

# Wind Power Forecasting Based on CNN and LSTM Models

Ge Dai \*

Department of Mathematics, Beijing Normal University, Beijing, 100091, China

\* Corresponding author: 1910700522@mail.sit.edu.cn

**Abstract.** Nowadays more and more sustainable and environmentally friendly energy sources are demanded and wind power is playing an increasingly significant role in the energy market. For example, according to the Chinese Academy of Sciences, China plans to generate more than ten times the current scale of wind power in the next decade. Accurate prediction of wind generation is essential for the scheduling, operation, and energy trading of power systems. In this study, a method combining long and short-term memory network (LSTM) and convolutional neural network (CNN) is proposed for wind power generation prediction. In this study, a wind power generation data set from Longyuan Power Group is used to adapt to the input requirements of the model. The data are entered into the LSTM model to capture long-term trends in the data. Next, this paper uses the CNN to process the local spatial features to improve the prediction accuracy. Combining the outputs of these two models, this paper obtains an efficient wind power generation prediction method. In order to evaluate the performance of the proposed method, root mean square error is used to analyze the prediction data. The results show that the combined method of LSTM and CNN is better than individual methods in prediction accuracy and stability. In conclusion, this study proposes a wind power forecasting method combining LSTM and CNN. This method helps to improve the dispatching and operation efficiency of the power system and provides strong support for the promotion of wind power in the energy market.

**Keywords:** Wind power forecast, LSTM, CNN, RMSE.

## 1. Introduction

Wind power is an important environmentally friendly power generation method, but it has not been widely used. The reason why wind power has not been well developed is that the instability and unpredictability of power generation still plague the relevant power sector and enterprises [1]. How to make a reasonable and effective prediction of wind power is a vital issue in the development of new energy. The most direct data basis for wind power generation is the wind turbine itself and related parameter data, climate, and generation time series diagram. For wind turbines that have been built and put into production, the former has changed relatively little over time. The latter two are closely related to time. Therefore, this paper hopes to practice and study the power prediction of wind power generation through the knowledge and model of time series.

There are many studies on wind power forecasting. The first type is prediction research based on the statistical correlation method. Such as the multiple linear regression model, autoregressive moving average model (ARIMA), and so on. Eldali et al. used the ARIMA model to forecast and improve the wind power data set of the Texas Electric Reliability Council (ERCOT) [2]. However, the environmental conditions of wind turbines are less considered, and the prediction accuracy is significantly reduced over a long period of time. Tatiyana et al. studied wind speed and direction [3]. They also used the ARIMA model to forecast wind power data, but again only looked at the forecast over a shorter period of time. Most of these basic methods are suitable for short-term forecasting, but when the amount of data is increasing and the time span of prediction is increasing, the accuracy will decline significantly. With the continuous development of neural network machine learning, many neural network-based models have been applied to the research of wind power prediction, such as the long short-term memory model (LSTM) [4], convolutional neural network (CNN) [5], artificial neural network (ANN) [6], support vector machine (SVM) [7] and other basic and widely used models. At the same time, there are many more complex synthetic models based on neural networks that are specifically applied to forecast wind power prediction. For example, the Statistical hybrid wind power

technology model (SHWIP) proposed by Ozkan et al., mainly forecasts different weather forecasts of weather forecasting systems [8]. Compared with ANN and SVM, this model requires only a small amount of historical data, which can reduce the cost of data storage and improve the efficiency of prediction.

To sum up, the existing prediction models have their own advantages and disadvantages and have not yet reached a high enough accuracy. Only if the forecasts are reasonable and accurate enough can wind power be integrated into a realistic grid with multiple generation methods to serve the public. Therefore, this paper will try to establish a model based on the combination of CNN and LSTM through Python programming language, and use this model to analyze and process the wind power data of Longyuan Power, the world's largest wind power operator under China National Energy Group [9]. This paper attempts to predict its subsequent power generation and gives the forecast results of wind power based on the length of time required for other power generation, such as the start-up of a thermal power plant takes 6 hours or forecast in advance according to the requirements of power generation planning [10].

## 2. Methods

### 2.1. Data Sources

The data set for this literature is collected from Longyuan Power Group Corp. Ltd to select the data of wind power data (SDWPF) for a week with 1008 observations.

### 2.2. Variable Selection

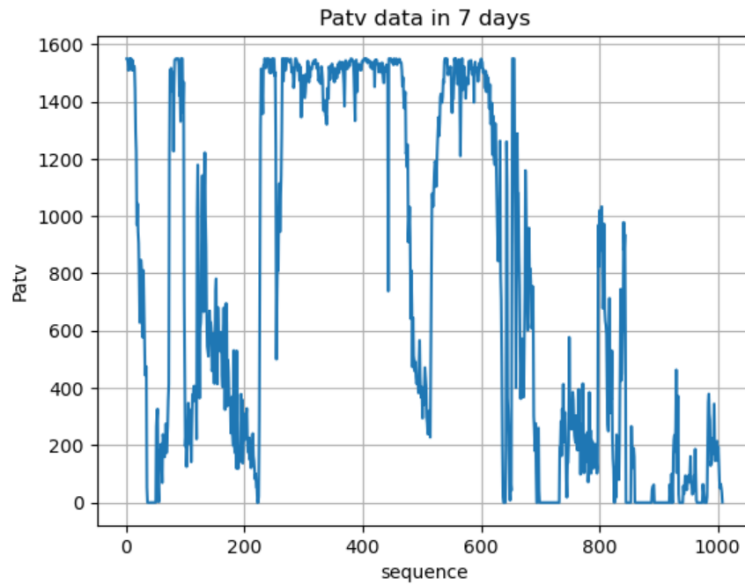
There are thirteen variables contained in each piece of data, which are arranged in Table 1.

**Table 1.** Variables of SDWPF and explanation

Variable	Explanation
TurbID	Wind turbine ID
Day	Record date
Timestamp	Record creation time
Wspd (m/s)	The wind speed recorded by the anemometer
Wdir (°)	The Angle between the wind direction and the turbine position
Etmp (°C)	Environmental temperature
Ndir (°)	Engine room direction, that is, the yaw Angle of the engine room
Prtv (kW)	Reactive power
Patv (kW)	Active power

Among these variables are two powers, the reactive power and the active power. reactive power refers to the electrical energy flowing through the power system that is not used to produce power and is therefore not the target power. The active power, also known as actual power, refers to the electrical energy used in the power system for actual energy transmission and to perform actual power work. Therefore, the active power is the target variable, and this variable is used to analyze and predict future wind power, which is plotted below in Figure 1. Each number in sequence reflects a record creation time (Timestamp).

This figure shows some simple conclusions. First of all, the maximum active power of wind turbines is about 1550, and it can be 0 when there is no wind or when wind turbines stop. Moreover, within a week, there are periodic changes, but there is a large gap between different cycles, and there is no obvious trend.



**Figure 1.** Active Power in 7 Days

### 2.3. Data Processing

There are three types of abnormal data that need to be processed in this dataset. The first is the zero value, there are some active and reactive powers less than zero, and the absolute value is less than 1. This is small for the positive span 0-1550, so we treat it as 0. The second is missing value. Some values cannot be extracted from the data collection system due to errors or errors at the time of data collection. These missing values take the error to 0 in the model evaluation, rather than allowing the error to be affected only by the true value, that is, the resulting error remains 0. Finally, the unknown value, for a period of time due to mechanical equipment or power network needs to stop external power generation, and the actual wind turbine generation power is unknown, these data should not be included in the model evaluation.

### 2.4. Model Selection

Convolutional Neural Network (CNN) is a deep learning model widely used in computer vision and image processing tasks. CNN automatically learns the features of images or other two-dimensional data through multi-layer convolution and pooling operations and performs efficient feature extraction and classification on the input data. Its core idea is to use the convolution layer and pooling layer to extract local features of input data and use these features for further processing and analysis.

Long Short-Term Memory (LSTM) is a special type of recurrent neural network (RNN) specifically designed to process serial data and time series data. Compared with the traditional RNN model, the LSTM model can deal with the long-term dependence better, and solve the problem of gradient disappearing and gradient explosion in the traditional RNN model. Therefore, we choose the combined model of CNN and LSTM to predict the wind power.

### 2.5. Model Evaluation

The model can be evaluated by various statistical methods, and the root mean square error is obtained to evaluate the wind power forecast after a comprehensive analysis of the advantages and disadvantages of various evaluation models. Root mean square error (RMSE) is one of the most commonly used evaluation metrics, which measures the difference between the predicted value of a model and the actual observed value. RMSE is calculated as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum (Patv\_pred - Patv\_true)^2} \quad (1)$$

Patv\_pred is the forecast value of active power and Patv\_true is the true value of active power. After clearing the 0 data and the wrong data in the data description, the structure and parameters of the neural network were modified and adjusted in this paper, and each parameter adjustment was carried out under the condition that other parameters were relatively reasonable. The specific results are shown below. The following table 2 shows the important parameters in the experiment and their explanations:

**Table 2.** Parameter directory

variable	Explanation
Percentage	The percentage of training set data to total data
conv_input	The number of input channels for each convolution layer
num_layers	Layers of the LSTM model
hidden_size	LSTM hidden status dimension
num_epochs	Training cycle
batch_size	The number of each training session
time_step	The training step of LSTM

In the experiment in table 3, we adjusted the time step in the LSTM model, that is, predicted the next data through the historical data of how many time points, and took the next new data as the latest data element of the data column, removed the initial element of the original data column, and formed a data column with the same number of new elements for the next round of prediction. The number of data columns is equal to the time step. We find that RMSE has the smallest value and the highest accuracy when the time step is 7.

**Table 3.** Time\_step

time_step	CNN layers	num_layers	hidden_size	num_epochs	RMSE
10	3	5	128	2000	27.62
9	3	5	128	2000	24.39
8	3	5	128	2000	20.68
7	3	5	128	2000	15.44
6	3	5	128	2000	17.96
5	3	5	128	2000	18.64

In table 4 and 5, we adjusted the number of layers of the LSTM model and the number of dimensions in each hidden layer. The number of layers in an LSTM network refers to the number of stacked LSTM cells. Increasing the number of layers helps the network capture more complex patterns and features, which improves the predictive power of the model. However, excessive increases in the number of layers will increase the calculation cost, gradient disappearance or explosion, overfitting risk, and other problems. Second, the hidden layer dimension refers to the number of neurons in each LSTM unit. The increase of hidden layer dimension can increase the expressiveness of the model and thus improve the prediction performance. However, increasing the hidden layer dimension does not overfit, but it also has the associated problems of increasing computational costs and gradients We find that RMSE has the smallest value and the highest accuracy when the num\_layer is 5 and the hidden\_size is 128.

**Table 4.** Num\_layer

time_step	CNN layers	num_layers	hidden_size	num_epochs	RMSE
7	3	8	128	2000	29.32
7	3	7	128	2000	23.44
7	3	6	128	2000	16.75
7	3	5	128	2000	15.44
7	3	4	128	2000	24.37
7	3	3	128	2000	24.93
7	3	8	128	2000	29.32

**Table 5.** Hidden\_size

time_step	CNN layers	num_layers	hidden_size	num_epochs	RMSE
7	3	5	8	2000	561.78
7	3	5	16	2000	145.66
7	3	5	64	2000	65.42
7	3	5	128	2000	15.44
7	3	5	256	2000	36.45
7	3	5	512	2000	43.74

We also tried to adjust the number of layers of CNN and finally determined the three-layer one-dimensional 128\*128 convolutional layer. The experimental results are shown in Table 6.

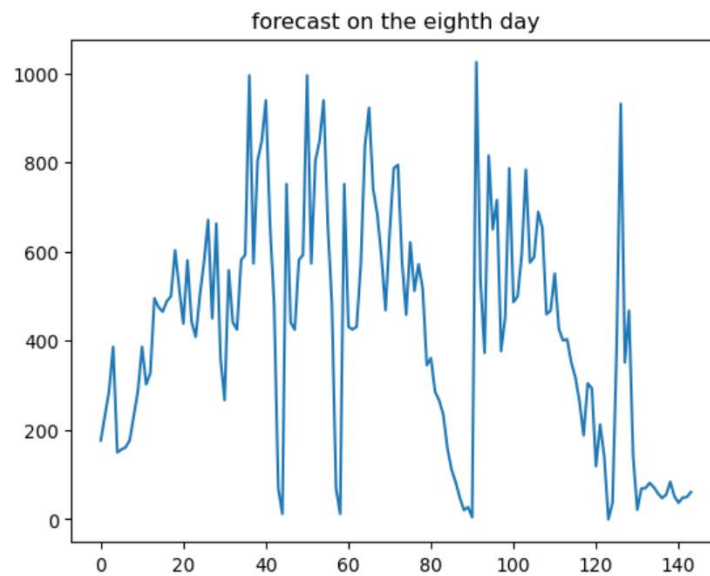
**Table 6.** CNN layers

time_step	CNN layers	num_layers	hidden_size	num_epochs	RMSE
7	4	5	128	2000	16.48
7	3	5	128	2000	15.44
7	2	5	128	2000	29.45

### 3. Results and Discussion

#### 3.1. Forecasting

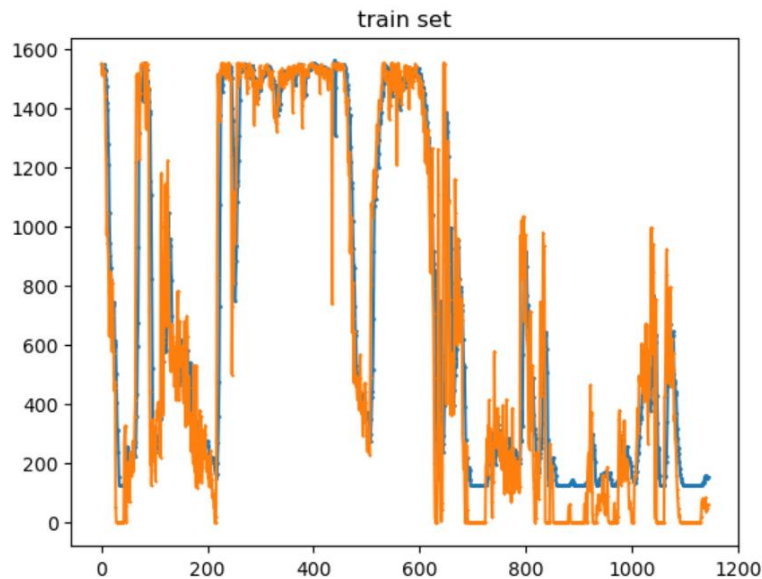
The main results show that there is a good prediction fit near the non-extreme value. With the help of LSTM, the active power in the next 24 hours can be predicted well, and the accuracy rate can reach 70.83% within the allowable error range from the latter part of this passage. Part of the forecast consequences are as follows in the figure 2:



**Figure 2.** Active power on the eighth day

#### 3.2. Evaluation

We used the RMSE mentioned above for model evaluation, and the fitting results in the training set are shown in Figure 3 below:



**Figure 3.** Active power forecast in train set

The same as active power on the eighth day, when we cut down the zero value of the data, it can be believed that this model has a great evaluation ability with its RSME value in this data set is 15.44 which is lower than another model parameter we try in chapter 2.5.

### 3.3. Critical Thinking

There is a certain interpretability that the model cannot predict well in the zero-value region. In real power plants, if the speed of wind that be cut out is exceeded, the power plant will avoid excessive damage to the generator speed. If the wind speed is lower than the speed of wind that be cut in, the generator cannot generate enough power to keep itself running. They also include extreme weather and special circuit failures or limitations. In these special cases, power plants artificially shut off wind power. These influencing factors, such as weather, are not considered in the time series model in this paper, LSTM+CNN. Therefore, in order to solve such situations, it may also be necessary to forecast the local weather, so that zero values can be predicted in a certain period of time. Of course, this passage doesn't include this condition so zero values are excluded when evaluating.

## 4. Conclusion

In this paper, LSTM and CNN models are used to analyze the data within 7 days and to analyze the data the next day. Through this passage, the results show that the combined model of LSTM and CNN can better predict and analyze wind power.

For further research, this paper can refer to the weather forecast data of the Meteorological Bureau, and use the weather prediction model and the spatial prediction model of the geographical location of the generator to make more accurate data prediction. If the user needs to predict the wind power of different time lengths, the parameter needs to be automatically adjusted, which can be further studied. For wind power prediction, human prediction models will become more and more accurate, which brings reliable data and model support for a wide range of wind power development in the future.

## References

- [1] Dai J, Yang X, Wen L. Development of wind power industry in China: A comprehensive assessment. *Renewable and Sustainable Energy Reviews*, 2018, 97: 156 - 164.
- [2] Eldali F A, Hansen T M, Suryanarayanan S, et al. Employing ARIMA models to improve wind power forecasts: A case study in ERCOT. *2016 North American Power Symposium (NAPS)*. IEEE, 2016, 1 - 6.

- [3] Yatiyana E, Rajakaruna S, Ghosh A. Wind speed and direction forecasting for wind power generation using ARIMA model. 2017 Australasian universities power engineering conference (AUPEC). IEEE, 2017, 1 - 6.
- [4] Sherstinsky A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Physica D: Nonlinear Phenomena*, 2020, 404: 132306.
- [5] Chua L O, Roska T. The CNN paradigm. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 1993, 40 (3): 147 - 156.
- [6] Agatonovic-Kustrin S, Beresford R. Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research. *Journal of pharmaceutical and biomedical analysis*, 2000, 22 (5): 717 - 727.
- [7] Jakkula V. Tutorial on support vector machine (svm). School of EECS, Washington State University, 2006, 37 (5): 3.
- [8] Ozkan M B, Karagoz P. A novel wind power forecast model: Statistical hybrid wind power forecast technique (SHWIP). *IEEE Transactions on industrial informatics*, 2015, 11 (2): 375 - 387.
- [9] Alobaid F, Mertens N, Starkloff R, et al. Progress in dynamic simulation of thermal power plants. *Progress in energy and combustion science*, 2017, 59: 79 - 162.
- [10] Chen N, Qian Z, Nabney I T, et al. Wind power forecasts using Gaussian processes and numerical weather prediction. *IEEE Transactions on Power Systems*, 2013, 29 (2): 656 - 665.