Article

A Novel Hybrid Method for Short-Term Wind Speed Prediction Based on Wind Probability Distribution Function and Machine Learning Models

Rabin Dhakal [1], Ashish Sedai [2], Suhas Pol [2], Siva Parameswaran [1], Ali Nejat [3] and Hanna Moussa [1,*]

1 Department of Mechanical Engineering, Texas Tech University, Lubbock, TX 79409, USA
2 National Wind Institute, Texas Tech University, Lubbock, TX 79409, USA
3 Department of Civil, Environmental and Construction Engineering, Texas Tech University, Lubbock, TX 79409, USA
* Correspondence: hanna.moussa@ttu.edu; Tel.: +1-806-632-8141

Abstract: The need to deliver accurate predictions of renewable energy generation has long been recognized by stakeholders in the field and has propelled recent improvements in more precise wind speed prediction (WSP) methods. Models such as Weibull-probability-density-based WSP (WEB), Rayleigh-probability-density-based WSP (RYM), autoregressive integrated moving average (ARIMA), Kalman filter and support vector machines (SVR), artificial neural network (ANN), and hybrid models have been used for accurate prediction of wind speed with various forecast horizons. This study intends to incorporate all these methods to achieve a higher WSP accuracy as, thus far, hybrid wind speed predictions are mainly made by using multivariate time series data. To do so, an error correction algorithm for the probability-density-based wind speed prediction model is introduced. Moreover, a comparative analysis of the performance of each method for accurately predicting wind speed for each time step of short-term forecast horizons is performed. All the models studied are used to form the prediction model by optimizing the weight function for each time step of a forecast horizon for each model that contributed to forming the proposed hybrid prediction model. The National Oceanic and Atmospheric Administration (NOAA) and System Advisory Module (SAM) databases were used to demonstrate the accuracy of the proposed models and conduct a comparative analysis. The results of the study show the significant improvement on the performance of wind speed prediction models through the development of a proposed hybrid prediction model.

Keywords: forecasting; machine learning; Weibull distribution; wind speed

1. Introduction

Wind energy is a variable renewable energy source [1] and the power produced by the wind turbine hence fluctuates with the variation of wind speed [2]; therefore, in wind farms, unexpected variations of wind power output may increase the operating costs of the electricity system. So, intermittency of wind is the biggest challenge for a wind farm to implement wind energy as a reliable autonomous source of electric power [3]. Moreover, a wind speed forecasting (WSF) system based on an accurate model that reflects the variation of wind speed is critical to effective wind energy harvesting, integration of available wind power into the electrical power grid, and analyzing the efficiency and performance of wind-turbine-based electrical generation systems [4]. Despite the development of various WSF methods, accurately predicting wind speed still remains a challenge. Furthermore, the length of the forecast horizons correlates with the accuracy of forecasting techniques. Wind speed prediction can have various implications requiring different time scales. For example, turbine control often necessitates a response time of seconds or fractions of seconds, whereas grid integration production planning and market response require longer time horizons. The time scale of prediction also differs according to the energy markets.

The real-time energy market requires a response in minutes, whereas the day-ahead energy market requires the prediction up to 24 h ahead as it requires information for energy trading for the next day [5,6]. There may be a requirement for different time scale forecast horizons in between these two-time scales. For example, economic load dispatching and load increment/decrement decisions require a time scale of 30 min to 6 h ahead [3].

Wind speed prediction models have been classified mainly into four categories in the literature: (a) The persistence model, in which future wind speed is deemed to be equal to the wind speed at the forecasting time [7]. It is an economical and simple method that can be adopted by almost everyone to serve as the base model for comparing forecasted values by other methods and its main drawback is its unsuitability for forecasting more than a one-time step of forecast horizons; (b) The physical method, in which numerical weather prediction (NWP) is used by incorporating complex atmospheric characteristics, including temperature, pressure, and wind shear into wind speed predictions [8]. For long-term forecasts, NWP produces precise estimates that are generally applied over vast areas. However, since numerical weather prediction models are memory- and time-intensive, they are not ideal candidates for short time horizons forecasting; (c) Statistical methods, in which one can explore the mathematical relationship between the various features of the wind time series data. This method includes the following models: Weibull-probability-density-based forecasting (WEB), autoregressive integrated moving average (ARIMA), and the Bayesian probability density function (BBM) approach. These models are mostly used for short forecast horizons and are not suitable for longer forecast horizons due to their non-linearity assumption for wind data; and (d) Artificial intelligence, which includes neural networks (ANN) [9,10], regression or decision trees (RT) [11,12], support vector regression (SVR) [13–16], and recurrent neural networks (RNN) [17–19].

This study was inspired by the work of Kadhem et al., Kaplan et al., and Ding et al. [20–22], where the idea of a probability-density-function-based wind speed prediction model was introduced. In this study, the performance of various univariate models is compared, and an error correction algorithm is proposed for the probability-density-based wind speed prediction model. The contribution of this research is twofold: firstly, the proposed error correction method is a novel method that improves the performance of the previously introduced wind-probability-density-based wind speed prediction; and secondly, it introduces a novel hybrid method that is capable of integrating all the studied methods with an optimized weighted coefficient for both the classical time series method and artificial intelligence methods.

This manuscript is divided into the following sections: literature review, methodology, results and discussion, and conclusion. The literature review section consists of a description of the wind speed prediction methods found in literature, which forms the basis of the proposed methodology. The methodology sections consist of a description of all the methods of wind speed prediction, the method of developing error correction, and making a hybrid wind speed prediction model. The results tables that were obtained are presented under the results and discussion section, along with discussion. Finally, the conclusion of the research and future works on the research area are presented under the conclusion section.

2. Literature Review

This section is dedicated to the theory related to the current study, the methods developed, and the description of the time scale of WSP, Weibull and Rayleigh probability distribution function, support vector regressions, and LSTM networks.

2.1. Time Scale in Wind Speed Prediction

One of the important subjects in wind speed prediction is the time scale requirement of forecast horizons; since the different application of wind speed prediction requires different types of time scale, the classification of the time scale of forecast horizons in wind speed prediction methods is an ambiguous subject [3]. Turbine control often necessitates

a response time of seconds or fractions of seconds, whereas grid integration production planning and market response require longer time horizons. The time scale of prediction also differs according to the energy markets; the real-time energy market requires the response in minutes, whereas the day-ahead energy market requires the prediction up to 24 h ahead as it requires information for energy trading for the next day [5,6]; also, there may be a requirement of different time scale forecast horizons in between these two time scales. For example, economic load dispatching and load increment/decrement decisions require a time scale of 30 min to 6 h ahead [3]. In this study, we focus on short-term wind speed prediction (a few hours ahead prediction, not exceeding 12 h).

2.2. Wind-Probability-Distribution-Function-Based Wind Speed Prediction Model

Wind-probability-distribution-function-based WSP models were developed by assuming that wind speed follows the same distribution for the next time period. This approach is aligned with the concept behind the persistence model (PM), according to which, any future wind speed value is equal to its last known value of wind speed due to the high autocorrelation on the behavior of the wind speed [7]. Despite its simplicity, the PM produces excellent WSP results and is used to assess the quality of new WSP methods [23]. PMs forecast wind speed (u_{t+h}) at any future time, as $t + h$, $h > 0$ is the same as wind speed (u_t) at current time t.

$$u_{t+h} = u_t, \quad h > 0 \tag{1}$$

Wind speed follows non-negative and right-skewed distribution rather than the normal distribution [24]. Several probabilities distribution functions are right-skewed and non-negative and are used for modeling wind speed. Weibull distribution and Rayleigh distribution are the most common probability distribution functions for wind speed modeling [25]. Although the Weibull distribution function is the most widely used function, there is no consensus on which best describes wind speed data for a specific case study site. Therefore, in this study, we consider both Weibull- and Rayleigh-based wind speed modeling and forecasting.

2.2.1. Weibull-Distribution-Based WSP (WEB)

The Weibull probability density function is a two-parameter distribution with a dimensionless shape parameter k and a velocity scale parameter c in m/s [26].

$$f(u) = \frac{k}{c}\left(\frac{u}{c}\right)^{k-1} \exp\left[-\left(\frac{u}{c}\right)^k\right] \tag{2}$$

where $f(u)$ denotes the probability distribution of wind speed u. The quality of wind resources can be evaluated from the parameters c and k. The parameter c is proportional to the wind speed, and k characterizes the shape of the Weibull distribution. Variable wind speeds are indicated by smaller values of k, whilst constant wind speeds are indicated by greater values. Typical values of k are between 1 and 3 [27]. Even though there are numerous methods to derive Weibull parameters, such as the graphical method, method of moments, maximum likelihood method, standard deviation method, modified maximum likelihood method, power density method, and equivalent energy method, the maximum likelihood method is deemed to be the best fit [28]; this method employs the following expression to calculate shape parameters (k) using an iterative process [26]:

$$k = \left(\frac{\sum_{i=1}^{N} u_i^k \ln(u_i)}{\sum_{i=1}^{N} u_i^k} - \frac{\sum_{i=1}^{N} \ln(u_i)}{N}\right)^{-1} \tag{3}$$

where u_i is the wind speed at the time step of i and the number of time steps is given by N. After getting the shape parameter k, the expression below is used to measure the scale parameter c [26].

$$c = \left(\frac{1}{N}\sum_{i=1}^{N} u_i^k\right)^{1/k} \tag{4}$$

Similar to the PM method, the Weibull-distribution-based wind speed prediction model (WEB) assumes that wind speed follows the same distribution for the next time period. So, the mean speed (u') can be used as a point forecast in the WEB [25].

$$u' = c\ \Gamma\left(1 + \frac{1}{k}\right)$$
(5)

where Γ is the gamma function, defined as: $\int_0^\infty \exp(-u)(u)^{x-1}dx$.

Median and mode can also be used for forecasting purposes [22]. However, mean speed might not provide accurate predictions due to the skewness in the Weibull probability density function.

$$median = c\ (\ln 2)^{1/k}$$
(6)

$$mode = c\left(1 - \frac{1}{k}\right)^{1/k}$$
(7)

where *mode* = 0 when $k \leq 1$. It is practically not possible to have a scale factor less than 1 at commercial wind farms [22]. In our study, we used mean speed as point forecast in WEB.

The cumulative distribution *F(u)* is an integral of the probability distribution function given by Equation (8), which gives the probability of getting wind speed *u* or less.

$$F(u) = 1 - \exp\left[-\left(\frac{u}{c}\right)^k\right]$$
(8)

$$\text{Assume,}\ R = F(u) = 1 - \exp\left[-\left(\frac{u}{c}\right)^k\right]$$
(9)

Using inverse transform, we get:

$$u = c\left[-ln(1-R)^{\frac{1}{k}}\right]$$
(10)

where *R* is a random variable with values between 0 and 1 and, as shown in Equation (9), representing the cumulative distribution function [21]. The values of random variable *R* between 0 and 1 should be uniformly distributed. Hence, in this way, we have used Equation (10) to simulate a wind speed using parameters of the Weibull distribution function and the method is represented as WEBS.

2.2.2. Rayleigh-Distribution-Based WSP (RYM)

The Rayleigh probability distribution function is a special case of the Weibull distribution function where k = 2. Therefore, in this case, the scale parameter can be determined using the following expression:

$$c = \frac{2u'}{\sqrt{\pi}}$$
(11)

Therefore, the probability density function represented by Rayleigh distribution and its cumulative distribution function are given as:

$$f(u) = \frac{2u}{c^2} \exp\left[-\left(\frac{u}{c}\right)^2\right]$$
(12)

$$F(u) = 1 - \exp\left[-\left(\frac{u}{c}\right)^2\right]$$
(13)

Similar to Equations (9) and (10), using an inverse transform of Equation (12), we get,

$$u = -c \times ln\ (1 - R)$$
(14)

Hence, Equation (14) can be used to simulate wind speed using the parameter of the Rayleigh distribution function.

2.3. Autoregressive Integrated Moving Average (ARIMA) Model

In ARIMA, a time series model reproduces the patterns of a variable's previous movements across time and uses this information to forecast its future movements [29]. Wind speed measurements obtained over time tend to be positively correlated. Many parametric time series models that consider the autoregressive (AR) process exist to account for this autocorrelation [30]. In an autoregressive model, we forecast the wind speed using a linear combination of past wind speed values.

$$u_t = \gamma_1 u_{t-1} + \gamma_2 u_{t-2} + \cdots + \gamma_p u_{t-p} + \epsilon_t \tag{15}$$

Equation (15) provides the AR model of order p, where γ is the autoregression coefficient and ϵ_t is the noise in time t. A moving average term is added in the autoregressive model, and the autoregressive moving average model is developed and described as follows:

The autoregressive moving average (ARMA) model is a type of autoregressive model that also adapts the moving average model. It is a statistical model that could be used for time series prediction of future wind speed values using past values and lagged forecast error. A general ARMA is denoted by ARMA (p, q) and can be expressed by the following expressions:

$$u_t = \delta + \sum_{i=1}^{p} \gamma_i\, u_{t-i} + \sum_{j=1}^{q} \varnothing_j e_{t-j} + e_t \tag{16}$$

where the second term from the right in Equation (16) is the moving average (MA) part of the ARMA model, δ is the constant, $\varnothing_j$ is the jth moving average coefficient, e_t is the error term at time period t, and u_t is the value of wind speed predicted at time step t. If the differencing is added to the ARMA model, the model is transformed into the ARIMA model. Therefore, the ARIMA model, introduced by Box and Jenkins, includes autoregression (AR), a moving average (MA), and differencing [31]. The non-seasonal model structure of ARIMA is expressed in the form of ARIMA (p, d, q), where d is the order of differencing (I) to make the model stationary. Hence, the seasonal time series is stationery in nature and becomes zero, and the ARIMA model is converted to the ARMA model [32].

2.4. Support Vector Regression (SVR)

Support vector regression is an extension of a support vector machine and was proposed by Drucker et al. [33]. A support vector machine was initially developed for the classification problem. SVR is based on the structural error minimization principle and consists of the 'Kernel Trick' and other optimization features that allow it to perform a noise-robust and non-linear regression [34,35]. Its stability and accuracy depend on several aspects, such as parameter tuning and feature selection. Parameter tuning is a procedure consisting of properly selecting the kernel function and its parameters and penalization term [36]. Feature selection consists of the selection of the most important variables of the model to describe the behavior of the trend [37]. SVR does the best trade-off between Field's empirical error and complexity [38].

2.5. Long Short-Term Memory (LSTM) Model

Recurrent neural networks (RNN) are a suitable model for time series forecasting problems. However, RNN are not suitable for long-term dependency tasks due to the vanishing/exploding gradient decent issues [39]. Therefore, the LSTM neural network arises, which can learn the long-term dependency jobs very efficiently compared to the general RNN model [40]. The LSTM model solves the vanishing/exploding gradient decent issues with gates present within each cell of an LSTM network [41–43]. LSTM is one of the popular artificial recurrent neural network architectures used in wind speed prediction [44]. The LSTM neural network was first proposed by Hochreiter and Schmidhuber [45]. An LSTM cell's internal state memory offers the internal storing of pertinent historical information. The flow of information through the cell is controlled by the cell's input, output, and forget gates and the mathematical implementation of each LSTM cell is described using Equations

(17) to (22). With the help of these gates, LSTM analyzes and saves pertinent data [46]. The phrase stacked/deep LSTM is often used to denote the LSTM network, referring to an LSTM network with two or more hidden layers. An LSTM network with a detail structure of an LSTM cell is shown in Figure 1.
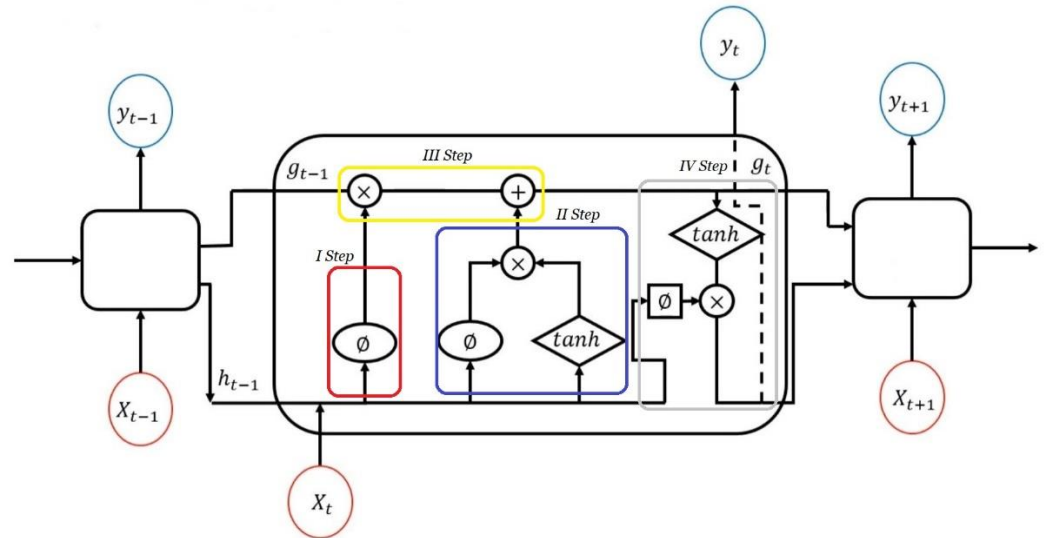
Figure 1. LSTM network and detail structure of an LSTM cell.

I Step: Forget layer

$$f_t = \varnothing(w_f.[y_{t-1}, x_t] + \beta_f \tag{17}$$

II Step: Update of new values (I_t) and creation of a vector of new information ($\widetilde{g}$) to add to the cell state.

$$I_t = \varnothing(w_i.[y_{t-1}, x_t] + \beta_i) \tag{18}$$

$$\widetilde{g} = \tanh(w_i.[y_{t-1}, x_t] + \beta_s) \tag{19}$$

III Step: Final cell state

$$g_t = f_t.g_{t-1} + I_t . \widetilde{g} \tag{20}$$

IV Step: Last stage using a sigmoid function and a tanh, regenerating values between -1 and 1.

$$\partial_t = \varnothing(w_\partial.[y_{t-1}, x_t] + \beta_\partial) \tag{21}$$

$$y_t = \partial_t.\tan h(g_t) \tag{22}$$

In Equations (17) to (22), x_t is the input, $\widetilde{g}$ is the state of the network, g_t is the temporary state and y_t is the output state at time step t. I_t denotes the input gate, ∂_t represents the output gate and f_t denotes the forget gate. The weight corresponding to the hidden layer, input layer, and output layer is denoted by w_f, w_i, and w_∂, respectively. β_f, β_i β_s, and β_∂ represent bias corresponding to the input, sate of network, temporary state, and output layer of the network. $\varnothing$ and tanh represent the sigmoid and tanh activation function, which is defined by the following expressions:

$$\varnothing(z) = \frac{1}{1 + e^{-z}} \tag{23}$$

$$\tan h(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}} \tag{24}$$

3. Methodology

3.1. Proposed Hybrid Method for WSP

In the hybrid wind speed prediction model, all five models are used, i.e, PM, WEB, ARIMA, SVR, and LSTM. For each time step of a forecast horizon, a weight parameter is assigned. The weight parameter is then optimized using linear optimization by minimizing the loss function. Any one of the performance parameters, such as MAPE, MAE, and RMSE, is taken as a loss function to be minimized. After deriving weight parameters for each time step for each model, the hybrid WSP model is applied to predict the future time step wind speed up to given forecast horizons.

3.2. Data Acquisition

Data for this study comprise wind speed data at a hub height of 80 m and were extracted from the National Renewable Energy Laboratory System Advisor Model (SAM) database [46]. The hub height of 80 m was chosen as the commercial larger scale wind turbine operates mostly at this hub height. We have collected data from four different regions (South Plains region of Texas, Southern Offshore region of Texas, Hills region of Arizona, and Hills region of West Virginia) in the United States to encompass various weather conditions. The variations of wind speed and direction, being dominant features for WSP, are shown in Figure 2. The dataset contains pressure and temperature data as well. The data were sampled at hourly intervals.
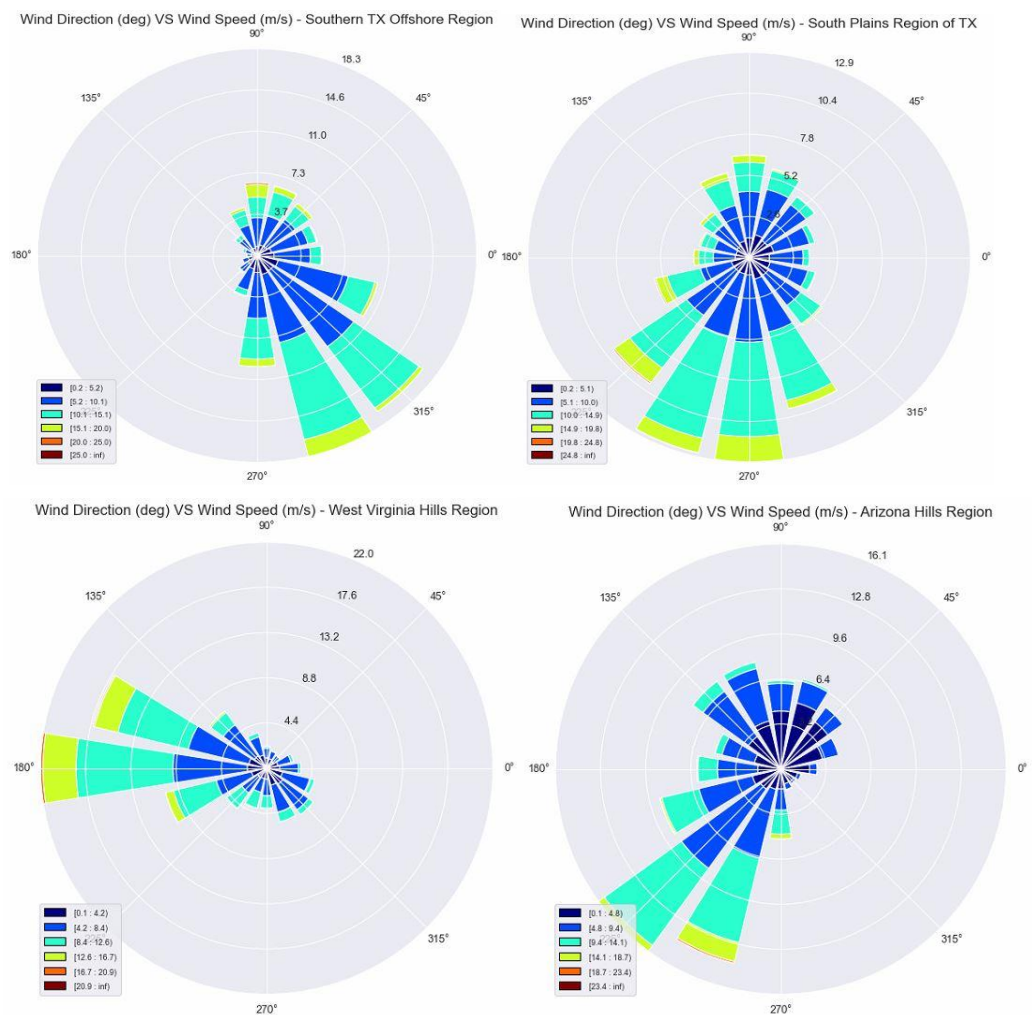
Figure 2. Distribution of wind speed and direction at four different sites.

3.3. Statistical Analysis

Basic statistical analysis is performed using the method presented in Section 2.2.1. The scale and shape parameters are calculated using Equations (3) and (4) for the WEB. The mean speed is calculated using Equation (5) for point forecast using WBM. Scale and shape parameters are used in Equation (10) to simulate wind speed data for WSP using WEBS. Similarly, the scale parameter for Rayleigh probability distribution is calculated using Equation (11) for the RYM. In addition, Equation (14) is used to simulate wind speed data for WSP using RYMS.

3.4. Error Correction and Wind Speed Generation

The flow diagram in Figure 3 shows the proposed algorithm for error correction for the simulated wind speed using Equations (10) and (14) for Weibull and Rayleigh distribution functions, respectively. Initially, the sequential variation data of wind speed is calculated, and mean variation is recorded. We assume that the wind distribution follows the persistence model. Hence, the predicted wind speed should also have the same sequential variation. Therefore, the simulated wind speed is checked each time to ensure that it is within the acceptable range ($-\alpha$ and α). If it is not within the limits, the algorithm of Gaussian filtering is applied [21]; this algorithm would prevent unnecessary deviations of the predicted value from the range extracted from historical data. Thus, error-corrected wind speed is generated and is used as the predicted value for WEBS and RYMS and input for other machine learning models. The error-corrected model based on WEBS is abbreviated as WEBSEC and the error-corrected model based on RYMS is abbreviated as RYMSEC.
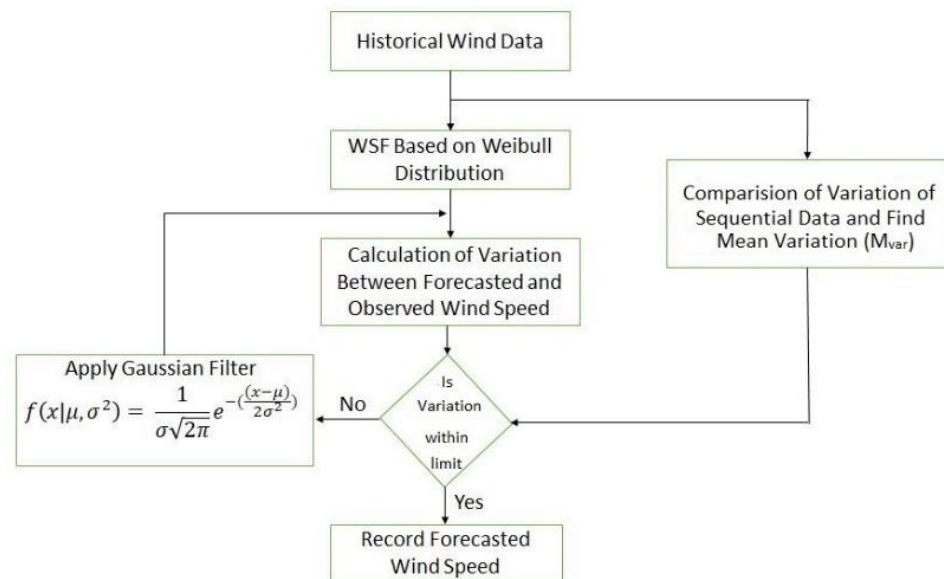
Figure 3. Error correction algorithm for simulated wind speed using Weibull and Rayleigh distribution function.

3.5. Data Preprocessing

Many machine learning algorithms compare attributes of data points to detect trends in the data. However, problems would arise when features were on different scales. Therefore, data are normalized before being sent for training in machine learning models. One of the most prevalent methods of data normalization is the min–max method in which values are transformed to possess values between 0 and 1. Therefore, if x and x' are the actual and normalized values of the feature and max and min are maximum and minimum values of the feature, then the normalization can be represented by Equation (25)

$$x' = \frac{x - min}{max - min} \tag{25}$$

After the normalization process, the dataset is divided into training and testing datasets. Usually, 70% of the data are used for the training of the model and the remaining 30% are used for testing the model. Out of the testing dataset, 10% of the total dataset is used for the validation of the model performance. However, we have train and test datasets for short forecast horizons of 6 h. For this, we have taken the last six time-step data as a test set and the remaining data as a training dataset.

3.6. Performance Evaluation

The evaluation of the performance of the individual models and the hybrid model proposed in this study is performed using popular statistical error indicators such as MAE, RMSE, and MAPE [47]. If x_j, x_j^p, and x_j' indicate the actual, predicted, and mean value of the wind speed, respectively, and n is the number of samples, each error indicator can be expressed using Equations (26) to (28) as follows:

$$MAE = \frac{1}{n} \sum_{j=1}^{n} \left| x_j - x_j^p \right| \tag{26}$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^{n} \left(x_j - x_j^p \right)^2 x_j - x_j^p} \tag{27}$$

$$MAPE = \frac{1}{n} \sum_{t=1}^{n} \left| \frac{x_j - x_j^p}{x_j} \right| \times 100\% \tag{28}$$

4. Results and Discussion

4.1. Probability Distribution Function Parameter Result

The estimation of Weibull and Rayleigh parameters for year 2010 at the South Plains region of Texas, Southern Offshore region of Texas, Hills region of West Virginia, and Hills region of Arizona are presented in this section. The shape and scale parameters were estimated for the whole year's data for all locations. Details are shown in Table 1.

Table 1. Scale and shape parameters for four locations for the year 2010.

Location	Weibull Parameter		Rayleigh Parameter
	k	**c**	**c**
South Plains TX	2.4	10.2	10.16
Southern Offshore TX	2.45	10	10.05
Hills region WV	2.19	8.22	8.23
Hills region AR	1.84	7.39	7.4

4.2. Comparative Study of Probability-Distribution-Function-Based WSP Models

First, Weibull and Rayleigh parameters were calculated as shown in Table 1. It is clearly seen in the table that the probability distribution for both Weibull and Rayleigh distribution functions were almost the same as the shape parameters of all three sites, close to two. We then investigated the forecasting accuracy using both Weibull- and Rayleigh-distribution-based models.

The results from the analysis of the probability-density-based wind prediction model are presented in Tables 2–5. The performance of the models was assessed through various performance metrics such as RMSE, MAE, and RMSE. This analysis was performed for short-term forecast horizons. While evaluating the performance of the next seven hours of time step, each hour of the time step was also evaluated to know about the detail of the forecasting model.

Table 2. Performance results of different WSP models for the South Plains TX region.

				RMSE				
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	1.02	1.1	0.61	2.18	3.48	1.1	1.07	1.51
WEBS	0.33	4.15	0.22	4.48	1.89	3.008	0.4	2.07
WEBSEC	1.77	1.28	0.19	3.82	0.62	2.81	0.02	1.50
RYM	1.02	1.1	0.61	2.19	3.48	3.16	1.1	1.81
RYMS	5.62	4.54	2.4	4.69	1.99	6.39	1.24	3.84
RYMSEC	1.3	0.52	0.4	4.67	5.36	1.02	5.26	2.65
				MAE				
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	1.02	1.1	0.61	2.18	3.48	3.16	1.1	1.81
WEBS	0.32	4.17	0.22	4.48	1.89	3.008	0.4	2.07
WEBSEC	1.77	1.28	0.19	3.82	0.62	2.81	0.01	1.50
RYM	1.02	0.107	0.61	2.19	3.48	3.16	1.1	1.67
RYMS	5.62	4.54	2.4	4.69	1.99	6.39	1.24	3.84
RYMSEC	1.36	0.52	0.4	4.67	5.36	1.02	5.26	2.66
				MAPE				
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	0.12	0.14	0.064	0.19	0.27	0.25	0.1	0.16
WEBS	0.04	0.52	0.023	0.4	0.15	0.24	0.04	0.20
WEBSEC	0.22	0.16	0.02	0.34	0.05	0.23	0.001	0.15
RYM	0.12	0.14	0.06	0.19	0.27	0.26	0.1	0.16
RYMS	0.7	0.57	0.25	0.41	0.16	0.52	0.12	0.39
RYMSEC	0.17	0.06	0.04	0.41	0.42	0.08	0.52	0.24

Table 3. Performance results of different WSP models for Southern Offshore TX region.

				RMSE				
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	3.47	1.68	1.85	2.02	2.75	2.94	2.43	2.45
WEBS	0.35	5.91	2.54	0.55	0.68	1.51	1.04	1.80
WEBSEC	0.19	0.25	3.7	0.63	0.7	2.65	2.32	1.49
RYM	3.45	1.69	1.86	2.03	2.76	2.94	2.44	2.45
RYMS	4.64	4.96	4.65	3.13	0.75	5.59	0.78	3.50
RYMSEC	1.41	0.75	3.28	1.21	0.66	2.7	0.37	1.48
				MAE				
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	3.47	1.68	1.85	2.02	2.75	2.94	2.43	2.45
WEBS	0.35	5.91	2.54	0.54	0.68	1.51	1.4	1.85
WEBSEC	0.19	0.25	3.7	0.63	0.7	2.68	2.32	1.50
RYM	3.46	1.69	1.86	2.03	2.76	2.94	2.44	2.45
RYMS	4.64	4.69	4.65	3.13	0.75	5.59	0.78	3.46
RYMSEC	1.41	0.75	3.28	1.21	0.66	2.7	0.37	1.48
				MAPE				
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	0.63	0.15	0.17	0.18	0.23	0.24	0.21	0.26
WEBS	0.06	0.55	0.23	0.04	0.05	0.12	0.12	0.17
WEBSEC	0.04	0.24	0.344	0.06	0.06	0.22	0.2	0.17
RYM	0.63	0.16	0.17	0.18	0.23	0.24	0.21	0.26
RYMS	0.85	0.46	0.43	0.28	0.06	0.47	0.06	0.37
RYMSEC	0.26	0.07	0.3	0.11	0.05	0.22	0.03	0.15

Table 4. Performance results of different WSP models for West Virginia Hills.

RMSE								
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	0.05	1.28	1.77	2.24	2.49	2.48	2.54	1.84
WEBS	0.94	3.03	5.56	0.2	0.1	2.47	2.9	2.17
WEBSEC	2.07	0.26	1.39	0.41	1.84	1.69	1.44	1.30
RYM	0.05	1.28	1.76	2.24	2.48	2.47	2.54	1.83
RYMS	0.91	1.63	0.02	4.29	8.11	2.8	0.6	2.62
RYMSEC	0.2	2.29	1.65	0.06	0.07	1.06	0.59	0.85
MAE								
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	0.05	1.28	1.77	2.22	2.49	2.48	2.54	1.83
WEBS	0.94	3.03	5.56	0.2	0.1	2.47	2.9	2.17
WEBSEC	2.07	0.26	1.39	0.411	0.84	1.69	1.44	1.16
RYM	0.05	1.28	1.76	2.24	2.48	2.47	2.54	1.83
RYMS	0.91	1.63	0.29	4.29	8.11	2.8	0.6	2.66
RYMSEC	0.2	2.29	1.65	0.06	0.07	1.06	0.59	0.85
MAPE								
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	0.007	0.15	0.19	0.23	0.25	0.25	0.26	0.19
WEBS	0.13	0.35	0.61	0.02	0.01	0.25	0.29	0.24
WEBSEC	0.28	0.03	0.15	0.04	0.18	0.17	0.14	0.14
RYM	0.008	0.14	0.19	0.23	0.25	0.25	0.25	0.19
RYMS	0.12	0.19	0.003	0.45	0.82	0.28	0.06	0.27
RYMSEC	0.02	0.26	0.18	0.006	0.007	0.1	0.06	0.09

Table 5. Performance results of different WSP models for Arizona Hills region.

RMSE								
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	4.42	3.007	3.92	3.9	2.6	1.75	2.07	3.10
WEBS	1.5	1.6	2.2	1.5	2.4	1.7	2.2	1.87
WEBSEC	1.16	1.69	1.23	1.29	0.38	1.66	0.35	1.11
RYM	4.42	3.007	3.92	3.9	2.6	1.7	2.07	3.09
RYMS	0.53	2.04	1.03	7.06	1.35	1.09	2.84	2.28
RYMSEC	0.36	1.66	1.97	0.23	0.6	1.64	0.77	1.03
MAE								
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	4.42	3.007	3.92	3.9	2.6	1.75	2.07	3.10
WEBS	1.15	1.6	2.2	1.5	2.4	1.7	2.2	1.82
WEBSEC	0.17	0.7	1.2	1.3	0.4	0.7	0.4	0.70
RYM	4.42	3.007	3.92	3.9	2.6	1.75	2.07	3.10
RYMS	0.53	2.04	1.03	7.06	1.35	1.09	2.84	2.28
RYMSEC	0.36	1.66	1.97	0.23	0.606	1.64	0.77	1.03
MAPE								
Model	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**t = 7**	**Average**
WEB	2.06	0.84	1.47	1.46	0.65	0.36	0.46	1.04
WEBS	0.72	0.45	0.84	0.54	0.6	0.35	0.5	0.57
WEBSEC	0.54	0.47	0.46	0.48	0.09	0.3	0.07	0.34
RYM	2.06	0.84	1.47	1.46	0.65	0.36	0.46	1.04
RYMS	0.24	0.57	0.39	2.64	0.3	0.22	0.63	0.71
RYMSEC	0.16	0.46	0.74	0.08	0.15	0.34	0.17	0.30

The results from Tables 2–5 imply that for all the regions, the error-corrected simulated probability-distribution-based model is more accurate than the general probability-distribution-based model and the simulated probability-distribution-based model. In the South Plains region of Texas, the error-corrected model can improve the general Weibull-based model to achieve a MAPE as low as 15%, RMSE as low as 1.5, and MAE as low as 1.5. In the Southern Texas Offshore region, the error-corrected model can improve the general Weibull-based model to achieve a MAPE as low as 17%, RMSE as low as 1.49, and MAE as low as 1.5. In the West Virginia Hills region, the error-corrected model can improve the general Rayleigh-based model to achieve a MAPE as low as 9%, RMSE as low as 0.85, and MAE as low as 0.85. Similarly, in the Arizona Hills region, the error-corrected model can improve the general Rayleigh-based model to achieve a MAPE as low as 30%, RMSE as low as 1.03, and MAE as low as 1.03.

The results show that for the South Plains Texas region and Southern Offshore Texas region, the Weibull-based model gives better results. In contrast, for the West Virginia Hills and Arizona Hills region, the Rayleigh-based model gives a better result. Therefore, it can be concluded that the wind distribution of the first two regions, i.e., South Plains TX and Southern Offshore region, can be more accurately described using the Weibull probability density function than the Rayleigh probability distribution function. In contrast, for the last two regions, i.e., West Virginia Hills and Arizona Hills region, wind distribution can be accurately described using the Rayleigh probability distribution function. Hence, the wind speed prediction model based on the probability distribution function also depends on how accurately that model describes the region's wind speed.

4.3. Comparative Analysis of Univariate Models

In this section, short-term forecasting was performed for six hours forecast horizon using seven different models based on the persistence model, classical time series model, and machine learning models. The results of short-term forecasting for four different case study sites are presented in Tables 6–9. The LSTM model emerged as the clear winner for the short-term wind speed prediction for all four case study sites. The LSTM model can produce a result with a MAPE as low as 3.53%, MAE as low as 0.4, and RMSE as low as 0.51. WEBSEC is also competitive with the LSTM model with a MAPE as low as 9.82, MAE as low as 1.09, and RMSE as low as 1.26. However, for site IV, whose wind speed distribution is not well described by the Weibull probability distribution function, the WEBSEC also does not give a good result. The SVR model is also competitive compared to the LSTM and WEBSEC models.

4.4. Development of Univariate Hybrid Model

The hybrid model based on a univariate wind forecasting model using persistence, classical time series, and machine learning models was developed after analyzing the performance of the model for predicting each time step of the total forecast horizon. For this, wind data of four case study sites were analyzed. After analyzing the performance of the individual model, a liner optimization was performed to minimize the performance metrics values to obtain thevweight function for each model contributing to the hybrid wind speed prediction model. The test data were divided again into the test and train data to evaluate the performance of the individual model and thus determine the weight function on the forecasted value on the test data. Then, the weight function was applied to the wind speed prediction value on the training dataset, thus determining the wind speed prediction value from the hybrid model. The performance results of the individual model of each time step of the forecast horizon are presented in Tables 6–9.

Table 6. Performance results of different univariate WSP models for South Plains TX region.

				RMSE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	0.68	0.98	1.28	0.47	0.14	2.88	1.07
WEBSEC	0.05	2.28	0.53	0.97	1.2	0.97	1.00
ARIMA	0.13	0.44	0.81	1.12	1.33	1.45	0.88
SVR	2.49	2.97	2.97	3.29	2.03	1.49	2.54
LSTM	0.05	2.28	0.53	0.97	1.2	0.97	1.00
				MAE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	0.68	0.98	1.28	0.474	0.146	2.88	1.07
WEBSEC	0.39	2.3	0.78	1.52	2.37	2.83	1.70
ARIMA	0.2	0.488	0.56	1.67	2.51	0.4	0.97
SVR	0.2	0.38	0.008	1.21	0.57	2.99	0.89
LSTM	0.39	2.35	0.78	1.52	2.37	2.83	1.71
				MAPE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	6.75	9.94	13.38	4.56	1.32	36.17	12.02
WEBSEC	3.92	23.52	0.08	0.14	0.21	0.35	4.70
ARIMA	2.004	4.93	5.89	16.16	22.81	5.05	9.47
SVR	1.9	3.8	0.9	11.7	5.2	37.57	10.18
LSTM	3.9	23.52	8.22	14.7	21.56	0.35	12.04

Table 7. Performance results of different univariate WSP models for Southern Offshore TX region.

				RMSE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	2.65	2.65	2.65	2.65	2.65	2.65	2.65
WEBSEC	0.13	0.45	2.81	3.75	1.05	0.58	1.46
ARIMA	3.27	3.17	2.97	2.75	2.6	2.48	2.87
SVR	4.08	2.85	1.44	0.56	0.07	0.38	1.56
LSTM	0.03	0.04	0.06	0.07	0.09	0.1	0.07
				MAE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	0.29	1.63	2.59	3.27	3.813	4.3	2.65
WEBSEC	2.21	1.47	2.75	4.37	2.21	1.06	2.35
ARIMA	0.92	2.15	2.91	3.37	3.76	4.13	2.87
SVR	0.94	1.04	0.59	0.39	0.44	0.47	0.65
LSTM	2.3	0.96	0.01	0.69	1.2	1.75	1.15
				MAPE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	3.14	20.12	36.23	50.55	64.35	79.1	42.25
WEBSEC	23.45	18.13	38.54	67.64	37.43	19.64	34.14
ARIMA	9.7	26.5	40.7	52.16	63.59	76.03	44.78
SVR	10.05	12.83	8.31	6.1	7.5	8.7	8.92
LSTM	24.5	11.9	0.014	10.8	21.2	32.2	16.77

Table 8. Performance results of different univariate WSP models for West Virginia Hills region.

				RMSE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	2.23	2.23	2.23	2.23	2.23	2.23	2.23
WEBSEC	4.07	0.31	2.61	1.78	1.91	0.11	1.80
ARIMA	1.9	1.5	1.2	0.96	0.77	0.61	1.16
SVR	0.58	1.96	3.39	3.49	3.009	2.7	2.52
LSTM	0.22	0.19	0.16	0.14	0.13	0.12	0.16
				MAE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	0.99	2.29	2.73	2.6	2.35	2.34	2.22
WEBSEC	2.84	0.25	3.18	2.16	1.79	0.003	1.70
ARIMA	0.7	1.64	1.79	1.33	0.89	0.72	1.18
SVR	0.77	0.68	0.24	0.53	0.3	0.04	0.43
LSTM	1.007	0.25	0.72	0.51	0.25	0.23	0.49
				MAPE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	11.64	31.53	41.17	37.31	32.53	32.37	31.09
WEBSEC	33.16	3.46	46.89	30.99	24.84	0.05	23.23
ARIMA	8.2	22.5	26.4	19.1	12.3	10.007	16.42
SVR	8.9	9.4	3.6	7.6	4.1	0.6	5.70
LSTM	11.7	3.5	10.7	7.4	3.4	3.2	6.65

Table 9. Performance results of different univariate WSP models for Arizona Hills region.

				RMSE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	0.79	0.79	0.79	0.79	0.79	0.79	0.79
WEBSEC	2.23	0.44	0.28	2.01	0.59	2.29	1.31
ARIMA	1.008	1.23	1.42	1.59	1.7	1.87	1.47
SVR	0.018	0.24	1.001	0.15	0.26	4.01	0.95
LSTM	2.08	2.05	2.03	2.06	2.02	2.01	2.04
				MAE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	0.22	0.89	0.421	0.08	3.81	2.38	1.30
WEBSEC	1.21	2.13	0.93	1.3	3.61	3.89	2.18
ARIMA	0.009	0.45	0.2	0.88	4.76	3.47	1.63
SVR	0.18	0.61	0.6	0.26	3.56	1.59	1.13
LSTM	3.1	3.73	3.24	2.73	0.99	0.41	2.37
				MAPE			
Models	**t = 1**	**t = 2**	**t = 3**	**t = 4**	**t = 5**	**t = 6**	**Average**
PM	4.7	16.4	8.4	1.9	52.6	11.1	15.85
WEBSEC	25.5	39.2	18.7	29.3	49.9	18.1	30.12
ARIMA	0.2	8.3	4.2	20	65.7	16.19	19.10
SVR	3.8	11.3	12.2	6.05	49.2	74.3	26.14
LSTM	65.1	68.8	65.5	61.3	137.8	19.5	69.67

The observation of the performance of the individual model for each time step of forecast horizons provides insights into the weight function that needs to be assigned to the individual model for each time step. Clearly, an individual model is not the best fit for all the time steps of the given forecast horizon. The $t = 1$ and $t = 2$ persistence model and ARIMA model work well for short-term wind speed prediction. The Weibull-based model is worse by a noticeable margin for this time step. However, the Weibull-based

model holds steady performance as forecasting horizon projects into the future, while the performance of all other models deteriorates. The Weibull-based model is best for $t = 6$. The persistence model is worse when the time steps of forecast horizons increase. The two machine learning models, SVR and LSTM, perform similarly; however, SVR performance is better by a small margin, making it difficult to choose one over the other for a given time step of a forecast horizon. This analysis suggests that the persistence model is best for wind speed prediction one hour or two hours ahead, the Weibull-based model is best for six hours ahead or longer forecast horizons, and ARIMA or machine learning models are best for in between these two forecast horizons. However, in this research, we have proposed to use all these models by providing weight functions for each model in each time step. The results after providing the weight function for the hybrid model are shown below in Tables 10–13.

Table 10. Performance results of different univariate hybrid WSP models for South Plains TX region.

	t = 1	t = 2	t = 3	t = 4	t = 5	t = 6	Average
RMSE	0.913	1.4775	1.2345	1.59	1.335	1.1225	1.27875
MAE	0.478	1.0877	0.557	1.4275	2.0095	2.492	1.34195
MAPE	4.7416	10.9635	4.4375	10.893	7.6055	8.429	7.845016667

Table 11. Performance results of different univariate hybrid WSP models for Southern Offshore region TX.

	t = 1	t = 2	t = 3	t = 4	t = 5	t = 6	Average
RMSE	2.4385	2.2565	1.812	1.7275	1.0715	0.8665	1.695416667
MAE	0.8795	1.511	1.557	2.15	2.10565	1.6975	1.650108333
MAPE	9.38	18.651	21.7755	33.3205	35.776	31.3145	25.03625

Table 12. Performance results of different univariate hybrid WSP models for West Virginia Hills region.

	t = 1	t = 2	t = 3	t = 4	t = 5	t = 6	Average
RMSE	1.7235	1.678	1.821	1.615	1.65285	0.681	1.528558333
MAE	1.00855	1.543	1.46	1.1545	1.2285	0.267	1.110258333
MAPE	11.798	21.248	21.6115	16.5885	17.0165	3.71455	15.32950833

Table 13. Performance results of different univariate hybrid WSP models for Arizona Hills region.

	t = 1	t = 2	t = 3	t = 4	t = 5	t = 6	Average
RMSE	0.9724	0.945	1.20825	1.3915	0.9315	2.368	1.302775
MAE	0.66385	1.27	1.21705	1.2315	3.392	2.8845	1.776483333
MAPE	13.99	23.42	24.635	27.7925	65.485	26.1035	30.23766667

Results from our four case study sites indicate performance improvement for each time step of the forecast horizon as well as overall accuracy. The hybrid model can produce a better result with a MAPE as low as 7.8%, MAE as low as 1.1, and RMSE as low as 1.2.

5. Conclusions

In this research, the performance of wind speed prediction models including persistence-forecasting-based models, classical time-series-based models, and machine-learning-based models for short-term forecast horizons was analyzed using four different case study sites. The results indicated the need for a hybrid wind speed prediction model that can incorporate all three types of wind speed prediction models to accurately predict wind speed for short-term forecast horizons. The proposed hybrid wind speed prediction was

developed using multivariate time series data, together with all of the other three models. This research highlights the competitive performance of the probability-density-based wind speed prediction method; the probability distribution of wind speed, which best describes the wind speed distribution of a given location, is also the best model for the wind speed prediction of the given location. Moreover, a novel method of error correction for wind speed forecasting based on the Weibull-distribution-based WSP model was proposed; this error correction can forecast wind speed accurately with MAPE of 4.7%, MAE of 1.7, and RMSE of 1, which is comparable to the best model out of the five models studied, i.e., MAPE of 5.7%, MAE of 0.43, and RMSE of 1.52. The wind speed of a region can be simulated based on the Weibull distribution parameters.

After analyzing the model's performance in predicting each time step of the whole forecast horizon, a hybrid model based on univariate wind forecasting was developed by incorporating persistence, traditional time series, and machine learning. The weight function that needs to be allocated to each model for each time step was determined by observing how well each model performed for each time step of the forecast window.

This study shows the competitive performance of the univariate model, which can be used where only univariate data are available for wind speed prediction. In this analysis, we suggested that the persistence model is the most accurate for predicting wind speed one or two hours in the future, a Weibull-based model for forecast horizons six hours or longer in the future, and an ARIMA or machine learning model is a good choice for forecast horizons in between two and six hours. As a result, a univariate model based on weight functions performs better and gives more weight to the most accurate approach for each time step. As this method used five different models for wind speed prediction, it might be more time-consuming when compared to a single model.

Author Contributions: Conceptualization, R.D., S.P. (Suhas Pol), S.P. (Siva Parameswaran) and H.M.; methodology, R.D., A.N. and A.S.; software, R.D.; resources, R.D., S.P. (Suhas Pol) and S.P. (Siva Parameswaran); writing—original draft preparation, R.D.; writing—review and editing, R.D., A.N. and H.M.; supervision, H.M. All authors have read and agreed to the published version of the manuscript.

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

References

1. Sinsel, S.R.; Riemke, R.L.; Hoffmann, V.H. Challenges and solution technologies for the integration of variable renewable energy sources—A review. *Renew. Energy* **2020**, *145*, 2271–2285. [CrossRef]
2. Wang, R.; Li, W.; Bagen, B. Development of wind speed forecasting model based on the Weibull probability distribution. In Proceedings of the 2011 International Conference on Computer Distributed Control and Intelligent Environmental Monitoring, Changsha, China, 19–20 February 2011; pp. 2062–2065.
3. Soman, S.S.; Zareipour, H.; Malik, O.; Mandal, P. A review of wind power and wind speed forecasting methods with different time horizons. In Proceedings of the North American Power Symposium 2010, Arlington, TX, USA, 26–28 September 2010; pp. 1–8.
4. Zhou, J.; Shi, J.; Li, G. Fine tuning support vector machines for short-term wind speed forecasting. *Energy Convers. Manag.* **2011**, *52*, 1990–1998. [CrossRef]
5. Keles, D.; Scelle, J.; Paraschiv, F.; Fichtner, W. Extended forecast methods for day-ahead electricity spot prices applying artificial neural networks. *Appl. Energy* **2016**, *162*, 218–230. [CrossRef]
6. Cervantes, J.; Dai, T.; Qiao, W. Optimal wind power penetration in the real-time energy market operation. In Proceedings of the 2013 IEEE Power & Energy Society General Meeting, Vancouver, BC, Canada, 21–25 July 2013; pp. 1–5.
7. Abdel-Aal, R.E.; Elhadidy, M.A.; Shaahid, S. Modeling and forecasting the mean hourly wind speed time series using GMDH-based abductive networks. *Renew. Energy* **2009**, *34*, 1686–1699. [CrossRef]
8. Chen, N.; Qian, Z.; Nabney, I.T.; Meng, X. Wind power forecasts using Gaussian processes and numerical weather prediction. *IEEE Trans. Power Syst.* **2013**, *29*, 656–665. [CrossRef]

9. Nair, K.R.; Vanitha, V.; Jisma, M. Forecasting of wind speed using ANN, ARIMA and Hybrid models. In Proceedings of the 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), Kerala, India, 6–7 July 2017; pp. 170–175.

10. Filik, Ü.B.; Filik, T. Wind speed prediction using artificial neural networks based on multiple local measurements in Eskisehir. *Energy Procedia* **2017**, *107*, 264–269. [CrossRef]

11. AKINCI, T.Ç.; NOĞAY, H.S. Application of decision tree methods for wind speed estimation. *Eur. J. Tech.* **2019**, *9*, 74–83. [CrossRef]

12. Irfan, A.; Bhuiyan, N.H.; Hasan, M.; Khan, M.M. Performance Analysis of Machine Learning Techniques for Wind Speed Prediction. In Proceedings of the 2021 12th International Conference on Computing Communication and Networking Technologies (ICCCNT), West Bengal, India, 6–8 July 2021; pp. 1–6.

13. Zhao, P.; Xia, J.; Dai, Y.; He, J. Wind speed prediction using support vector regression. In Proceedings of the 2010 5th IEEE Conference on Industrial Electronics and Applications, Taichung, Taiwan, 15–17 June 2010; pp. 882–886.

14. Kong, X.; Liu, X.; Shi, R.; Lee, K.Y. Wind speed prediction using reduced support vector machines with feature selection. *Neurocomputing* **2015**, *169*, 449–456. [CrossRef]

15. Heidari, A.; Navimipour, N.J.; Unal, M. Applications of ML/DL in the management of smart cities and societies based on new trends in information technologies: A systematic literature review. *Sustain. Cities Soc.* **2022**, *85*, 104089. [CrossRef]

16. Heidari, A.; Jafari Navimipour, N.; Unal, M.; Toumaj, S. Machine learning applications for COVID-19 outbreak management. *Neural Comput. Appl.* **2022**, *34*, 1–36. [CrossRef]

17. Duan, J.; Zuo, H.; Bai, Y.; Duan, J.; Chang, M.; Chen, B. Short-term wind speed forecasting using recurrent neural networks with error correction. *Energy* **2021**, *217*, 119397. [CrossRef]

18. Cao, Q.; Ewing, B.T.; Thompson, M.A. Forecasting wind speed with recurrent neural networks. *Eur. J. Oper. Res.* **2012**, *221*, 148–154. [CrossRef]

19. Heidari, A.; Navimipour, N.J.; Unal, M.; Toumaj, S. The COVID-19 epidemic analysis and diagnosis using deep learning: A systematic literature review and future directions. *Comput. Biol. Med.* **2021**, *141*, 105141. [CrossRef] [PubMed]

20. Kadhem, A.A.; Wahab, N.I.A.; Aris, I.; Jasni, J.; Abdalla, A.N. Advanced wind speed prediction model based on a combination of weibull distribution and an artificial neural network. *Energies* **2017**, *10*, 1744. [CrossRef]

21. Kaplan, O.; Temiz, M. A novel method based on Weibull distribution for short-term wind speed prediction. *Int. J. Hydrog Energy* **2017**, *42*, 17793–17800. [CrossRef]

22. Ding, Y. *Data Science for Wind Energy*; CRC Press: Boca Raton, FL, USA, 2019.

23. 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]

24. Zhu, X.; Genton, M.G. Short-term wind speed forecasting for power system operations. *Int. Stat. Rev.* **2012**, *80*, 2–23. [CrossRef]

25. Dhakal, R.; Yadav, B.K.; Koirala, N.; Kumal, B.B.; Moussa, H. Feasibility study of distributed wind energy generation in Jumla Nepal. *Int. J. Renew. Energy Res.* **2020**, *10*, 1501–1513.

26. Stevens, M.; Smulders, P. The estimation of the parameters of the Weibull wind speed distribution for wind energy utilization purposes. *Wind Eng.* **1979**, 132–145.

27. Shaban, A.H.; Resen, A.K.; Bassil, N. Weibull parameters evaluation by different methods for windmills farms. *Energy Rep.* **2020**, *6*, 188–199. [CrossRef]

28. Rehman, S.; Mahbub Alam, A.; Meyer, J.P.; Al-Hadhrami, L. Wind speed characteristics and resource assessment using Weibull parameters. *Int. J. Green Energy* **2012**, *9*, 800–814. [CrossRef]

29. Cadenas, E.; Rivera, W.; Campos-Amezcua, R.; Heard, C. Wind speed prediction using a univariate ARIMA model and a multivariate NARX model. *Energies* **2016**, *9*, 109. [CrossRef]

30. Brown, B.G.; Katz, R.W.; Murphy, A.H. Time series models to simulate and forecast wind speed and wind power. *J. Appl. Meteorol. Climatol.* **1984**, *23*, 1184–1195. [CrossRef]

31. Yatiyana, E.; Rajakaruna, S.; Ghosh, A. Wind speed and direction forecasting for wind power generation using ARIMA model. In Proceedings of the 2017 Australasian Universities Power Engineering Conference (AUPEC), Melbourne, VIC, Australia, 19–22 November 2017; pp. 1–6.

32. Singh, S.; Mohapatra, A. Repeated wavelet transform based ARIMA model for very short-term wind speed forecasting. *Renew. Energy* **2019**, *136*, 758–768.

33. Drucker, H.; Burges, C.J.; Kaufman, L.; Smola, A.; Vapnik, V. Support vector regression machines. *Adv. Neural Inf. Process. Syst.* **1996**, *9*, 155–161.

34. Santamaría-Bonfil, G.; Reyes-Ballesteros, A.; Gershenson, C. Wind speed forecasting for wind farms: A method based on support vector regression. *Renew. Energy* **2016**, *85*, 790–809. [CrossRef]

35. Shahani, N.M.; Zheng, X.; Guo, X.; Wei, X. Machine Learning-Based Intelligent Prediction of Elastic Modulus of Rocks at Thar Coalfield. *Sustainability* **2022**, *14*, 3689. [CrossRef]

36. Santamaría-Bonfil, G.; Frausto-Solís, J.; Vázquez-Rodarte, I. Volatility forecasting using support vector regression and a hybrid genetic algorithm. *Comput. Econ.* **2015**, *45*, 111–133. [CrossRef]

37. Guyon, I.; Elisseeff, A. An introduction to variable and feature selection. *J. Mach. Learn. Res.* **2003**, *3*, 1157–1182.

38. Schölkopf, B.; Smola, A.J.; Bach, F. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, USA, 2002.

39. Bengio, Y.; Simard, P.; Frasconi, P. Learning long-term dependencies with gradient descent is difficult. *IEEE Trans. Neural Netw.* **1994**, *5*, 157–166. [CrossRef]

40. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.

41. Jaseena, K.; Kovoor, B.C. Decomposition-based hybrid wind speed forecasting model using deep bidirectional LSTM networks. *Energy Convers. Manag.* **2021**, *234*, 113944. [CrossRef]

42. Shahani, N.M.; Kamran, M.; Zheng, X.; Liu, C. Predictive modeling of drilling rate index using machine learning approaches: LSTM, simple RNN, and RFA. *Pet. Sci. Technol.* **2022**, *40*, 534–555. [CrossRef]

43. Khatiwada, A.; Kadariya, P.; Agrahari, S.; Dhakal, R. Big Data Analytics and Deep Learning Based Sentiment Analysis System for Sales Prediction. In Proceedings of the 2019 IEEE Pune Section International Conference (PuneCon), Pune, India, 18–20 December 2019; pp. 1–6.

44. Bali, V.; Kumar, A.; Gangwar, S. Deep learning based wind speed forecasting-A review. In Proceedings of the 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence), Noida, India, 10–11 January 2019; pp. 426–431.

45. Hochreiter, S.; Schmidhuber, J. Long short-term memory. *Neural Comput.* **1997**, *9*, 1735–1780. [CrossRef] [PubMed]

46. Freeman, J.M.; DiOrio, N.A.; Blair, N.J.; Neises, T.W.; Wagner, M.J.; Gilman, P.; Janzou, S. *System Advisor Model (SAM) General Description (Version 2017.9. 5)*; National Renewable Energy Lab (NREL): Golden, CO, USA, 2018.

47. Willmott, C.J.; Matsuura, K. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Clim. Res.* **2005**, *30*, 79–82. [CrossRef]