

# ***Short-term Electricity Price Forecast of Electricity Market Based on E-BLSTM Model***

**Wang Baoyi<sup>a,\*</sup>, Ji Xiaoqiong<sup>b</sup>, Zhang Shaomin<sup>c</sup>**

*School of Control and Computer Engineering, North China Electric Power University, Baoding, 071003, China*

*<sup>a</sup>email: wangbaoyiqj@126.com, <sup>b</sup>email: jxq2314@126.com, <sup>c</sup>email: zhangshaomin@126.com*

**Keywords:** long-term and short-term memory; deep learning; electricity price forecast; electricity market

**Abstract:** The price changes in the electricity market will adversely affect the utility revenue and user cost. The current electricity price forecasting method has a low degree of utilization of its periodic variation law and a short forecast step size, which makes the electricity price forecast have large errors. A two-way LSTM model based on ELU activation function is proposed to predict the short-term electricity price change on the supply side of the power market. The gradient disappearance problem in the back propagation calculation process is optimized by ELU, and the accuracy of electricity price prediction is improved. Experiments on models and algorithms in the electricity price database of the PJM power market in the United States show that compared with the ARIMA and ARMA models, the proposed model has higher accuracy, and the algorithm converges to a lower loss rate, which can provide greater fluctuations in the supply side of the power market. The electricity price is accurately predicted.

## **1. Introduction**

Electricity price is the fulcrum of the electricity market, and the high and low price changes will have a direct impact on the power industry and related industries [1-2]. For the entire grid system, accurate electricity price forecasting can guide users to use electricity behaviour, adjust power consumption and power consumption time to reduce system peak-to-valley difference, improve system load rate, reduce network loss, and ensure system operation stability [3].

Since electricity price changes can be viewed as time series analysis, there are certain changes in each cycle such as season and week, so electricity price prediction can be based on some statistical analysis models [4], such as Autoregressive Integrated Moving Average Model (ARIMA) and Auto-Regressive and Moving Average Model (ARMA). However, the short-term electricity price change on the supply side of the electricity sales market is an unstable stochastic process. There are spikes, zero electricity prices, negative electricity prices, etc. The current forecasting models mostly use neural networks, support vector regression and integrated multi-model methods to predict electricity market electricity prices [5].

Literature compared the Recurrent Neural Network (RNN) with traditional statistical regression models, fuzzy inference systems and neural network methods, and proved the effectiveness of the

model in electricity price forecasting. RNN is difficult to predict long-term dependencies [6], and the predictable step size is shorter. The Long Short-Term Memory (LSTM) improves the structure of the cyclic neural network, making predictable steps longer and more capable of time series prediction [7]. Literature proposes a method for tracking state of conversations in interview dialogues using long-term and short-term memory and artificial neural networks. The artificial neural network based method is used to construct the dialog state detection model, and the model has acceptable accuracy.

In this paper, the price change on the supply side of the electricity sales market is characterized by time series, periodicity and volatility. The ELU-based Bi-directional Long Short-Term Memory (E-BLSTM) is constructed. The BLSTM model is based on the cyclic neural network and can effectively solve the long-term dependence problem in the cyclic neural network. It can be used to predict time series and accurately predict the price of electricity with large fluctuations. The related experiments prove the feasibility of the model.

## 2. Short-term Electricity Price Forecast Based on E-BLSTM(ELU-based bidirectional LSTM)

### 2.1. Model Design

LSTM maintains long-term memory and effectively utilizes the timing characteristics of electricity prices [8]. There are three types of such valve nodes: the forget gate, the input gate, and the output gate. At time  $t$ , the model receives a new input vector  $x_t$  and receives the output value  $h_{t-1}$  of the hidden layer at time  $t-1$ . The input gate determines how much information the input vector  $x_t$  at the current time can be input to the cell state, forgotten. The gate represents the cell state  $h_{t-1}$  at time  $t-1$  and how much information can be input to the cell state at time  $t$ , and the output gate represents the output information  $o_t$  at time  $t$ . The three gates are multiplied by the matrixes  $U$  and  $W$  by  $x_t$  and  $h_{t-1}$ , respectively, and multiplied by the sigmoid function to convert the value between 0-1 to the value of 0-1 as the gated state. The cell structure is shown in Figure 1.

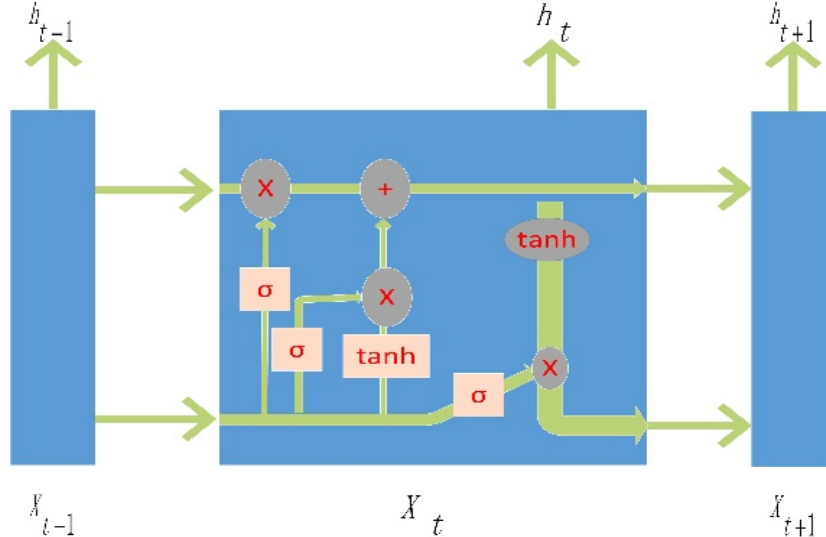


Figure 1: Long short-term memory structure

There is a correlation between the past and the future of electricity price data. The period of change is long and the regularity is weak. It is easy to generate the above problems by using the traditional LSTM structure to perform back propagation calculation multiple times. The characteristics of the electricity price prediction model constructed in this paper are as follows: (1) Using the long-length LSTM neurons and constructing a two-way LSTM model, focusing on the

regularity of the electricity price time series from the past to the future, and paying attention to the electricity price time series never coming to the past; (2) The activation function of neurons and gating units is changed to the improved ReLU activation function - ELU to improve the learning effect, effectively avoiding the "gradient disappearance" problem easily generated by the back propagation calculation model.

The sigmoid does not completely solve the "gradient disappearance" problem caused by saturation. In view of this, the linear unsaturation of ReLU can be used to effectively alleviate the "gradient disappearance" problem, but if the learning rate is too large, ReLU will cause the neuron state to always be 0, showing a "death" state. And because the negative part of the ReLU function is 0, directly using ReLU as the activation function of the forgetting gate will reduce its effect and seriously affect its "forgetting" effect. ELU is an improvement of ReLU. ELU inherits the advantages of ReLU and adds a non-zero output to the negative input, thus preventing the occurrence of silent neurons, and the derivative convergence is zero, thus improving the learning efficiency [9-10].

The traditional RNN calculation state is of the form:

$$h_t = \tanh(Wx_t + Wh_{t-1} + b) \quad (1)$$

And the LSTM can add the above-mentioned three gates to include the addition of the addition in the state accumulation, and solve the disappearance of the gradient. When calculating the state of the gating unit, set t time, the new electricity price input vector is  $x_t$ , the input vector input to t time at time t-1 is  $h_{t-1}$ , and the activation functions are ELU,  $f_t$ ,  $i_t$ ,  $c_t$ ,  $o_t$  and  $h_t$  represent the forgetting gate, input gate, neuron state, output gate, and output to the next moment,  $a_t$  is the amount of information controlled by the input gate,  $U_*$ ,  $W_*$  are the corresponding weight, and \* indicates a wildcard,  $\odot$  Hadamard Product, the relevant formula is as follows:

$$f_t = \text{elu}(U_f x_t + W_f h_{t-1} + b_f) \quad (2)$$

$$i_t = \text{elu}(U_i x_t + W_i h_{t-1} + b_i) \quad (3)$$

$$a_t = \text{elu}(U_c x_t + W_c h_{t-1} + b_c) \quad (4)$$

$$c_t = i_t \odot a_t + f_t \odot c_{t-1} \quad (5)$$

$$o_t = \text{elu}(U_o x_t + W_o h_{t-1} + b_o) \quad (6)$$

$$h_t = o_t \odot \text{elu}(c_t) \quad (7)$$

## 2.2. Model Initialization

The E-BLSTM model involves many parameters, and the initial parameter settings are not the final optimal results. This paper uses real-time electricity price data released by a company. The real-time electricity price is recorded every hour. The time window is  $24*n$ . According to experience, the number of neurons in each layer of the hidden layer is initially set to 4,  $n = 1$ , the input is a one-dimensional electricity price floating point number, the output is a one-dimensional electricity price floating point number, and the output is a one-dimensional electric value. The data set selects data from January to November 2017.

In the initial setting of hyper parameters, this paper uses the common LSTM and the E-BLSTM model designed in this paper to predict the same electricity price data set. The activation function of the former input and output layer is tanh, the activation function of the gating unit is sigmoid, and the activation function of the latter is ELU. To prevent over fitting, the rejection rate per layer was 0.2. We use MSE to calculate the loss and Adam for gradient optimization.

### 2.3. Algorithm Steps

The core steps of model training are three steps: positive calculation of gate and neuron states, inverse update of gradients, and calculation of loss values. Data preprocessing and building models are required before model training. During training, the results of each step are calculated previously, then the loss value of each step is calculated, and then the back propagation calculation is performed based on the result of forward propagation, and the gradient and weight matrix  $W$  are updated. The core algorithm steps are as follows:

Step 2-1: Model construction, initializing hyper parameters. The core of the model construction is the initialization parameter matrix  $W$  and the bias matrix  $b$ .

Step 2-2: Calculate the state of the gating unit forward, calculate the neuron status and output. The data needs to be scaled into the  $[0, 1]$  interval before input. The training and test data need to set the dimension according to the time window  $t$ .

Step 2-3: Back propagation solves the gradient. For bidirectional LSTM, the two layers of neurons are subjected to back propagation calculations, that is, two back propagation from 1 to  $t$  and from  $t$  to 1.

In the prediction, the time parameter  $x$  is directly input, the forward calculation method is called, and the trained parameter matrix is used to output the predicted value  $y$ .

### 3. Experimental Analysis

According to the US PJM power market, the price trend of electricity prices is obtained from the electricity price data of a region from November 6 to 10, 2017. Under the influence of many factors, the previous ARMA ( $p=24, d=2, q=1$ ) model and ARIMA ( $p=24, q=3$ ). The model has lost its regularity and accuracy in predicting electricity price. The comparison between the predicted values of the two models and the real values is shown in Figures 2.

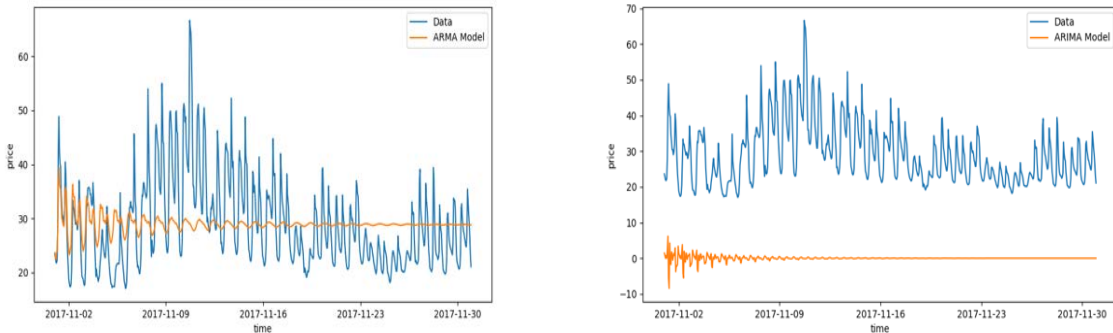


Figure 2: The predicted value using ARMA and ARIMA compared with the real value

From the prediction results of Figure 2, when the prediction interval is long, the ARMA and ARIMA models have been difficult to predict the experimental data, and its accuracy has no reference, and the prediction effect on the volatile price data is even worse.

As form LSTM model, the optimal time window size is 1 day, that is, the step size is 24, that is, the data of the previous day predicts the data of the next day, and the optimal number of neurons is 8, the most excellent rejection rate is 0.1. Although the cross-entropy function is slightly faster than the mean square error calculation, in the current electricity price forecast, the time difference between the two is negligible, and the mean square error accuracy is 5% higher than the cross-entropy. The optimal weight parameter update method is Adam. Figure 3 shows the variation of the model training times and model errors for the common LSTM and E-BLSTM. Figure 4

shows the comparison between the predicted values and the real values of the two models. It can be seen that the model can converge to a very low loss rate, that is, the model is stable.

The program concluded that the final RMSE of the LSTM model was 3.96 with an accuracy of 91%, the RMSE of the E-BLSTM was 3.06, and the accuracy was 95%. Both models converge at around 20 training times, and the latter loses less. In the comparison of the real electricity price and the predicted electricity value, the triangle is the predicted value, the origin is the real value, and the abscissa is the hour (a total of 720 hours in November, so there are 720 data points). It can be seen from the graph that the predicted value is roughly consistent with the true value and the model is accurate. In contrast, the E-BLSTM model can achieve a stable low error rate after a limited number of iterations and can converge. It can be seen that the ELU-BLSTM model can effectively predict the supply side price of the volatile electricity market.

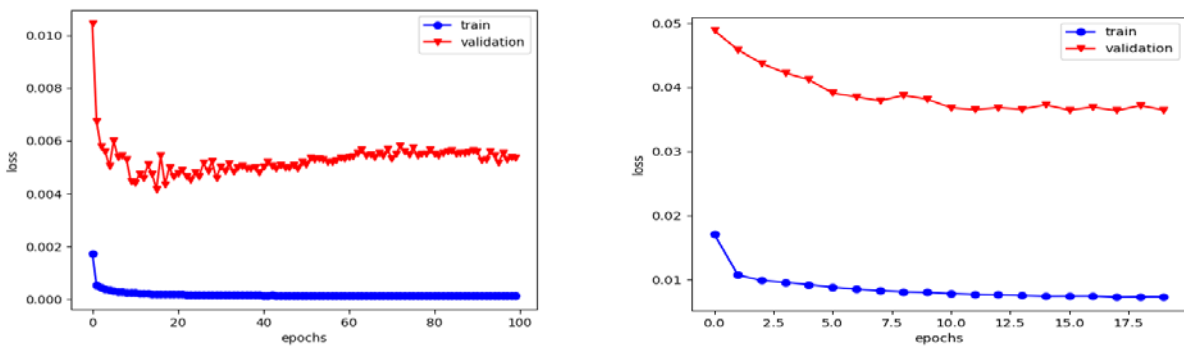


Figure 3: The training error varies with the training frequency of LSTM and E-BLSTM

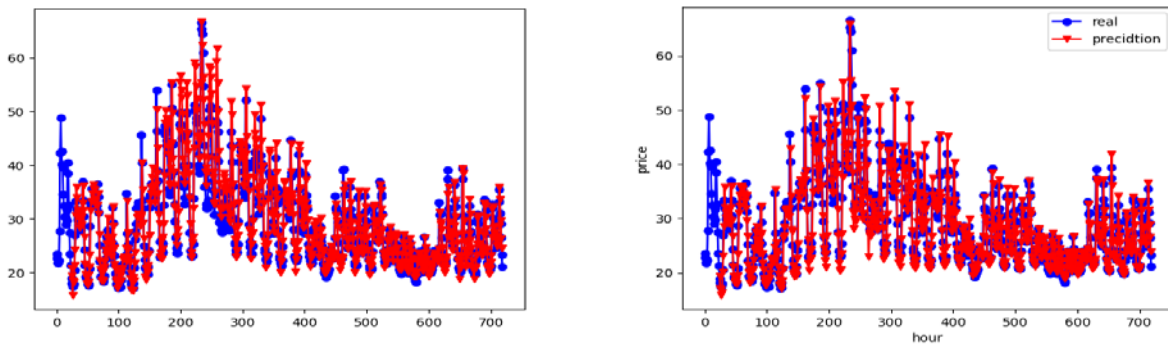


Figure 4: The predicted value using LSTM and E-BLSTM compared with the real value

#### 4. Conclusions

Accurate forecasting of the supply side electricity price in the electricity sales market is beneficial for market participants and managers to formulate corresponding strategies for the trend of electricity price changes in a timely manner. Improve interest and optimize market resource allocation. Based on the characteristics of time series and instability of short-term electricity price changes, based on the LSTM deep learning model based on cyclic neural network, this paper designs a bidirectional LSTM model based on ELU activation function to predict the electricity price of electricity supply side and use it. The model design and algorithm for the current electricity price data provided by the PJM power market in the United States validate the model and algorithm. Compared with the ARIMA model, the E-BLSTM model has higher accuracy and can converge to a lower loss rate. For the power company, an accurate bidding plan can be formulated to prevent large

deviations when the forecast deviation is large, and to avoid risks under the factors with more variables; the user can effectively control the purchase cost when the power consumption fluctuates greatly. The method can also be extended to the field of energy consumption to improve the profitability of utility investment.

## References

- [1] Li M A, Menghua F, Lei G., et al. Latest Development Trends of International Electricity Markets and Their Enlightenment [J]. *Automation of Electric Power Systems*, 2014, 38(13):1-9.
- [2] Yin J, Wang S, Gong L. The Effects of Factor Market Distortion and Technical Innovation on China's Electricity Consumption [J]. *Journal of Cleaner Production*, 2018, 188:195-202.
- [3] Ziel F, Weron R. Day-ahead Electricity Price Forecasting with High-dimensional Structures: Univariate vs. Multivariate Modeling Frameworks [J]. *Energy Economics*, 2018, 70:396-420.
- [4] Zhao Z , Wang C , Nokleby M , et al. Improving Short-Term Electricity Price Forecasting Using Day-Ahead LMP with ARIMA Models[J]. 2018.
- [5] He Y, Liu R, Li H, et al. Short-term Power Load Probability Density Forecasting Method Using Kernel-based Support Vector Quantile Regression and Copula Theory [J]. *Applied Energy*, 2017, 185:254-266.
- [6] Xin W, Ji W U, Chao L, et al. Exploring LSTM Based Recurrent Neural Network for Failure Time Series Prediction [J]. *Journal of Beijing University of Aeronautics & Astronautics*, 2018.
- [7] Anbazhagan S, Kumarappan N. Day-Ahead Deregulated Electricity Market Price Forecasting Using Recurrent Neural Network [J]. *IEEE Systems Journal*, 2013, 7(4):866-872.
- [8] Karim F., Majumdar S., Darabi H., et al. LSTM Fully Convolutional Networks for Time Series Classification [J]. *IEEE Access*, 2017, 6(99):1662-1669.
- [9] Djork-Arné Clevert, Unterthiner T., Hochreiter S. Fast and Accurate Deep Network Learning by Exponential Linear Units (ELUs) [J]. *Computer Science*, 2015.
- [10] Ide H., Kurita T. Improvement of Learning for CNN with ReLU Activation by Sparse Regularization[C]// *International Joint Conference on Neural Networks*. IEEE, 2017.