# Power Flow Prediction: A Case in Ningxia Electricity Market

Bin Yan*
School of Mechanical Engineering
Southeast University
Nanjing, China
yanbin@goldwind.com.cn

Yifan Zhou
School of Mechanical Engineering
Southeast University
Nanjing, China

Dongsheng Yu
Beijing Goldwind Science & Creation Windpower Equipment Co., Ltd..
Beijing, China

Xianpeng Wang
Beijing Goldwind Science & Creation Windpower Equipment Co., Ltd..
Beijing, China

*Abstract*—**With the further opening of the bidding market in China, the accuracy of electricity price prediction directly affects the operational decisions and profits of power producers. The core factor that affects electricity price is power flow. In the early stage of electric power reform, the data of electricity price was too insufficient to support the forecasting analysis. This paper assists electric power traders to fill in the appropriate amount of electricity during the transaction process by predicting the relevant cross-section power flow. Computational methods are complex and require data of many variables at present. Therefore, this paper uses autoregressive integrated moving average (ARIMA) model and long short-term memory (LSTM) model to predict the power flow. The prediction error of the model is less than 5%. Furthermore, the conclusion shows that there is no difference between weekdays and weekends, and the power flow is a stationary time series. Based on the result of this research, some decision-making suggestions that can maximize the profit of the manufacturer are given.**

***Keywords- Power flow prediction; ARIMA; LSTM.***

## I. Introduction

On May 24, 2018, the National Energy Administration of China issued the "Notice on the Relevant Requirements for Wind Power Construction Management in 2018" (hereinafter referred to as the Notice) [1]. The notice stipulates that the newly-added centralized onshore wind power projects and the offshore wind power projects without determinate investment subject in the provinces that have not yet issued the 2018 wind power construction plan should obtain feed-in tariff through competition. In general, the focus of wind power construction management in 2018 is to promote wind power consumption and control the wind curtailment rate, as well as reduce the feed-in tariff and subsidy intensity. The Notice clearly stated that consumption of wind power is the primary task, and the power transmission and consumption should be strictly implemented. In addition, wind power should be deployed in a competitive manner, and the investment of wind power construction should be optimized.

“Grid parity” is a general guideline for power generation enterprises in various provinces and cities, which refers to accessing to grid at the benchmark price of local thermal power generation. However, under the new deal, “bidding policy” has already sounded the assembly of competition in the wind power industry. As the WeChat public account “The Wind Power Front Observation” said, “The difference between the benchmark price of thermal power is no longer important, because everyone is competing for their own status in the industry other than the external standard.”

In fact, renewable energy bidding policy has been implemented for many years in Europe, e.g., the UK (from 2015), France (from 2015), Netherlands (from 2016), Denmark (from 2016) and Germany (from 2017). Extensive literature established electricity price forecasting models and provided decision-making strategies for enterprises in the bidding environment. For example, Prabavathi and Gnanadass developed a power market trading model using the customer aggregate demand and the supply curve in [2]. Chaâbane predicted the electricity price via a mixed autoregressive fractionally integrated moving average (ARFIMA) and neural network model in [3]. Florian and Rick developed a probabilistic model to predict mid-term and long-term electricity prices in [4]. Gligorić et al. proposed an autoregressive model to predict short-term electricity prices in [5]. LSTM with the differential evolution (DE) algorithm, denoted as DE–LSTM, is used for electricity price prediction by Lu et al. in [6]. The ARIMA models have been developed for different months from January to December as most suitable for simulating and forecasting the venture capital (VC) over the observation site, the Stationary R-squared, R-squared, Root Mean Square Error and Normalized Bayesian information criterion (BIC) etc. are used to test the validity and applicability of the developed ARIMA models revealing significant accuracy in the model performance [7].

The above method requires a large amount of electricity price sequence data when making electricity price forecast (current domestic spot transaction is 96 points quote). However, since January 2018, the spot transaction of the Ningxia electric power trading market has entered the bidding mode, and the accumulated transaction price data is too insufficient, which cannot support the direct prediction of the above algorithm. Based on these, this paper focuses on the core factor of the power grid - the trend value - to analyze the "grid" structure of Ningxia.

It is generally known that power flow straight affect on price and electricity load. Xiao et al. [8] put forward a forecasting method of time-dependent correlation-based conditional probabilistic power flow for increasing the accuracy of power flow in distributed power network. Chen et al. [9] calculate the usage shares of the common network by Guangdong and

978-1-7281-0860-5/19/$31.00 ©2019 IEEE

Guangxi by employing the power flow tracing method under given power supply paths, and could be used in the transmission pricing. Guo et al. [10] based on the practical research and development experience of a provincial power grid security early warning system, the realization algorithm, technical characteristics and application occasions of ultra-short-term and short-term power flow forecasting in power grid security early warning system are introduced. The methods mentioned above are more physical to estimate power flow, which is computationally intensive and difficult to obtain data. In this paper, ARIMA and LSTM methods are used, and no other variable data are needed. This paper focuses on the prediction of power flow, which provides a reference for the quotation. The traditional time series method (autoregressive moving average (ARMA) model, etc.) and deep learning algorithm (LSTM) are applied to predict the power flow value. The comparison between the predicted value and the maximum power flow value of the "cross section" can assist the decision-making of power consumption reporting in the spot trading stage of the power station at the relevant location. Traditional power flow calculation or prediction focuses on the "steady" probability distribution of power flow for power grid fault diagnosis. While in this paper, the time series method is mainly applied to predict the maximum power flow value of different sections at the "next time". Electricity markets are more unpredictable than other commodities referred to as extreme volatile. So the choice of the forecasting model has become even more important.

The remaining part of this paper is organized as follows: Section 2 presents the problem statement. Section 3 describes the data. Then Section 4 explains the principles of deep learning for forecasting. Section 5 uses ARIMA for forecasting. After that, Section 6 demonstrates the results discussion.

## II. Problem Statement

### A. Type of ecetricity transaction

Current electricity transaction in China is divided into four types: contract trading, spot trading, futures trading, and auxiliary service trading.

A contract transaction is a medium-term and long-term contract transaction, which refers to the energy transaction conducted by the market entity through the signing of a power sale contract.

Spot trading, that is short-term trading, to compensate for the difference between contracted trading power and short-term load demand, including the day-to-day energy trading formed by the power company bidding and the real-time energy trading organized to ensure the immediate balance of power supply and demand.

Futures trading, long-term contract trading in accordance with certain rules and regulations, futures contracts in the variety, specifications, quantity, duration and payment of transactions.

Auxiliary service transactions, power generation companies to provide services such as system frequency modulation, accident backup and voltage support, through competition to ensure the market value of quality goods.

However, Since Ningxia has just entered the power trading market, the cumulative number of transactions is small and the price of electricity is small. The fluctuation of electricity prices cannot be predicted from the perspective of “electricity price”, and the flow in the power grid is to measure the amount of power distribution of the grid. In this paper the current load situation of the power grid, and then help to determine the electricity and electricity prices in the power trading.

## III. Data Wrangling

### A. Data Description

The flow data of power grid used in this paper are from the Ningxia power grid dispatch in the year-2018. The data is daily based which is used for day-ahead power market. The data set has a total of 322 sample data, of which the missing data accounts for 2.5% (that is, 8 data points), and the missing data is filled by the quadratic interpolation method.

The original data scale can be as large as 900. This is not favorable for forecasting due to Section unwinding. Thus, the next step is preprocessing, i.e. standardization.

### B. Preprocessing

- Statistical properties of dataset

Considering that the data may vary from Monday to Friday and two days of the week, the Kruskal-Wallis test is applied to explain the weekend effect of the data. The Kruskal-Wallis rank sum test constructs the H statistic to test the difference of Different interval data, the result show that it has a P value of 0.7021, which is greater than the significance level of 0.05. Therefore, the null hypothesis cannot be rejected, that is, there is no weekend effect in the data. The daily characteristics, seasonal changes and changes in operating modes of the section are matched with the predicted output of the power generation enterprise. When the power generation and the cross-section are reversed, the medium- and long-term priority power generation rights should be increased to ensure the power generation capacity of the power generation enterprise when the section capacity is limited. The largest, the highest customer yield.

## IV. Principles of Deep Learning for Forecasting

### A. Standardization of Datasets – LSTM

The standardization of datasets is a common requirement for many machine-learning procedures. A usual way is normalization, i.e. to transform the data to its center by removing the mean value of each feature, then scale it by dividing their standard deviation (std.) for each feature.

$$x_{std} = \frac{x - x_{mean}}{\sigma_x}, \tag{1}$$

where $x_{mean}$ represents the mean of the corresponding feature data column x, and $\sigma_x$ denotes the standard deviation.

An alternative way is to scale the feature in between a given minimum and maximum value, e.g. in [0,1], so that the maximum absolute value of each feature is scale to one:

$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{2}$$

where $x_{min}$ and $x_{max}$ stand for the minimum and maximum value for each feature (column data) respectively.

Note that, during testing phrase in the LSTM model, the standardization is applied on both training and testing data; however, the information used for (e.g. the min and max values in (2)) are only calculated based on the training dataset to guarantee a genuine forecasting, because in reality the testing data can never be known in advance.

In this paper, the following way is adopted. Motivations to use this scaling for the flow of power grid are:

- Robust to very small standard deviations of features.
- The flow time series in power grid is not necessarily stationary; thus, scaling the test data using (1) based on the mean and std values from the training data set may incur large errors when transforming them back to the original scale; on the other hand, the flow values is regulated by Ningxia power grid dispatch based on certain grid-operation-codes, which typically confine the flow values to some reasonable ranges, i.e. the min and max values can be assumed stationary in the local power grid.

### B. *LSTM Introduction*

LSTM was introduced by Sepp Hochreiter and Jürgen Schmidhuber in 1997. Unlike traditional recurrent neural network (RNNs), an LSTM network is well-suited to learn from experiences to identify and predict the time series when there are very longtime lags with unknown size. This is one of the main reasons why LSTM outperforms alternative RNNs and other sequence learning methods like hidden Markov models (HMM) in certain types of problems.

The innovations of LSTM are the two concepts: "gate" ($\sigma$) and cell state ($s_t$). Gates are used to optionally allow information to pass through. Mathematically, it is an activation function (e.g. sigmoid) with pointwise multiplication operations. The gates output values between 0 and 1, describing the portion of information to be remembered. A value of 0 means "forget", while a value of 1 means "remember". An LSTM has three kinds of these gates: a forget gate $f_t$, an input gate $i_t$ and an output gate $o_t$, to control the cell states of step-t. Fig.1 shows the internal structure of one LSTM cell unit.

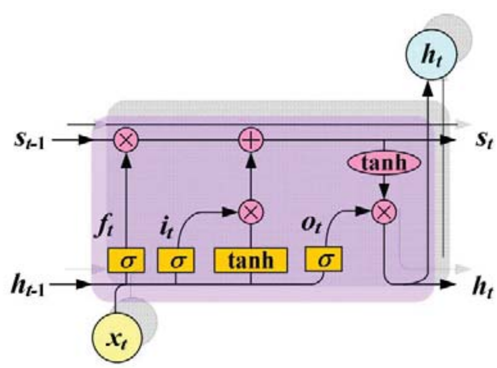


Figure 1. The internal structure of one LSTM cell unit

The updating equations used in the above diagram are:

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \tag{3}$$

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \tag{4}$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \tag{5}$$

$$s_t = f_t \cdot s_{t-1} + i_t \cdot \sigma(W_s x_t + U_s h_{t-1} + b_s) \tag{6}$$

$$h_t = o_t \cdot tanh(s_t) \tag{7}$$

where, "·" stands for the inner product operation and

$x_t$: input of step $t$, the value of flow at time $t$

$s_t$: cell state of step $t$

$h_t$: output of step $t$, the value of flow at time $t$

$\sigma$: activation function (hyperbolic tangent)

$W_k, U_k, b_k$: weight matrices or vectors ($k = f, i, o, s$)

$f_t, i_t, o_t$: different gate outputs of step $t$

### C. *Forecasting Performance of the Deep Learning Method*

In this part, the experiment investigating the forecasting performance of the deep learning method under different input lengths of features is carried out. All the experiments here are implemented in python and pytorch. 2/3 of this year-2018 data is used for training and the remaining 1/3 is served for testing. The forecasting workflow is as follows. The data is depicted in Fig. 2.

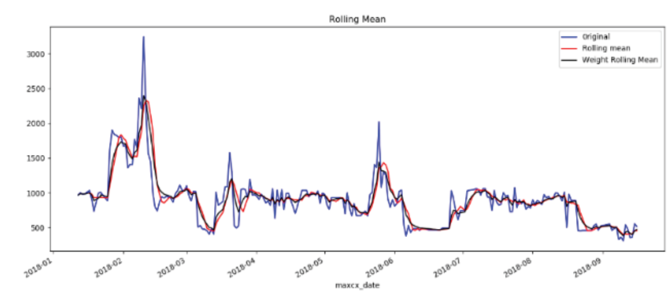


Figure 2. The year-2018 power grid flow values of NINGXIA

- Feature Engineering

Network inputs: the inputs are the past flow values in a certain time window of length $L$ (look back window), i.e. $x(t-1), x(t-2), \cdots, x(t-L)$ ; in this study, length $L = 2,4,8,16$ are compared;

Where,

$t$: time interval

$x(t)$: the maxflow value in time $t$

Network outputs: the forecasted flow value at time $t$, i.e. $x(t)$.

- Network design

Hidden Layer number:2; LSTM cell units: 4

Batch size: 1 (the number of training examples in one forward/backward process. The higher the batch size, the more memory space is needed).

Learning rate = 0.001.

Loss Function: The $L_2$ loss is considered.

- Training and Testing

Training epochs: 100 (more epochs might bring potential accuracy improvement but longer time)

Sample set partition: 2/3 for training; 1/3 for testing.

Hence, the designed LSTM network has an input layer with L inputs, a hidden layer with four LSTM blocks and an output layer of dimension one. The default sigmoid activation function is used for each LSTM block. The network is trained over 100 epochs where a batch size of one is used.

TABLE I. RESULT OF EACH LSTM BLOCK

| | Different Input Lengths | | | |
|---|---|---|---|---|
| | *LSTM* $L = 2$ | *LSTM* $L = 4$ | *LSTM* $L = 8$ | *LSTM* $L = 16$ |
| Test RMSE | 139.25 | 132.25 | 111.75 | 119.5 |
| Train RMSE | 108.75 | 105.25 | 83.75 | 98.5 |

In the process of power trading, the daily maximum power flow value in the power grid is accumulated into a time series data set. After preliminary statistical analysis, it is known that the stability of this time series data is not very good, and according to the physical characteristics of the power grid, it is easy to know the trend. The value is a "random" value, which is one of the measures of the power distribution of the entire grid. The maximum power flow value has a long-term dependence on each influencing factor affecting the power flow, and after giving a small time increment. The change of the tidal current value is continuous change; the early use of the traditional time series method to predict the maximum tidal current value requires the most smoothness test first, and then the relevant differential processing is needed to meet the needs of the ARIMA series of models, and the traditional time series model. The long-term dependence of sequences is often ignored, so the LSTM and ARIMA series of models are compared in this article.

## V. PRINCIPLEA OF ARIMA FOR FORECASTING

### *A. ARIMA Model Introduction*

ARIMA is a statistical analysis model that uses time series data to predict the future trends. Augmented Dickey–Fuller (ADF) test of time series must be used in this model. The difference between ARIMA model and ARMA model is that, ARMA model is established for stationary time series, while ARIMA model is used to model non-stationary time series. In other words, in order to build an ARMA model for non-stationary time series, the first step is to transform the difference into a stationary time series, and then to build an ARMA model.

### *B. Forecasting Performance of ARIMA Model*

The process mainly contains four parts: stationarity test, autocorrelation function (ACF) and partial autocorrelation function (PACF) graph, weekend effect test, and F test. The details are as follows:

- Stationarity test - Augmented Dickey-Fuller Test

Dickey-Fuller = -3.9304, Lag order = 6, p-value = 0.01338

alternative hypothesis: stationary

It can be seen from the ADF test results that the time series is stable.

- Autocorrelation graph & Partial autocorrelation graph

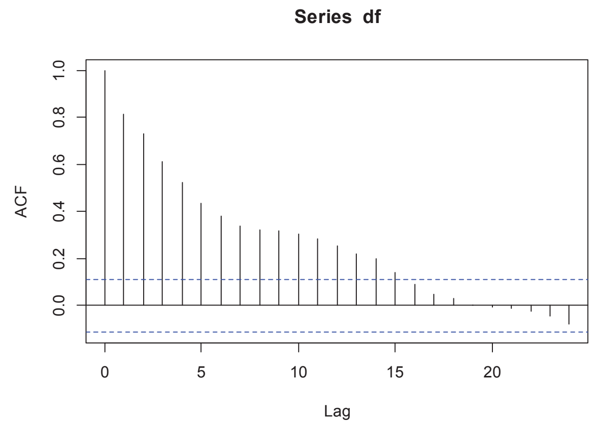


Figure 3. Autocorrelation graph

As can be seen from the ACF graph, the autocorrelation coefficient is trailed.

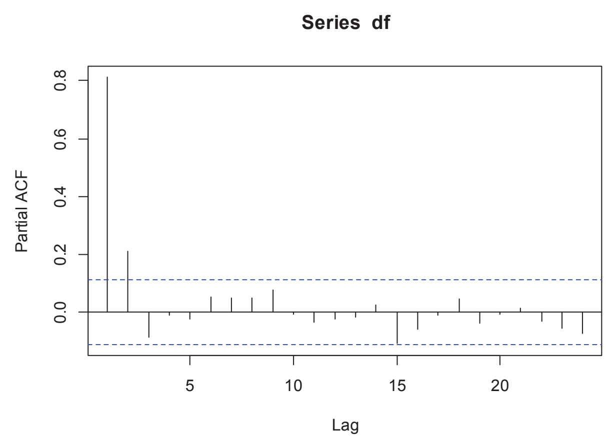


Figure 4. Partial autocorrelation graph

As can be seen in Fig. 4, the partial autocorrelation coefficient is censored.

It can be seen from the autocorrelation and partial autocorrelation graph that the autocorrelation coefficient is trailed and the partial autocorrelation coefficient is censored, so we initially use the auto regressive (AR) model for prediction. As for the selection of several orders of AR model, we use the F test criteria to select it.

- F test

An F-test is performed on the 0, 1, and 2 order AR models to determine the order of the AR model.

TABLE II. MODEL SIGNIFICANCE TEST

| Test statistics | AR(p) model | | |
|---|---|---|---|
| | *0* | *1* | *2* |
| Sum of residuals | 356.36 | 189.24 | 165.06 |
| F value | 397.05 | 12.91 | 0.05 |

TABLE III. RESULT UNDER EACH PARAMETER

| | Different parameters | | |
|---|---|---|---|
| | *(0,1)* | *(1,1)* | *(2,1)* |
| Test RMSE | 113 | 91 | 56 |
| Train RMSE | 239.78 | 221.76 | 212.82 |

According to the F-tested order method, when $\alpha = 0.05$, $N = 226$, the F distribution table of the Appendix F is found to be $F = 3.88$. When $P = 1$, the F value is found to be 3.90 > 3.88, which indicates that the F test is significant and the AR (1) model is not suitable. When $P = 2$, the F value is 0.05 < 3.88, which indicates that the F test is not significant and the AR (2) model is suitable.

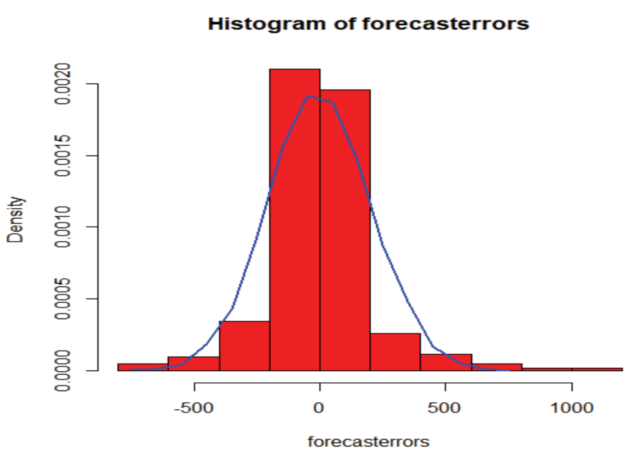

Figure 5. Prediction error histogram

The histogram shows that the prediction errors are roughly normal and the average is close to 0 (subject to the normal distribution of zero mean). Therefore, it is reasonable to consider the prediction error as a normal distribution with a mean value of zero variance (a normal distribution subject to zero mean and variance).

The ARMA model was constructed and compared. It was found that the sum of the residuals of the time series model added to the moving average (MA) part is larger than the sum of the residuals of the AR(2) model. Therefore, the AR(2) model is finally used for modeling analysis.

## VI. RESULTS DISCUSSION

In this section, the experiment is carried out to test the forecasting performance on different methods (LSTM, ARMA, ARIMA). The comparison results are given as follow:

- The comparison of root mean squared error (RMSE) between two methods

TABLE IV. COMPARE THE RESULTS OF LSTM AND AR(2)

| | RMSE | |
|---|---|---|
| | *LSTM L = 8* | *AR(2)* |
| Train | 83.75 | 212.82 |
| Test | 111.75 | 56 |

- The comparison of the flow values in the test dataset between two methods

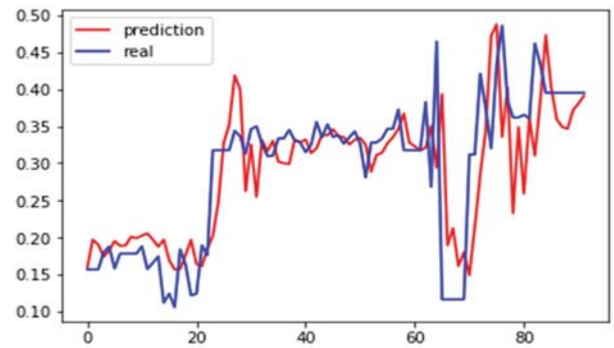

Figure 6. Compare the flow values in test dataset by LSTM

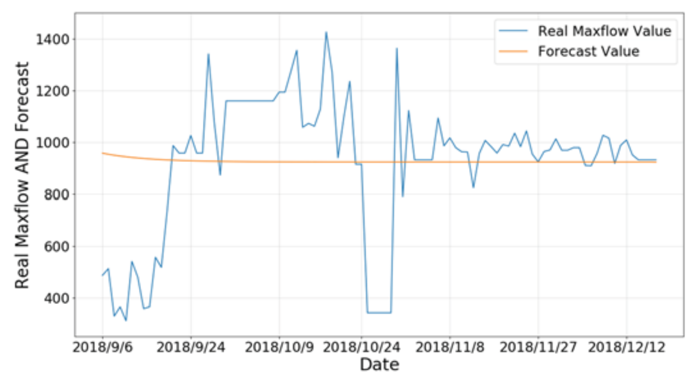

Figure 7. Compare the forecast value with actual value in AR(2)

The AR model is simple to use, which has a good performance on the linearity of the data performs better on the stationary sequence data. The LSTM model requires a higher amount of data, but can find long-term data dependence laws although the sequence is not stable.

To sum up, the LSTM-based deep learning framework cannot forecast the flow values of power grid with satisfactory accuracy under different input lengths, forecasting horizons and data sizes, due to the size of dataset, with data accumulation we believe that the LSTM algorithm will perform better and better.

## VII. CONCLUSION

Use the cross-section data to evaluate the spot transaction report before reporting, to avoid the failure to generate electricity due to the cross-section problem, which will result in the loss of the base power of the enterprise. At the same time, the cross-sectional distribution time is used to reasonably match the electricity price and electricity of the previous transaction, and the power generation enterprise is upgraded. At the same time, the real-time evaluation of the cross-sectional data with the distribution of intraday data increases the rational distribution of intraday transactions, reducing the impact of a large number of low electricity prices on medium and long-term high electricity prices, and achieving different benefits for power generation.

Comparing the predicted values of ARMA algorithm and LSTM algorithm in power flow sequence prediction, the LSTM algorithm performs relatively smoothly. The specific predicted values are as follows:

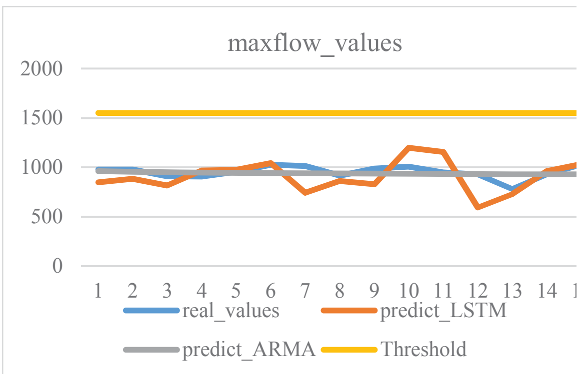


Figure 8. Compare the predicted values of ARMA and LSTM

Compare the predicted maximum current value of the next moment with the section limit. If the gap is large, the spot trading volume on the next day can be appropriately increased. With the accumulation of data, it is believed that the proportional coefficient can be found through analysis.

Although the prediction errors of LSTM and ARMA models are not ideal at present, with the accumulation of power transaction data, it is believed that these two typical time series data analysis algorithms will have better space in the future.

## Acknowledgment

The research work is supported by the National Natural Science Foundation of China (Grant Nos. 71671041).

## References

[1] China National Energy Board, Notice on the relevant requirements for wind power construction management in 2018, May 24, 2018.

[2] M. Prabavathi and R. Gnanadass, “Electric power bidding model for practical utility system,” Alexandria Engineering Journal, vol. 57, pp. 277-286, March 2018.

[3] N. Chaabane, “A hybrid ARFIMA and neural network model for electricity price prediction,” International Journal of Electrical Power & Energy Systems, vol. 55, pp. 187-194, February 2014.

[4] Z. Florian and S. Rick, “Probabilistic mid- and long-term electricity price forecasting,” Renewable and Sustainable Energy Reviews, vol. 94, pp. 251-266, October 2018.

[5] Z. Gligorić, S. Š. Savić, A. Grujić, M. Negovanović, M. Negovanović, and O. Musić, “Short-term electricity price forecasting model using Interval-Valued autoregressive process,” Energies, vol. 11, pp. 1911, July 2018.

[6] L. Peng, S. Liu, R. Liu, and L. Wang, “Effective long short-term memory with differential evolution algorithm for electricity price prediction,” Energy, November 2018.

[7] Saha D., et al., Long-term trend of ventilation coefficient over Delhi and its potential impacts on air quality. Society and Environment, vol. 15, August 2019.

[8] T. Xiao, W. Pei, H. Ye, and G. Niu, “Time-dependent correlation-based ultra-short term forecasting of conditional probabilistic power flow in distribution networks,” High Voltage Engineering, vol. 44, pp. 2362-2371, July 2018.

[9] Z. Chen, et al., “A power flow tracing based approach for determining the usage share of west-to-east power transmission common network,” Journal of North China Electric Power University, vol. 41, pp. 22-26, January 2014.

[10] Y. Guo, W. C. Wu, B. M. Zhang, “Power flow forecasting method and its application in early warning and security countermeasure system,” Automation of Electric Power Systems, vol. 34, pp. 108-111, May 2010.