Research on Short-term Wind Power Forecasting Based on Stacking Integrated Model

Haoyue Wang\textsuperscript{1a*}, Hongtao Zhao\textsuperscript{1b*}

\textsuperscript{1}School of Mathematics and Physics, North China Electric Power University, Beijing 102206, China

\textsuperscript{a*}573615575@qq.com, \textsuperscript{b*}565668212@qq.com

Abstract—With the large-scale use of wind farms, wind power has become an indispensable energy source. However, due to the influence of internal instability factors of wind, it poses a huge threat to the economic efficiency and reliability of power system operation. In order to solve this problem, according to the random volatility of wind power, this paper establishes a wavelet threshold noise reduction model to process wind speed characteristics. Because the wind power is very different when the wind speed is the same, many of the wind turbine's own attributes are added to the input characteristics. Based on the strong generalization ability of decision tree models, the gradient descent tree, extreme gradient boost tree, and lightGBM algorithm are used as the base learner, and the multiple linear regression model is used as the meta-learner to construct a stacking integrated learning model. Compared with methods such as long and short-term memory neural networks, the performance and generalization ability of the cascading generalization algorithm combination model is relatively good.

1. Introduction
Under the background of accelerating global carbon reduction and China's "carbon neutral" and "carbon peak" goals, wind power has become an indispensable and important energy source. Large-scale wind power integration into the power grid brings a series of uncertainties to the reasonable scheduling and safe and stable operation of wind power system\textsuperscript{[2]}. The realization of the accurate short-term prediction of wind power not only reduces the operating cost of the power system, but also provides a reliable basis for the dispatching and stable operation of the power system.

Considering the random variation of wind speed, the wavelet threshold denoising method is used to process the data feature of wind speed, and the appropriate threshold is determined by mean square error and signal-to-noise ratio. Several decision tree models were selected as base learner, considering that GBDT and other models have high generalization ability and are not easily affected by outliers and missing values. Finally, the stacking integrated model was constructed, and the linear regression model was used as the meta-learner. Through comparative tests, the stacking model was determined to have good performance.

2. Model building

2.1. Data Source Description
This paper collected the data of a wind farm fan in western China in the past year, including its external environmental factors: wind speed, blade speed, wind direction, yaw Angle, gear box oil temperature, gear box bearing oil temperature, environmental temperature, engine room temperature,
engine temperature, and internal environmental factors: Blade speed, yaw Angle, gearbox oil temperature, gearbox bearing oil temperature, engine room temperature, engine temperature.

2.2. Model design

2.3. Data preprocessing

2.3.1. Wind power curve clustering
Using historical power data as the clustering object, the wind power data before 24 hours were clustered into two categories, the purpose of clustering is to find similar days. Due to too much data, wind power data of December was only used as the demonstration sample[1].

![Wind power curve clustering](image)
It can be seen from Fig. 2 that the peak value of wind power shown in (a) is significantly lower than that shown in (b), so the clustering curve in (a) is called low fluctuation process and that in (b) is called high fluctuation process in this paper. After clustering, there are 9792 low-fluctuation data, and the rest are 31968 high-fluctuation data.

2.3.2. The wavelet noise reduction

The wavelet threshold denoising method is used to preprocess the wind speed data. The principle is to decompose the signal and separate the signal into different wavelet coefficients. The appropriate threshold is selected to denoise the signal, and the signal larger than the threshold will be retained, and then the signal can be reconstructed. Because the hard threshold method will lead to a sharp peak in the curve behind the small focus, and the signal obtained by the soft threshold method will have oscillation phenomenon, resulting in a large deviation, so the compromise method of soft and hard thresholds is adopted to achieve noise reduction[3-4].

![Fig. 3 Comparison of wind speed before and after denoising](image)

It can be seen from the figure that after noise reduction, the wind speed signal becomes much smoother, and its basic trend remains unchanged, so the compromise method of soft and hard thresholds achieves the expected effect. The processed wind speed is replaced by the original wind speed to complete the data pretreatment.

3. Wind power Prediction based on stacking improved model

The common stacking integration model is a two-layer model. This article uses a three-layer model as an improved stacking integrated learning model. The stacking model integrates the four models of GBDT, xgboost, lightbm, and Extratrees. GBDT and xgboost are used as the first layer of base learners, and then the prediction results of the first layer of base learners are used as the input features of the second layer of learners. Lightbm and Extratrees are used as the base learner of the second layer, and finally the linear model is used as the meta-learner to output the final prediction value[5].
4. The empirical analysis

4.1. Error indicators
According to the requirements of "QGDW10588-2015 Wind Power Forecast Function Specification", the root mean square error of the short-term forecast value should be less than 15%, and the actual power during the curtailment period is replaced by the theoretical power.

4.2. Results analysis
In the previous article, the data set has been divided into a high volatility process and a low volatility process. The training set and the test set are established under each volatility process, and the ratio of the training set to the test set is 20:1. In the prediction of the next 24 hours, next 1 hour and next 10 minutes, the test set uses the data of the last ten days of the high volatility process data set during the high volatility process, and the data of the last two days of the low volatility process data set for the test set during the low volatility process. By comparing the LSTM, GRU, and CNN models with the stacking model, it is illustrated that the stacking model is suitable for short-term wind power forecasting. The following will show the predictions of 144 samples on the last day.

Fig. 5 Wind power forecast for the next 24 hours
As shown in the figure, several lines represent the prediction results of stacking, GRU, LSTM, and CNN models. In the comparison model, the performance of CNN is relatively good, the performance of LSTM model is poor, and the effect of GRU model is slightly better than that of LSTM model. Obviously, none of the three can meet the accuracy requirements. Although the three-layer stacking ensemble learning model is still not satisfactory in prediction, the RMSE in both the high volatility and low volatility processes is less than 15%, which is a better method.

5. Conclusion
In this paper, data processing is carried out first. After clustering the wind power curve of Japan, the clustering curve is divided into high fluctuation process and low fluctuation process. Then the wavelet threshold noise reduction model is established to process the wind speed characteristics. Because the wind power is very different under the same wind speed, many of the fan's own attributes are added to the input characteristics. Secondly, in the prediction model, the stacking integrated learning model is built with gradient descent tree, extreme gradient lift tree, LightGBM algorithm as base learner and multiple linear regression model as meta-learner according to the strong generalization ability of decision tree model. Finally, the wind power of the next 24h, 1h and 10min is predicted by comparing the neural network models such as LSTM, GRU and CNN with the stacking improved model.
In a word, it is concluded that the prediction effect of stacking improved model is significantly better than that of individual prediction, and its performance and generalization ability are better.

In this paper, the point prediction of wind power is carried out, corresponding to the single value of sampling points, which contains less information. In future research, the interval prediction of wind power can be considered to obtain the fluctuation range of wind power, which provides a solid foundation for the stable operation of the power grid.

Acknowledgments
The work is partly supported by the National Social Science Foundation of China (No. 19ZDA081).

References
[1] Zhang Yixiao, Guo Wenpu, Kang Kai, Yao Yunlong, Wang Pan. TDOA/DOA joint positioning method based on clustering and grid search[J]. Tactical Missile Technology, 2020(01): 105-112.
[2] REN Y, SUGANTHAN P, SRIKANTH N. Ensemble methods for wind and solar power forecasting: a sate-of-art review[J]. Renewable and Sustainable Energy Reviews, 2015, 50: 82–91.
[3] Mou Xuejiao, LI Huailiang, Tuo Xianguo. An improved wavelet threshold denoising algorithm and its simulation research [J]. Automation instrumentation, 2020, 41(08): 46-50.
[4] Qian Qiuliang, Dong Baowei, Shao Xinye, Shao Jianlong. Application of Improved Wavelet Threshold Denoising in Bearing Fault Diagnosis [J]. Data Communication, 2020(04): 17-20.
[5] Li Yonggang, Wang Yue, LIU Fengrui, Wu Binyuan. Combination model of short-term wind speed prediction based on Stacking fusion [J]. Power grid technology, 2020, 44(08): 2875-2882.