



## Article

# Wind Speed Forecast Based on Post-Processing of Numerical Weather Predictions Using a Gradient Boosting Decision Tree Algorithm

Wenqing Xu <sup>1</sup>, Like Ning <sup>2,3</sup>  and Yong Luo <sup>1,\*</sup> 

<sup>1</sup> Ministry of Education Key Laboratory for Earth System Modeling, Department of Earth System Science, Tsinghua University, Beijing 100084, China; xwq15@mails.tsinghua.edu.cn

<sup>2</sup> Key Laboratory of Ecosystem Network Observation and Modeling, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Sciences, Beijing 100101, China; ninglk@igsnrr.ac.cn

<sup>3</sup> Yucheng Comprehensive Experiment Station, Chinese Academy of Science, Beijing 100101, China

\* Correspondence: yongluo@tsinghua.edu.cn

Received: 18 March 2020; Accepted: 9 July 2020; Published: 12 July 2020



**Abstract:** With the large-scale development of wind energy, wind power forecasting plays a key role in power dispatching in the electric power grid, as well as in the operation and maintenance of wind farms. The most important technology for wind power forecasting is forecasting wind speed. The current mainstream methods for wind speed forecasting involve the combination of mesoscale numerical meteorological models with a post-processing system. Our work uses the WRF model to obtain the numerical weather forecast and the gradient boosting decision tree (GBDT) algorithm to improve the near-surface wind speed post-processing results of the numerical weather model. We calculate the feature importance of GBDT in order to find out which feature most affects the post-processing wind speed results. The results show that, after using about 300 features at different height and pressure layers, the GBDT algorithm can output more accurate wind speed forecasts than the original WRF results and other post-processing models like decision tree regression (DTR) and multi-layer perceptron regression (MLPR). Using GBDT, the root mean square error (RMSE) of wind speed can be reduced from 2.7–3.5 m/s in the original WRF result by 1–1.5 m/s, which is better than DTR and MLPR. While the index of agreement (IA) can be improved by 0.10–0.20, correlation coefficient be improved by 0.10–0.18, Nash–Sutcliffe efficiency coefficient (NSE) be improved by −0.06–0.6. It also can be found that the feature which most affects the GBDT results is the near-surface wind speed. Other variables, such as forecast month, forecast time, and temperature, also affect the GBDT results.

**Keywords:** wind speed forecast; numerical weather prediction; post processing; gradient boosting decision tree

## 1. Introduction

Among the renewable energy technologies currently developed, wind power is a renewable energy with mature technology and large-scale development prospects. One of the key technologies for the development of wind power is forecasting of the amount of power generated by wind farms. As the output power of a wind turbine is directly related to wind speed, wind power forecasts strongly depend on wind speed forecasts.

In the development of wind power forecast technology for wind farms, mesoscale model simulation is a useful method for wind speed forecast. Rife et al. [1] used the mesoscale numerical weather prediction (NWP) model MM5 to predict the low-level wind in the boundary layer. Storm et al. [2] evaluated the performance of Weather Research and Forecast (WRF) model in predicting the low-level

jet (LLJ) and pointed out that LLJ results simulated by WRF model were similar to observations. This result indicates that the mesoscale model has the ability to capture some characteristics of boundary layer wind. Marquis et al. [3] pointed out that the wind speed predicted by the WRF model can result in wind resources being better developed and utilized. Foley et al. [4] discussed the methods in wind power generation forecasting and mentioned that the use of mesoscale models for dynamic downscaling of large-scale forecasts is the basis of wind power forecasting for wind farms. Zhao et al. [5] presented a day-ahead wind power forecasting system. In the wind power forecasting system, the WRF model was used to forecast the wind speed. Mahoney et al. [6], Stathopoulos et al. [7], and Wyszogrodzki et al. [8] did similar work, mesoscale models are widely used in wind power forecasting.

Based on the NWP model, some research focuses on the improvement of wind speed forecasting results. Deppe et al. [9] used aggregated results of WRF model simulations with different planetary boundary layer (PBL) schemes to improve the turbine height wind speed forecasts. Tateo et al. [10] also used the ensemble method of combining different PBL scheme results. Cheng et al. [11] and Marjanovic et al. [12] explored the impact of the choice of physical schemes on the wind forecasts, and the data assimilation is also an effective way to improve the wind speed forecasts [13–20].

In the current status of wind speed forecasting, combining a mesoscale numerical model with post-processing algorithms is an efficient method [21,22]. Such mesoscale numerical models mainly include WRF model or other NWP models, and the post-processing algorithms mainly include statistical methods and machine learning methods.

The statistical methods used have been developed over many years. The Model Output Statistics (MOS) method has been widely used for a long time [23–27]. In addition to the MOS method, there have also been some studies related to error correction by analyzing the systematic errors of the numerical model. Stensurd et al. [28] compared the grid point temperature data of the mesoscale model with observed data and obtained the systematic deviation of the model. After subtracting the average temperature deviation from the model results, the corrected temperature was closer to the observed value. Another work, by Stensured et al. [29], integrated ensemble predictions into this error correction method. The work used the results of 23 ensemble members, a simple seven-day continuous average was used to correct the deviation, and the deviation correction value of the ensemble result was compared with the temperature at 2 m above ground level (AGL). The corrected result was better than that of the MOS method. Eckel et al. [30] and Hacker et al. [31] also carried out similar studies related to ensemble forecasting and the reduction of systematic error. In addition, other statistical methods, such as the Kalman filter, have also been widely used in numerical weather model post-processing [32–36].

In recent years, machine learning algorithms have been applied to wind forecasting in wind farms. Li et al. [37] used three typical neural networks to predict the 1-h-ahead wind speed. Ishak et al. [38] obtained a mesoscale weather forecast using the MM5 (fifth-generation Penn State/NCAR Mesoscale Model) model, and multiple linear regression (MLR), support vector machine (SVM), artificial neural network (ANN) were used to process the mesoscale model results and output the wind speed. The results showed that the SVM algorithm performed the best due to its ability to capture non-linear relations. Sweeney et al. [39] combined several post-processing methods (Short-term bias-correction, Diurnal cycle correction, Linear least-square, Kalman filter, Mean and variance correction, Directional-bias forecast, and ANN) and proposed a combined post-processing method to reduce the error in the wind speed forecast. Zjavka et al. [40] used a polynomial neural network method to process the output of mesoscale numerical meteorological model and obtained a more accurate wind speed forecast. Zhao et al. [41] built a wind speed forecast system based on WRF ensemble results and post-processing algorithms. Zhao et al. [42] combined non-linear and non-parameter algorithms to correct the wind speed output of the numerical model. Papayiannis et al. [43] proposed a new method based on optimal transportation theory to fit the observed wind speed and model results.

From the previous works, it can be seen that the mainstream method for wind speed prediction uses a numerical weather model to predict the meteorological features and select some of the features

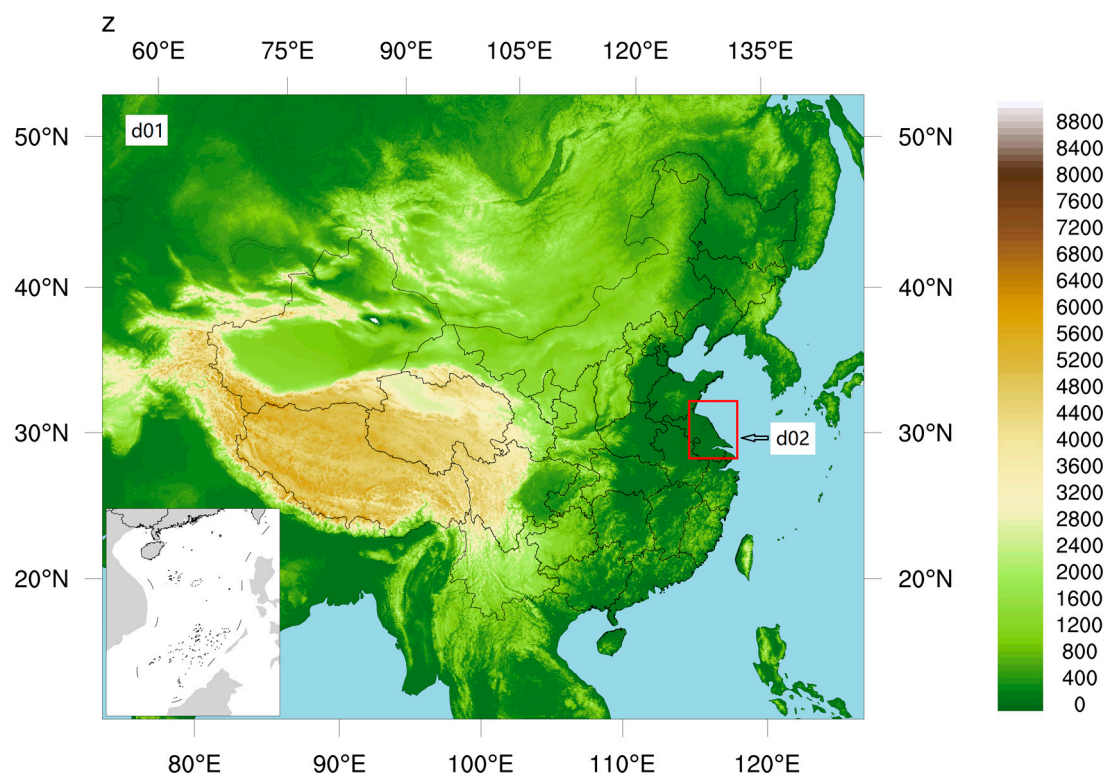
as the input of a post-processing algorithm, then it uses the post-processing algorithm to output the corrected wind speed.

In this paper, the Weather Research and Forecast (WRF) model was used for numerical weather prediction in wind farm areas. The gradient boosting decision tree (GBDT) method was used to perform the post-processing task and output the post-processed wind speed results. Two additional machine learning models were used for comparison. By comparing the RMSE of the WRF model's wind speed output with the RMSE of the post-processing models' wind speed output, we evaluate the performance of GBDT as a post-processing method. Section 2 mainly introduces the WRF model, observation data, and GBDT algorithm. Section 3 analyzes the results of GBDT and its feature importance distribution. Section 4 presents our main conclusions.

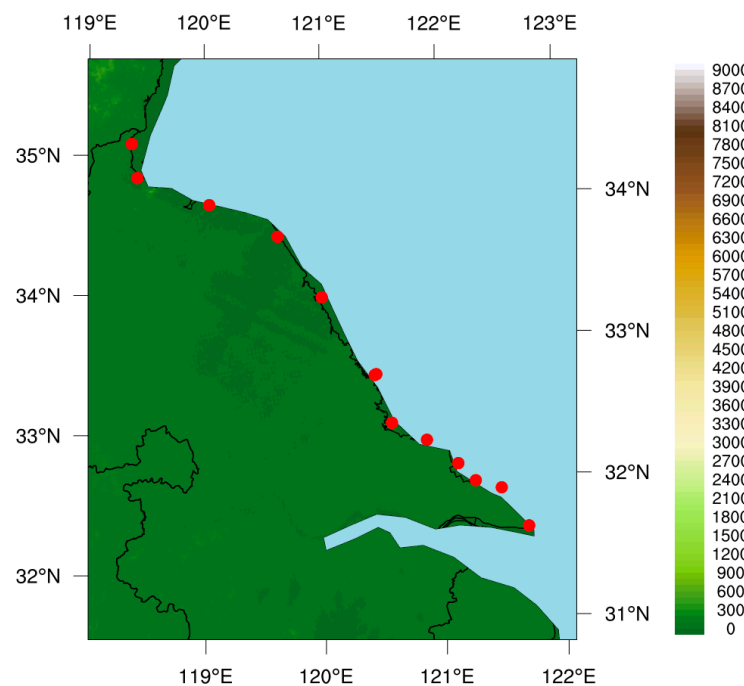
## 2. Experiments

### 2.1. Numerical Weather Model

The numerical weather prediction model we used in our tests is the Weather Research and Forecast (WRF) model (Version 3.9.1) [44]. We designed a nested-grid simulation of the WRF model, where the outer grid has 25 km horizontal resolution and the inner grid has 5 km horizontal resolution. Figure 1 shows domains d01 and d02 of our simulation. Figure 2 shows the domain d02 area, which contains some wind observation towers; these towers were used to evaluate the forecasting results of WRF and measure the wind speed forecast improvement of the post-processing algorithm.

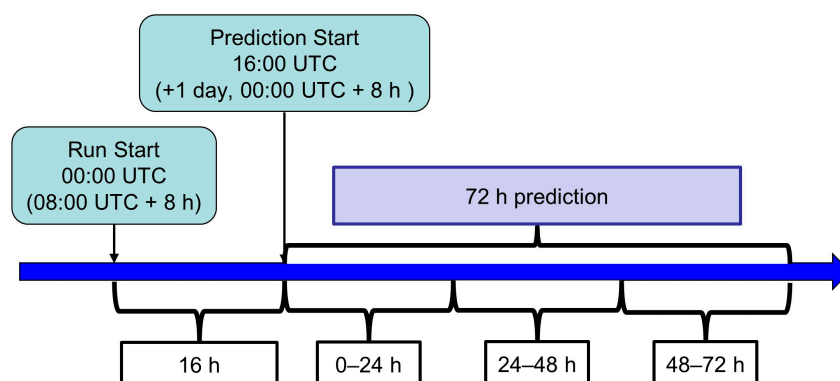


**Figure 1.** Nested grid of WRF models. The d01 grid contains most of China and the d02 grid contains the area of wind towers.



**Figure 2.** Domain 02 of the WRF model. There are 14 wind observation towers (red dots) distributed along the coastal area of Jiangsu province; each tower records 10 m above ground level (AGL), 30 m AGL, 50 m AGL, and 70 m AGL wind observations.

In China, every wind farm needs to provide wind power forecasts to the power grid. Every morning (UTC + 8 h), the wind farm needs to make a forecast of its own power generation for the next few days. The forecast time starts at 00:00 (UTC + 8 h) of the next day and lasts for several days. Therefore, our WRF model starts running at 00:00 (UTC) every day, and the result starting at 16:00 UTC (+1 day, 00:00 UTC + 8 h) every day is taken as the result of wind speed forecast of the wind farm, where the output interval of the WRF model is 10 min. Figure 3 shows the WRF running time configuration. We start the WRF model every day and make 24, 48, and 72 h forecasts. The model runs from 1 June 2009 and continues until 27 June 2010. From the results of the WRF model, we have 24, 48, and 72 h wind speed forecasts every day over the time range of one year.



**Figure 3.** WRF model running time for the wind power forecasting of wind farms in China.

The driving data of the WRF model are from the final operational global analysis data (FNL) [45], which provide the initial field and boundary forcing of the WRF model. The FNL data were provided by the National Centre for Environmental Prediction (NCEP), which have 1 degree of spatial resolution and a 6-h temporal resolution.

Table 1 shows the domain configuration and parameter settings of the WRF model; in our study period, the models that were run every day used the same set of configurations.

**Table 1.** Domain configuration and parameter settings of the WRF model.

Domain	01	02
Grid number	252 × 207	81 × 96
Grid resolution	25 km	5 km
Vertical levels	41	41
Microphysics	Thompson graupel	
Longwave radiation	RRTMG	
Shortwave radiation	RRTMG	
Land-surface	Noah	
Cumulus convention	Kain–Fritsch	
PBL	YSU	

After obtaining the model results, we interpolated the wind field into heights of 10, 30, 50, and 70 m using linear interpolation. We also interpolated the wind field into the wind tower's location using bilinear interpolation. We used the results of the interpolation to compare with the observed wind speed.

The ETA values of the near-surface layers were 1.0000, 0.9960, 0.9920, 0.9900, 0.9851 ... The average altitudes of near-surface layers in model domain were 16.07, 48.22, 72.36, 101.16, and 134.84 m. It can be seen that there are models layer near 50 m and 70 m. The height of the wind tower data is 10, 30, 50, and 70 m. For the data of 10 m, we use the output result of the wind speed of 10 m in WRF model. For data at other heights, we use linear interpolation in the vertical direction for interpolation. Despite the nonlinear characteristics of the wind in the boundary layer atmosphere, due to the existence of model layers at the heights of 50 m and 70 m, the nonlinear changes in the boundary layer have no significant effect on the results of 50 m and 70 m.

## 2.2. Wind Observation Data

In order to evaluate the results of the post-processing algorithm, we used wind speed observation data from 14 wind observation towers to evaluate the model results and the post-processing results. The geographical distribution of wind towers is shown in Figure 2. These wind towers are distributed in the coastal areas of Jiangsu, China, where wind power has been widely developed.

Table 2 shows the terrain height, location and sensor parameters of 14 wind towers. The time interval of the observation data of the wind tower is 10 min. The observation period lasted from June 2009 to May 2010. Each wind tower had observation data at different heights near the ground, including 10 m above ground level (AGL), 30 m AGL, 50 m AGL, and 70 m AGL.

**Table 2.** Wind observation tower locations, terrain height, and sensor parameters.

Tower ID	Terrain Height (m)	Longitude (E)	Latitude (N)	Sampling Frequency	Sensor Bias
10001	1	119.2167	35.0175	1 s	±1%
10002	1	119.2044	34.7666	1 s	±1%
10003	1	119.7784	34.4695	1 s	±1%
10004	2	120.3096	34.142	1 s	±1%
10005	1	120.5754	33.6442	1 s	±1%
10006	0.5	120.8807	33.0107	1 s	±1%
10007	0.5	120.8904	33.0131	1 s	±1%
10008	0.5	120.8955	33.0145	1 s	±1%
10009	2	120.9377	32.6452	1 s	±1%
10010	1	121.1993	32.47	1 s	±1%
10011	1	121.4183	32.2547	1 s	±1%
10012	2	121.5318	32.1059	1 s	±1%
10013	2	121.7346	32.0139	1 s	±1%
10014	1.5	121.8894	31.7003	1 s	±1%

### 2.3. Results Measurements

In order to evaluate the results of the different tests, the following evaluation metrics were calculated to evaluate the WRF model results and post-processing results.

#### 2.3.1. Root Mean Square Error

The root mean square error (RMSE) is widely used in NWP to evaluate the error of wind speed and other meteorological variables. The calculation of RMSE is

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (M_i - O_i)^2} \quad (1)$$

where  $n$  is the number of observations,  $M_i$  is the wind speed in the model results or post-processing results, and  $O_i$  is the wind speed observation.

#### 2.3.2. Index of Agreement

Index of agreement (IA) is a standardized measure of the degree of model prediction error [46–48]. It can be calculated by

$$\text{IA} = 1 - \frac{\sum_{i=1}^n (M_i - O_i)^2}{\sum_{i=1}^n (|M_i - \bar{O}| + |O_i - \bar{O}|)^2} \quad 0 \leq \text{IA} \leq 1 \quad (2)$$

where  $\bar{O}$  is the average value of the wind speed observation.

The Index of Agreement varies between 0 and 1, where a value of IA close to 1 indicates well-matched results, and 0 indicates no agreement at all.

The index of agreement can detect additive and proportional differences in the observed and simulated means and variances [49].

#### 2.3.3. Correlation Coefficient

The Pearson correlation coefficient (R) is also widely used to evaluate the performance of wind speed simulation of NWP. It reflects the correlation between wind speed simulation series and observation series.

$$R = \frac{\text{Cov}(M, O)}{\sqrt{\text{Var}[M] \text{Var}[O]}} \quad (3)$$

Here,  $\text{Cov}(M, O)$  represents the covariance of the model results and observation wind speed, and  $\text{Var}[M]$  and  $\text{Var}[O]$  represent the variance of the model results and observation wind speed. These variables can be calculated as

$$\text{Cov}(M, O) = \sum_{i=1}^n (M_i - \bar{M})(O_i - \bar{O}) \quad (4)$$

$$\text{Var}[M] = \sum_{i=1}^n (M_i - \bar{M})^2 \quad (5)$$

$$\text{Var}[O] = \sum_{i=1}^n (O_i - \bar{O})^2 \quad (6)$$

### 2.3.4. Nash–Sutcliffe Efficiency Coefficient

The Nash–Sutcliffe efficiency coefficient (NSE) is used to assess the predictive ability of numerical model [50]. NSE can be calculated as

$$\text{NSE} = 1 - \frac{\sum_{i=1}^n (M_i - O_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2} \quad (7)$$

NSE ranges from  $-\infty$  to 1. The result of 1 ( $\text{NSE} = 1$ ) means that the model results matches the observation perfectly. The result of 0 ( $\text{NSE} = 0$ ) means that the model results are as accurate as the mean of the observed data, and the less-than-zero results ( $\text{NSE} < 0$ ) means that.

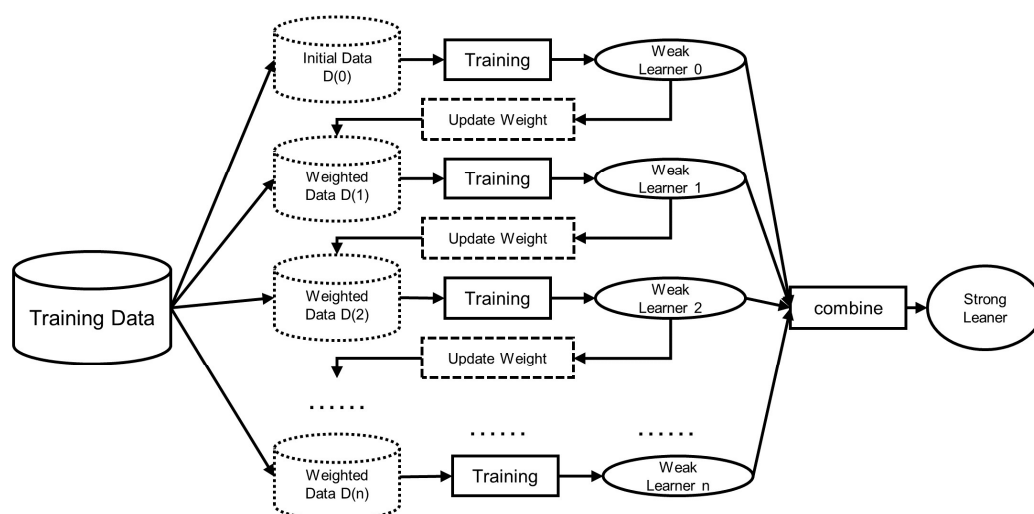
### 2.4. Gradient Boosting Decision Tree

In our study, gradient boosting decision tree (GBDT) is used to conduct the post-processing of WRF model output. The original results of the WRF model were processed by horizontal and vertical interpolation into values of meteorological variables at different heights of each wind tower. In the GBDT model training process, these meteorological variables are used as input, and the wind speed observations are used as output.

#### 2.4.1. Ensemble Learning Approach: Boosting

The GBDT algorithm is a kind of ensemble learning algorithm. Ensemble learning is not a specific machine learning algorithm but completes the learning task by building and combining multiple machine learners; we call these learners “weak learners” and the combined learner a “strong learner” [51].

Boosting is one such ensemble method [52,53]. At the beginning of boosting training, a weak learner is trained with a training data set with initial weights. The weights of the training data set are updated, according to the learning error performance of weak learner. The update of the weights makes the weights of the training samples with high learning errors higher. These poor samples will get more attention in the weak learners, later in the training process. After updating the weights, the GBDT algorithm continues to train new weak learners, based on the updated training data set. This is repeated until the number of weak learners reaches a predetermined number; finally, these weak learners are integrated through a collection strategy to obtain the final strong learner. Figure 4 shows the total process of the boosting algorithm.



**Figure 4.** Model training process of the boosting algorithm in ensemble learning.

In the history of the development of the boosting algorithm, AdaBoost (adaptive boosting) [54,55] is a common algorithm. In each iteration, AdaBoost calculates the error of each sample, calculates a new weight based on the error of each sample, and performs the next iteration. Unlike AdaBoost, GBDT determines the new weight distribution by calculating the error gradient in each iteration. Therefore, the GBDT algorithm can achieve more accurate results than the AdaBoost algorithm.

#### 2.4.2. Classification and Regression Tree (CART)

The gradient boosting decision tree (GBDT) method is a kind of boosting algorithm. GBDT uses a CART (classification and regression tree) decision tree as its weak learner. A CART decision tree can take categorical and numerical features as its input and can be used for classification and regression tasks.

The decision tree algorithm is a classic algorithm in machine learning and was proposed by Quinlan [56]. In the history of decision tree algorithm development, ID3 decision trees and C4.5 decision trees are important types of decision tree algorithms [57]. The C4.5 algorithm improves the deficiencies of the ID3 algorithm. At the same time, C4.5 also has shortcomings such as inability to perform regression and low calculation efficiency. Compared with the C4.5 algorithm, CART has higher computational efficiency and can handle regression problems [58].

For categorical features, the CART decision tree calculates the Gini coefficient [58] to select split features and decide how to split the features in the tree branches. The ID3 and C4.5 algorithms use information gain to select features. CART uses Gini coefficients instead of information gain ratios. The Gini coefficient represents the impurity of the model. The smaller the Gini coefficient, the lower the impurity and the better the characteristics.

Suppose there are  $K$  categories in a feature, and the probability of the  $k$ th category is  $p_k$ . Then, the expression of the *Gini* coefficient is

$$Gini(p) = \sum_{k=1}^K p_k(1 - p_k) = 1 - \sum_{k=1}^K p_k^2 \quad (8)$$

For a given sample  $D$ , suppose there are  $K$  categories, and the number of the  $k$ th category is  $C_k$ . Then, the expression of the *Gini* coefficient of the sample  $D$  is

$$Gini(D) = 1 - \sum_{k=1}^K \left( \frac{|C_k|}{D} \right)^2 \quad (9)$$

For a sample  $D$ , if  $D$  is divided into two parts,  $D_1$  and  $D_2$ , by a value  $a$  of a feature  $A$ , then, under the condition of feature  $A$ , the *Gini* coefficient of  $D$  is

$$Gini(D, A) = \frac{|D_1|}{|D|} Gini(D_1) + \frac{|D_2|}{|D|} Gini(D_2) \quad (10)$$

For the numerical features in the regression problem, the CART decision tree uses a measurement of the sum of variance. The measurement goal of the CART regression tree is: for feature  $A$ , a partition point  $s$  divides the data set into  $D_1$  and  $D_2$ ; we need to find which feature and feature value division points minimize the mean square error of the respective sets  $D_1$  and  $D_2$ , and also minimize the sum of the mean square errors of  $D_1$  and  $D_2$ . The expression is

$$\min_{(A,s)} \left[ \min_{c_1} \sum_{x_i \in D_1(A,s)} (y_i - c_1)^2 + \min_{c_2} \sum_{x_i \in D_2(A,s)} (y_i - c_2)^2 \right] \quad (11)$$

where  $c_1$  is the mean value of the data set  $D_1$ , and  $c_2$  is the mean value of the data set  $D_2$ .

### 2.4.3. Training Process of GBDT

The other boosting algorithm, GBDT, iterates by calculating gradients [59,60]. Using the CART decision tree, we can perform multiple training iterations, where each iteration trains a new CART decision tree based on the training data after the updating of the weights. The GBDT algorithm can be expressed by the formula

$$f(x) = \sum_{t=1}^m \gamma_t h_t(x) \quad (12)$$

where  $f(x)$  is the final strong learner,  $h_t(x)$  is the weak learner of each iteration, and  $\gamma_t$  is the weight of each weak learner in the strong learner. We suppose that the strong learner is  $f_t(x)$  when iterating to the  $t^{\text{th}}$  round during the GBDT iteration process and that the loss function is  $L(y, f_t(x))$ , where  $x$  is the input data (WRF output) and  $y$  is the label (wind speed observation).

When iterating to a new step, GBDT builds the strong learner in a greedy way

$$f_t(x) = f_{t-1}(x) + \gamma_t h_t(x) \quad (13)$$

The newly added CART decision tree  $h_t(x)$  tries to minimize the loss  $L$ , where the new loss function is

$$L(y, f_t(x)) = L(y, f_{t-1}(x) + h(x)) \quad (14)$$

and the new CART decision tree is

$$h_t(x) = \arg \min_h \sum_{i=1}^n L(y_i, f_{t-1}(x_i) + h(x_i)) \quad (15)$$

GBDT's iteration process intends to solve this minimization problem, numerically, using steepest descent: the steepest descent direction is the negative gradient of the loss function evaluated at the current model  $f_{t-1}$ , which can be calculated for any differentiable loss function

$$f_t(x) = f_{t-1}(x) - \gamma_t \sum_{i=1}^n \nabla_f L(y_i, f_{t-1}(x_i)) \quad (16)$$

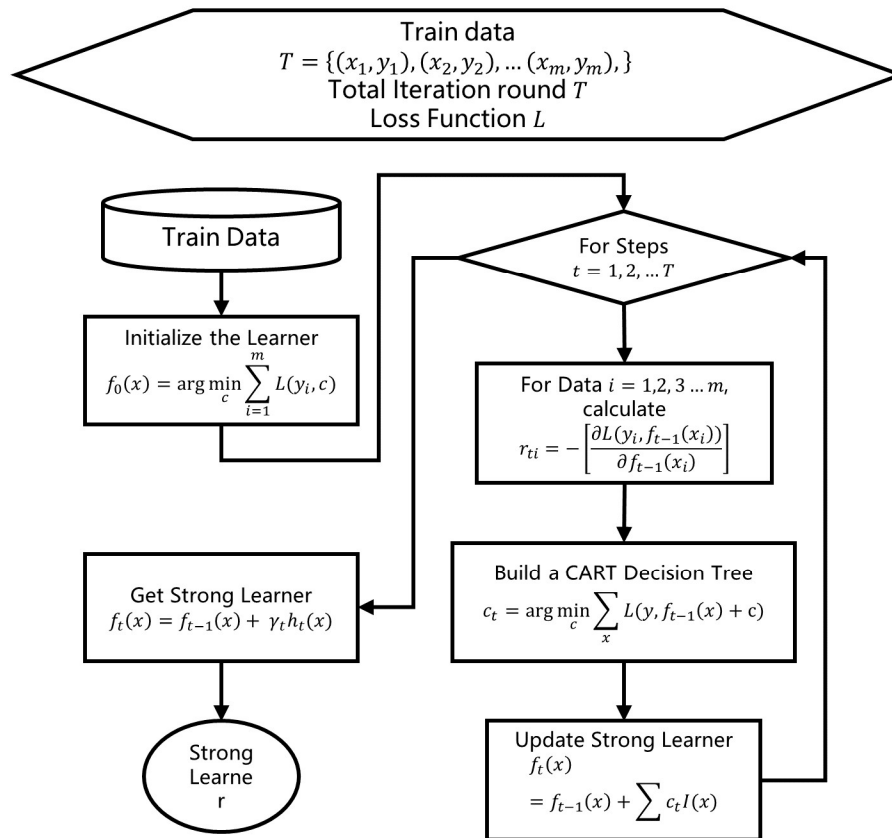
where  $\gamma_t$  is chosen using line search

$$\gamma_t = \arg \min_{\gamma} \sum_{i=1}^n L\left(y_i, f_{t-1}(x_i) - \gamma \frac{\partial L(y_i, f_{t-1}(x_i))}{\partial f_{t-1}(x_i)}\right) \quad (17)$$

where  $\gamma$  is the weight of the weak learner.

Figure 5 shows the training process of GBDT. When performing the  $t$ th round of CART decision tree training, the loss of the sample is used to calculate the negative gradient of decision tree, and the negative gradient of the loss function of the  $i$ th sample can be expressed as

$$r_{ti} = -\left[ \frac{\partial L(y_i, f_{t-1}(x_i))}{\partial f_{t-1}(x_i)} \right] \quad (18)$$



**Figure 5.** Training process of the gradient boosting decision tree method.

By using  $(x_i, r_{ti}) (i = 1, 2, 3 \dots n)$ , a CART decision tree can be trained, and the  $t$ th CART decision tree in the integrated model is obtained as

$$c_t = \arg \min_c \sum_x L(y, f_{t-1}(x) + c) \quad (19)$$

where  $c$  is the combination of leaves of the decision tree and  $c_t$  is the leaf combination of the  $t$ th decision tree.

After we obtain the  $t$ th decision tree, we can update the strong learner

$$f_t(x) = f_{t-1}(x) + \sum c_t I(x) \quad (20)$$

where  $I(x)$  is the leaf output of input  $x$ .

#### 2.4.4. Feature Importance of GBDT

After training the GBDT model, we can calculate the feature importance distribution of the GBDT model to obtain each feature's importance. During the branching of the decision tree, the variable to be split and the split value of the variable are determined by calculating the information gain. The information gain can be expressed as  $I(A, D)$ , where  $A$  is the features and  $D$  is the data samples. After all decision trees are constructed, the importance (or contribution value) of a feature can be obtained by calculating the information gain of the feature for the decision tree and dividing by the total frequency of the feature in all of the trees of the GBDT strong learner

$$S(a) = \frac{\sum I(a, D)}{N_a} \quad (21)$$

where  $N_a$  is the total frequency of the feature  $a$  in all trees.

### 2.5. Features Selection and Parameters Setting of GBDT

After we obtain the WRF model forecast results for the 0–24 h, 24–48 h, and 48–72 h forecasts, we need to decide which variables to use as the input features of GBDT. The output data interval of the WRF model is 10 min. Therefore, we take the output of each moment of the WRF model as one record of input data for GBDT.

In fitting the near-surface wind speeds, we used multiple variables at different heights and pressure layers as the input features of the GBDT model. We used the linear interpolation method to interpolate the model results into different layers. These variables consist of wind speed, wind direction, temperature, pressure, height, absolute vorticity (avo), and potential vorticity (pvo). Table 3 shows the variables at different pressure layers, the pressure layers being 850, 700, 500, and 300 hPa. Table 4 shows the variables at different height layers, with 29 height layers from 10 m to 5000 m; as our post-processing output result is the near-surface wind speed, we used more height layers in the lower atmosphere.

**Table 3.** GBDT input features at different pressure layers.

Variables	Pressure Layers
Wind speed, wind direction, temperature, height, avo, pvo.	850 hPa, 700 hPa, 500 hPa, 300 hPa

**Table 4.** GBDT input features at different height layers.

Variables	Height Levels
Wind speed, wind direction, temperature, pressure, avo, pvo	10 m, 30 m, 50 m, 70 m, 90 m, 100 m, 120 m, 150 m, 200 m, 250 m, 300 m, 350 m, 400 m, 450 m, 500 m, 600 m, 700 m, 800 m, 1000 m, 1250 m, 1500 m, 1750 m, 2000 m, 2500 m, 3000 m, 3500 m, 4000 m, 4500 m, 5000 m.

As the GBDT model is based on the CART decision tree, which can effectively deal with categorical features, we added several categorical features into the input of the GBDT model. Table 5 shows the categorical features used in GBDT. For the date of each forecast record, we took ‘month’ as a categorical feature, the month feature ranging from January to December. As for the time of each forecast record, we took ‘hour’ as a categorical feature, where this feature ranged from 1–24, which indicated the forecast record at different hours of the day.

**Table 5.** GBDT input categorical features.

Feature	Categories
Month	January, February, March, . . . . . , December
Hour	1, 2, 3, . . . . . , 24.
Wind direction	N, S, E, W, NW, NE, SW, SE

When we completed the feature engineering, we built the GBDT ensemble model to carry out the post-processing work. We used LightGBM [61] to build our GBDT models, where the input of LightGBM was the features we obtained from the output of each forecast time of the WRF model and the output of LightGBM during the training process is the wind speed observations from wind towers.

We divided the total forecast records into two parts: one was training data, used to train the GBDT model, and the other was testing data, used to evaluate the error of the GBDT model.

In the WRF results and observation data, the interval between adjacent data records is 10 min. The state of the atmosphere is unlikely to change a lot within 10 min, the adjacent data records are very similar. Therefore, if all the data are randomly divided into training data and test data according to the traditional machine learning method, the results cannot reflect the true performance of the model. Based on this, we chose to use the entire day of data as a training set or test set.

In each month, we chose the data in date 3, 7, 11, 15, 19, 23, and 27 as test data and data in other date as train data. Table 6 shows the train and test data split in one month. For each wind tower, we trained the GBDT model on the WRF output observation speed at different forecast times (0–24 h, 24–48 h, and 48–72 h) and different heights (10, 30, 50, and 70 m).

**Table 6.** Train and test data split in each month.

Date Used as Train Data	Date Used as Test Data
1, 2, 4, 5, 6, 8, 9, 10, 12, 13, 14, 16, 17, 18, 20, 21, 22, 24, 25, 26, 28, (29), (30), (31)	3, 7, 11, 15, 19, 23, 27

(29), (30), (31) means that some month may not have these date.

Before the training of the GBDT models, we had to set the parameters of the GBDT models. Our GBDT models mainly needed to tune two parameters: the number of leaves and the minimum number of data in each leaf. These two parameters were both related to the effect of fitting and could reduce overfitting. Table 7 shows the parameter configurations of LightGBM and the tuning ranges of the two parameters.

**Table 7.** LightGBM parameters configuration and parameter tuning ranges.

Param Name	Value/Value Range
Number of iterations	2000
Learning rate	0.1
Number of leaves	10, 20, 40, 80, 160
Minimum data in leaf	10, 20, 40, 80
Bagging fraction	0.8
Bagging frequency	5
Feature fraction	0.9
Metric	Mean square error

In Table 7, two parameters—number of leaves and minimum data in leaf—needed to be tuned. We set a pairwise combination of the values of two parameters, where the number of leaves contained 10, 20, 40, 80, and 160; and the minimum data in leaf contained 10, 20, 40, and 80.

## 2.6. Models Used for Comparison

Most previous machine learning based post-processing algorithms used artificial neural network (ANN) as its' machine learning model or a part of ensemble model. Also, the basic model of our GBDT algorithm is CART decision tree regression (DTR). Based on these, we chose the multi-layer perceptron regression (MLPR) and DTR as the model used for comparison to show the performance improvement of GBDT over traditional machine learning models. The DTR and MLPR models used the same train data and test data as GBDT and the RMSEs of each model was calculated to compare the performance of post-processing results of WRF. Table 8 is the parameters setting of MLPR model and Table 9 is the parameters setting of DTR model.

**Table 8.** Parameters setting of MLPR model.

Parameter Name	Value
Hidden layer sizes	100
Activation function	Relu
Optimization method	Adam
Iterations	200
Loss function	Mean Square Error (MSE)
Learning rate init	0.001

**Table 9.** Parameters setting of DTR model.

Parameter Name	Value
Criterion	Mean Square Error (MSE)
Split method	Best split
Max depth	No limit
Iterations	200
Loss function	Mean Square Error (MSE)
Learning rate init	0.001

### 2.7. Significance Test

When we obtain the statistical variables of all test results, we need to perform significance tests on these statistical variables to verify whether the results of different tests are significantly different. Among the statistical results of all tests, we need to test whether the following statistical variables are significant in different tests:

1. Whether the statistical variables (RMSE, IA, R, NSE) of GBDT results have changed significantly compared to WRF results.
2. Whether the statistical variables (RMSE, IA, R, NSE) of GBDT results have changed significantly compared to the comparison models (DTR, MLPR).

For the significance test, two-sample Student's *t*-test is used to calculate the *p*-value of the following hypothesis

$$H(test1 \& test2) : S_{test1} - S_{test2} = 0 \quad (22)$$

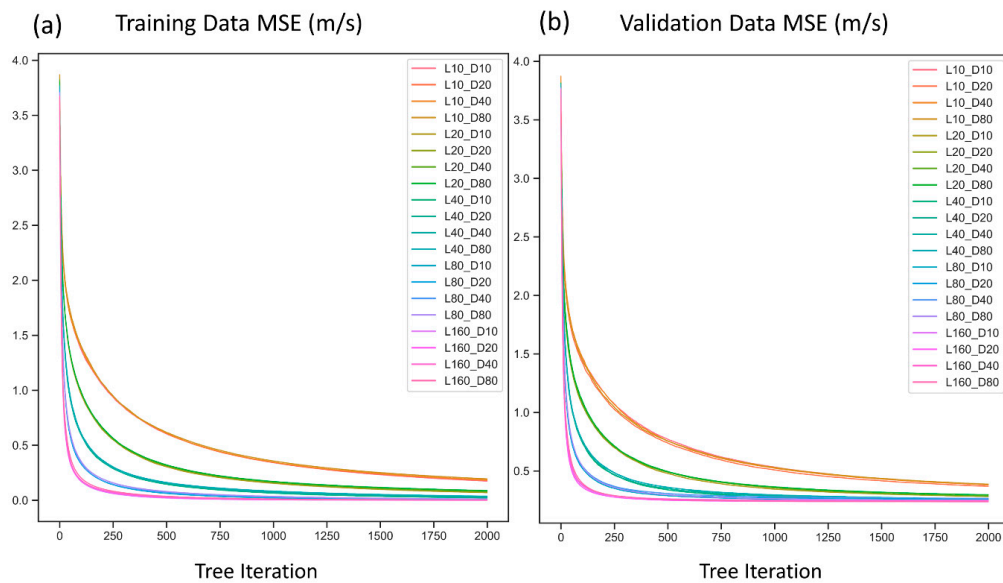
where  $H(test1 \& test2)$  is the *t*-test hypothesis of the two tests,  $S_{test1}$  is the statistical variable of test 1 and  $S_{test2}$  is the same statistical variable of test 2. If the *p*-value is less than 0.01, then it can be proved that two results are significantly different at a level of 99% confidence, while a *p*-value less than 0.05 means that the 'significant difference' passed the 95% confidence level.

## 3. Results

### 3.1. GBDT Parameters Tuning Results

We used the WRF model output and 10 m wind speed observations of tower '10001' to perform the GBDT parameter tuning. The wind speed forecast time of the tuning work was 0–24 h. Figure 6 shows the parameter tuning results of the parameters 'number of leaves' and 'minimum data in leaf'. From Figure 6, we can see that when 'L (number of leaves)' is the same, the curves of different 'D (Minimum data in leaf)' are very close. With larger 'L', the faster the MSE curve decreases. Thus, a larger number of leaves can reduce MSE and achieve better post-processing results.

Table 10 is the result of parameter tuning after 2000 iterations. For each L and D, we have the results of training data (train) and test data (val). It can be seen from Table 7 that, although increasing 'L' can reduce the MSE of training data and test data, it will cause overfitting. In the case of keeping 'L' unchanged and changing 'D', we can see that 'val' has no obvious change, and the MSE of 'train' increases, indicating that increasing 'D' can weaken overfitting. Combining these results, we set the Number of leaves to 80 and the Minimum data in leaf to 80 in the GBDT wind speed post-processing model.



**Figure 6.** Parameter tuning results for ‘number of leaves’ and ‘minimum data in leaf’. Different lines represent different combinations of parameters (e.g., ‘L10\_D20’ represents Number of leaves set to 10 and minimum data in leaf set to 20: (a) is the training data mean-square error (MSE) change with iteration step; and (b) is the test data MSE change with iteration step.

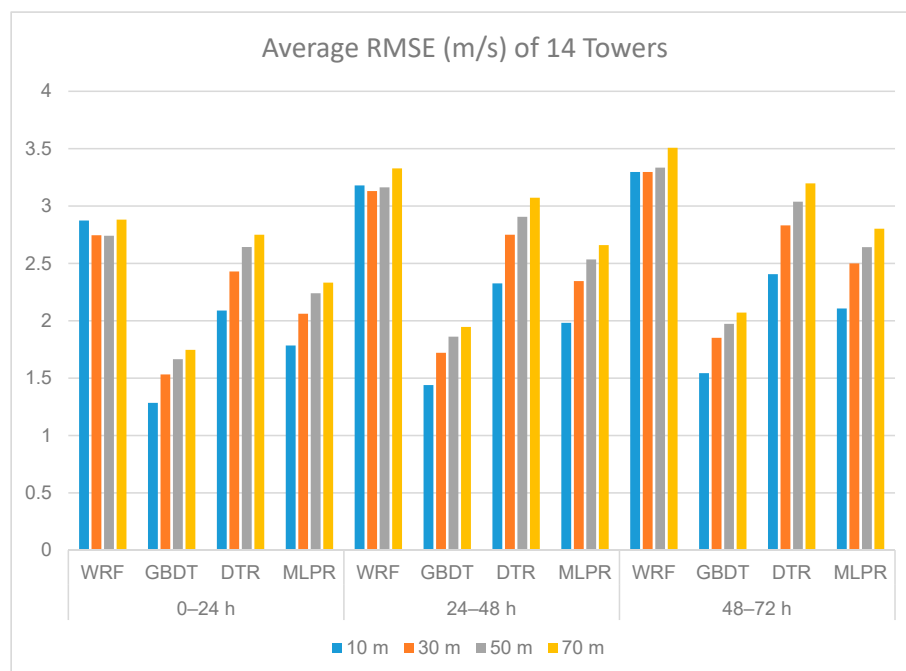
**Table 10.** Mean square error (MSE) of number of leaves (L) and minimum data in leaf (D) pairs after 2000 iterations.

L \ D	10		20		40		80	
	Train	Val	Train	Val	Train	Val	Train	Val
10	0.174	0.380	0.179	0.369	0.185	0.382	0.195	0.387
20	0.069	0.281	0.074	0.285	0.080	0.287	0.088	0.296
40	0.022	0.248	0.025	0.257	0.029	0.263	0.037	0.264
80	0.004	0.250	0.005	0.255	0.007	0.266	0.011	0.250
160	0.000	0.249	0.001	0.237	0.001	0.243	0.002	0.238

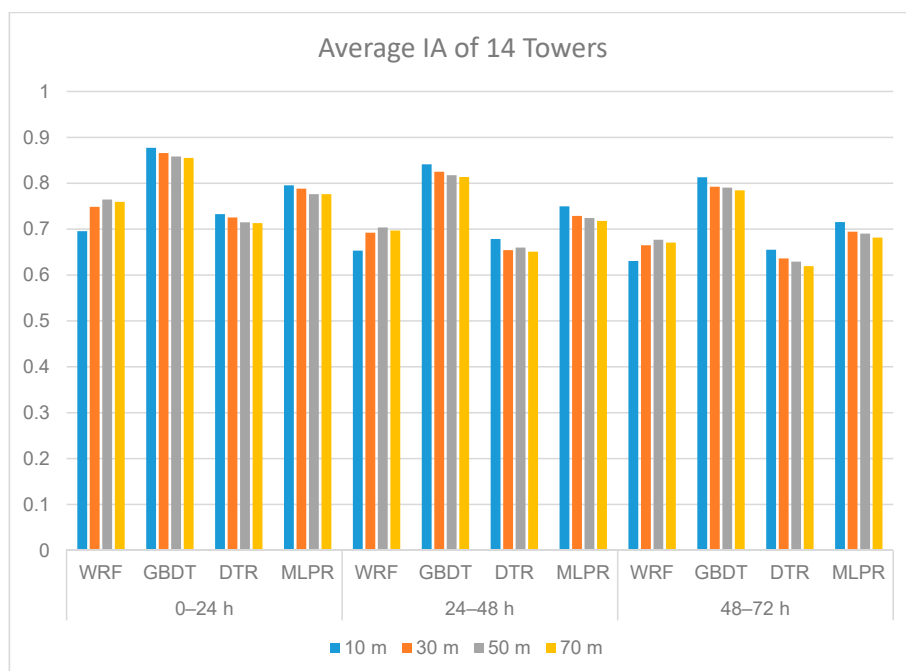
### 3.2. Post-Processing Results

After parameter tuning, we used the train data in Table 6 to train the GBDT model for 0–24 h, 24–48 h, and 48–72 h wind speed forecasts at different heights of different towers and evaluated the results with the test data. We calculated the RMSE, IA, R, and NSE of the wind speed output of WRF model results and post-processing results using the test data sets. Appendix A contains the RMSE results of all wind towers, including 0–24 h (Table A1), 24–48 h (Table A2), and 48–72 h (Table A3). Appendix B contains the IA results of all wind towers, including 0–24 h (Table A4), 24–48 h (Table A5), and 48–72 h (Table A6). Appendix C contains the NSE results of all wind towers, including 0–24 h (Table A7), 24–48 h (Table A8), and 48–72 h (Table A9). Appendix D contains the R results of all wind towers, including 0–24 h (Table A10), 24–48 h (Table A11), and 48–72 h (Table A12).

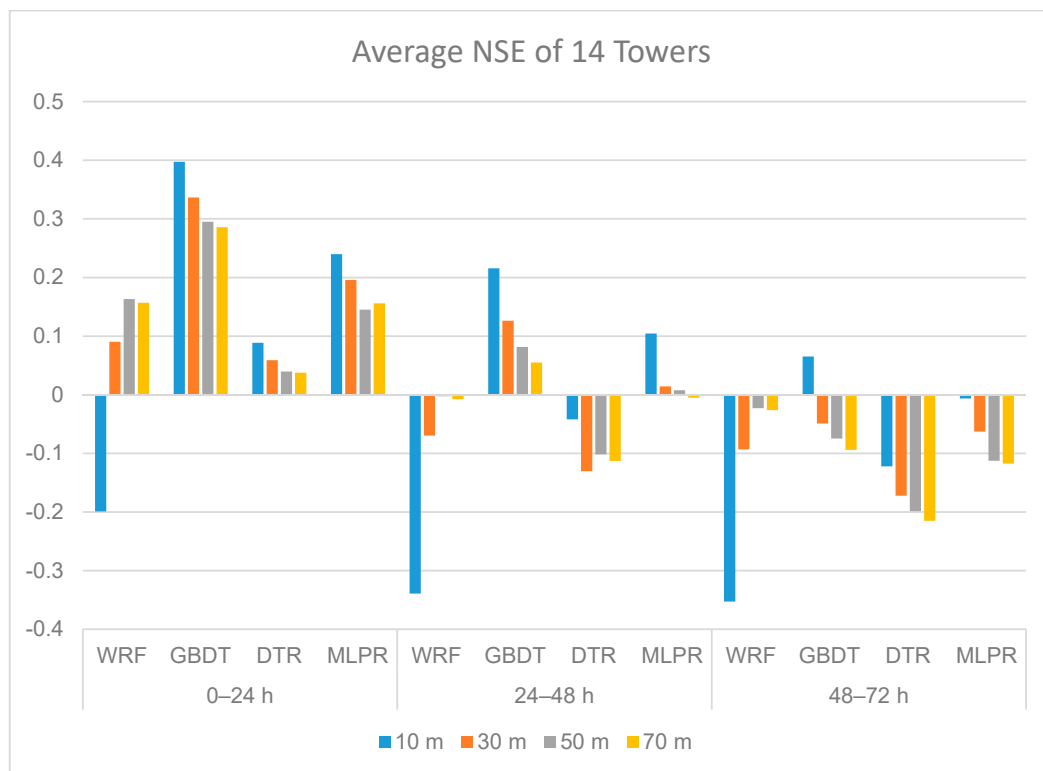
We calculate the average RMSE, IA, NSE, and R value of 14 towers and obtain Figures 7–10. These figures contain the results of WRF, GBDT, DTR, and MLPR in different height (10, 30, 50, and 70 m) and different forecast time (0–24 h, 24–48 h, and 48–72 h). We also conduct the significance tests using the statistical variables of different towers, the significance test results are shown in Table 11.



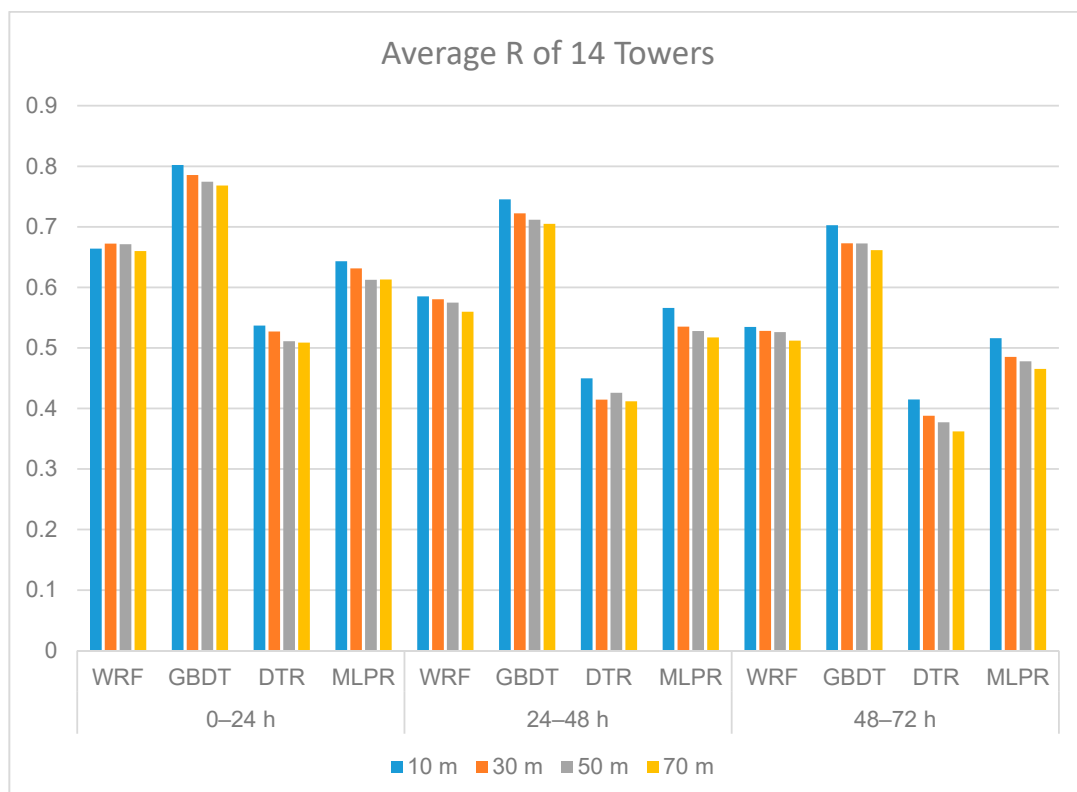
**Figure 7.** RMSEs of the WRF original output (WRF), GBDT output (GBDT), decision tree regression output (DTR), and multi-layer perceptron regression output (MLPR).



**Figure 8.** IAs of the WRF original output (WRF), GBDT output (GBDT), decision tree regression output (DTR), and multi-layer perceptron regression output (MLPR).



**Figure 9.** NSEs of the WRF original output (WRF), GBDT output (GBDT), decision tree regression output (DTR), and multi-layer perceptron regression output (MLPR).



**Figure 10.** Rs of the WRF original output (WRF), GBDT output (GBDT), decision tree regression output (DTR), and multi-layer perceptron regression output (MLPR).

**Table 11.** *p*-value of each significant tests.

Indices		10 m			30 m			50 m			70 m		
		WRF and GBDT	DTR and GBDT	MLPR and GBDT	WRF and GBDT	DTR and GBDT	MLPR and GBDT	WRF and GBDT	DTR and GBDT	MLPR and GBDT	WRF and GBDT	DTR and GBDT	MLPR and GBDT
RMSE	0–24 h	$1.8 \times 10^{-10}$	$3.1 \times 10^{-12}$	$5.32 \times 10^{-8}$	$4.09 \times 10^{-13}$	$1.76 \times 10^{-14}$	$2.31 \times 10^{-10}$	$2.17 \times 10^{-13}$	$3.26 \times 10^{-18}$	$1.51 \times 10^{-10}$	$4.15 \times 10^{-14}$	$1.47 \times 10^{-15}$	$1.23 \times 10^{-9}$
	24–48 h	$9.96 \times 10^{-11}$	$8.1 \times 10^{-13}$	$4.32 \times 10^{-9}$	$4.55 \times 10^{-15}$	$3.67 \times 10^{-14}$	$2.11 \times 10^{-11}$	$1.87 \times 10^{-17}$	$3.57 \times 10^{-16}$	$1.64 \times 10^{-12}$	$3.19 \times 10^{-18}$	$1.24 \times 10^{-13}$	$1.53 \times 10^{-11}$
	48–72 h	$8.39 \times 10^{-12}$	$1.01 \times 10^{-10}$	$8.8 \times 10^{-8}$	$2.4 \times 10^{-16}$	$3.45 \times 10^{-13}$	$7.56 \times 10^{-10}$	$7.03 \times 10^{-19}$	$1.56 \times 10^{-14}$	$2.39 \times 10^{-12}$	$4.79 \times 10^{-19}$	$1.56 \times 10^{-13}$	$1.24 \times 10^{-11}$
IA	0–24 h	$1.68 \times 10^{-8}$	$1.18 \times 10^{-9}$	$1.57 \times 10^{-6}$	$2.42 \times 10^{-7}$	$6.31 \times 10^{-11}$	$8.75 \times 10^{-6}$	$7.36 \times 10^{-6}$	$3.02 \times 10^{-10}$	$1.09 \times 10^{-5}$	$4.85 \times 10^{-6}$	$7.99 \times 10^{-11}$	$9.79 \times 10^{-6}$
	24–48 h	$4.54 \times 10^{-9}$	$4.89 \times 10^{-9}$	$3 \times 10^{-7}$	$3.98 \times 10^{-9}$	$4.74 \times 10^{-12}$	$3.82 \times 10^{-7}$	$1.48 \times 10^{-8}$	$3.21 \times 10^{-10}$	$4.22 \times 10^{-8}$	$5.23 \times 10^{-9}$	$1.05 \times 10^{-10}$	$4.03 \times 10^{-7}$
	48–72 h	$3.6 \times 10^{-9}$	$3.65 \times 10^{-14}$	$4.78 \times 10^{-9}$	$1.27 \times 10^{-9}$	$6.3 \times 10^{-11}$	$3.43 \times 10^{-8}$	$2.14 \times 10^{-10}$	$1.37 \times 10^{-15}$	$1.01 \times 10^{-8}$	$1.7 \times 10^{-10}$	$3 \times 10^{-15}$	$3.77 \times 10^{-9}$
R	0–24 h	$2.87 \times 10^{-6}$	$1.18 \times 10^{-10}$	$1.2 \times 10^{-7}$	$3.35 \times 10^{-6}$	$1.78 \times 10^{-12}$	$3.79 \times 10^{-7}$	$2.63 \times 10^{-5}$	$6.68 \times 10^{-12}$	$4.53 \times 10^{-7}$	$1.28 \times 10^{-5}$	$2.84 \times 10^{-12}$	$4.77 \times 10^{-7}$
	24–48 h	$2.14 \times 10^{-8}$	$1.28 \times 10^{-9}$	$1.43 \times 10^{-8}$	$1.17 \times 10^{-9}$	$1.1 \times 10^{-12}$	$1.92 \times 10^{-8}$	$1.38 \times 10^{-9}$	$1.06 \times 10^{-11}$	$1.67 \times 10^{-9}$	$1.86 \times 10^{-9}$	$1.81 \times 10^{-12}$	$7.65 \times 10^{-9}$
	48–72 h	$1.57 \times 10^{-10}$	$1.04 \times 10^{-15}$	$7.08 \times 10^{-11}$	$2.31 \times 10^{-10}$	$2.08 \times 10^{-12}$	$1.05 \times 10^{-9}$	$3.28 \times 10^{-11}$	$5.16 \times 10^{-17}$	$2.25 \times 10^{-10}$	$2.42 \times 10^{-10}$	$2.08 \times 10^{-16}$	$7.3 \times 10^{-11}$
NSE	0–24 h	$3.76 \times 10^{-7}$	$1.44 \times 10^{-6}$	0.003782	0.000748	$9.14 \times 10^{-6}$	0.01537	0.0939	0.000184	0.021304	0.086145	0.000173	0.036681
	24–48 h	$1.5 \times 10^{-6}$	$4.24 \times 10^{-5}$	0.024931	0.00354	$5.56 \times 10^{-5}$	0.067188	0.221716	0.001957	0.207006	0.366434	0.007758	0.347349
	48–72 h	$4.62 \times 10^{-5}$	$5.15 \times 10^{-6}$	0.063338	0.410048	0.016471	0.757702	0.311537	0.00555	0.435889	0.198474	0.014306	0.65035

‘WRF and GBDT’ means the significant test between WRF and GBDT results. ‘DTR and GBDT’ means the significant test between DTR and GBDT results. ‘MLPR and GBDT’ means the significant test between MLPR and GBDT results.

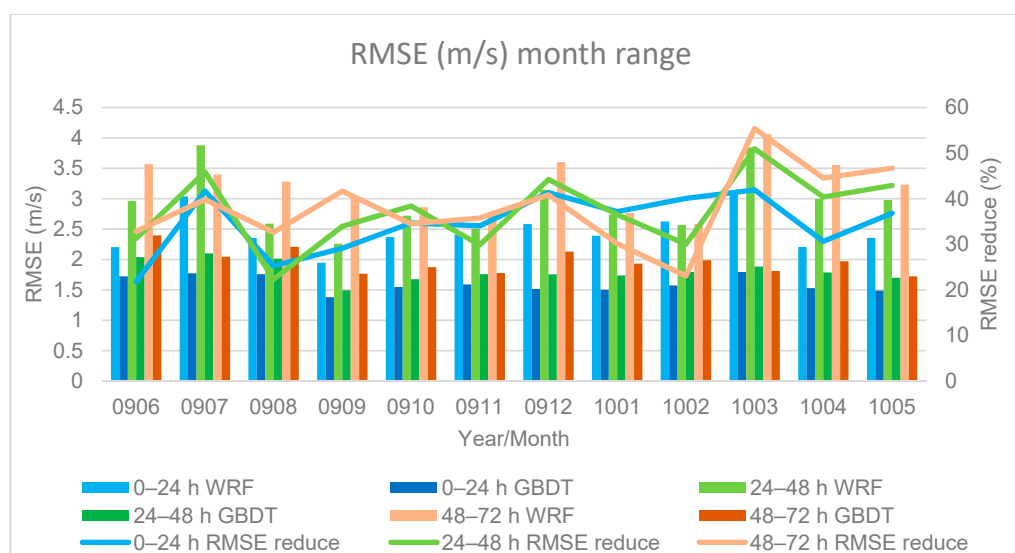
Figure 9 shows the average NSE value of 14 towers in each test. From Figure 9 we can find out that the NSE results of WRF model are close to zero, which means that the simulation results are close to the average level of observations, that is, the overall results are credible, but the simulation errors are large. Therefore, it is very necessary to post-process the results of the model.

From Figures 7–10 we can find that the RMSE of WRF output is about 2.7–3.5 m/s, the IA is about 0.61–0.75, the NSE is about  $-0.35$ – $0.15$  and the R is about 0.51–0.67. The increasing RMSE, as well as the decreasing IA, R and NSE between 0–24 h, 24–48 h, and 48–72 h indicates that the error of WRF forecast increases with the forecast time.

From Table 11 we can find that all the significance test results of RMSE, IA and R passed the confidence level of 99%, which means the improvement of GBDT model in RMSE, IA, and R is significant. However, some significance test results of NSE did not pass the 95% confidence level, especially in 24–48 h and 48–72 h, meaning that the NSE improvement in some cases cannot be trusted.

By comparing the results of WRF and post-processing models, it can be found that each post-processing method can reduce RMSE within a certain range. The RMSE of GBDT results is smaller than MLPR and DTR, which shows that GBDT can reduce more RMSE. The degree of reduction of RMSE by GBDT model is between 1–1.5 m/s, compared with 0.2–1 m/s of DTR and 0.7–1.2 m/s of MLPR. For IA, R, and NSE results, GBDT achieved the highest IA, R, and NSE among the four tests. Compared with WRF, GBDT has greatly improved these three metrics (IA between 0.10–0.20, R between 0.10–0.18, and NSE between  $-0.06$ – $0.6$ ). The reduction of RMSE as well as the improvement of IA and R indicate that GBDT can fit the near-surface wind speed with a smaller error than other two post-processing models. Thus, GBDT can be used to perform post-processing of WRF model to forecast the near-surface wind speed.

In order to further study the error changes of WRF and GBDT in different months, we calculated the average RMSE of wind speed forecast for each month. Figure 11 is the RMSE results in different months, including the monthly changes in the RMSE of WRF and GBDT and the percentage reduction in RMSE of GBDT relative to the WRF results.



**Figure 11.** RMSE results in different month, contains the WRF original output (WRF), GBDT output (GBDT), and the RMSE reduction (%) of GBDT.

From Figure 11, we can find that the RMSE of the WRF forecast results varies greatly between different months, and the RMSE in March, April, June, July, and December is large. However, at the same time, the RMSE of the GBDT results change less in different months. In the months when the RMSE of the WRF result is large, GBDT can reduce more RMSE, so that the final RMSE is roughly the same in each month.

In order to compare the post-processing effects of the GBDT model at different times, we calculate RMSE and IA for different hours in different months. Figure 12 is the RMSE in different month and hour, Figure 13 is the IA in different month and hour. From Figures 12 and 13 we can find that in July, September, October, and December, the forecast results at 12–24 are worse than those at 0–12. Compared with WRF, the results of GBDT did not have larger errors during the above-mentioned poor forecast time. It can be seen from Figures 12c and 13c that when the forecast results have larger error, GBDT reduce more RMSE and improve more IA than other time.

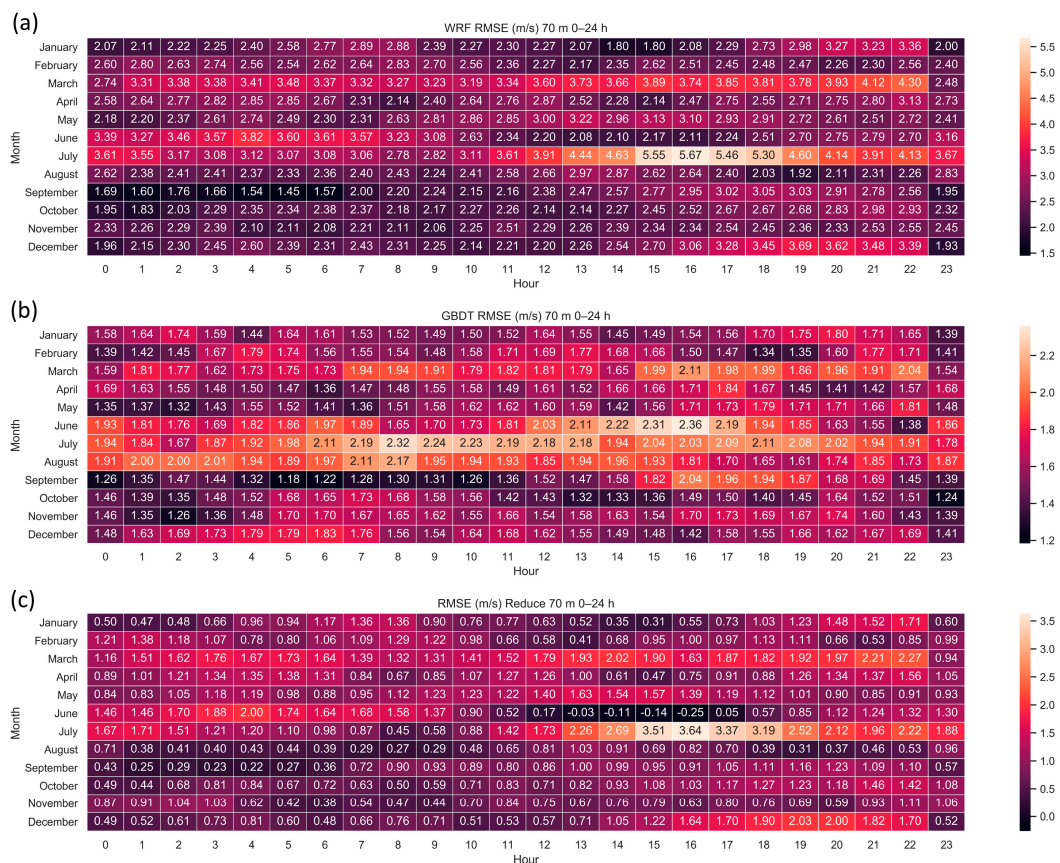
### 3.3. Weibull Distributions

In general cases, the distribution of near surface wind speed can be fitted using Weibull distribution [62]. The density function of Weibull distribution is

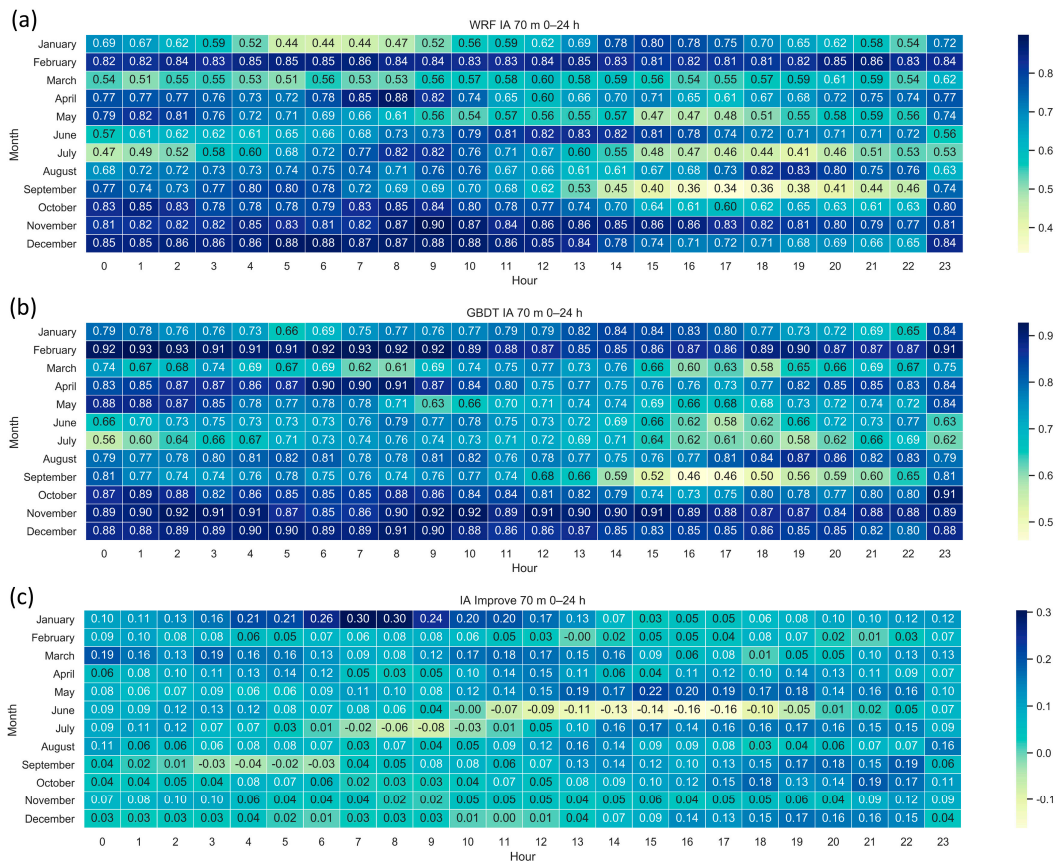
$$f(x; \lambda, k) = \frac{k}{\lambda} \left( \frac{x}{\lambda} \right)^{k-1} e^{-\left( \frac{x}{\lambda} \right)^k} \quad (23)$$

Here  $x$  is the wind speed,  $k > 0$  is the shape parameter, and  $\lambda > 0$  is the scale parameter of the Weibull distribution.

The two parameters of Weibull distribution can be used to determine whether the distribution of NWP results or the post-processing results is similar to the observations. We used the 10 m wind speed results of test data to fit the Weibull distribution of WRF, GBDT, DTR, and MLPR, the Weibull distributions are shown in Figure 8, and the parameters of these distributions are listed in Table 12.



**Figure 12.** RMSE results of 14 towers in different month and different hour, (a) the WRF results, (b) the GBDT results, and (c) the reduced value of GBDT relative to WRF results.

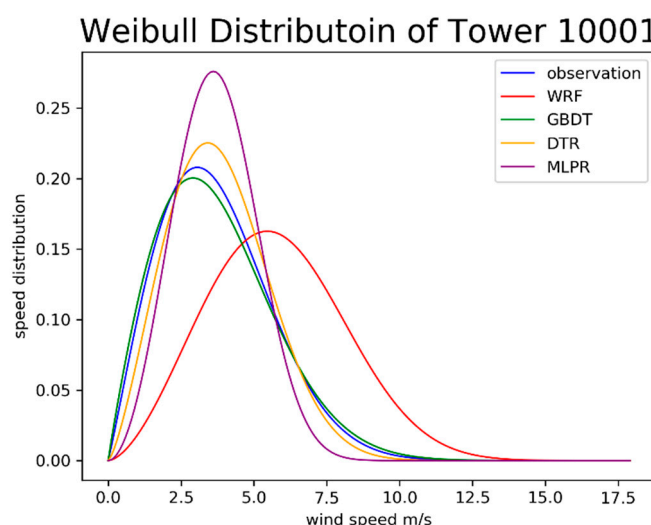


**Figure 13.** IA results of 14 towers in different month and different hour, (a) the WRF results, (b) the GBDT results, and (c) the improved value of GBDT relative to WRF result.

**Table 12.** Shape and scale parameters of 10 m wind speed Weibull distributions.

Tower	Observation		WRF		GBDT		DTR		MLPR	
	K (Shape)	Lambda (Scale)	K (Shape)	Lambda (Scale)	K (Shape)	Lambda (Scale)	K (Shape)	Lambda (Scale)	K (Shape)	Lambda (Scale)
10001	2.07	4.21	2.66	6.52	1.95	4.21	2.37	4.30	2.92	4.16
10002	2.17	4.61	2.55	6.55	1.88	4.48	2.38	4.55	2.73	4.46
10003	2.27	4.44	2.43	6.58	2.00	4.23	2.14	4.17	3.02	4.33
10004	2.14	5.22	2.53	6.77	2.01	5.05	2.10	4.85	2.63	5.01
10005	2.26	5.77	2.60	7.56	2.24	5.75	2.71	5.94	2.95	5.66
10006	2.02	4.44	2.54	7.52	1.93	4.49	2.30	4.82	2.65	4.42
10007	2.05	4.47	2.54	7.55	1.94	4.48	2.33	4.79	2.66	4.43
10008	2.20	5.11	2.54	7.58	2.22	5.04	2.62	5.38	2.95	5.05
10009	2.05	5.15	2.51	7.27	2.03	5.15	2.38	5.34	2.75	5.16
10010	1.76	5.37	2.50	6.44	1.85	5.34	1.94	5.37	2.19	5.42
10011	2.05	4.99	2.57	7.38	2.08	4.95	2.19	5.01	2.63	4.89
10012	1.97	4.93	2.50	6.77	1.91	5.05	2.04	4.92	2.49	4.94
10013	2.03	4.42	2.47	6.96	1.94	4.37	2.37	4.46	2.51	4.35
10014	2.32	4.56	2.57	7.69	2.12	4.47	2.47	4.50	2.92	4.45

From Figure 14 and Table 12 we can find that, relative to the original WRF results, the Weibull distribution curves of post-processing models are closer to the observations. Among the post-processing models, the curves of GBDT models are closest to the observations, both in shape and peak value. The results above show that the GBDT post-processing model can capture the distribution better than other post-processing models and original WRF output.



**Figure 14.** 10 m wind speed Weibull distribution of observation, WRF, GBDT, DTR, and MLPR on test data of Tower 10001.

### 3.4. GBDT Feature Importance Results

After training the GBDT models, we calculated the feature importance of each GBDT model. We first calculated the feature importance distribution of each GBDT model, and then calculated the average value of the feature importance distribution over the 14 towers. As the results in Figure 7 showed that the forecast results for 0–24 h, 24–48 h, and 48–72 h were not significantly different, we calculated the average of the results of the above three forecast times at different heights (10 m, 30 m, 50 m, and 70 m).

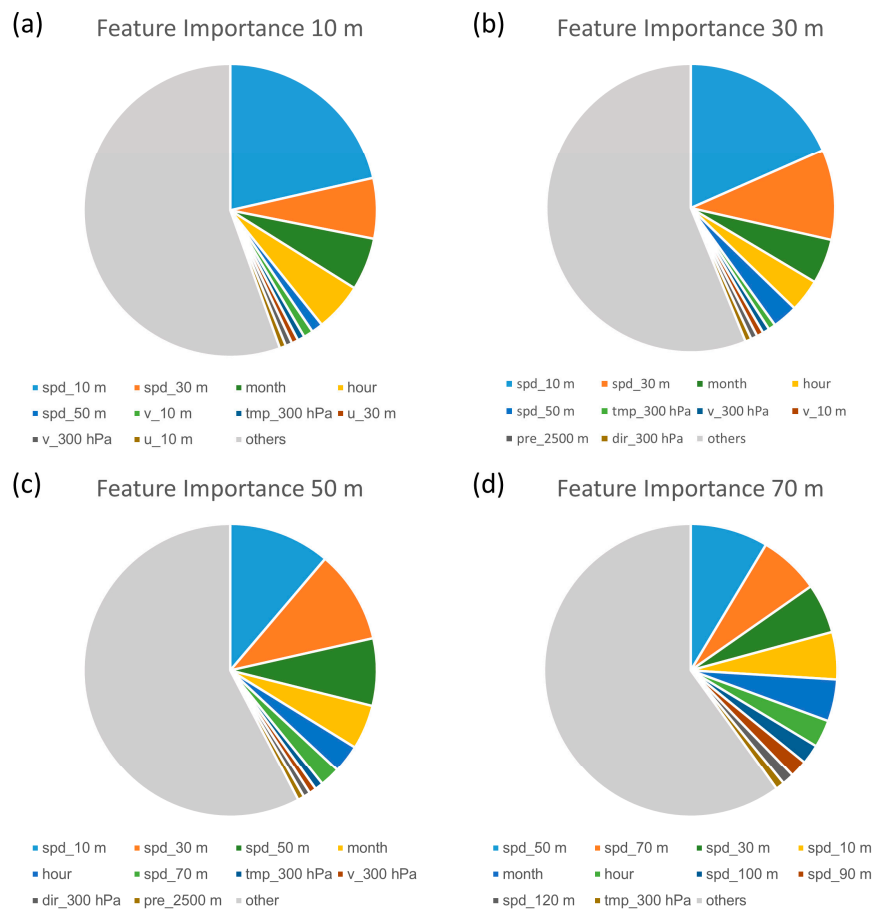
Figure 15 shows the feature importance results. It can be seen from Figure 15 that the 10 m wind speed output of the WRF model had the greatest effect on the GBDT post-processing results at 10 m, 30 m, and 50 m; the 30 m WRF output also had a large contribution. The 50 m and 70 m wind speed outputs of the WRF model could strongly affect the 50 m and 70 m GBDT results, and the 50 m wind speed output of WRF model was the largest contributor to the 70 m GBDT results. This means that the most important components of the near-surface wind speed post-processing model are the near-surface wind speeds in the WRF model.

At the same time, we can also see that the two features ‘month’ and ‘hour’ were also very important. In Figure 15, ‘month’ and ‘hour’ are in the top six features of importance. This means that changes in near-surface wind speeds were related to the forecast month and the forecast hour. We input ‘month’ and ‘hour’ into GBDT as two categorical features, such that these two features could contribute to the GBDT result.

Although the near-surface wind speed features importantly contributed to the GBDT results, the ‘other’ component still comprised about a half of the importance in Figure 15. This means the rest of the features also contributed to the final post-processing results. Thus, even if the effect of a single feature is limited, a large number of features can still have a strong effect on the result.

### 3.5. Feature Importance Sensitivity Tests

In order to further investigate the effect of different features on the results, we set up sensitivity tests on the input features. For sensitivity tests, we set up three sets of tests for different input features. We kept the GBDT model parameters of each sensitivity test the same. Table 13 shows the input features in different sensitivity tests. Test 1 uses all the features as input, Test 2 uses ‘other’ features as input and Test 3 uses the near surface wind speed, ‘hour’, and ‘month’ features as input.



**Figure 15.** Feature importance at 10 m (a), 30 m (b), 50 m (c), and 70 m (d). ‘spd\_10 m’ represents the 10 m wind speed output of the WRF model; ‘tmp’ represents the temperature features; ‘dir’ represents the direction features; and ‘u’ and ‘v’ represent the U and V components of wind speed, respectively, rotated to earth coordinates.

**Table 13.** Feature importance sensitivity test settings.

Test	Features
Test 1	All features
Test 2	‘Other’ features
Test 3	10 m speed, 30 m speed, 50 m speed, 70 m speed, hour, month

Figures 16 and 17 are the results of each tests, Figure 16 is the average RMSE and Figure 17 is average IA. We also did the significance test between Tests 1–3. Table 14 shows the significance test results. The  $p$ -values of each significance tests are less than 0.01, means that all the results passed the confidence level of 99%. From Figures 16 and 17 we can find that Test 1 has the smallest error and the highest IA, which means that we can obtain the best post-processing wind speed when we input all the features into GBDT model. Compared to Test 2, Test 3 has less RMSE and higher IA, the near-surface wind speed, ‘hour’ and ‘month’ features has a greater impact on the post-processing results than the ‘other’ feature. The significant improvement between Tests 1 and 3 indicate that it is necessary to add ‘other’ features to the input of GBDT post-processing model.

### 3.6. GBDT Feature Split Value Distributions

In the process of feature split for numerical features, each split has a split value. The distribution of feature split values depends on the distribution of feature values and feature’s contribution distribution.

If the feature split values are highly distributed over an interval, it means that: (1) the feature value has wide distribution in that interval; and (2) a change of feature value in that interval will have a great effect on the GBDT result.

In Section 3.4, we found that WRF model's wind speed at 10 m had the most importance in the 10 m wind speed GBDT output. We plot the distributions of 10 m wind speed observation and the distributions of 10 m WRF wind speed feature split values in 10. For the wind speed observation distributions, we used the Weibull distribution function to fit the 10 m wind speed observations.

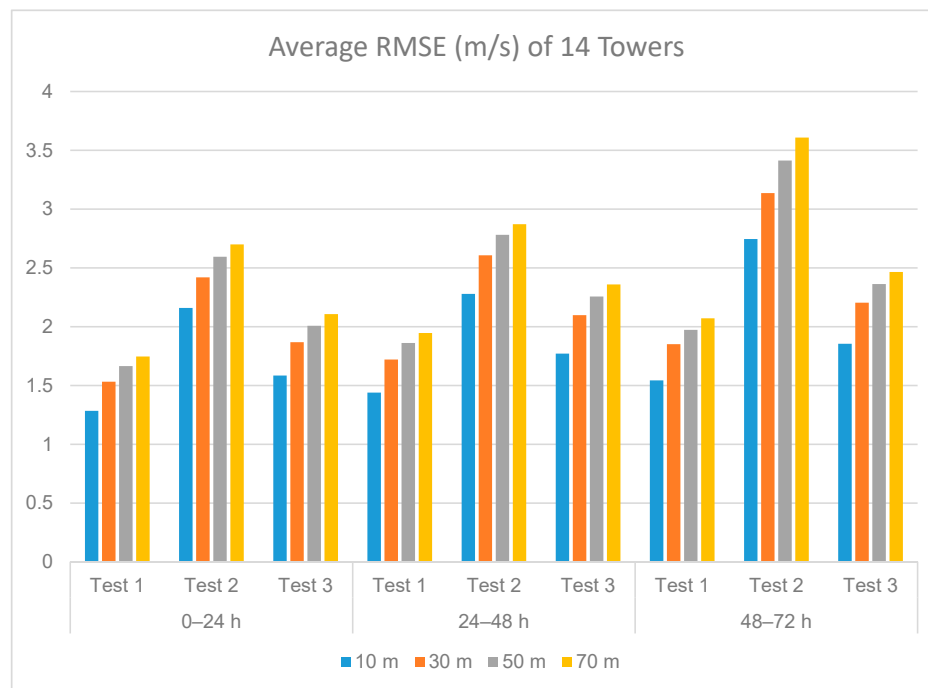


Figure 16. Wind speed RMSE of sensitivity test results.

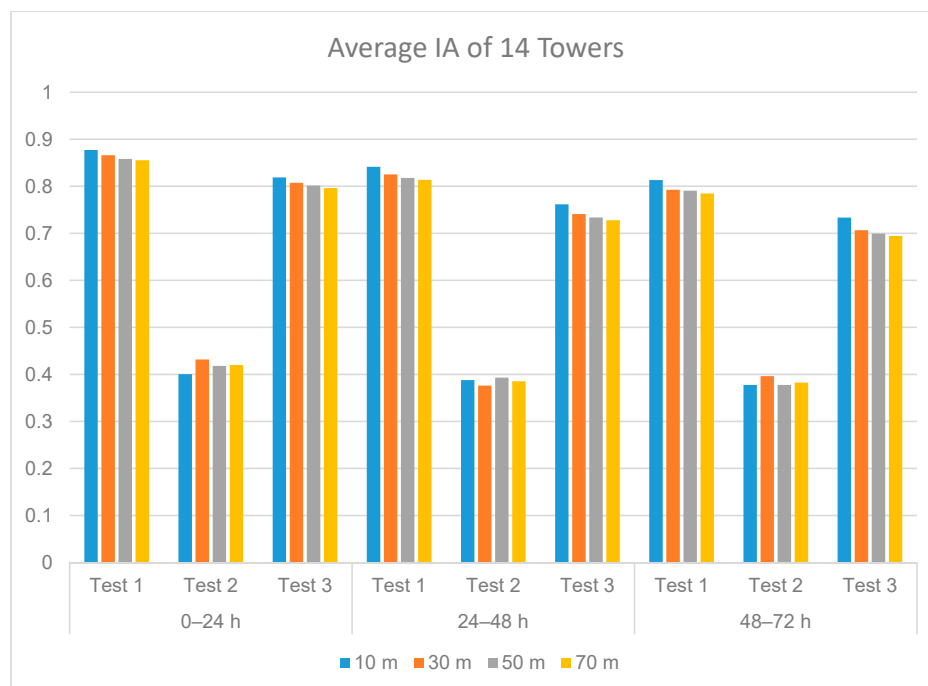


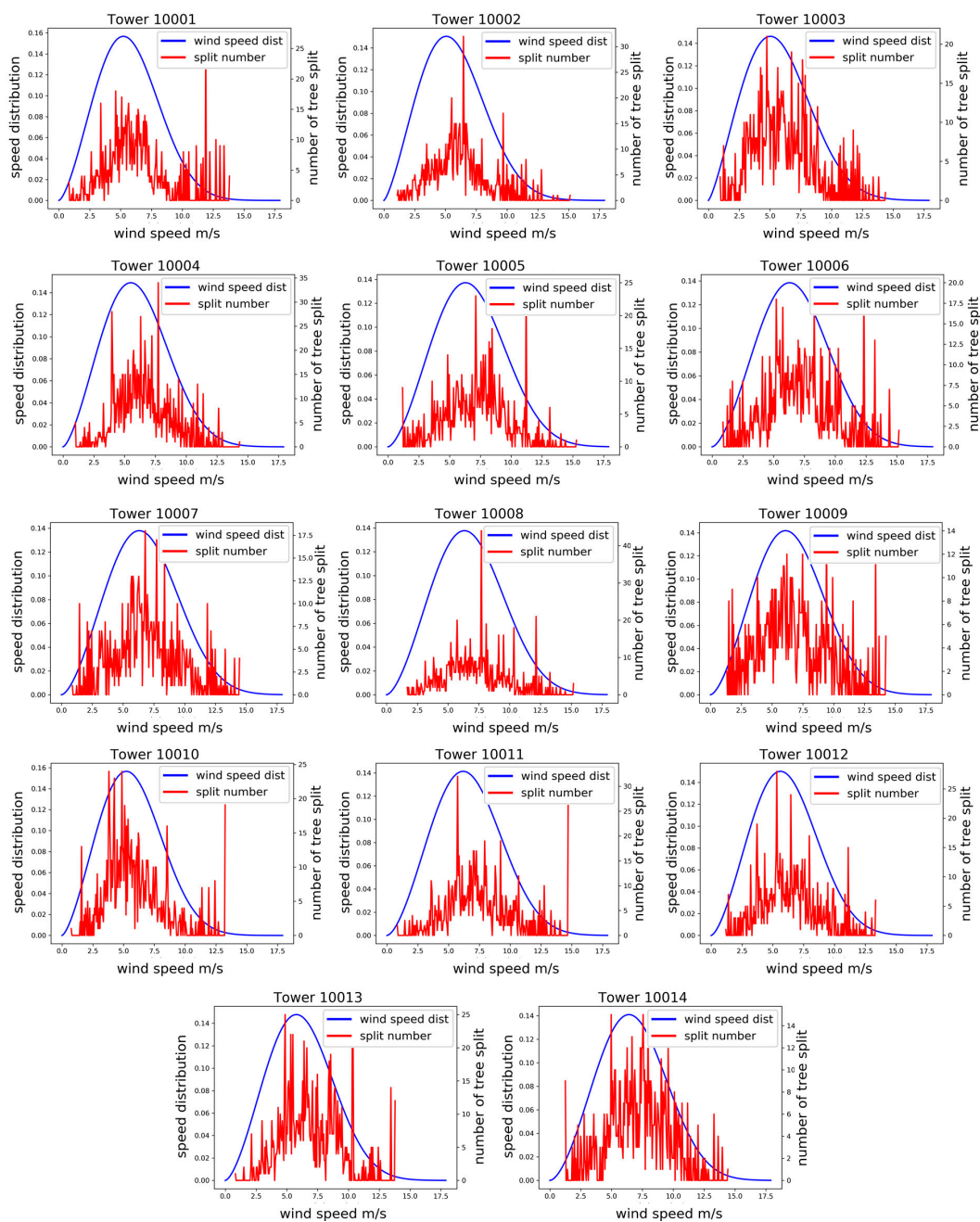
Figure 17. Wind speed IA of sensitivity test results.

**Table 14.** *p*-value of each significance test.

Indices		10 m		30 m		50 m		70 m	
		Tests 1 and 2	Tests 1 and 3	Tests 1 and 2	Tests 1 and 3	Tests 1 and 2	Tests 1 and 3	Tests 1 and 2	Tests 1 and 3
RMSE	0–24 h	$4.35 \times 10^{-7}$	$3.81 \times 10^{-5}$	$2.48 \times 10^{-8}$	$3.43 \times 10^{-8}$	$1.62 \times 10^{-10}$	$1.58 \times 10^{-8}$	$3.49 \times 10^{-8}$	$1.13 \times 10^{-7}$
	24–48 h	$6.23 \times 10^{-6}$	$5.28 \times 10^{-5}$	$3.43 \times 10^{-5}$	$2.34 \times 10^{-6}$	$1.73 \times 10^{-6}$	$2.8 \times 10^{-7}$	$4.01 \times 10^{-6}$	$3.74 \times 10^{-6}$
	48–72 h	$9.36 \times 10^{-6}$	0.000277	$5.95 \times 10^{-5}$	$1.58 \times 10^{-5}$	$1.84 \times 10^{-6}$	$7.75 \times 10^{-7}$	$6.45 \times 10^{-5}$	$5.63 \times 10^{-6}$
IA	0–24 h	$1.92 \times 10^{-13}$	$6.21 \times 10^{-7}$	$3.55 \times 10^{-13}$	$2.43 \times 10^{-6}$	$4.01 \times 10^{-16}$	$4.17 \times 10^{-5}$	$4.78 \times 10^{-15}$	$2.49 \times 10^{-5}$
	24–48 h	$1.44 \times 10^{-12}$	$6.21 \times 10^{-9}$	$1.51 \times 10^{-12}$	$1.22 \times 10^{-8}$	$1.2 \times 10^{-13}$	$5.85 \times 10^{-9}$	$3.82 \times 10^{-17}$	$8.59 \times 10^{-8}$
	48–72 h	$8.02 \times 10^{-12}$	$8.47 \times 10^{-11}$	$4.87 \times 10^{-10}$	$3.24 \times 10^{-10}$	$9.33 \times 10^{-10}$	$3.25 \times 10^{-11}$	$6.24 \times 10^{-9}$	$8.31 \times 10^{-10}$

‘Test 1 and 2’ means *t*-test between Test 1 and Test 2. ‘Test 1 and 3’ means *t*-test between Test 1 and Test 3.

Figure 18 shows the distributions of all the towers (towers 10001–10014). From Figure 18, we find that the distribution of the wind speed observation at 10 m is roughly similar to the distribution of the feature split value; this is because the distribution of WRF output wind speed was similar to the wind speed observations. However, in the high speed regions (wind speed  $> 8$  m/s), the Weibull distribution decreases significantly but the feature split value distribution remains high. These distributions indicate that a change of feature value in the regions with high wind speeds still has an impact on the GBDT results. It also indicated that, if the WRF model is inaccurately simulated in high wind speed regions, large errors will occur in the GBDT results. This indicates that, in the area we studied, improvement of the WRF model's simulation performance for high wind speeds can improve the overall wind speed prediction performance.



**Figure 18.** Distribution of 10 m speed feature split values and Weibull distribution of 10 m wind speed observation of all 14 towers.

#### 4. Conclusions

In this work, based on WRF model, we conducted a one-year wind speed forecast of the coastal area of Jiangsu, China. According to the power grid's requirements for wind power forecasting in wind farms, we obtained 0–24 h, 24–48 h, and 48–72 h wind speed forecasts for wind power forecasting. Based on the NWP forecast results, we extracted multiple variables from WRF output, at different height and pressure levels, for each moment. We built a GBDT regression model to correct the output near-surface wind speed of the WRF model and compared the performance with other two post-processing methods. Finally, we analyzed feature importance in the GBDT model and found which features had a greater impact on the results of the GBDT model. Our main conclusions are as follows:

The Weibull distributions of 10 m wind speed results shows that after the post-processing results, the wind distributions were closer to the observations and the GBDT model had the best performance to fit the near surface wind.

The root mean square error (RMSE) of the wind speed forecast for the wind farm was approximately 2.7–3.5 m/s for different wind towers and at different levels. Wind speed errors will cause greater errors in wind power forecasting, which will affect the operations of wind farms and power grids. After GBDT model correction, the RMSE of the wind speed was reduced, between a range of 1–1.5 m/s, on the test data sets. Also, the IA and R results shows that GBDT can improve these indices, the IA can be improved by 0.10–0.20, and R can be improved by 0.10–0.18. These improvements indicate that the GBDT model, using a large number of features as input, can reduce the wind speed forecast error of the wind farm.

In different months and at different times, the error of the WRF results varies greatly. In the month and time with large error, the GBDT model can reduce a larger error, so that the error distribution of the final wind speed forecast results in different months and different times does not have significant difference.

By analyzing the feature importance of each GBDT model, we found that the distribution of feature importance is different for the correction models at different heights. From the 10, 30, 50, and 70 m wind speed correction results, we found that the near-surface wind speed distribution has a strong impact on the correction results. The 10 m model output wind speed can greatly affect the correction results at 10, 30, and 50 m. The 30 m model output wind speed can affect 10, 30, 50, and 70 m correction results. The 50 m model output wind speed can affect the 50 m and 70 m correction results. The 70 m model output wind speed can affect the 70 m correction results.

At the same time, we found that two categorical features, 'month' and 'hour', also had a great impact on the result of the correction. This shows that WRF simulation errors have some characteristics in different months and at different hours of the day.

From the feature split distribution of the 10 m WRF output wind speed, it could be seen that the distribution of the feature split and the Weibull distribution of wind speed do not completely coincide, but that the distribution of the split value in the high wind speed region is greater than the Weibull distribution of wind speed. This result shows that the decision tree has frequent branching when the wind speed value is high. Thus, high wind speeds simulated by the WRF model have a great impact on the GBDT results and can easily cause errors.

There have been many studies which used a numerical weather model to make weather forecasts and performed post-processing algorithms to correct the forecast. Such post-processing algorithms only used a few features as their input for two main reasons: (1) in some machine learning algorithms, such as ANN and SVM, too many feature inputs will negatively affect the algorithm's performance, by some less-effective features. (2) In some algorithms, if the input number of features is too large, the amount of calculation increases squarely or exponentially, which will make the model impossible to train in a short time. The GBDT algorithm we used did not have the above problems: it can input a huge number of features, pick out the important features, and ignore the less important features. Our results showed that even the most important feature (near-surface wind speed), only takes up a

small portion of the entire importance. The GBDT results show that smaller errors can be achieved with more features.

For the categorical features, such as forecast month and forecast hour, it is very difficult to input them into post-processing algorithms such as neural networks and other statistical methods. However, with a decision tree, categorical features can be easily processed. Therefore, the GBDT algorithm also has an advantage over other algorithms in being able to deal with categorical features. Our results showed that the feature importance of forecast month and forecast hour was high and, so, categorical features like forecast month and forecast hour have improved our model's performance.

Another advantage of GBDT is that it can calculate feature importance by analyzing the gain of information while the decision tree is split. Therefore, we can find which features are important and which are less important.

In summary, a more effective method for wind speed forecasting of wind farms is to use a mesoscale meteorological model to forecast wind speed, followed by use of the GBDT algorithm to correct the model simulation results.

**Author Contributions:** Conceptualization, Y.L.; Methodology, W.X.; Software, W.X.; Validation, W.X.; Formal analysis, W.X.; Investigation, W.X.; Resources, Y.L.; Data curation, W.X.; Writing—original draft preparation, W.X.; Writing—review and editing, Y.L. and L.N.; Visualization, W.X.; Supervision, Y.L. and L.N.; Project administration, Y.L.; Funding acquisition, Y.L. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by National Key Research and Development Program of China (2018YFB1502803) and Scientific Research Program of Tsinghua University “Research on Wind Farm Weather Forecasting Technology for Power Grid”.

**Acknowledgments:** The FNL data was provided by CISL Research Data Archive (RDA) web site (<https://rda.ucar.edu/datasets/ds083.2/>). The wind observation data of 14 wind towers in Jiangsu was provided by National Climate Center (China Meteorological Administration). These data played a key role in our research work, we are grateful to the providers of the above data.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

**Table A1.** The 0–24 h RMSE (m/s) of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	3.00	1.30	2.08	1.78	3.10	1.46	2.30	1.99	3.33	1.69	2.67	2.37	3.36	1.62	2.58	2.23
2	2.64	1.29	2.20	1.92	2.73	1.53	2.41	2.19	2.83	1.65	2.65	2.19	3.08	1.71	2.76	2.41
3	2.82	1.20	2.01	1.65	2.70	1.50	2.25	2.08	2.70	1.66	2.76	2.16	2.89	1.78	2.73	2.37
4	2.28	1.35	2.17	1.89	2.42	1.65	2.54	2.15	2.57	1.79	2.70	2.34	2.76	1.91	2.88	2.59
5	2.69	1.52	2.31	2.00	2.63	1.72	2.67	2.27	2.65	1.83	2.80	2.53	2.75	1.92	2.94	2.59
6	3.41	1.17	1.88	1.68	2.82	1.54	2.41	2.09	2.76	1.67	2.72	2.25	2.94	1.76	2.68	2.41
7	3.45	1.18	1.97	1.69	2.80	1.59	2.63	2.16	2.72	1.75	2.70	2.28	2.83	1.83	2.85	2.38
8	3.04	1.35	2.14	1.87	2.87	1.68	2.60	2.27	2.76	1.83	2.73	2.48	3.03	1.91	2.93	2.53
9	2.93	1.41	2.37	2.13	2.72	1.57	2.57	2.07	2.72	1.72	2.62	2.32	2.81	1.80	3.04	2.39
10	2.30	1.51	2.37	2.02	2.47	1.62	2.51	2.10	2.64	1.71	2.74	2.23	2.98	1.81	2.79	2.39
11	2.93	1.29	2.16	1.71	2.82	1.43	2.39	1.92	2.82	1.51	2.50	2.03	2.91	1.58	2.66	2.03
12	2.39	1.25	1.97	1.65	2.38	1.46	2.37	1.99	2.36	1.49	2.53	2.14	2.43	1.57	2.65	2.18
13	2.87	1.06	1.75	1.42	2.76	1.35	2.20	1.78	2.65	1.55	2.44	1.95	2.76	1.63	2.53	2.10
14	3.49	1.11	1.89	1.58	3.22	1.35	2.15	1.80	2.85	1.47	2.44	2.09	2.81	1.63	2.48	2.06
Ave	2.87	1.28	2.09	1.78	2.75	1.53	2.43	2.06	2.74	1.66	2.64	2.24	2.88	1.75	2.75	2.33

**Table A2.** The 24–48 h RMSE (m/s) of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	3.13	1.34	2.24	2.00	3.26	1.50	2.36	2.18	3.50	1.70	2.79	2.40	3.54	1.66	2.80	2.45
2	2.78	1.40	2.34	2.01	2.92	1.61	2.80	2.36	3.05	1.77	2.93	2.43	3.27	1.85	2.97	2.63
3	3.03	1.26	2.15	1.78	2.94	1.60	2.56	2.30	2.98	1.75	2.87	2.68	3.19	1.84	3.11	2.72
4	2.61	1.48	2.40	2.06	2.82	1.87	3.07	2.42	3.00	2.02	3.07	2.68	3.23	2.14	3.20	2.84
5	3.01	1.62	2.50	2.17	3.03	1.91	2.88	2.53	3.07	2.03	3.22	2.70	3.20	2.13	3.39	2.76
6	3.82	1.39	2.19	1.90	3.37	1.80	2.74	2.40	3.33	1.94	2.96	2.60	3.54	2.05	3.18	2.74
7	3.83	1.39	2.23	1.90	3.32	1.85	2.88	2.45	3.28	2.01	3.02	2.72	3.42	2.06	3.19	2.77
8	3.47	1.53	2.40	2.06	3.37	1.89	2.93	2.55	3.31	2.03	2.99	2.66	3.58	2.15	3.54	2.87
9	3.33	1.64	2.46	2.24	3.19	1.83	2.76	2.39	3.24	1.94	3.05	2.55	3.35	2.03	3.00	2.81

Table A2. Cont.

	10 m				30 m				50 m				70 m			
Tower	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
10	2.57	1.69	2.79	2.26	2.85	1.80	2.80	2.46	3.09	1.93	2.83	2.60	3.47	2.07	3.18	2.85
11	3.33	1.46	2.29	2.02	3.28	1.61	2.72	2.19	3.30	1.71	2.71	2.35	3.40	1.82	2.92	2.45
12	2.77	1.46	2.24	1.96	2.84	1.73	2.93	2.42	2.86	1.80	2.90	2.45	2.97	1.85	2.86	2.47
13	3.15	1.23	1.95	1.62	3.10	1.51	2.45	2.13	3.03	1.71	2.64	2.34	3.18	1.76	2.72	2.41
14	3.69	1.26	2.37	1.76	3.54	1.59	2.62	2.06	3.26	1.73	2.69	2.32	3.26	1.85	2.97	2.47
Ave	3.18	1.44	2.33	1.98	3.13	1.72	2.75	2.35	3.16	1.86	2.91	2.53	3.33	1.95	3.07	2.66

Table A3. The 48–72 h RMSE (m/s) of WRF model original output and post-processing model results.

	10 m				30 m				50 m				70 m			
Tower	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	3.35	1.44	2.34	2.08	3.51	1.60	2.62	2.33	3.70	1.75	2.74	2.48	3.81	1.77	2.76	2.49
2	3.00	1.46	2.37	2.19	3.19	1.70	2.59	2.43	3.32	1.81	2.91	2.52	3.57	1.90	3.04	2.74
3	3.31	1.38	2.10	1.87	3.33	1.74	2.76	2.42	3.37	1.87	2.91	2.57	3.59	2.01	2.96	2.79
4	2.80	1.61	2.44	2.18	3.07	1.98	2.97	2.60	3.24	2.07	3.20	2.66	3.47	2.19	3.37	2.89
5	3.10	1.66	2.69	2.21	3.16	1.95	3.02	2.64	3.22	2.06	3.13	2.54	3.35	2.11	3.48	2.84
6	3.83	1.53	2.39	1.99	3.42	2.00	2.88	2.79	3.39	2.16	3.16	2.74	3.59	2.25	3.43	2.84
7	3.83	1.49	2.33	1.93	3.39	2.05	3.26	2.58	3.37	2.19	3.24	2.77	3.51	2.31	3.42	2.98
8	3.51	1.68	2.50	2.24	3.46	2.07	3.03	2.75	3.40	2.21	3.37	2.90	3.64	2.33	3.45	3.07
9	3.34	1.69	2.54	2.36	3.23	1.89	2.89	2.64	3.26	2.01	3.17	2.70	3.37	2.11	3.36	2.92
10	2.68	1.93	2.99	2.59	2.95	2.03	2.97	2.61	3.18	2.02	3.07	2.77	3.53	2.18	3.22	3.00
11	3.36	1.53	2.36	2.05	3.35	1.71	2.67	2.24	3.38	1.79	2.78	2.45	3.49	1.95	3.00	2.59
12	2.89	1.55	2.42	2.24	3.01	1.81	2.77	2.55	3.06	1.92	3.14	2.74	3.16	1.96	3.12	2.85
13	3.26	1.34	2.01	1.76	3.30	1.68	2.51	2.17	3.25	1.87	2.77	2.58	3.45	1.92	2.96	2.67
14	3.90	1.32	2.21	1.82	3.79	1.69	2.73	2.25	3.55	1.87	2.94	2.57	3.58	1.99	3.20	2.57
Ave	3.30	1.54	2.41	2.11	3.30	1.85	2.83	2.50	3.34	1.97	3.04	2.64	3.51	2.07	3.20	2.80

## Appendix B

**Table A4.** The 0–24 h IA of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	0.59	0.83	0.66	0.74	0.62	0.80	0.65	0.69	0.62	0.77	0.61	0.65	0.62	0.78	0.64	0.69
2	0.67	0.85	0.69	0.73	0.70	0.83	0.68	0.71	0.70	0.82	0.66	0.75	0.68	0.82	0.65	0.71
3	0.65	0.85	0.68	0.76	0.73	0.85	0.75	0.78	0.75	0.85	0.68	0.79	0.74	0.85	0.71	0.77
4	0.78	0.88	0.74	0.79	0.80	0.86	0.73	0.79	0.79	0.85	0.72	0.77	0.77	0.84	0.71	0.75
5	0.74	0.86	0.72	0.79	0.78	0.86	0.70	0.77	0.80	0.85	0.72	0.75	0.79	0.85	0.71	0.78
6	0.65	0.89	0.76	0.79	0.76	0.86	0.73	0.79	0.78	0.86	0.70	0.77	0.77	0.86	0.75	0.78
7	0.64	0.88	0.74	0.79	0.76	0.87	0.69	0.80	0.79	0.86	0.73	0.80	0.78	0.86	0.71	0.79
8	0.69	0.86	0.71	0.80	0.75	0.85	0.71	0.77	0.78	0.84	0.71	0.75	0.75	0.84	0.69	0.74
9	0.70	0.87	0.72	0.77	0.76	0.87	0.72	0.79	0.77	0.86	0.72	0.78	0.78	0.86	0.68	0.79
10	0.80	0.91	0.80	0.86	0.80	0.89	0.76	0.84	0.78	0.88	0.74	0.81	0.75	0.87	0.73	0.80
11	0.71	0.89	0.74	0.82	0.75	0.89	0.73	0.81	0.76	0.89	0.74	0.81	0.77	0.89	0.74	0.82
12	0.78	0.91	0.81	0.86	0.82	0.91	0.80	0.84	0.83	0.91	0.78	0.81	0.83	0.90	0.76	0.82
13	0.70	0.91	0.80	0.85	0.76	0.89	0.76	0.84	0.79	0.88	0.74	0.83	0.79	0.87	0.73	0.80
14	0.63	0.88	0.71	0.78	0.70	0.88	0.75	0.81	0.77	0.89	0.75	0.80	0.79	0.88	0.77	0.84
Ave	0.70	0.88	0.73	0.80	0.75	0.87	0.73	0.79	0.76	0.86	0.71	0.78	0.76	0.86	0.71	0.78

**Table A5.** The 24–48 h IA of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	0.56	0.81	0.64	0.67	0.59	0.78	0.61	0.62	0.58	0.76	0.56	0.62	0.59	0.75	0.56	0.62
2	0.65	0.82	0.63	0.71	0.67	0.80	0.60	0.67	0.67	0.79	0.59	0.70	0.66	0.78	0.60	0.66
3	0.62	0.83	0.63	0.72	0.69	0.83	0.66	0.70	0.71	0.83	0.66	0.71	0.70	0.83	0.63	0.70
4	0.73	0.85	0.70	0.77	0.74	0.81	0.60	0.74	0.72	0.79	0.62	0.71	0.70	0.79	0.63	0.70
5	0.70	0.84	0.69	0.75	0.73	0.82	0.67	0.76	0.74	0.82	0.63	0.74	0.74	0.82	0.66	0.75
6	0.60	0.84	0.68	0.74	0.68	0.81	0.66	0.72	0.70	0.81	0.67	0.74	0.68	0.81	0.66	0.72
7	0.60	0.84	0.69	0.76	0.69	0.82	0.67	0.74	0.71	0.82	0.68	0.72	0.70	0.82	0.66	0.74
8	0.63	0.82	0.67	0.73	0.68	0.81	0.64	0.74	0.71	0.80	0.68	0.72	0.68	0.78	0.58	0.71

Table A5. Cont.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
9	0.64	0.83	0.68	0.72	0.69	0.82	0.67	0.72	0.70	0.82	0.67	0.74	0.70	0.82	0.69	0.71
10	0.77	0.88	0.74	0.81	0.75	0.86	0.73	0.79	0.72	0.84	0.72	0.76	0.69	0.82	0.66	0.74
11	0.66	0.86	0.70	0.78	0.69	0.85	0.64	0.75	0.70	0.84	0.68	0.75	0.70	0.85	0.69	0.76
12	0.72	0.87	0.74	0.79	0.75	0.86	0.69	0.76	0.76	0.86	0.69	0.76	0.76	0.85	0.70	0.79
13	0.67	0.87	0.74	0.81	0.71	0.86	0.70	0.76	0.73	0.84	0.70	0.75	0.73	0.84	0.71	0.72
14	0.60	0.83	0.55	0.73	0.64	0.82	0.61	0.73	0.70	0.83	0.69	0.73	0.72	0.83	0.67	0.73
Ave	0.65	0.84	0.68	0.75	0.69	0.83	0.65	0.73	0.70	0.82	0.66	0.72	0.70	0.81	0.65	0.72

Table A6. The 48–72 h IA of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	0.56	0.79	0.62	0.66	0.58	0.77	0.58	0.64	0.59	0.76	0.59	0.60	0.59	0.73	0.59	0.60
2	0.65	0.81	0.63	0.68	0.66	0.79	0.62	0.64	0.66	0.78	0.61	0.65	0.65	0.78	0.59	0.65
3	0.59	0.80	0.66	0.70	0.65	0.80	0.63	0.69	0.67	0.81	0.65	0.71	0.66	0.80	0.67	0.68
4	0.71	0.82	0.68	0.74	0.71	0.80	0.63	0.72	0.70	0.80	0.61	0.71	0.69	0.79	0.62	0.69
5	0.69	0.83	0.64	0.76	0.72	0.81	0.66	0.71	0.73	0.81	0.68	0.76	0.72	0.81	0.62	0.72
6	0.58	0.80	0.63	0.72	0.66	0.76	0.62	0.64	0.67	0.76	0.63	0.68	0.66	0.77	0.58	0.69
7	0.58	0.81	0.65	0.73	0.66	0.77	0.58	0.70	0.68	0.78	0.62	0.70	0.68	0.77	0.61	0.67
8	0.61	0.78	0.64	0.67	0.65	0.76	0.60	0.65	0.68	0.76	0.58	0.68	0.66	0.74	0.59	0.64
9	0.62	0.81	0.64	0.68	0.67	0.79	0.64	0.70	0.68	0.79	0.62	0.68	0.68	0.79	0.61	0.68
10	0.73	0.83	0.65	0.73	0.72	0.81	0.67	0.74	0.69	0.80	0.65	0.69	0.66	0.79	0.64	0.68
11	0.63	0.83	0.70	0.74	0.65	0.81	0.67	0.74	0.67	0.81	0.66	0.72	0.67	0.81	0.65	0.72
12	0.69	0.84	0.70	0.73	0.71	0.84	0.72	0.73	0.71	0.82	0.63	0.70	0.72	0.82	0.66	0.72
13	0.63	0.84	0.70	0.76	0.67	0.81	0.68	0.73	0.69	0.79	0.66	0.67	0.68	0.80	0.63	0.67
14	0.56	0.80	0.62	0.71	0.60	0.78	0.59	0.70	0.65	0.79	0.62	0.69	0.67	0.79	0.61	0.72
Ave	0.63	0.81	0.66	0.72	0.66	0.79	0.64	0.69	0.68	0.79	0.63	0.69	0.67	0.78	0.62	0.68

## Appendix C

**Table A7.** The 0–24 h NSE of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	−0.67	0.11	−0.09	0.11	−0.46	−0.01	−0.07	−0.12	−0.51	−0.19	−0.14	−0.21	−0.39	−0.20	−0.07	−0.04
2	−0.20	0.28	0.01	0.05	−0.06	0.18	−0.01	−0.02	−0.05	0.13	−0.02	0.14	−0.13	0.16	−0.05	0.03
3	−0.22	0.24	−0.02	0.13	0.07	0.26	0.16	0.23	0.14	0.28	−0.02	0.21	0.10	0.29	0.07	0.18
4	0.18	0.43	0.06	0.22	0.26	0.33	0.07	0.22	0.24	0.30	0.03	0.13	0.22	0.24	0.04	0.09
5	0.04	0.31	0.04	0.18	0.23	0.30	−0.05	0.08	0.28	0.28	0.03	0.03	0.29	0.30	0.01	0.23
6	−0.49	0.46	0.15	0.16	0.14	0.28	0.04	0.22	0.23	0.28	−0.02	0.08	0.20	0.30	0.12	0.12
7	−0.51	0.43	0.16	0.21	0.14	0.34	−0.10	0.26	0.27	0.31	0.06	0.18	0.26	0.30	−0.08	0.15
8	−0.17	0.31	−0.02	0.29	0.11	0.25	−0.04	0.14	0.25	0.18	−0.08	0.03	0.16	0.15	−0.03	−0.09
9	−0.16	0.37	0.04	0.24	0.16	0.37	0.04	0.18	0.22	0.26	−0.02	0.16	0.23	0.29	−0.05	0.21
10	0.11	0.58	0.22	0.45	0.19	0.49	0.08	0.34	0.17	0.43	0.11	0.25	0.05	0.40	0.02	0.19
11	−0.16	0.47	0.06	0.27	0.09	0.46	0.08	0.24	0.17	0.45	0.15	0.28	0.19	0.45	0.14	0.25
12	0.13	0.56	0.37	0.45	0.31	0.57	0.32	0.35	0.38	0.56	0.24	0.21	0.40	0.55	0.18	0.29
13	−0.16	0.58	0.26	0.42	0.13	0.49	0.16	0.36	0.26	0.39	0.07	0.31	0.28	0.35	0.01	0.19
14	−0.50	0.43	0.01	0.17	−0.05	0.42	0.13	0.28	0.24	0.48	0.17	0.23	0.32	0.43	0.22	0.39
Ave	−0.20	0.40	0.09	0.24	0.09	0.34	0.06	0.20	0.16	0.29	0.04	0.15	0.16	0.29	0.04	0.16

**Table A8.** The 24–48 h NSE of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	−0.76	−0.04	−0.07	−0.13	−0.57	−0.22	−0.26	−0.39	−0.63	−0.36	−0.33	−0.40	−0.50	−0.44	−0.26	−0.27
2	−0.22	0.14	−0.13	0.01	−0.09	0.05	−0.16	−0.12	−0.10	−0.05	−0.20	−0.03	−0.15	−0.07	−0.18	−0.08
3	−0.31	0.10	−0.16	0.02	−0.02	0.19	−0.11	−0.07	0.04	0.15	−0.08	0.11	0.00	0.17	−0.16	0.04
4	0.02	0.22	0.02	0.16	0.08	0.01	−0.35	0.08	0.07	−0.04	−0.31	−0.01	0.03	−0.09	−0.19	−0.08
5	−0.11	0.24	−0.02	0.10	0.07	0.12	−0.11	0.16	0.12	0.14	−0.21	0.09	0.14	0.15	−0.00	0.16
6	−0.71	0.26	−0.01	0.07	−0.10	0.08	−0.14	−0.05	0.01	0.10	−0.06	0.10	−0.03	0.05	−0.06	0.00
7	−0.69	0.27	0.06	0.18	−0.07	0.14	0.02	0.07	0.05	0.12	−0.02	−0.02	0.05	0.11	−0.15	0.08
8	−0.38	0.12	−0.08	0.04	−0.09	0.08	−0.18	0.13	0.03	−0.02	−0.05	−0.07	−0.04	−0.15	−0.23	−0.02

Table A8. Cont.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
9	−0.34	0.16	−0.12	−0.02	−0.04	0.12	−0.05	−0.09	0.02	0.15	−0.04	0.07	0.03	0.12	−0.03	−0.05
10	0.03	0.41	0.15	0.25	0.07	0.32	0.08	0.26	0.02	0.18	0.03	0.16	−0.11	0.12	−0.20	0.11
11	−0.30	0.31	−0.02	0.22	−0.06	0.25	−0.20	0.09	0.01	0.21	−0.09	0.09	0.03	0.24	0.00	0.09
12	−0.04	0.36	0.14	0.24	0.13	0.35	0.03	0.12	0.20	0.30	−0.04	0.09	0.21	0.29	−0.01	0.25
13	−0.31	0.37	0.10	0.29	−0.04	0.26	−0.04	0.10	0.10	0.17	−0.02	0.05	0.11	0.16	0.02	−0.20
14	−0.63	0.11	−0.44	0.03	−0.24	0.01	−0.36	−0.10	0.03	0.09	−0.00	−0.12	0.12	0.11	−0.14	−0.10
Ave	−0.34	0.22	−0.04	0.10	−0.07	0.13	−0.13	0.01	−0.00	0.08	−0.10	0.01	−0.01	0.05	−0.11	−0.01

Table A9. The 48–72 h NSE of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	−0.57	−0.00	−0.13	−0.12	−0.40	−0.14	−0.21	−0.12	−0.45	−0.26	−0.22	−0.40	−0.34	−0.36	−0.17	−0.29
2	−0.11	0.13	−0.14	0.01	−0.01	0.02	−0.19	−0.18	−0.00	−0.01	−0.11	−0.15	−0.06	−0.02	−0.21	−0.06
3	−0.19	0.01	−0.09	−0.01	0.01	0.05	−0.16	−0.05	0.07	0.09	−0.11	0.05	0.02	0.04	−0.05	−0.07
4	0.06	0.13	−0.08	0.13	0.11	0.05	−0.22	0.07	0.11	0.06	−0.21	−0.04	0.09	0.01	−0.16	−0.07
5	−0.04	0.15	−0.18	0.16	0.11	0.05	−0.06	−0.02	0.16	0.07	−0.01	0.12	0.17	0.11	−0.16	0.04
6	−0.72	0.03	−0.11	0.01	−0.13	−0.25	−0.23	−0.18	−0.02	−0.19	−0.17	−0.11	−0.05	−0.12	−0.36	−0.13
7	−0.71	0.09	−0.10	0.03	−0.11	−0.18	−0.23	−0.06	0.00	−0.11	−0.30	−0.13	0.01	−0.10	−0.29	−0.22
8	−0.44	−0.13	−0.13	−0.20	−0.15	−0.27	−0.39	−0.28	−0.02	−0.29	−0.35	−0.11	−0.08	−0.39	−0.29	−0.26
9	−0.39	0.05	−0.25	−0.16	−0.09	−0.09	−0.16	0.01	−0.03	−0.09	−0.23	−0.20	−0.00	−0.11	−0.23	−0.16
10	−0.12	0.12	−0.28	−0.08	−0.03	0.03	−0.18	0.05	−0.08	−0.08	−0.21	−0.18	−0.19	−0.16	−0.23	−0.14
11	−0.40	0.08	−0.01	0.00	−0.15	0.03	−0.08	0.04	−0.07	−0.04	−0.16	−0.00	−0.04	−0.05	−0.18	−0.01
12	−0.15	0.20	0.03	0.08	0.01	0.16	0.11	0.02	0.07	0.08	−0.27	−0.06	0.11	0.08	−0.07	0.11
13	−0.40	0.13	−0.05	0.09	−0.14	0.01	−0.07	−0.08	0.00	−0.14	−0.17	−0.27	−0.02	−0.10	−0.30	−0.32
14	−0.75	−0.06	−0.19	−0.04	−0.34	−0.17	−0.34	−0.11	−0.07	−0.13	−0.26	−0.10	0.01	−0.14	−0.29	−0.06
Ave	−0.35	0.07	−0.12	−0.01	−0.09	−0.05	−0.17	−0.06	−0.02	−0.07	−0.20	−0.11	−0.03	−0.09	−0.21	−0.12

## Appendix D

**Table A10.** The 0–24 h R of WRF model original output and post-processing model results.

	10 m				30 m				50 m				70 m			
Tower	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	0.48 *	0.73 *	0.43 *	0.56 *	0.52 *	0.69 *	0.41 *	0.49 *	0.52 *	0.64 *	0.36 *	0.41 *	0.52 *	0.66 *	0.39 *	0.47 *
2	0.60 *	0.76 *	0.45 *	0.53 *	0.61 *	0.71 *	0.45 *	0.50 *	0.60 *	0.71 *	0.43 *	0.56 *	0.58 *	0.70 *	0.41 *	0.50 *
3	0.59 *	0.76 *	0.45 *	0.59 *	0.63 *	0.76 *	0.55 *	0.61 *	0.64 *	0.76 *	0.45 *	0.63 *	0.63 *	0.74 *	0.50 *	0.60 *
4	0.72 *	0.81 *	0.54 *	0.64 *	0.69 *	0.78 *	0.53 *	0.64 *	0.67 *	0.76 *	0.51 *	0.60 *	0.65 *	0.74 *	0.51 *	0.57 *
5	0.66 *	0.77 *	0.53 *	0.63 *	0.68 *	0.78 *	0.50 *	0.61 *	0.68 *	0.77 *	0.52 *	0.58 *	0.67 *	0.76 *	0.50 *	0.62 *
6	0.70 *	0.82 *	0.58 *	0.64 *	0.71 *	0.80 *	0.53 *	0.64 *	0.70 *	0.79 *	0.49 *	0.61 *	0.67 *	0.79 *	0.56 *	0.61 *
7	0.67 *	0.82 *	0.56 *	0.64 *	0.69 *	0.79 *	0.47 *	0.64 *	0.69 *	0.79 *	0.54 *	0.65 *	0.68 *	0.79 *	0.50 *	0.64 *
8	0.68 *	0.79 *	0.51 *	0.65 *	0.67 *	0.78 *	0.50 *	0.61 *	0.68 *	0.76 *	0.50 *	0.57 *	0.65 *	0.75 *	0.48 *	0.56 *
9	0.62 *	0.80 *	0.51 *	0.61 *	0.67 *	0.79 *	0.51 *	0.65 *	0.68 *	0.78 *	0.52 *	0.61 *	0.68 *	0.78 *	0.46 *	0.63 *
10	0.70 *	0.85 *	0.64 *	0.74 *	0.70 *	0.83 *	0.58 *	0.71 *	0.68 *	0.80 *	0.55 *	0.67 *	0.65 *	0.80 *	0.54 *	0.64 *
11	0.69 *	0.83 *	0.54 *	0.69 *	0.70 *	0.81 *	0.53 *	0.67 *	0.70 *	0.81 *	0.56 *	0.67 *	0.69 *	0.82 *	0.55 *	0.70 *
12	0.75 *	0.85 *	0.67 *	0.74 *	0.74 *	0.85 *	0.65 *	0.71 *	0.74 *	0.85 *	0.60 *	0.67 *	0.74 *	0.83 *	0.58 *	0.68 *
13	0.75 *	0.85 *	0.63 *	0.74 *	0.73 *	0.83 *	0.58 *	0.71 *	0.72 *	0.81 *	0.55 *	0.70 *	0.72 *	0.80 *	0.54 *	0.66 *
14	0.68 *	0.80 *	0.49 *	0.61 *	0.67 *	0.81 *	0.56 *	0.67 *	0.70 *	0.82 *	0.57 *	0.65 *	0.72 *	0.81 *	0.61 *	0.71 *
Ave	0.66	0.80	0.54	0.64	0.67	0.79	0.53	0.63	0.67	0.77	0.51	0.61	0.66	0.77	0.51	0.61

\* The correlation coefficient has a confidence level of more than 99%.

**Table A11.** The 24–48 h R of WRF model original output and post-processing model results.

	10 m				30 m				50 m				70 m			
Tower	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	0.46 *	0.70 *	0.40 *	0.44 *	0.48 *	0.66 *	0.34 *	0.36 *	0.47 *	0.63 *	0.27 *	0.36 *	0.46 *	0.63 *	0.28 *	0.36 *
2	0.56 *	0.71 *	0.38 *	0.50 *	0.56 *	0.68 *	0.33 *	0.43 *	0.54 *	0.66 *	0.32 *	0.47 *	0.54 *	0.64 *	0.33 *	0.43 *
3	0.56 *	0.73 *	0.37 *	0.53 *	0.58 *	0.72 *	0.42 *	0.49 *	0.59 *	0.73 *	0.42 *	0.51 *	0.57 *	0.72 *	0.37 *	0.50 *
4	0.65 *	0.76 *	0.48 *	0.59 *	0.60 *	0.71 *	0.33 *	0.56 *	0.57 *	0.68 *	0.36 *	0.51 *	0.54 *	0.66 *	0.38 *	0.48 *
5	0.59 *	0.74 *	0.47 *	0.57 *	0.59 *	0.71 *	0.44 *	0.58 *	0.60 *	0.71 *	0.38 *	0.55 *	0.59 *	0.70 *	0.44 *	0.57 *
6	0.57 *	0.74 *	0.46 *	0.55 *	0.56 *	0.70 *	0.42 *	0.52 *	0.56 *	0.70 *	0.44 *	0.55 *	0.53 *	0.70 *	0.43 *	0.52 *
7	0.57 *	0.73 *	0.48 *	0.58 *	0.56 *	0.71 *	0.46 *	0.55 *	0.56 *	0.71 *	0.46 *	0.52 *	0.55 *	0.72 *	0.42 *	0.55 *

Table A11. Cont.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
8	0.56 *	0.72 *	0.43 *	0.54 *	0.55 *	0.70 *	0.40 *	0.55 *	0.56 *	0.69 *	0.46 *	0.51 *	0.53 *	0.67 *	0.32 *	0.50 *
9	0.52 *	0.72 *	0.45 *	0.52 *	0.56 *	0.71 *	0.45 *	0.53 *	0.56 *	0.71 *	0.44 *	0.55 *	0.56 *	0.71 *	0.47 *	0.51 *
10	0.64 *	0.81 *	0.55 *	0.66 *	0.62 *	0.78 *	0.53 *	0.63 *	0.59 *	0.74 *	0.51 *	0.59 *	0.54 *	0.72 *	0.42 *	0.55 *
11	0.60 *	0.77 *	0.49 *	0.61 *	0.60 *	0.76 *	0.39 *	0.58 *	0.60 *	0.75 *	0.46 *	0.57 *	0.59 *	0.75 *	0.47 *	0.58 *
12	0.67 *	0.78 *	0.55 *	0.64 *	0.65 *	0.77 *	0.48 *	0.59 *	0.64 *	0.76 *	0.47 *	0.58 *	0.63 *	0.76 *	0.49 *	0.62 *
13	0.68 *	0.80 *	0.55 *	0.67 *	0.65 *	0.78 *	0.48 *	0.58 *	0.63 *	0.76 *	0.49 *	0.57 *	0.62 *	0.76 *	0.50 *	0.54 *
14	0.59 *	0.73 *	0.23 *	0.53 *	0.56 *	0.72 *	0.33 *	0.54 *	0.58 *	0.74 *	0.48 *	0.54 *	0.60 *	0.75 *	0.43 *	0.55 *
Ave	0.59	0.75	0.45	0.57	0.58	0.72	0.41	0.54	0.57	0.71	0.43	0.53	0.56	0.70	0.41	0.52

\* The correlation coefficient has a confidence level of more than 99%.

Table A12. The 48–72 h R of WRF model original output and post-processing model results.

Tower	10 m				30 m				50 m				70 m			
	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR	WRF	GBDT	DTR	MLPR
1	0.44 *	0.66 *	0.36 *	0.43 *	0.47 *	0.62 *	0.30 *	0.40 *	0.46 *	0.61 *	0.32 *	0.34 *	0.45 *	0.58 *	0.32 *	0.35 *
2	0.57 *	0.68 *	0.37 *	0.47 *	0.56 *	0.65 *	0.36 *	0.40 *	0.55 *	0.64 *	0.36 *	0.42 *	0.53 *	0.62 *	0.31 *	0.41 *
3	0.50 *	0.67 *	0.41 *	0.49 *	0.52 *	0.66 *	0.37 *	0.46 *	0.53 *	0.68 *	0.40 *	0.51 *	0.51 *	0.66 *	0.43 *	0.45 *
4	0.60 *	0.72 *	0.46 *	0.56 *	0.56 *	0.67 *	0.38 *	0.52 *	0.54 *	0.66 *	0.36 *	0.50 *	0.52 *	0.65 *	0.37 *	0.48 *
5	0.57 *	0.72 *	0.39 *	0.58 *	0.57 *	0.70 *	0.43 *	0.51 *	0.57 *	0.70 *	0.45 *	0.59 *	0.56 *	0.70 *	0.37 *	0.53 *
6	0.52 *	0.67 *	0.40 *	0.52 *	0.52 *	0.62 *	0.37 *	0.41 *	0.52 *	0.62 *	0.37 *	0.47 *	0.50 *	0.63 *	0.30 *	0.48 *
7	0.52 *	0.69 *	0.41 *	0.54 *	0.51 *	0.63 *	0.31 *	0.49 *	0.52 *	0.64 *	0.36 *	0.49 *	0.51 *	0.63 *	0.35 *	0.45 *
8	0.51 *	0.65 *	0.40 *	0.45 *	0.50 *	0.63 *	0.33 *	0.42 *	0.51 *	0.62 *	0.31 *	0.46 *	0.49 *	0.60 *	0.32 *	0.41 *
9	0.47 *	0.70 *	0.40 *	0.47 *	0.51 *	0.68 *	0.40 *	0.49 *	0.52 *	0.68 *	0.36 *	0.47 *	0.52 *	0.68 *	0.35 *	0.46 *
10	0.58 *	0.74 *	0.41 *	0.54 *	0.56 *	0.71 *	0.44 *	0.56 *	0.52 *	0.71 *	0.41 *	0.49 *	0.49 *	0.68 *	0.40 *	0.47 *
11	0.54 *	0.74 *	0.48 *	0.56 *	0.53 *	0.71 *	0.44 *	0.56 *	0.53 *	0.72 *	0.42 *	0.53 *	0.52 *	0.71 *	0.41 *	0.53 *
12	0.57 *	0.75 *	0.49 *	0.54 *	0.56 *	0.75 *	0.52 *	0.54 *	0.55 *	0.72 *	0.37 *	0.50 *	0.55 *	0.72 *	0.43 *	0.53 *
13	0.59 *	0.75 *	0.49 *	0.59 *	0.55 *	0.72 *	0.46 *	0.54 *	0.55 *	0.70 *	0.43 *	0.46 *	0.52 *	0.70 *	0.37 *	0.45 *
14	0.50 *	0.70 *	0.35 *	0.50 *	0.47 *	0.68 *	0.31 *	0.49 *	0.50 *	0.68 *	0.36 *	0.48 *	0.52 *	0.69 *	0.33 *	0.52 *
Ave	0.53	0.70	0.41	0.52	0.53	0.67	0.39	0.49	0.53	0.67	0.38	0.48	0.51	0.66	0.36	0.47

\* The correlation coefficient has a confidence level of more than 99%.

## References

1. Rife, D.L.; Davis, C.A.; Liu, Y.; Warner, T.T. Predictability of low-level winds by mesoscale meteorological models. *Mon. Weather Rev.* **2004**, *132*, 2553–2569. [\[CrossRef\]](#)
2. Storm, B.; Dudhia, J.; Basu, S.; Swift, A.; Giammanco, I. Evaluation of the weather research and forecasting model on forecasting low-level jets: Implications for wind energy. *Wind Energy* **2009**, *12*, 81–90. [\[CrossRef\]](#)
3. Marquis, M.; Wilczak, J.; Ahlstrom, M.; Sharp, J.; Stern, A.; Smith, J.C.; Calvert, S. Forecasting the wind to reach significant penetration levels of wind energy. *Bull. Am. Meteorol. Soc.* **2011**, *92*, 1159–1171. [\[CrossRef\]](#)
4. Foley, A.M.; Leahy, P.G.; Marvuglia, A.; McKeogh, E.J. Current methods and advances in forecasting of wind power generation. *Renew. Energy* **2012**, *37*, 1–8. [\[CrossRef\]](#)
5. Zhao, P.; Wang, J.; Xia, J.; Dai, Y.; Sheng, Y.; Yue, J. Performance evaluation and accuracy enhancement of a day-ahead wind power forecasting system in China. *Renew. Energy* **2012**, *43*, 234–241. [\[CrossRef\]](#)
6. Mahoney, W.P.; Parks, K.; Wiener, G.; Liu, Y.; Myers, W.L.; Sun, J.; Delle Monache, L.; Hopson, T.; Johnson, D.; Haupt, S.E. A wind power forecasting system to optimize grid integration. *IEEE Trans. Sustain. Energy* **2012**, *3*, 670–682. [\[CrossRef\]](#)
7. Stathopoulos, C.; Kaperoni, A.; Galanis, G.; Kallos, G. Wind power prediction based on numerical and statistical models. *J. Wind Eng. Ind. Aerodyn.* **2013**, *112*, 25–38. [\[CrossRef\]](#)
8. Wyszogrodzki, A.A.; Liu, Y.; Jacobs, N.; Childs, P.; Zhang, Y.; Roux, G.; Warner, T.T. Analysis of the surface temperature and wind forecast errors of the NCAR-AirDat operational CONUS 4-km WRF forecasting system. *Meteorol. Atmos. Phys.* **2013**, *122*, 125–143. [\[CrossRef\]](#)
9. Deppe, A.J.; Gallus, W.A., Jr.; Takle, E.S. A WRF ensemble for improved wind speed forecasts at turbine height. *Weather Forecast.* **2013**, *28*, 212–228. [\[CrossRef\]](#)
10. Tateo, A.; Miglietta, M.M.; Fedele, F.; Menegotto, M.; Monaco, A.; Bellotti, R. Ensemble using different Planetary Boundary Layer schemes in WRF model for wind speed and direction prediction over Apulia region. *Adv. Sci. Res.* **2017**, *14*, 95. [\[CrossRef\]](#)
11. Cheng, W.Y.; Liu, Y.; Liu, Y.; Zhang, Y.; Mahoney, W.P.; Warner, T.T. The impact of model physics on numerical wind forecasts. *Renew. Energy* **2013**, *55*, 347–356. [\[CrossRef\]](#)
12. Marjanovic, N.; Wharton, S.; Chow, F.K. Investigation of model parameters for high-resolution wind energy forecasting: Case studies over simple and complex terrain. *J. Wind Eng. Ind. Aerodyn.* **2014**, *134*, 10–24. [\[CrossRef\]](#)
13. Liu, Y.; Warner, T.; Liu, Y.; Vincent, C.; Wu, W.; Mahoney, B.; Swerdlin, S.; Parks, K.; Boehnert, J. Simultaneous nested modeling from the synoptic scale to the LES scale for wind energy applications. *J. Wind Eng. Ind. Aerodyn.* **2011**, *99*, 308–319. [\[CrossRef\]](#)
14. Zhang, F.; Yang, Y.; Wang, C. The Effects of Assimilating Conventional and ATOVS Data on Forecasted Near-Surface Wind with WRF-3DVAR. *Mon. Weather Rev.* **2015**, *143*, 153–164. [\[CrossRef\]](#)
15. Ancell, B.C.; Kashawlic, E.; Schroeder, J.L. Evaluation of wind forecasts and observation impacts from variational and ensemble data assimilation for wind energy applications. *Mon. Weather Rev.* **2015**, *143*, 3230–3245. [\[CrossRef\]](#)
16. Ulazia, A.; Saenz, J.; Ibarra-Berastegui, G. Sensitivity to the use of 3DVAR data assimilation in a mesoscale model for estimating offshore wind energy potential. A case study of the Iberian northern coastline. *Appl. Energy* **2016**, *180*, 617–627. [\[CrossRef\]](#)
17. Che, Y.; Xiao, F. An integrated wind-forecast system based on the weather research and forecasting model, Kalman filter, and data assimilation with nacelle-wind observation. *J. Renew. Sustain. Energy* **2016**, *8*, 53308. [\[CrossRef\]](#)
18. Ulazia, A.; Sáenz, J.; Ibarra-Berastegui, G.; González-Rojí, S.J.; Carreno-Madinabeitia, S. Using 3DVAR data assimilation to measure offshore wind energy potential at different turbine heights in the West Mediterranean. *Appl. Energy* **2017**, *208*, 1232–1245. [\[CrossRef\]](#)
19. Cheng, W.Y.; Liu, Y.; Bourgeois, A.J.; Wu, Y.; Haupt, S.E. Short-term wind forecast of a data assimilation/weather forecasting system with wind turbine anemometer measurement assimilation. *Renew. Energy* **2017**, *107*, 340–351. [\[CrossRef\]](#)

20. Akish, E.; Bianco, L.; Djalalova, I.V.; Wilczak, J.M.; Olson, J.B.; Freedman, J.; Finley, C.; Cline, J. Measuring the impact of additional instrumentation on the skill of numerical weather prediction models at forecasting wind ramp events during the first Wind Forecast Improvement Project (WFIP). *Wind Energy* **2019**, *22*, 1165–1174. [\[CrossRef\]](#)
21. Costa, A.; Crespo, A.; Navarro, J.; Lizcano, G.; Madsen, H.; Feitosa, E. A review on the young history of the wind power short-term prediction. *Renew. Sustain. Energy Rev.* **2008**, *12*, 1725–1744. [\[CrossRef\]](#)
22. Jung, J.; Broadwater, R.P. Current status and future advances for wind speed and power forecasting. *Renew. Sustain. Energy Rev.* **2014**, *31*, 762–777. [\[CrossRef\]](#)
23. Glahn, H.R.; Lowry, D.A. The use of model output statistics (MOS) in objective weather forecasting. *J. Appl. Meteorol.* **1972**, *11*, 1203–1211. [\[CrossRef\]](#)
24. Carter, G.M.; Dallavalle, J.P.; Glahn, H.R. Statistical forecasts based on the National Meteorological Center's numerical weather prediction system. *Weather Forecast.* **1989**, *4*, 401–412. [\[CrossRef\]](#)
25. Jacks, E.; Bower, J.B.; Dagostaro, V.J.; Dallavalle, J.P.; Erickson, M.C.; Su, J.C. New NGM-based MOS guidance for maximum/minimum temperature, probability of precipitation, cloud amount, and surface wind. *Weather Forecast.* **1990**, *5*, 128–138. [\[CrossRef\]](#)
26. Hart, K.A.; Steenburgh, W.J.; Onton, D.J.; Siffert, A.J. An evaluation of mesoscale-model-based model output statistics (MOS) during the 2002 Olympic and Paralympic Winter Games. *Weather Forecast.* **2004**, *19*, 200–218. [\[CrossRef\]](#)
27. Wilks, D.S.; Hamill, T.M. Comparison of ensemble-MOS methods using GFS reforecasts. *Mon. Weather Rev.* **2007**, *135*, 2379–2390. [\[CrossRef\]](#)
28. Stensrud, D.J.; Skindlov, J.A. Gridpoint predictions of high temperature from a mesoscale model. *Weather Forecast.* **1996**, *11*, 103–110. [\[CrossRef\]](#)
29. Stensrud, D.J.; Yussouf, N. Short-range ensemble predictions of 2-m temperature and dewpoint temperature over New England. *Mon. Weather Rev.* **2003**, *131*, 2510–2524. [\[CrossRef\]](#)
30. Eckel, F.A.; Mass, C.F. Aspects of effective mesoscale, short-range ensemble forecasting. *Weather Forecast.* **2005**, *20*, 328–350. [\[CrossRef\]](#)
31. Hacker, J.P.; Rife, D.L. A practical approach to sequential estimation of systematic error on near-surface mesoscale grids. *Weather Forecast.* **2007**, *22*, 1257–1273. [\[CrossRef\]](#)
32. Homleid, M. Diurnal corrections of short-term surface temperature forecasts using the Kalman filter. *Weather Forecast.* **1995**, *10*, 689–707. [\[CrossRef\]](#)
33. Roeger, C.; Stull, R.; McClung, D.; Hacker, J.; Deng, X.; Modzelewski, H. Verification of mesoscale numerical weather forecasts in mountainous terrain for application to avalanche prediction. *Weather Forecast.* **2003**, *18*, 1140–1160. [\[CrossRef\]](#)
34. McCollor, D.; Stull, R. Hydrometeorological accuracy enhancement via postprocessing of numerical weather forecasts in complex terrain. *Weather Forecast.* **2008**, *23*, 131–144. [\[CrossRef\]](#)
35. Delle Monache, L.; Nipen, T.; Liu, Y.; Roux, G.; Stull, R. Kalman filter and analog schemes to postprocess numerical weather predictions. *Mon. Weather Rev.* **2011**, *139*, 3554–3570. [\[CrossRef\]](#)
36. Cassola, F.; Burlando, M. Wind speed and wind energy forecast through Kalman filtering of Numerical Weather Prediction model output. *Appl. Energy* **2012**, *99*, 154–166. [\[CrossRef\]](#)
37. Li, G.; Shi, J. On comparing three artificial neural networks for wind speed forecasting. *Appl. Energy* **2010**, *87*, 2313–2320. [\[CrossRef\]](#)
38. Ishak, A.M.; Remesan, R.; Srivastava, P.K.; Islam, T.; Han, D. Error correction modelling of wind speed through hydro-meteorological parameters and mesoscale model: A hybrid approach. *Water Resour. Manag.* **2013**, *27*, 1–23. [\[CrossRef\]](#)
39. Sweeney, C.P.; Lynch, P.; Nolan, P. Reducing errors of wind speed forecasts by an optimal combination of post-processing methods. *Meteorol. Appl.* **2013**, *20*, 32–40. [\[CrossRef\]](#)
40. Zjavka, L. Wind speed forecast correction models using polynomial neural networks. *Renew. Energy* **2015**, *83*, 998–1006. [\[CrossRef\]](#)
41. Zhao, J.; Guo, Z.; Su, Z.; Zhao, Z.; Xiao, X.; Liu, F. An improved multi-step forecasting model based on WRF ensembles and creative fuzzy systems for wind speed. *Appl. Energy* **2016**, *162*, 808–826. [\[CrossRef\]](#)
42. Zhao, X.; Liu, J.; Yu, D.; Chang, J. One-day-ahead probabilistic wind speed forecast based on optimized numerical weather prediction data. *Energy Convers. Manag.* **2018**, *164*, 560–569. [\[CrossRef\]](#)

43. Papayiannis, G.I.; Galanis, G.N.; Yannacopoulos, A.N. Model aggregation using optimal transport and applications in wind speed forecasting. *Environmetrics* **2018**, *29*, e2531. [CrossRef]
44. Skamarock, W.C.; Klemp, J.B.; Dudhia, J.; Gill, D.O.; Barker, D.M.; Wang, W.; Powers, J.G. A Description of the Advanced Research WRF Version 3. NCAR Technical Note-475+ STR. 2008. Available online: <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.484.3656> (accessed on 11 November 2019).
45. National Centers for Environmental Prediction/National Weather Service/NOAA/U.S. Department of Commerce. NCEP FNL Operational Model Global Tropospheric Analyses, Continuing from July 1999. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. 2000. Available online: <https://doi.org/10.5065/D6M043C6> (accessed on 11 December 2018).
46. Willmott, C.J. On the validation of models. *Phys. Geogr.* **1981**, *2*, 184–194. [CrossRef]
47. Willmott, C.J. On the evaluation of model performance in physical geography. In *Spatial Statistics and Models*; Gaile, G.L., Willmott, C.J., Eds.; Springer: Dordrecht, The Netherlands, 1984; pp. 443–460.
48. Willmott, C.J.; Ackleson, S.G.; Davis, R.E.; Feddema, J.J.; Klink, K.M.; Legates, D.R.; O'Donnell, J.; Rowe, C.M. Statistics for the evaluation and comparison of models. *J. Geophys. Res. Oceans* **1985**, *90*, 8995–9005. [CrossRef]
49. Legates, D.R.; McCabe, G.J., Jr. Evaluating the use of “goodness-of-fit” measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **1999**, *35*, 233–241. [CrossRef]
50. Nash, J.E.; Sutcliffe, J.V. River flow forecasting through conceptual models part I—A discussion of principles. *J. Hydrol.* **1970**, *10*, 282–290. [CrossRef]
51. Zhou, Z. *Ensemble Methods: Foundations and Algorithms*, 1st ed.; Chapman & Hall/CRC: New York, NY, USA, 2012.
52. Kearns, M.J.; Valiant, L.G. Cryptographic limitations on learning Boolean formulae and finite automata. In *Machine Learning: From Theory to Applications*; Hanson, S.J.E.A., Ed.; Springer: Berlin/Heidelberg, Germany, 1993; pp. 29–49.
53. Schapire, R.E. The Strength of Weak Learnability. *Mach. Learn.* **1990**, *5*, 197–227. [CrossRef]
54. Friedman, J.H.; Hastie, T.; Tibshirani, R. Additive logistic regression: A statistical view of boosting. *Ann. Stat.* **2000**, *28*, 337–407. [CrossRef]
55. Kegl, B. The return of AdaBoost.MH: Multi-class Hamming trees. In Proceedings of the International Conference on Learning Representations, Banff, AB, Canada, 14–16 April 2014.
56. Quinlan, J.R. Induction of Decision Trees. *Mach. Learn.* **1986**, *1*, 81–106. [CrossRef]
57. Quinlan, J.R. Improved Use of Continuous Attributes in C4.5. *J. Artif. Int. Res.* **1996**, *4*, 77–90. [CrossRef]
58. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; Wadsworth & Brooks/Cole Advanced Books & Software: Monterey, CA, USA, 1984.
59. Mason, L.; Baxter, J.; Bartlett, P.L.; Frean, M. Boosting Algorithms as Gradient Descent. *Adv. Neural Inf. Process. Syst.* **1999**, 512–518. [CrossRef]
60. Friedman, J.H. Greedy function approximation: A gradient boosting machine. *Ann. Stat.* **2001**, *29*, 1189–1232. [CrossRef]
61. Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. In *Advances in Neural Information Processing Systems 30*; Guyon, I., Luxburg, U.V., Bengio, S., Wallach, H., Fergus, R., Vishwanathan, S., Garnett, R., Eds.; Curran Associates, Inc.: Red Hook, NY, USA, 2017; pp. 3146–3154.
62. Lun, I.Y.F.; Lam, J.C. A study of Weibull parameters using long-term wind observations. *Renew. Energy* **2000**, *20*, 145–153. [CrossRef]

