Short-term Wind Power Forecasting Based on Convolutional Neural Networks

Jinli Dou a, *, Chun Liu b, Bo Wang c
State Key Laboratory of Operation and Control of Renewable Energy & Storage Systems (China Electric Power Research Institute), Beijing, 100192, China

*, a Corresponding author: victordou@126.com, b liuchun@epri.sgcc.com.cn
C wangbo@epri.sgcc.com.cn

Abstract. Traditional methods of short-term wind power prediction are mostly based on NWP (Numerical Weather Prediction) data on a single station in single-time cross-section and lack in spatiotemporal correlation mining of data. Therefore, a CNN-based prediction method is proposed. Firstly, based on the theoretical analysis of convolution neural network, the input was modelled considering the time correlation, and a variety of convolution neural network structures were designed. Then, a variety of error evaluation criteria were used to evaluate the correlation between single-layer and multi-layer feedforward neural networks as well as the convolutional neural networks prediction method. The error and the actual prediction results were analyzed. Prediction error analysis shows that the convolution neural network model can effectively mine the time correlation between data and improve the accuracy of short-term wind power predictions.

Keywords. Convolutional Neural Networks(CNN), Short-term Wind Power Forecasting, Multidimensional spatial-temporal data.

1. Introduction

With the gradual depletion of fossil fuels and the aggravation of environmental pollution, it is of great significance to make full use of clean energy such as wind energy, solar energy and nuclear energy. China's total installed wind power capacity in 2016 reached 150 million kilowatts, ranking first in the world. It is estimated that by 2020, wind power installed capacity will exceed 230 million kilowatts. However, wind power itself is characterized by randomness, volatility and intermittency, wind abandoning phenomenon is quite serious. Forecasting wind power can provide important reference for power scheduling, correspondingly, it is very important to improve wind power forecasting accuracy. Current wind power prediction methods are mostly based on NWP data, using a variety of mathematical methods to predict. Therefore, research of improvement in wind power prediction accuracy is mainly focused on the following directions:

1) Input data preprocessing. The method generally uses data mining to preprocess the input data for abnormal recognition, then improves the prediction accuracy. In [2], weather types are clustered by data preprocessing and modeled separately according to different weather types. Literature [3] identified
abnormal NWP data by k-means clustering and neural network methods and corrected the data to a certain extent before entering the predictive model.

2) Prediction method improvement. The core of wind power prediction itself lies in establishing a nonlinear relationship between historical data and predicted power. With the advent of the artificial intelligence represented by deep learning, many new algorithms for machine learning are also proposed. In [4], the historical power data and the predicted wind speed are input to a deep neural network, which improves the prediction accuracy to a certain extent. Literature [5] made use of long short-term memory networks to complete the ultra-short-term wind power prediction, while literature [6] improved the accuracy of multi-scale wind speed predictions through deep Boltzmann machines.

3) Combined forecasting method. Combined forecasting method refers to a combination of multiple types of methods or inputting data, it acquires a variety of classification according to different technical routes. There are many combinations of prediction methods, such as literature [7] used a combination of wavelet transform and FA neural network wind power prediction method, and the use of support vector machine to reduce the prediction error.

At present, the research direction lacks the consideration of the correlation of multiple spatial-temporal elements at the input level, mostly taking few elements of a single time section as input, lacking the spatiotemporal correlation mining of multiple elements, and on the other hand, the data input is limited, using data within one year for modeling can easily lead to over-fitting, resulting in poor model generalization ability. The convolutional neural networks prediction method proposed in this paper can take multi-dimensional time series data as an input to mine the temporal correlation of multidimensional data. On the other hand, it can be fully trained in the background of big data to improve the prediction accuracy of wind power.

2. Convolutional neural networks

2.1. Defects of Traditional Feedforward Neural Networks

Traditional feedforward neural networks have a natural disadvantage in dealing with multi-dimensional ordered data such as images and videos. Assuming that pixels in a RGB three-channel image are 200 * 200 * 3, the network is trained with a single fully connected hidden layer containing 100 neurons, training of the weight of 200 * 200 * 3 * 100 is needed, the actual training process is slow and can simply lead to over-fitting phenomenon. This phenomenon is caused by the following two reasons:

1) Multi-dimensional ordered data itself possesses spatial-temporal correlation, for instance, the pixels in each frame in the video are related to both the points around the frame and the pixels at the same position in the nearby frame. Similarly, the wind power is not only related to many weather elements at the predicted time point, but also related to the elements near the predicted time point, directly inputting the single-layer neural network disrupts the space-time correlation between the data.

2) With the deepening of the neural network structure and the increase of training weights, the training time and training data needed will also increase, otherwise the overfitting phenomenon may occur easily. The historical NWP and power data of wind farms that can be extracted in actual operation are less than three years. Under the data resolution of 15 minutes, there will be some data missing and it can hardly meet the training requirements.

2.2. Theory of Convolutional Neural Networks

Convolutional neural network is the first deep neural networks model that has been successfully trained in the field of computer vision and has achieved considerable application achievements. The convolutional neural networks layer is still an indispensable part of the deep neural network and is widely applied in the classification and recognition of the graphics and video processing. A typical convolutional neural network architecture shown in Fig.1:
The basic convolutional neural network is mainly composed of the following network layers:

1) Input layer, suitable for multidimensional ordered data, the input layers of the convolutional neural network are mostly arranged in width * height * depth structure. Corresponding to the long, wide and channel coordinates of the image data, the wind power prediction data was constructed into the form of time, weather elements and space positions, which can greatly preserve the space-time correlation between data.

2) Convolutional layer, the convolutional layer mainly uses convolutional kernels to convolve on the input data, that is, each neuron is connected with a number of local area of the previous layer to calculate the inner product of the weight and small area pixels, a typical convolutional operation is shown in the following Fig.2. The size of the convolution kernel in the graph is 2 * 2. When the convolution operation is performed on a 5 * 5 image, the generated feature map size is (5-2 +1) * (5-2 + 1), the feature map needs further feature extraction to achieve the desired effect. In general, the convolutional neural network selects different numbers of convolution kernels according to the specific situation, so as to extract multiple features of the data.

3) Activation layer, the traditional sigmoid activation function possesses bilateral suppression, the signal feature space mapping has a very good effect, but its application in multilayer neural networks prone to gradient disappear problem, thus rectifier excitation function was used in CNN, which has unilateral suppression, wide excitation boundary and sparse activation characteristics, is widely applied in the construction of deep neural networks.

4) Pooling layers, the main purpose of the pooling layer is to down-sample the data obtained from the previous layer. Currently, the maximum value pooling layer is widely used.

5) Fully connected layer, fully connected layers have different usages depending on which activation function is connected. For instance, SoftMax can be used for classification, while sigmoid and relu can get fixed values. The neurons in this layer will interact with all neurons are connected.
It should be noted that only convolutional and fully connected neurons in the CNN possess weight and offset parameters and need to be trained by data.

CNN has the following two characteristics with comparison to the traditional neural networks:

1) Local connection. During the convolution operation, each convolution kernel is connected with the local receptive field, the parameters in the receptive field are extracted and calculated as a parameter according to the fixed formula, which can avoid parameter explosion growth problem compared with the fully connected way.

2) Parameter sharing, CNN is mainly used to identify displacement, scaling, and other forms of distorted multidimensional data, the convolution kernel parameters in the convolution operation is the same, so with the local connection it can extremely lessen training parameters, which is closer to the actual biological network, reducing the complexity of data reconstruction in feature extraction and classification, and at the same time ensuring the feature extraction ability of the network.

3. Input MODELS suitable for convolution neural network

3.1. Data Screening and Processing
There are many meteorological elements in the NWP that can be screened by calculating the correlation coefficient, and the parameters with large correlation coefficient are selected as the input of the wind power prediction algorithm.

And the continuous data is normalized, in particular, for the wind angle between 0-360 degrees, respectively, can take the sine and cosine values, making the wind direction value change smoothly with the continuous change of wind direction, the final elements screened out were wind speed and direction at 30 meters and 100 meters, humidity and temperature, sea level and ground pressure and other eight elements.

3.2. Inputting model structure
Since the initial application area of convolutional neural networks is image processing, it is necessary to process the data into a certain spatial structure when applied to wind power prediction. Similar to the image of the RGB three channels, the wind speed value, wind direction sine, wind direction cosine, temperature and humidity parameters are relatively independent parameters, they should be placed in different channels, so the final data input structure used in this article is showed in Fig.3, in actual wind power forecast, on the one hand different meteorological elements can be inputted, on the other hand the input time interval can also be changed.
4. Example analysis

According to wind power prediction norms, wind power prediction error evaluation criteria are as follows:

Root mean square error:

$$E_{rms} = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (P_{hi} - P_{si})^2} = \frac{1}{S_{op} \sqrt{n}}, i = 1, L, n$$  \hspace{1cm} (1)

Average absolute error:

$$E_{ma} = \frac{1}{n} \sum_{i=1}^{n} \left(\frac{P_{hi} - P_{si}}{S_{op}}\right), i = 1, L, n$$  \hspace{1cm} (2)

Using the NWP and the actual power data collected in a wind farm in Shandong among two and a half years, a variety of convolutional neural networks and BP neural networks were modeled. CNN (2 hours) takes data in two hours before the predicted time, the prediction error evaluation is shown in Table 1, it can be concluded that single-layer convolution neural network’s performance under various error evaluation criteria are better than the feedforward neural network, which shows that CNN can effectively mine the temporal correlation between NWP data. The convolution neural networks with two layers slightly reduced errors in some aspects, but the forecast of two hours before the time point data as the input is not ideal.

| Table 1. Error evaluation parameters in different neural networks |
|---------------------------------------------------------------|
| | FNN | CNN (1 layer) | CNN (2 layers) | CNN (2 hours) |
|---|---|---|---|---|
| $E_{rms}$ | 15.01% | 13.79% | 13.63% | 14.02% |
| $E_{ma}$ | 10.71% | 9.67% | 9.15% | 9.76% |
| $r$ | 0.7231 | 0.7601 | 0.7659 | 0.7497 |
| $Q_R$ | 0.943 | 0.9474 | 0.9474 | 0.9386 |

Among the actual application of convolution neural network for wind power prediction, there are several points to note and directions to improve:

1. CNN is with layers, neurons, the number of iterations and other adjustable parameters, the input also has a variety of parameters and the choice of the time window, it is necessary to choose a more scientific way to adjust in order to achieve the best the forecast result.

2. With the gradual use of deep learning and artificial intelligence algorithms, CNN can be used as the basic algorithm for deep learning in wind power prediction and other fields. The input model based on time correlation proposed in this paper has very important meaning for reference, it is necessary to mine the spatial correlation based on this model and further use the deep neural network, which is also the next research direction in the background of big data.

5. Conclusion

Aiming at the multidimensional time dependence of wind power forecasting data, this paper proposes a short-term wind power forecasting method based on CNN, and draws the following conclusions:

1. NWP data in a variety of weather elements are time-dependent, through a special arrangement modeled, you can more easily through data mining to improve wind power prediction accuracy.
2. The convolution neural network can effectively extract the correlation between the weather elements over a period of time, so that the wind power prediction can be carried out more accurately. It also has important reference significance for the future deep learning algorithms.

Acknowledgements
This article is supported by Key Technology Project of State Grid Corporation of China (NY1717020) and Beijing Science and Technology Project (Z161100002626039).

References
[1] Electric Power Planning &Electric Institute. Annual Report on China’s Energy Development (2016). Beijing: EPPEI, 2016.
[2] Wang Bo, Feng Shuanglei, Liu Chun. Study on weather typing based wind power prediction[J]. Power System Technology, 2014, 38(1): 93-98(in Chinese).
[3] Xu Qianyao, He Dawei, Zhang Ning. A Short-Term Wind Power Forecasting Approach With Adjustment of Numerical Weather Prediction Input by Data Mining. IEEE Transactions on Sustainable Energy, 2015, 6(4): 1283-1291.
[4] Wu Wenzu, Chen Kunjin, Qiao Ying, et al. Probabilistic Short-term Wind Power Forecasting Based on Deep Neural Networks[C]. PMAPS, 2016.
[5] Zhu Qiaomu, Li Hongyi, Wang Ziqi, et al. Short-Term Wind Power Forecasting Based on LSTM. 2017, 41(12):3797-3802.
[6] C. Y. Zhang, C. L. P. Chen, M. Gan, et al. Predictive Deep Boltzmann Machine for Multiperiod Wind Speed Forecasting." IEEE Transactions on Sustainable Energy, 2015, 6(4): 1416-1425.
[7] Ashraf U1 Haque, M. Hashem Nehrir, Paras Mandal. A Hybrid Intelligent Model for Deterministic and Quantile Regression Approach for Probabilistic Wind Power Forecasting. IEEE Transactions on Power Systems, 2014,29(4):1663-1672.