

Article

# Ultra-Short-Term Prediction of Wind Power Based on Error Following Forget Gate-Based Long Short-Term Memory

Pei Zhang <sup>1,2,\*</sup>, Chunping Li <sup>1</sup>, Chunhua Peng <sup>1</sup> and Jiangan Tian <sup>1</sup>

<sup>1</sup> School of Electrical and Automation Engineering, East China Jiaotong University, Nanchang 330013, China; 2018028085207014@ecjtu.edu.cn (C.L.); chpeng@ecjtu.edu.cn (C.P.); 2018028085207009@ecjtu.edu.cn (J.T.)

<sup>2</sup> School of Electrical Engineering, Beijing Jiaotong University, Beijing 100044, China

\* Correspondence: 8618@ecjtu.edu.cn

Received: 6 June 2020; Accepted: 2 October 2020; Published: 16 October 2020



**Abstract:** To improve the accuracy of ultra-short-term wind power prediction, this paper proposed a model using modified long short-term memory (LSTM) to predict ultra-short-term wind power. Because the forget gate of standard LSTM cannot reflect the correction effect of prediction errors on model prediction in ultra-short-term, this paper develops the error following forget gate (EFFG)-based LSTM model for ultra-short-term wind power prediction. The proposed EFFG-based LSTM model updates the output of the forget gate using the difference between the predicted value and the actual value, thereby reducing the impact of the prediction error at the previous moment on the prediction accuracy of wind power at this time, and improving the rolling prediction accuracy of wind power. A case study is performed using historical wind power data and numerical prediction meteorological data of an actual wind farm. Study results indicate that the root mean square error of the wind power prediction model based on EFFG-based LSTM is less than 3%, while the accuracy rate and qualified rate are more than 90%. The EFFG-based LSTM model provides better performance than the support vector machine (SVM) and standard LSTM model.

**Keywords:** error following forget gate-based long short-term memory; long short-term memory; ultra-short-term prediction; wind power

## 1. Introduction

Renewable energy is increasingly being discussed to phase out fossil fuel power generation to address changes in conditions, such as the new climate system and serious air pollution. Wind power, as one of major source of renewable energies, varies through time and space due to various factors, such as wind speed, wind direction, and temperature. When large-scale wind power is integrated into the grid, fluctuations of wind power bring challenges to the system operations of power systems [1]. Accurate wind power prediction enables secure and economic operation, as well as better utilization of wind power [2]. According to the forecast time scale, wind power forecast is divided into ultra-short-term forecast within 4 h ahead, short term forecast within 3 days, medium-term forecast ranges from 1 week to 1 month, and long-term forecast within 1 year [3]. Ultra-short-term wind power prediction refers to forecast wind power in the next 15 min to 4 h. The prediction interval is 15 min, and the predicted wind power is rolled in next time [4]. Existing wind power forecasting methods can be divided into two categories: physical methods and data-driven methods (statistical and artificial intelligence methods) [5].

The physical models first compute the wind speed at the hub height of wind turbines based on terrain, wind farm layout, numerical weather prediction (NWP) data, and environmental characteristics

(topography, land use, etc.). The physical models then calculate the wind power according to the wind speed power curve or meteorological power parameter model. This method can better reflect the physical essence of meteorological factors. References [6,7] used weather (wind, temperature, lightning density, humidity, barometric pressure, etc.) data to predict wind power. However, physical models rely heavily on meteorological data. However, the meteorological data typically provided once or twice a day, the physical model has to wait for hours between two iterations, which limit their application for ultra-short term forecasting.

To solve the problems of the physical model, researchers have shifted to a data-driven model that can deal with a large number of multivariate data. Data-driven methods usually use historical data of wind power and weather for prediction. The auto-regression integrated moving average (ARIMA), artificial neural network (ANN) method, and support vector machine (SVM) method are commonly used at present [8,9]. Some researchers decompose the wind power and then establish a prediction method based on the time series model [10,11]. However, the decomposition error will be generated by decomposing the wind speed and power sequence. This error will be transferred to the prediction model, which will reduce the prediction accuracy. Reference [12] used wind speed, wind direction, air temperature, and air pressure as input variables, and established an SVM model for short-term wind power forecasting. Its prediction accuracy is better than the artificial neural network. Wind power has time series characteristics. However, the wind power at the predicted time is related, not only to the state at the last moment, but also to the past moments. The existing time series models (such as ARMA) and commonly used neural networks (ANN, SVM, etc.) cannot learn the correlation between wind power and wind speed, wind direction, etc. Therefore, these prediction models are difficult to further improve the prediction accuracy.

To solve the problems of the traditional neural network model, researchers shifted to the based on deep learning model which can deal with time series data. Long short-term memory (LSTM) is a neural network based on deep learning. Compared with the traditional neural network, LSTM has obvious advantages in dealing with a large number of samples and nonlinear data. LSTM has recently received researchers' attention and has been used in the field of power system prediction [13,14]. Some researchers have applied it to the field of wind power prediction [15–17]. For example, there are prediction methods based on traditional LSTM or LSTM combined with optimization algorithm [18,19]. It shows that the prediction accuracy of LSTM is better than ANN, support vector regression (SVR), back propagation (BP), and Bayesian network. Because the LSTM network has short-term memory capability, it can model the influence of the wind power at the previous time on the wind power at the current time. Reference [20] used empirical mode decomposition and vibration mode decomposition to decompose the wind power sequence, then applied the LSTM model to predict wind power. However, the method using multiple decompositions will increase the prediction error. Because decomposition of the original sequence will produce decomposition errors, these errors will be transferred to the prediction model. Reference [21] used principal component analysis to select input variables and established a short-term wind power forecasting model based on the LSTM network. The input data is NWP data without considering the impact of historical data on the prediction moment [22]. Most of these methods do not improve LSTM for wind power prediction. Because the wind power and meteorological data are dynamic time series with large randomness.

To solve the above problems of the traditional LSTM model for wind power prediction, this paper proposes a modified LSTM, called error following forget gate (EFFG)-based LSTM, which updates the output of forget gate using the difference between the predicted value and the actual value. The input data of this paper is the integrated data of NWP and historical wind power. First, Spearman rank correlation analysis is performed between wind power and NWPs to select influential weather factors on wind power. Second, Spearman rank correlation analysis is performed among historical wind powers to determine the timestep of the prediction network. Finally, the EFFG-based LSTM is developed for ultra-short-term wind power prediction. Therefore, this paper provides a kind of one-step ahead forecasting on a 15-min resolution.

## 2. Correlation Analysis to Determine Input Variables for LSTM

There are many factors affecting wind power. Redundant information will be introduced if all meteorological factors are used as input variables. Nevertheless, few factors will result in insufficient information. Wind power has time series characteristics. The wind power at the present moment is related to historical wind power. The time step in the LSTM model determines how long historical wind power data needs to be used as input variables. If the time step is too small, it will cause a lack of prediction information. Otherwise, it will decrease model performance.

Because the probabilistic distribution of NWP data and wind power data is not a normal distribution, Spearman rank correlation coefficient analysis is used to select meteorological factors and determine the time step. The Spearman correlation coefficient is calculated as follows

$$\rho_s = 1 - \frac{6 \sum_{i=1}^n d_i^2}{n(n^2 - 1)} \quad (1)$$

$$d_i = rg(X_i) - rg(Y_i)$$

where  $rg(X_i)$  and  $rg(Y_i)$  are the ranks of each sequence of  $X$  and  $Y$ , and  $n$  is the number of samples.  $X_i$  and  $Y_i$  are two sequences of wind power.

Figure 1 shows that the correlation analysis between wind power and NWP data and the correlation analysis among wind power time series are carried out using the Spearman rank coefficient to determine the input variables of LSTM. A time series, including real wind power (RWP) at the historical time and NWP at forecast time is reconstructed, which is used as the input of each step of prediction for wind power prediction.

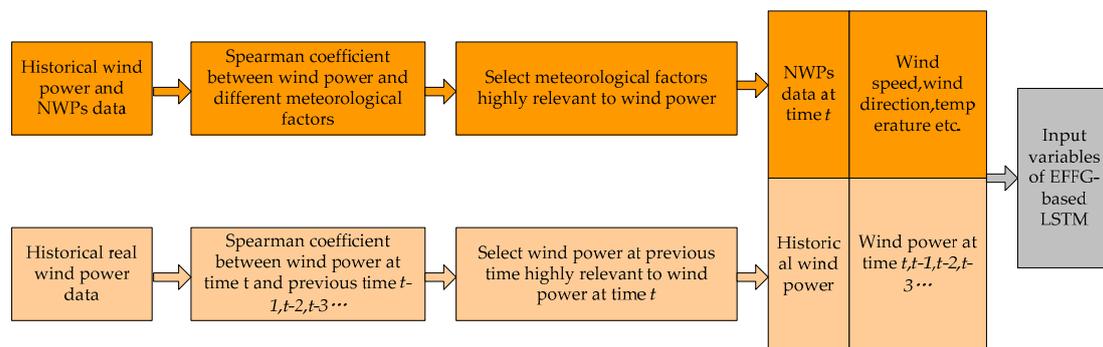


Figure 1. Correlation analysis to determine input variables for long short-term memory (LSTM).

## 3. Error Following Forget Gate-Based LSTM

LSTM has three gates: input gate, output gate, and forget gate, to protect and control the state of LSTM cell [23]. The inputs of the LSTM cell at time  $t$  include the inputs variables at time  $t$ , the outputs at time  $t-1$ , and the state variable at time  $t-1$ . The outputs of the LSTM cell include the output value at time  $t$  and LSTM cell state at time  $t$ . Figure 2 shows the LSTM model and its internal structure.

The forget gate update of traditional LSTM uses the output of the previous moment and the input of the prediction moment. According to the output at time  $t-1$  and the input at time  $t$ , the forget gate  $f(t)$  determines how much cell information is saved to  $t$  time cell state. The mathematical expression of the forget gate is:

$$f_t = S(w_f [h_{t-1}, x_t] + b_f) \quad (2)$$

where  $S(\cdot)$  is the sigmoid activation function;  $W_f$  is the weight matrix of the forgetting gate, and  $b_f$  is the bias of the forgetting gate.

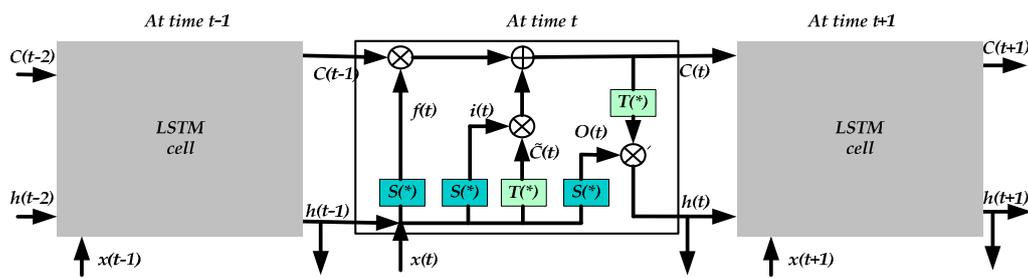


Figure 2. LSTM timing prediction model and its internal structure.

The ultra-short-term wind power prediction is to predict the wind power in the next 15 min. Therefore, when the prediction at time  $t$  is carried out, the actual power value at the  $t$  time and the predicted value are obtained from the model output. The error between the actual value and the predicted value determines how much historical information is forgotten. From Equation (2), it can be seen that the traditional LSTM forgetting gate cannot update with the error between the actual value and the predicted value. The forget gate of traditional LSTM cannot take into account the adjustment effect of the deviation between the predicted value and the actual value at the next predicted moment.

In the ultra-short-term prediction of wind power, the deviation between the actual and the predicted value at the previous moment cannot only reflect the prediction ability of the model but also reflect the positive effect of historical information on the output. If the deviation is large, the influence of the previous moment should have little impact on the prediction moment. The deviation between the predicted value and the actual value at the previous moment is proposed in this paper to update the input for the forget gate. Therefore, the newly proposed forget gate is called error following the forget gate. The calculation formula of the new forget gate is as follows

$$f_i = S(W_f k + b_f) \tag{3}$$

where  $k = h_{t-1} - h_{t-1}$ ,  $h'_{(t-1)}$  is the actual value at time  $t-1$ ;  $h_{(t-1)}$  is the predicted value at time  $t-1$ .  $W_f$  and  $b_f$  are the weight matrix and bias of the forget gate. The weight  $W_f$  and bias  $b_f$  will be optimized by the optimization algorithm in the model training stage, such as Adam.

In comparison with the traditional LSTM network, the forget gate input is enhanced to take into account of the deviation between the predicted value and the actual value at the previous moment. The input gate and output gate remain the same as the ones in traditional LSTM. Figure 3 shows the new structure of the EFFG-based LSTM network at time  $t$ .

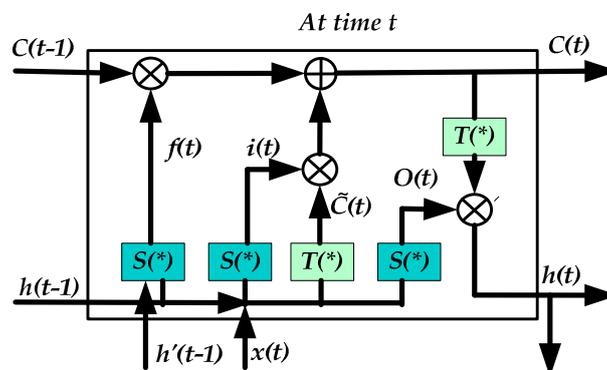


Figure 3. Error following forget gate (EFFG)-based LSTM network internal structure.

The update formula of the input gate is as Formula (4).

$$i_t = s(w_i \cdot [h_{t-1}, x_t] + b_i) \tag{4}$$

where  $S(\cdot)$  is the sigmoid activation function;  $W_i$  is the weight matrix of the input gate, and  $b_i$  is the offset term of the input gate.

The update formula of the current time memory is as Formula (5)

$$\tilde{c}_t = T(w_c \cdot [h_{t-1}, x_t] + b_c) \tag{5}$$

where  $T(\cdot)$  is the tanh activation function;  $W_c$  is the weight matrix of the input gate, and  $b_c$  is the offset term of the input gate.

The update formula of the new cell state is as Formula (6).

$$c_t = f_t \cdot c_t + i_t \cdot \tilde{c}_t \tag{6}$$

The update formula of the output gate is as Formula (7).

$$o_t = s(w_o \cdot [h_{t-1}, x_t] + b_o) \tag{7}$$

where  $S(\cdot)$  is the sigmoid activation function;  $W_o$  is the weight matrix of the input gate, and  $b_o$  is the offset term of the input gate.

The output of the EFFG-based LSTM network is shown in Formula (8).

$$h_t = o_t \cdot \tanh(c_t) \tag{8}$$

#### 4. Ultra-Short-Term Wind Power Prediction Model Based on EFFG-Based LSTM

The prediction model of wind power based on EFFG-based LSTM is shown in Figure 4. Assume a wind farm, according to Spearman correlation analysis, the RWP takes power at the first three moments of the predicted moment, and NWP takes the wind speed and direction at the prediction moment. The deviation between the predicted value and the actual value at the last moment is used as the input of EFFG-based LSTM model to update the forget gate.

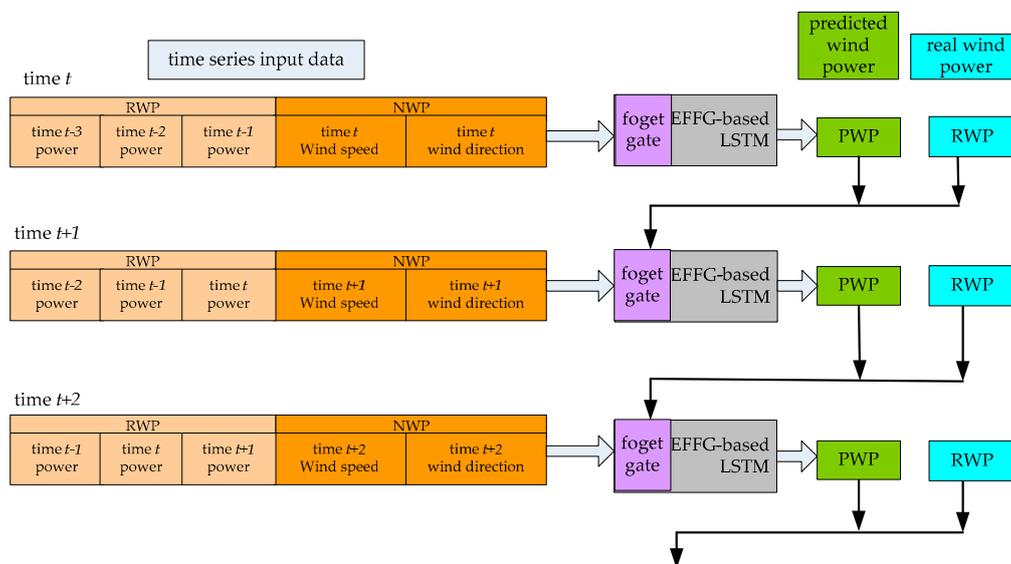
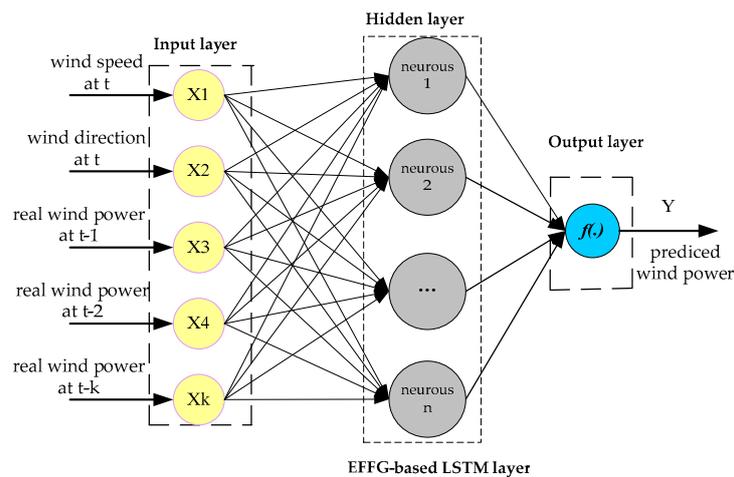


Figure 4. The dynamic structure of wind power forecasting model based on EFFG-based LSTM.

The structure of EFFG-based LSTM is composed of one input layer, one hidden layer, and one output layer. The static network structure is shown in Figure 5.



**Figure 5.** The static structure of wind power prediction model based on EFGG-based LSTM.

**Input layer:**  $X$  is the historical wind power of the past time and the wind speed and direction at the predicted time. The input layer normalizes the input data according to Equation (9).

$$X' = \frac{x - (x_{\max} - x_{\min})/2}{(x_{\max} - x_{\min})/2} \quad (9)$$

where  $x_{\max}$  and  $x_{\min}$  are the maximum and minimum values of the variable, respectively.

**Hidden layer:** The hidden layer is the EFGG-based LSTM network. One hidden layer can ensure the faster prediction calculation speed and avoid overfitting phenomenon caused by too many hidden layers.

**Output layer:**  $f(\cdot)$  is the activation function of the output layer.  $Y$  is the predicted value of wind power at the next time. The output layer weights and biases the output of the hidden layer, and outputs one-dimensional predicted wind power. The inverse normalization is calculated according to Equation (10). Finally, the predicted value of wind power is obtained.

$$Y = \frac{1}{2} [y'(y_{\max} - y_{\min}) + (y_{\max} + y_{\min})] \quad (10)$$

## 5. Case Study

### 5.1. Data Description and Test Design

The data (NWP and historical real wind power) is from a practical wind farm in northwest China. The period is from January to December 2017. The sampling time interval of wind power is 15 min. the time interval of NWP data is 15 min. NWP data includes wind speed and direction of 170 m, 100 m, and 30 m, temperature, pressure, humidity. The NWP data is every 15 min as a rapid refresh model. The model is a method of weather forecast based on the mathematical model of atmospheric movement and using the current weather conditions as input data. The forecast is computed for the next 15 min time step.

Firstly, the correlation coefficient between the wind power and NWP features, RWP historical moments is calculated by using the Spearman correlation coefficient, and then the input variables of the prediction model are selected.

It can be seen from Table 1 that the correlation coefficient average value of the wind speed and wind direction at 100 m is 0.57 and 0.52. The correlation coefficient average value of the wind speed at 30 m is 0.51. The two variables are strongly correlated when the correlation coefficient is greater than 0.5 [24]. When forecasting the wind power at time  $t$ , the wind speed and wind direction at 100 m, wind speed 30 m at time  $t$  are put into the newly constructed input time series.

**Table 1.** Spearman correlation coefficient between wind power and numerical weather prediction (NWP) features. RWP = real wind power.

NWPs	RWP	January to March	April to June	July to September	October to December	Average Value
wind speed at 170 m		0.45	0.38	0.40	0.48	0.43
wind speed at 100 m		0.53	0.57	0.61	0.55	0.57
wind speed at 30 m		0.56	0.46	0.54	0.49	0.51
wind direction at 170 m		0.33	0.45	0.38	0.41	0.39
wind direction at 100 m		0.48	0.55	0.58	0.49	0.52
wind direction at 30 m		0.26	0.36	0.33	0.34	0.32
temperature		0.15	0.09	0.05	0.17	0.12
pressure		0.42	0.37	0.28	0.36	0.36
humidity		0.03	0.05	0.11	0.08	0.07

It can be seen from Table 2 that the correlation coefficient of the power at time  $t-1$ ,  $t-2$ , and  $t-3$  are greater than 0.5. Therefore, the current wind power at time  $t$  has a strong correlation with the power at time  $t-1$ ,  $t-2$ , and  $t-3$ . When forecasting the wind power at time  $t$ , the wind power at the past three times is put into the newly constructed input time series.

**Table 2.** Spearman correlation coefficient between power at  $t$  time and at time  $t$  before.

RWP at Time $t$	15 min Interval (t-1)	30 min Interval (t-2)	45 min Interval (t-3)	60 min Interval (t-4)	75 min Interval (t-5)
RWP at 8:00	0.83	0.72	0.51	0.37	0.21
RWP at 10:15	0.78	0.70	0.48	0.23	0.11
RWP at 13:30	0.81	0.67	0.50	0.18	0.07
RWP at 16:00	0.68	0.60	0.52	0.22	0.03
RWP at 21:45	0.73	0.66	0.49	0.15	0.06

EFFG-based LSTM network parameter set timestep is set to 4, which is the three historical moments of RWP and at  $t$  time of NWP. The number of hidden layer neurons (EFFG-based LSTM cell) is not linearly related to the forecast accuracy (such as root mean square error (RMSE)). The number of neurons in the hidden layer should be determined according to the number of input features and the model training accuracy. The number of neurons in the hidden layer is set to 12 to achieve the best prediction accuracy. The EFFG-based LSTM gate activation function remains at the default value.

Finally, use the historical data of March, June, September, and December 2017 of the wind farm as the training data set of the model. Then, the four groups of prediction tests were conducted as follows:

- (1) Group 1: to predict the wind power within 24 h of February 22.
- (2) Group 2: to predict the wind power within 24 h of May 13.
- (3) Group 3: to predict the wind power within 24 h of September 30.
- (4) Group 4: to predict the wind power within 24 h of November 17.

## 5.2. Forecast Error Computation

Because the actual wind power value has zero value, the mean absolute percentage error (MAPE) in the forecast effect evaluation index will be meaningless [25]. Therefore, root mean square error (RMSE), accuracy  $R1$ , and qualification  $R2$ , are used to evaluate the prediction results.

The calculation formula of RMSE is:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - P'_i)^2} \quad (11)$$

where  $n$  is the number of forecast samples;  $P$  is the actual value of wind power and  $P'$  is predicted value of wind power;  $i$  is the serial number of the actual value and the predicted value.

The calculation formula of the accuracy rate  $R_1$  is:

$$R_1 = \left[ 1 - \sqrt{\frac{1}{N} \sum_{K=1}^N \left( \frac{P_{MK} - P_{PK}}{P_{cap}} \right)^2} \right] \times 100\% \quad (12)$$

In the formula,  $P_{MK}$  is the average value of the actual power in the  $K$  period,  $P_{PK}$  is the average value of the predicted power in the  $K$  period,  $P_{cap}$  is the starting capacity of the wind farm in the corresponding period, and  $N$  is the total number of predicted periods.

The calculation formula for the pass rate  $R_2$  is:

$$R_2 = \frac{1}{N} \sum_{K=1}^N B_K \times 100\% \quad (13)$$

In the formula,  $\left( 1 - \frac{P_{MK} - P_{PK}}{P_{cap}} \right) \times 100\% \geq 75\%$ ,  $B_K = 1$ ;  $\left( 1 - \frac{P_{MK} - P_{PK}}{P_{cap}} \right) \times 100\% < 75\%$ ,  $B_K = 0$ .

### 5.3. Comparison of Prediction Results

To verify that EFGG-based LSTM model has higher prediction accuracy, SVM and standard LSTM wind power prediction model are used for prediction and comparative analysis. Therefore, the same input data for all three methods for the prediction test. The predicted wind power and prediction error curve of Group 1 as shown in Figures 6 and 7. The predicted wind power and prediction error curve of Group 2 as shown in Figures 8 and 9. The predicted wind power and prediction error curve of Group 3 as shown in Figures 10 and 11. The predicted wind power and prediction error curve of Group 4 as shown in Figures 12 and 13. The time in the x-axis is the local time in all figure.

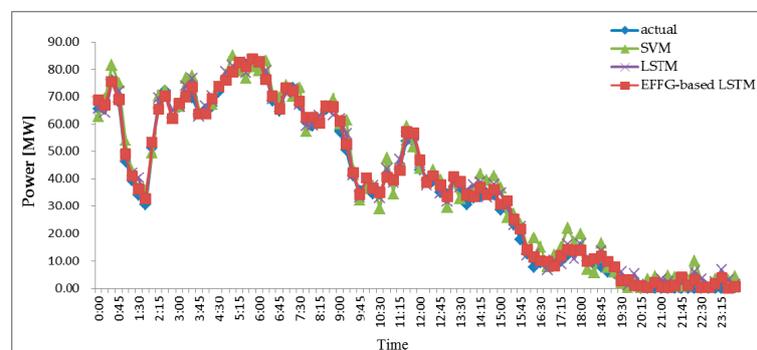


Figure 6. Each model predicts wind power of 24 h of Group 1.

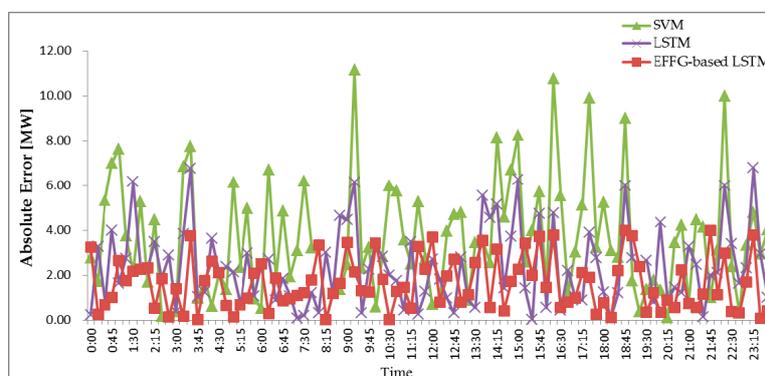


Figure 7. Each model prediction error of 24 h of Group 1.

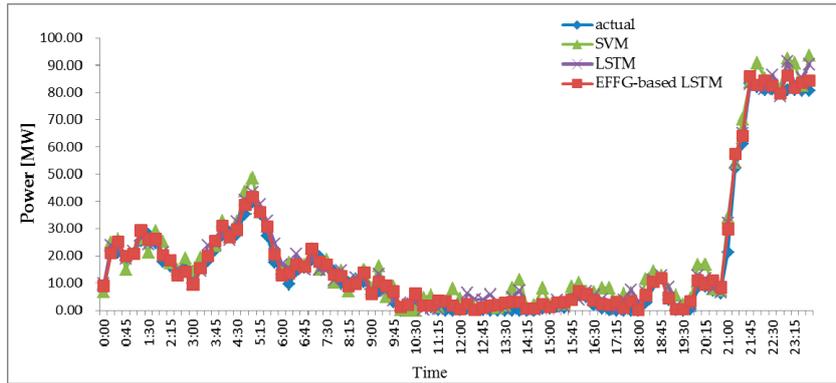


Figure 8. Each model predicts wind power of 24 h of Group 2.

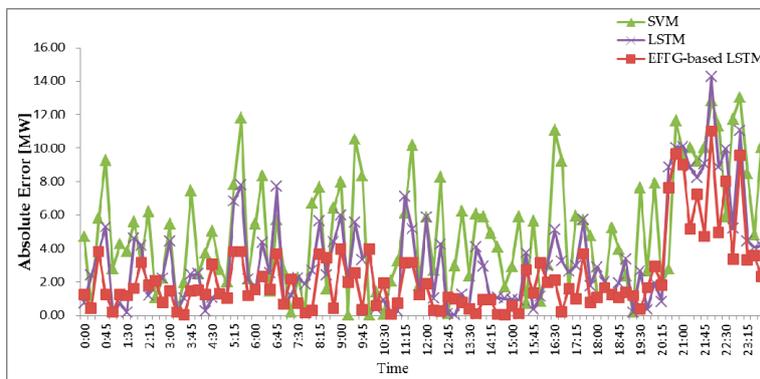


Figure 9. Each model prediction error of 24 h of Group 2.

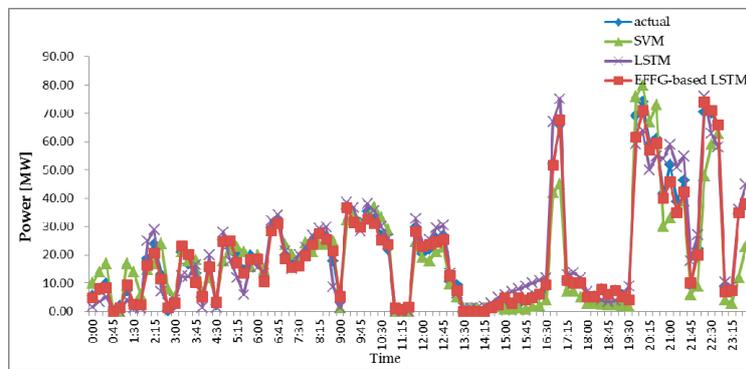


Figure 10. Each model predicts wind power of 24 h in Group 3.

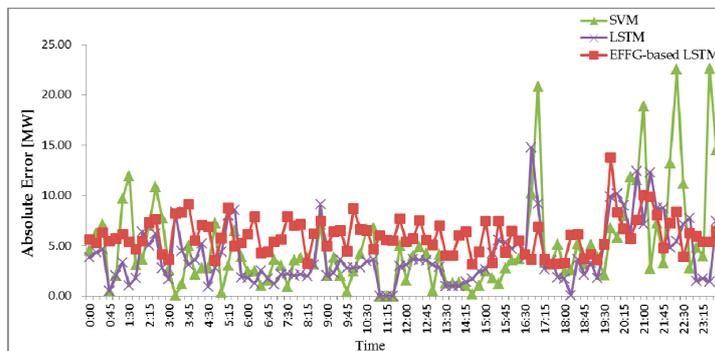
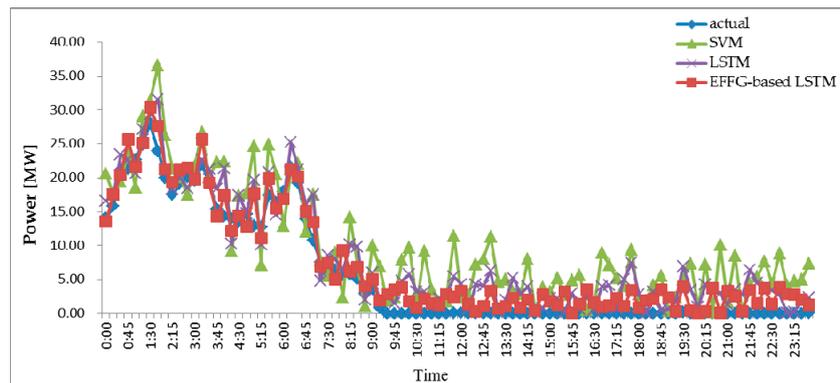
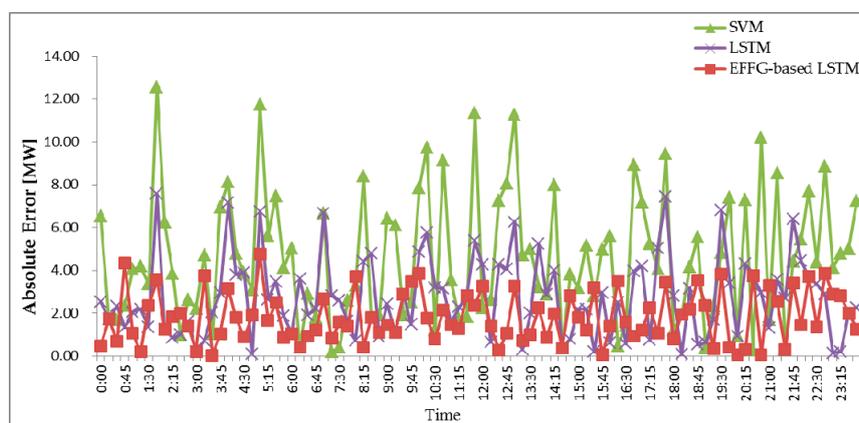


Figure 11. Each model prediction error of 24 h of Group 3.



**Figure 12.** Each model predicts wind power of 24 h of Group 4.



**Figure 13.** Each model prediction error of 24 h of Group 4.

It can be seen from Figures 6–13 that the SVM prediction model is the worst of the three models. The reason is that SVM is not a prediction model suitable for time series and cannot process time series information. LSTM model and EFFG-based LSTM model not only use deep learning technology to optimize network parameters, but also deal with the correlation information and time correlation between wind power time series as a time series model. When the wind power fluctuates, the accuracy of EFFG-based LSTM model is better than traditional LSTM model. The reason is that the traditional LSTM forget gate is updated by the last time output and input data, which cannot reflect the influence of the error between the predicted value and the actual power value on the forget gate. The EFFG-based LSTM forget gate is updated by the error between the predicted value and the actual power value. When the error is large, the forgetting coefficient of the model is large. Then the historical output value will be forgotten more, and the effect of historical value on the model will be smaller. When the wind power suddenly changes, there is no correlation between the next moment's wind power and the previous moment's power, it is necessary to reduce the role of historical value. Therefore, EFFG-based LSTM prediction model has the highest prediction accuracy.

It is difficult to predict the wind power slope, this paper uses the historical wind power slope data to train the prediction model. However, the following Figure 14 shows the up ramp period of wind power in Figure 8 of our paper. It can be seen from the figure that there is a prediction delay and the prediction error is relatively large. The maximum prediction error of the better model (EFFG-based LSTM) is over 10 MW. The same problem appears in Figure 10 of the up ramp period. This is one of the tasks we need to further study.

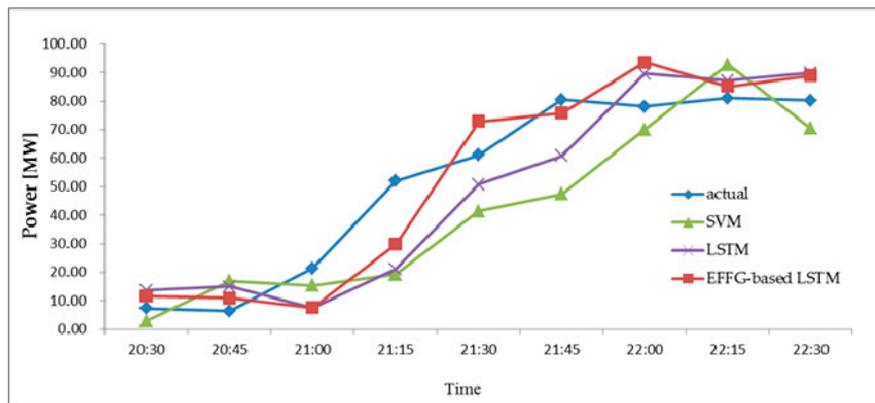


Figure 14. Detail of the ramp in Figure 8.

Tables 3–6 show *RMSE*, *R1*, and *R2* of 4 h and 24 h forecast results of Group 1, Group 2, Group 3, and Group 4.

Table 3. Root mean square error (*RMSE*), accuracy rate (*R1*), pass rate (*R2*) of prediction results of Group 1 of each model.

	SVM	LSTM	EFG-Based LSTM
<i>RMSE</i>	6.71	3.75	2.57
<i>R1</i>	73.19	83.54	90.81
<i>R2</i>	74.07	82.10	90.62

Table 4. Root mean square error (*RMSE*), accuracy rate (*R1*), pass rate (*R2*) of prediction results of Group 2 of each model.

	SVM	LSTM	EFG-Based LSTM
<i>RMSE</i>	7.81	3.23	2.97
<i>R1</i>	73.78	82.44	90.78
<i>R2</i>	75.04	82.22	90.82

Table 5. Root mean square error (*RMSE*), accuracy rate (*R1*), pass rate (*R2*) of prediction results of Group 3 of each model.

	SVM	LSTM	EFG-Based LSTM
<i>RMSE</i>	7.81	3.88	2.87
<i>R1</i>	73.78	81.02	90.11
<i>R2</i>	74.54	80.47	90.65

Table 6. Root mean square error (*RMSE*), accuracy rate (*R1*), pass rate (*R2*) of prediction results of Group 4 of each model.

	SVM	LSTM	EFG-Based LSTM
<i>RMSE</i>	7.01	3.13	2.44
<i>R1</i>	74.68	81.40	90.06
<i>R2</i>	73.04	81.77	90.20

It can be seen from Tables 3–6 that the *RMSE*, *R1*, and *R2* of the EFG-based LSTM prediction model are optimal compared with the SVM and LSTM prediction models. The *RMSE* of the EFG-based LSTM model changed small from Group 1 to Group 4 prediction, and the *R1* and *R2* were above 90% in 24 h prediction. The accuracy and qualified rate of SVM and LSTM models decreased from Group 1

to Group 4 prediction. SVM and traditional LSTM model can't deal with the period when wind power changes dramatically. This problem is solved by an improved forgetting gate, so the performance of the EFFG-based LSTM prediction model is the best.

## 6. Conclusions

This paper proposed the EFFG-based LSTM model for an ultra-short-term wind power forecasting method. The experimental outcomes are as follows: (1) Spearman correlation coefficient method can better find the relationship between predictive factors affecting wind power because the probabilistic distribution of NWP data and wind power data are not known distributions. (2) The input data is the integrated data of NWP and historical wind power. Compared with the monotype input data, it can better reflect the support effect of the wind power prediction. (3) The proposed method can realize an ultra-short-term wind power prediction, considering the influence of the error between the predicted value and the actual value on the prediction model. The forecast accuracy of the EFFG-based LSTM model is better than SVM and the traditional LSTM model.

**Author Contributions:** Conceptualization P.Z. and C.P.; software, C.L. and R.Y.; validation, C.L. and P.Z.; formal analysis, C.L. and R.Y.; writing—original draft preparation, J.T. and C.L.; writing—review and editing, P.Z. and C.P.; visualization, J.T. and C.L.; supervision, P.Z. and M.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The authors gratefully acknowledge the financial support of the research and application of new energy power prediction and evaluation based on micrometeorological big data (19214310D), National Natural Science Foundation of China (51867008), Natural Science Foundation of Jiangxi Province of China (20192ACBL20007).

**Acknowledgments:** Thanks to Ningxia Helanshan No.3 wind farm for providing us with wind farm data.

**Conflicts of Interest:** The authors declare no conflict of interest. The founding sponsors had no role in the design of the study, in the collection, analyses, or interpretation of data, in the writing of the manuscript, and in the decision to publish the results.

## References

1. Lujano-Rojas, J.M.; Osório, G.J.; Matias, J.C.O. A heuristic methodology to economic dispatch problem incorporating renewable power forecasting error and system reliability. *J. Renew. Energy* **2016**, *87*, 731–743. [[CrossRef](#)]
2. Carta, J.A.; Velazquez, S.; Cabrera, P. A review of measure-correlate-prediction (MCP) methods used to estimate long-term wind characteristics at a target site. *J. Renew. Sustain. Energy Rev.* **2013**, *4*, 362–400. [[CrossRef](#)]
3. Tascikaraoglu, A.; Uzunoglu, M. A review of combined approaches for prediction of short-term wind speed and power. *J. Renew. Sustain. Energy Rev.* **2014**, *34*, 785–792. [[CrossRef](#)]
4. Tan, L.; Han, J.; Zhang, H. Ultra-short-term wind power prediction by salp swarm algorithm-based optimizing extreme learning machine. *IEEE Access* **2015**, *8*, 44470–44484. [[CrossRef](#)]
5. Okumus, I.; Dinler, A. Current status of wind energy forecasting and a hybrid method for hourly predictions. *Energy Convers. Manag.* **2016**, *123*, 362–371. [[CrossRef](#)]
6. Sarwat, A.; Amini, M.; Domijan, A.; Damnjanovic, A.; Kaleem, F. Weather-based interruption prediction in the smart grid utilizing chronological data. *J. Mod. Power Syst. Clean. Energy* **2016**, *4*, 308–315. [[CrossRef](#)]
7. Zhao, W.; Weisheng, W.; Bo, W. Regional wind power forecasting model with NWP grid data optimized. *Front. Energy* **2017**, *11*, 175–183.
8. Giorgi, M.D.; Campilongo, S.; Ficarella, A.; Congedo, P.M. Comparison between wind power prediction models based on wavelet decomposition with least-squares support vector machine (LS-SVM) and artificial neural network (ANN). *Energies* **2014**, *7*, 5251–5272. [[CrossRef](#)]
9. Liu, H.; Xiwei, M.; Yanfei, L. Wind speed forecasting method based on deep learning strategy using empirical wavelet transform long short term memory neural network and Elman neural network. *J. Energy Convers. Manag.* **2018**, *156*, 498–514. [[CrossRef](#)]
10. Renani, E.; Elias, M.; Rahim, N.A. Using Data-driven approach for wind power prediction: A comparative study. *J. Energy Convers. Manag.* **2016**, *118*, 193–203. [[CrossRef](#)]

11. Liu, X.J.; Mi, Z.Q.; Lu, B. A Novel Approach for Wind Speed Forecasting Based on EMD and Time-Series Analysis. In Proceedings of the Asia-Pacific Power & Energy Engineering Conference, Denver, CO, USA, 18–20 September 2009.
12. Zeng, J.; Qiao, W. Short-term wind power prediction using a wavelet support vector machine. *IEEE Trans. Sustain. Energy* **2012**, *3*, 255–264. [[CrossRef](#)]
13. Zhang, S.; Wang, Y.; Liu, M.; Bao, Z. Data-based line trip fault prediction in power systems using LSTM networks and SVM. *IEEE Access* **2018**, *6*, 7675–7686. [[CrossRef](#)]
14. Zheng, C. A novel equivalent model of active distribution networks based on LSTM. *IEEE Trans. Neural Netw. Learn. Syst.* **2019**, *30*, 2611–2624. [[CrossRef](#)] [[PubMed](#)]
15. Xu, G.; Xia, L. Short-Term Prediction of Wind Power Based on Adaptive LSTM. In Proceedings of the 2018 2nd IEEE Conference on Energy Internet and Energy System Integration (EI2), Beijing, China, 20–22 October 2018.
16. Qiaomu, Z.; Hongyi, L.; Ziqi, W.; Jinfu, C.; Bo, W. Short-term wind power forecasting based on LSTM. *J. Power Syst. Technol.* **2017**, *41*, 3797–3802.
17. Shahid, F.; Zameer, A.; Mehmood, A.; Raja, M.A.Z. A novel wavelet's long short-term memory paradigm for wind power prediction. *Appl. Energy* **2020**, *269*, 1509–1517. [[CrossRef](#)]
18. Xiaoyun, Q.; Xiaoning, K.; Chao, Z.; Shuai, J.; Xiuda, M. Short-Term Prediction of Wind Power Based on Deep Long Short-Term Memory. In Proceedings of the 2016 IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), Xi'an, China, 25–28 October 2016; pp. 1148–1152.
19. Shi, X.; Lei, X.; Huang, Q.; Huang, S.; Ren, K.; Hu, Y. Hourly day-ahead wind power prediction using the hybrid model of vibrational model decomposition and long short-term memory. *Energies* **2018**, *11*, 3227. [[CrossRef](#)]
20. Hao, Y.; Zuhong, O.; Shengquan, H.; Anbo, M. A cascaded deep learning wind power prediction approach based on a two-layer of mode decomposition. *Energy* **2019**, *189*, 116316. [[CrossRef](#)]
21. Qu, X.Y.; Kang, X.N.; Zhang, C. Short-Term Prediction of Wind Power Based on Deep Long Short-Term Memory. In Proceedings of the Asia-Pacific Power & Energy Engineering Conference, Suzhou, China, 15–17 April 2016.
22. Giorgi, M.G.D.; Ficarella, A.; Tarantino, M. Assessment of the benefits of numerical weather predictions in wind power forecasting based on statistical methods. *Energy* **2018**, *36*, 3968–3978. [[CrossRef](#)]
23. Sepp, H.; Jü, S. Long short-term memory. *J. Neural Comput.* **1997**, *8*, 782–790.
24. Piantadosi, J.; Howlett, P.; Boland, J. Matching the grade correlation coefficient using a copula with maximum disorder. *J. Ind. Manag. Optim.* **2007**, *3*, 305–312. [[CrossRef](#)]
25. Sreelakshmi, K.; Ramakanth, K.P. Performance evaluation of short term wind speed prediction techniques. *IJCSNS Int. J. Comput. Sci. Netw. Secur.* **2008**, *8*, 572–581.

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).