time span, the impact of environment changes, date and time on electrical load is becoming more and more obvious. In order to achieve accurate short-term power load forecast, environmental data, date and time and other auxiliary data should be introduced.

3) *Medium-term load forecasting*: Medium-term load forecasting is the basis for many market participants to formulate strategies. In some power systems, large end-users must plan the load of their power systems one year in advance. In the medium-term load forecasting scenario, economic factors such as oil prices, seasonal weather forecasting, and predicted GDP changes will be used as input of the forecasting model to reduce forecasting errors.

Long-term load forecasting: If the power load can be forecasted several years or decades in advance, it will greatly enhance the scientific nature of power system planning and reduce the cost of social resource. However, in the long run, power load is affected by energy policy, economic development, population growth and decline, climate change, power investment and many other aspects. Therefore, long-term load forecasting has always been a difficult problem in the field of energy planning, lacking reliable and systematic forecasting methods.

III. RECURSIVE NEURAL NETWORKS AND LSTM

A. BP Neural Network

Traditional BP neural network consists of three parts: input layer, hidden layers and output layer. Each neuron in the latter layer is connected with all neurons in the upper layer, and there is no connection between neurons in the same layer.

BP neural network algorithm can be divided into two parts. In the process of forward transmission, Data is input from input layer, transferred to hidden layer and calculated, and then output through output layer. In this stage, the parameters in the network do not change, but the output is obtained by forward

transmission and the error is compared with the real value. The second is the reverse propagation of errors. In this stage, according to the errors between the output results and the real values, the network output can be close to the real value by propagating backwards from the output layer and modifying the weights and biases in the network layer by layer.

B. Recursive Neural Network

Compared with BP neural network, the hidden layer in RNN has the edge which connects to the next time step of the hidden layer. The output of the hidden layer not only transmits to the output layer, but also to the hidden layer of the next time step. Structure of RNN gives it capability to memorize historical information.

The traditional neural network cannot use the information contained in the previous data when processing an input data. RNN introduces the concept of time series. When processing data at each time, RNN can make use of the information at the previous time, which makes RNN possess ability to deal with time series problems very well. However, RNN also has some obvious shortcomings. RNN often fails to make full use of the information which is different from the current time due to gradient disappearance.

C. LSTM Networks

In order to overcome the disadvantage of RNN in dealing with long-interval strong correlation sequence data, Sepp Hochreiter and Jurgen Schmidhuber proposed long short-term memory in 1997 [15]. The original LSTM unit consists of a recursively connected memory unit and two control thresholds, i.e. input gate and output gate. In the memory unit, the current hidden state of neurons is stored. The input gate is essentially a common neuron, which is used to express the influence of the current input on the hidden state. The output gate is used to indicate the effect of the hidden state on the output of LSTM unit. Intuitively, they can be seen as regulators of numerical flows connected through LSTM. Hence, they are called as gates. In 2000, Gers and Schmidhuber added a forgetting gate to the recursively connected memory unit on the basis of the initial LSTM unit, which improved the ability to distinguish the effects of different intervals of data on the memory hidden state, and improved the neuron performance [16].

LSTM solves the problem of gradient disappearance in traditional RNN. In the process of back propagation, some neurons can pass through the gate without attenuation, which makes LSTM have the ability to remember long-term information. Based on the advantages of long-term memory, LSTM network has shown strong ability in machine translation, speech recognition and other fields.

Timestamp	Hour	Load	Temperat	Humidity	Pressure	Wind Dir	Wind Speed
01/01/2011 00:59	0	970	-4.17	0.74	1013.907	10.46429	5.42
01/01/2011 01:59	1	932	-4.43	0.731739	1013.8	7.695652	5.56
01/01/2011 02:59	2	884	-4.76	0.720385	1013.8	8.038462	5.01
01/01/2011 03:59	3	839	-5.13	0.73	1012.88	7.12	4.59
01/01/2011 04:59	4	813	-5.50	0.73	1012.8	0.4	4.30
01/01/2011 05:59	5	819	-5.99	0.73	1012.8	4.64	4.69
01/01/2011 06:59	6	853	-5.92	0.744615	1012.8	6.923077	4.76
01/01/2011 07:59	7	891	-5.74	0.7684	1011.788	3.44	5.38
01/01/2011 08:59	8	940	-5.64	0.776897	1011.7	4.344828	4.74
01/01/2011 09:59	9	1014	-5.40	0.761739	1011.7	2.782609	4.57
01/01/2011 10:59	10	1076	-4.89	0.769615	1011.315	5.307692	4.91
01/01/2011 11:59	11	1119	-4.36	0.780741	1010.544	2.481481	4.69
01/01/2011 12:59	12	1106	-3.68	0.782	1009.84	5.28	4.49
01/01/2011 13:59	13	1065	-2.34	0.768889	1008.815	8.444444	5.21
01/01/2011 14:59	14	1035	-1.66	0.737407	1008	7.111111	5.24
01/01/2011 15:59	15	1029	-1.90	0.70963	1008	9.777778	5.35
01/01/2011 16:59	16	1086	-2.72	0.71625	1008	7	4.76
01/01/2011 17:59	17	1185	-3.34	0.740435	1007.304	1.043478	4.30
01/01/2011 18:59	18	1250	-3.70	0.767143	1007	3.142857	5.99
01/01/2011 19:59	19	1234	-3.78	0.770741	1007	1.111111	5.37
01/01/2011 20:59	20	1223	-4.13	0.777037	1007	1.259259	6.18
01/01/2011 21:59	21	1162	-4.28	0.79	1007	1.416667	5.03
01/01/2011 22:59	22	1081	-4.66	0.788636	1007	0.909091	6.52
01/01/2011 23:59	23	996	-4.67	0.7996	1007	2	5.40

FIGURE I. PART OF EXPERIMENTAL DATA.

IV. POWER LOAD FORECASTING EXPERIMENT BASED ON LSTM NETWORK

A. Experimental Settings

The data set used in this experiment is electrical load data from Ljubljana for the whole year of 2011. Besides the power load, it also contains other information, including temperature, humidity, pressure, wind direction and wind speed. Therefore, except for the timestamp, the dimension of each data is 7. The granularity of the original data is several minutes and the measurement interval of each data point is slightly different, so the average method is adopted to integrate the data into one hour granularity. Part of experimental data is shown in the Figure 1 above.

Based on the analysis of the influential factors and classification of power load forecasting discussed above, the data set contains only time along with weather data, lacking economic and social factors, which is relatively short in time length. Therefore, the data set is more suitable for short-term load forecasting and this paper will carry out short-term load forecasting experiments for one hour and one week respectively. According to the ratio of 4:1, the first 80% data (7008 items) were used as training set, and the last 20% data (1752 items) were used as test set. In addition, due to the difference in size of different data, it is necessary to normalize the data into (0,1) before entering the network.

This experiment is based on Tensorflow with which users can design the structure of the neural network by building computational diagrams modify its structure intuitively. In order to judge the effect of the model, the average absolute percentage error (MAPE) and root mean square error (RMSE) are used to quantify the prediction performance of the model. Equations to compute these are shown below:

$$MAPE(\%) = \frac{1}{n} \sum_{i=1}^n \left(\frac{|y_i - pred_i|}{y_i} \right) \times 100\% \quad (1)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - pred_i)^2} \quad (2)$$

where y_t represents the actual load data and $prod_t$ represents the predicted value. According to these two equations, the lower MAPE and RMSE indicate higher the accuracy of the prediction model.

B. Experimental Results and Analysis

1) *Hourly power load forecasting based on data of previous day:* All the information of the previous day is used to predict the current load value. Therefore, the input layer contains 7*24 neurons, the output layer has one neuron, and the hidden layer in the middle is one layer, which contains 10 LSTM units. The square of error was used as loss function. The Adam algorithm, which is an extended method of random gradient descent, was adopted. Adam can design independent adaptive learning rates for different parameters by calculating the first and second moment estimates of gradients, which can make the parameters converge to better results and speed up the operation with good applicability. The initial learning rate is set at 0.0005.

According to Figure 2, it can be seen that when the number of iterations reaches 1000, the value of the loss function has basically remained unchanged and the model can be considered to have converged. Prediction of input test data is compared with the actual value, which is shown in Figure3. The blue line is the curve of the actual value and the orange line is the curve of the predicted value. From these experimental results, it can be seen that the one-step prediction based on the previous day's historical information has a good effect, and the trend of the predicted value is the same as the actual value with small absolute error. The calculated results show that the MAPE is 2.09% and the RMSE is 47.30%.

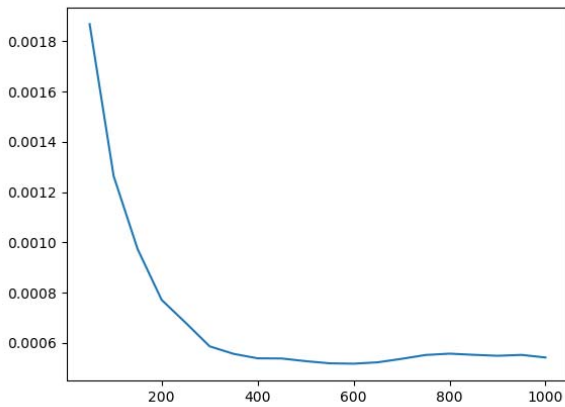


FIGURE II. LOSS CURVE VARYING THE NUMBER OF ITERATIONS

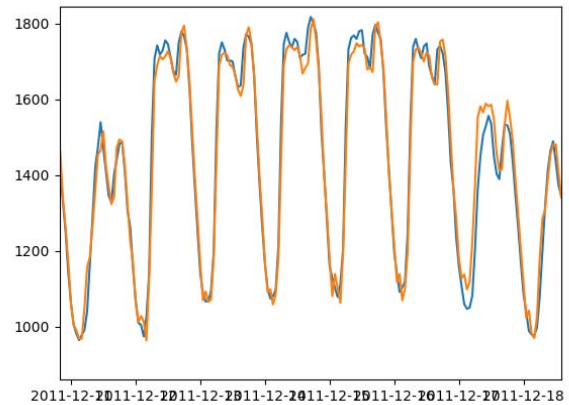


FIGURE III. FORECASTING RESULT WITHIN A WEEK

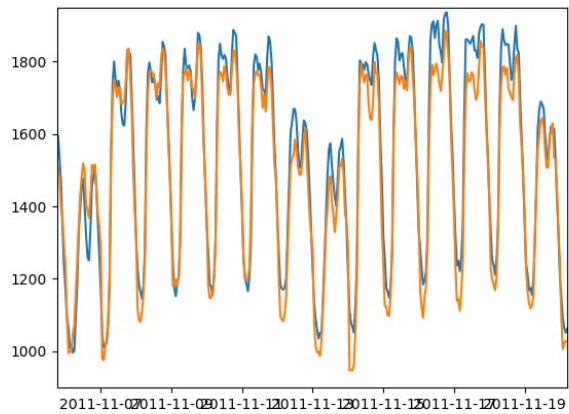


FIGURE IV. FORECASTING RESULT WITHIN TWO WEEKS

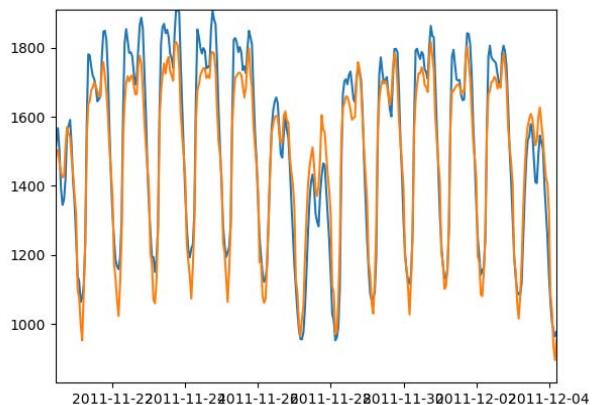


FIGURE V. FORECASTING RESULT WITHIN TWO WEEKS BASED ON HISTORY AND WEATHER DATA

2) *Load forecasting of current week based on last week's data:* All the information of one week was used to predict the power load value of the next week. The input is changed to 7×168 , and the output is 1×168 with other hyperparameters unchanged. The two-week forecast result is shown in Figure 4. It can be seen that the forecast result is relatively poor when the forecast time is one week. Although the absolute error is high, the trend of the predicted value is still consistent with the actual value, which shows that the network has limited ability to predict the load value after a long period of time, but it still can learn the periodic change weekly and daily. The RMSE was 140.71 and MAPE was 7.03%.

3) *Load forecasting current week based on history and weather data:* In order to improve the prediction accuracy, existence of weather forecast can be considered, which is advisable to incorporate into the input data. In addition, by observing the regularity of data, it is found that the electrical load data has obvious regularity in a week. The power consumption is higher in working days, but lower in rest days. However, the original data does not provide weekly information for every single day so that the datetime library was adopted to process the timestamp and get the weekly information of that day. After adding these into training data, the input dimension increases to $(7 + 5 + 1) \times 168$ with other settings unchanged.

The experimental results are shown in Figure 5. The RMSE is 130.19 and the MAPE is 5.99%. After adding the weather information and the week information, the values of the two indexes are reduced, which means the prediction accuracy is slightly improved. The importance of weather information for short-term power load forecasting and the help of weekly information for time series modeling are verified.

V. CONCLUSION

Power load forecast with high precision has important economic and social significance. This paper analyses the development of electrical load forecasting, summarizes the influential factors used in previous work and forecasting types of different time series, and simply analyses the relationship between these factors and forecasting types. The BP neural network, RNN and LSTM networks are analyzed progressively with the modeling ability of time series data. LSTM network which can memorize long-term and short-term information is selected as the prediction model for experiment.

According to the characteristics of the data set, this paper focuses on short-term load forecasting. The corresponding LSTM networks are constructed in two directions, one hour and one week, respectively. The one-hour prediction model is satisfactory. Although the prediction of one week is relatively ineffective because of its long forecast time, the prediction is improved by adding the input of weekly information and weather forecast information, which verifies the importance of weather information for short-term power load forecasting and the obvious help of weekly information for medium-term load forecasting.

ACKNOWLEDGMENT

This work was financially supported by the Science and Technology Project of State Grid Corporation (Research and Demonstration Application of Monitoring and Management Technology of City Energy System Based on Large Data and Artificial Intelligence, SGIT0000YXJS1800395).

REFERENCES

- [1] L. Zhou and XL. Yong, Electricity Demand Forecasting Based on AR IMA Model and Linear Neural Network, *Journal of Ludong University (Natural Science Edition)*, 2015, 31(03):277-282. [Chinese version]
- [2] Marcelo EspinozaM, Johan A.K.Suykens and Ronnie Belmans, Bart De Moor, Electric Load Forecasting, *IEEEControlSystems*, 2007, 27(5):43-57.
- [3] Luca Ghelardoni, Alessandro Ghio, Davide Anguita, Energy Load Forecasting Using Empirical Mode Decomposition and Support Vector Regression, *IEEE TRANSACTIONS ON SMART GRID*, VOL. 4, NO. 1, MARCH 2013.
- [4] W. Wu, K. Chen, Y. Qiao, Z. Lu, "Probabilistic short-term wind power forecasting based on deep neural networks", 2016 International Conference on Probabilistic Methods Applied to Power Systems (PMAPS), pp. 1-8, Oct 2016.
- [5] ZT. Jia, F. Zhao, Review and Prospect of Artificial Neural Network Applications in Smart Grid, *Papers Collection of the National Smart Grid Client Energy Management Academic Annual Meeting in 2015*, 2015:10. [Chinese version]
- [6] YY. Chen, ZJ. Bao and XG. Guo, Interval Load Forecasting Using Improved Scalar Method, *Industrial Control Computer*, 2017, 30(04):130-131. [Chinese version]
- [7] YF. Li, YQ. H and GL. Jiang, Short-term Load Forecasting Based on PCA-SVM, *Proceedings of the CSU-EPSA*, 2007(05):66-70. [Chinese version]
- [8] M.R. AlRashidi, K.M. EL-Naggar, Long term electric load forecasting based on particle swarm optimization, *Applied Energy*, 2009, 87(1).
- [9] Rio P D, Analysing the interactions between renewable energy promotion and energy efficiency support schemes: The impact of different instruments and design elements, *Energy Policy*, 2010, 38(9):4978-4989.
- [10] HZ. Li, S. Guo, CJ. Li, JQ. Sun, A hybrid annual power load forecasting model based on generalized regression neural network with fruit fly optimization algorithm, *Knowledge-Based Systems*, 2013, 37(2):378-387.
- [11] H.M.Al-Hamadi, S.A. Soliman, Long-term/mid-term electric load forecasting based on short-term correlation and annual growth, *Electric Power Systems Research*, 2004, 74(3).
- [12] X. Qiu, Y. Ren, PN. Suganthan, GAJ. Amaratunga, Empirical Mode Decomposition based ensemble deep learning for load demand time series forecasting, *Applied Soft Computing*, 2017, 54(C):246-255.
- [13] Marin Matijaš, Johan A.K.Suykens, Slavko Krajcar, Load forecasting using a multivariate meta-learning system, *Expert Systems With Applications*, 2013, 40(11).
- [14] MM. Tan, Short Term Traffic Flow Forecasting Based on Integration of LSTM Model and Grey Model, *Nanjing University of Posts and Telecommunications*, 2017. [Chinese version]
- [15] Hochreiter S, Schmidhuber J, Long short-term memory, *Neural Computation*, 1997, 9(8):1735-1780.
- [16] Gers F A, Schmidhuber J, Cummins F, Learning to Forget: Continual Prediction with LSTM. *Neural Computation* 12(10): 2451-2471, *Neural Computation*, 2000, 12(10):2451-2471.