Applied Energy 88 (2011) 1405–1414

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

ARMA based approaches for forecasting the tuple of wind speed and direction

Ergin Erdem, Jing Shi *

Department of Industrial and Manufacturing Engineering, North Dakota State University, Dept. 2485, PO Box 6050, Fargo, ND 58108, USA

ARTICLE INFO

Article history:
Received 16 May 2010
Received in revised form 15 October 2010
Accepted 24 October 2010

Keywords:
Combined forecasting
Wind speed
Wind direction
ARMA
Vector autoregression

ABSTRACT

Short-term forecasting of wind speed and direction is of great importance to wind turbine operation and efficient energy harvesting. In this study, the forecasting of wind speed and direction tuple is performed. Four approaches based on autoregressive moving average (ARMA) method are employed for this purpose. The first approach features the decomposition of the wind speed into lateral and longitudinal components. Each component is represented by an ARMA model, and the results are combined to obtain the wind direction and speed forecasts. The second approach employs two independent ARMA models – a traditional ARMA model for predicting wind speed and a linked ARMA model for wind direction. The third approach features vector autoregression (VAR) models to forecast the tuple of wind attributes. The fourth approach involves employing a restricted version of the VAR approach to predict the same. By employing these four approaches, the hourly mean wind attributes are forecasted 1-h ahead for two wind observation sites in North Dakota, USA. The results are compared using the mean absolute error (MAE) as a measure for forecasting quality. It is found that the component model is better at predicting the wind direction than the traditional-linked ARMA model, whereas the opposite is observed for wind speed forecasting. Utilizing VAR approaches rather than the univariate counterparts brings modest improvement in wind direction prediction but not in wind speed prediction. Between restricted and unrestricted versions of VAR models, there is little difference in terms of forecasting performance.

1. Introduction

Wind is a promising renewable energy source, and the energy generated from the wind has been growing at double digits in recent years. In 2009, the wind energy generated was 340 TWh worldwide which accounted for 2% of the worldwide electricity usage, and the capacity of the wind-powered generators reached 159.2 GW [1]. In the recent 3 years, the amount of energy produced from wind sources have been doubled, and this trend is expected to continue. In the US, the installed wind capacity was 11,575 MW at the end of 2006. In 2009, this figure was more than tripled and reached an installed capacity of 35,159 MW, which indicates an astonishing annual growth rate of 45% [1].

However, power generation from wind has some drawbacks. One major problem is that wind is an intermittent energy source which means that there exists large variability in the production of energy due to various factors, such as wind speed, air density, and turbine characteristics. Another problem is that wind is usually regarded as a non-dispatchable energy source. It is difficult to manage the energy production from wind based sources upon the demand. Usually, intermittency can be considered as a problem related with the dispatchability [2]. Wind speed plays an important role for the amount of electricity generated by the wind turbines. The theoretical amount of energy that might be generated from the wind is proportional to the cube of wind speed and slight changes in the wind speed might cause significant changes in the amount of the total electricity generated from the wind. On practical terms, the power output of wind turbine and the wind speed are linked with a power curve graph, which usually shows S-shape characteristics. Therefore, accurate forecasting the wind speed carries a particular importance. In literature, various researchers have used different techniques to forecast the wind speed, and those techniques can be classified into four different subcategories.

The first category incorporates physical properties such as terrain, obstacles, pressure and temperature to forecast the wind speed [3,4]. Usually, numerical weather prediction models (NWP) are used for the large scale weather prediction. To forecast the short-term wind speed at local sites, this approach might not give the reliable results. However, it can perform equally well or better with large forecasting horizons (e.g. more than 6 h), as compared with the other statistical models [5–7]. The second category approaches the forecasting problem using conventional statistical methods. Among the most notable approaches are the Auto Regressive Integrated Moving Average (ARIMA) based models introduced by Box–Jenkins [8]. To name a few here, Lalarukh and Yasmin incorporated autocorrelation, non-Gaussian distribution

* Corresponding author. Tel.: +1 701 231 7119; fax: +1 701 231 7195.
E-mail address: jing.shi@ndsu.edu (J. Shi).

0306-2619/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.apenergy.2010.10.031

and diurnal nonstationarity in ARIMA models to forecast the wind speed [9]. Brown et al. used square root transformation to obtain the wind speed distributions that are approximately Gaussian and established the link between the wind speed and power and analyzed the autocorrelations for the historical wind speed data [10]. Torres et al. compared the ARIMA based models with the persistence based models. It was found that the forecasts provided by the persistence model outperforms the ARIMA based models for shorter forecasting horizons, whereas for longer forecasting horizons, the ARIMA based models consistently perform better as compared to the persistence model [11]. The third category is the methods from the artificial intelligence and machine learning fields. For instance, Li and Shi compared three different artificial neural networks (ANNs) to compare the wind speed, and found that the selection of neural network type and parameters greatly affects the performance of wind speed forecasting [12]. Mabel and Fernandez used the wind speed, relative humidity and generation hours as input variables to train an ANN-based model, and obtained fairly satisfactory results [13]. Other methods adopted in literature include fuzzy logic, support vector machine, and some hybrid methods are employed to forecast wind speed for different applications [5,14,15]. The fourth category for the wind speed forecasting is the spatial–temporal models. Unlike the statistical methods that only take the time series data in consideration, as the name implies, the spatial models take the location information into consideration and predict the wind speed based on the speed information at the neighboring sites. One main assumption behind spatial models is that the wind speeds are correlated in the neighboring sites, and various researchers have employed spatial models to forecast the wind speed [12,16]. Morales et al. developed a joint methodology for constructing a spatial–temporal model to forecast the wind speed [17]. Xydis et al. developed a prognostic model and employed spatial correlation along with the exergy (i.e., measure of the maximum useful work that can be done by a system) analysis to estimate the wind potential for a particular region and predict the total amount of energy that can be obtained from that site [18].

Meanwhile, wind direction is an important factor for efficient turbine control for getting the most power with a given wind speed. Based on the forecasts on wind direction, it would be possible to align the turbine with the wind direction to get the most energy output. However, research on forecasting wind direction is far less common, and the related publications for energy conversion purpose are even scarcer. Zhang et al. took a Bayesian approach to develop models for predicting wind speed, wind direction, and ambient temperature, and combined them to assess the online conductor thermal overload risk [19]. Daniele et al. incorporated the forecast of the wind direction in their model to develop volcanic ash tracking model [20]. In a similar manner, wind speed and direction are also included in the hierarchical Bayesian based models model to predict the ozone levels [21]. Wind direction is also incorporated as a factor for forecasting the pollutant level [22]. In addition to that, research is conducted on characterizing wind direction along with the wind speed for long-term assessment of wind power generation. Burlando et al. employed statistical analysis for this purpose and presented the basic statistical measures to assess the wind potential. The wind potential was assessed by combining long term statistical analysis of wind speed and direction along with the numerical modeling of three dimensional wind flows over a complex terrain. The estimates were corrected using few available and reliable data measures [23]. Bao et al. developed a circular regression based approach along with Bayesian averaging method for bias correction of the forecasts obtained by numerical weather prediction models [24]. Kalsuner et al. developed an approach for predicting wind vector based on the "similar days" approach, in which the wind pattern is compared with the historical data by using set of criteria defined for assessing similarity [25]. In addition, Potter and Negnevitsky employed an adaptive neuro-fuzzy inference system to forecast the wind vector [26].

Since both predictions of wind speed and direction are critical for efficient wind energy harvesting, it is reasonable to assume that forecasting the wind speed in connection with the wind direction will differ from forecasting these two attributes independently. However, little can be found in the literature on how to forecast multiple wind attributes simultaneously. This research is carried out to fill the gap. In this paper, we use four different approaches to forecast the wind speed and direction. The first approach uses the component model where the prevailing wind direction is obtained and the wind speed is decomposed into lateral and longitudinal components based on the prevailing wind direction. Decomposing the wind speed into the lateral and longitudinal components has been used for characterization of the joint distribution of wind speed and direction by the researchers [27,28]. The second approach employs two independent ARMA models, namely, a traditional ARMA model for forecasting the wind speed, and a linked ARMA model for wind direction. Linked ARMA models are used for wind speed in that they have been proven effective to forecast the directional variables [29]. The third approach involves developing vector autoregressive models (i.e., VAR) that enable to combine the wind speed and direction by considering the correlation between the two attributes. VAR models have been used by various researchers to determine the business cycles, interest rates, and on similar research on economics [30,31]. The fourth approach is derived from third approach where some parameter values of the VAR model are restricted to have the value of 0. This approach is proposed in the literature to avoid the over-parameterized models [32,33].

The remainder of this paper is organized as follows. In the next section, we discuss about the methodology for the component approach, the traditional-linked ARMA model, and the vector AR approaches. In addition, forecasting performance will also be discussed. In the third section, the discussion focuses on the procedure to obtain the model parameters. In the fourth section, the forecasting results are presented, and discussion is provided on the results. In the last section, main findings are summarized.

2. Methodology

In this section, the basic concept of ARMA model is introduced first. We then furnish the mathematical models associated with the four different approaches, namely the component approach, the independent approach based on ARMA models, and the combined approach based on VAR models, for forecasting wind speed and wind direction.

2.1. Conventional ARMA model

A typical ARMA model which is denoted by ARMA(p, q) can be expressed as follows [34],

$$y_t = \delta + \sum_{i=1}^{p} \varphi_i y_{t-i} + \sum_{j=1}^{q} \phi_j e_{t-j} + e_t, \tag{1}$$

where δ is the constant term of the ARMA model, φ_i is the ith autoregressive coefficient, ϕ_j is the jth moving average coefficient, e_t is the error term at time period t, and represents the value of wind speed observed or forecasted at time period t.

The first step for establishing the ARMA model is identification. In order to identify and develop the model, the stationarity assumption should be checked. For this purpose, inspection of the run plots and Auto Correlation Function (ACF) plots can be used for deciding on the order of differencing. The autoregressive and moving average terms can be decided using the ACF and partial autocorrelation function (PACF) plots [35,36]. In addition, different tests such as Durbin–Watson, Dickey–Fuller, Augmented Dickey–Fuller, and root examination for univariate time series can also be employed [37–39].

If differencing is used, the model is converted to ARIMA model. The ACF and PACF graphs have been used for deciding on the autoregressive and moving average terms in our model. For the diagnostic checking properties, ACF and PACF graphs of the residuals are analyzed. In addition to that, Ljung–Box Q test can also be applied to check whether the residuals are randomly distributed or not. The test statistic can be expressed as follows [40],

$$Q_{LB} = (n)(n+2)\sum_{j=1}^{h}\frac{\rho^2(j)}{n-j}, \tag{2}$$

where n is the sample size, $\rho(j)$ is the autocorrelation at lag j, and h is the number of lags that is being tested. The null hypothesis of randomness is rejected, if $Q_{LB} > \chi^2_{1-\alpha;h}$, where χ^2 is the percent point function of the chi-square distribution.

2.2. Component approach

In the component model, the first step is to find out the prevailing wind direction. Based on the prevailing wind direction, the lateral and longitudinal components of the wind speed can be calculated. In order to find the prevailing wind direction, the following approach is utilized [41],

$$\bar{\theta} = \left\{\begin{array}{ll} \tan^{-1}(S/C) & S>0, C>0 \\ \tan^{-1}(S/C)+\pi & C<0 \\ \tan^{-1}(S/C)+2\pi & S<0, C>0 \end{array}\right\}, \tag{3}$$

where $\bar{\theta}$ is the mean direction. S is the summation of the sine values of the angles of wind vectors with respect to the north axis (i.e., $S=\sum_{i=1}^{n}\sin\theta_i$), and C is the summation of the cosine values of the angles of wind vectors with respect to the north axis (i.e., $C=\sum_{i=1}^{n}\cos\theta_i$).

After finding the mean direction, next step is to decompose the wind velocity vectors into lateral and longitudinal components. In order to do so, the following formulas are used.

$$v_{y'} = v\cos(\theta-\bar{\theta}), \tag{4}$$

$$v_{x'} = v\sin(\theta-\bar{\theta}), \tag{5}$$

where $v_{y'}$ is the longitudinal component of the wind speed, is the lateral component of the wind speed, θ is the angle of wind vector with respect to north axis, and $\bar{\theta}$ is the mean direction.

Using a single prevailing wind direction would help simplify the underlying component model thus resulting in more parsimonious models. Several researchers construct the component models based on the concept of a single prevailing wind direction [41,42]. One reason for decomposing the wind vector into the lateral and longitudinal components is to reduce the interaction between these two components. Decomposing the wind speed into two orthogonal components based on the prevailing wind direction could help to build two separate ARMA models to represent the lateral and longitudinal components. After finding the lateral and longitudinal components, the next step is the employing ARMA model for the forecast. After ARMA model is constructed and verified, forecasts are obtained for the lateral and longitudinal components respectively. Based on the forecasted components, the corresponding wind direction can be found as follows,

$$\theta_t = \tan^{-1}\left(\frac{v_{x'}}{v_{y'}}\right)+\bar{\theta}, \tag{6}$$

where θ_t is the forecasted wind direction for time period t. The forecast for the corresponding wind speed can be found by the following equation,

$$v = \sqrt{(v_{x'}^2 + v_{y'}^2)}, \tag{7}$$

using this approach, it is possible to obtain the forecasts for wind speed and wind direction simultaneously.

2.3. Traditional-linked ARMA modeling approach

In this approach, we consider two independent ARMA models for forecasting the wind speed and direction. For wind speed component, we employ the traditional ARMA model as described above. For the wind direction, we will adopt a link function. The purpose of the link function is to convert a linear variable to a circular variable by mapping the real line to the circle, and usually the monotone link functions $g(u)$, which range from $-\pi$ to π with the change of the linear variable u from $-\infty$ to ∞, are selected for this purpose. Using the inverse of a link function, we can convert the circular variable to a linear variable, and then apply the traditional ARMA methods to forecast the linear variable. After forecasting the linear variable, it is possible to convert it back to the circular variable by using the link function. By employing such a procedure, the forecasts on the circular variable can be obtained.

In the literature, there are several functions that might be used for converting a circular variable to a linear variable. Among those functions, the link function in the following form is adopted [43],

$$g(u) = 2\pi\left(\Phi(u)-\frac{1}{2}\right)+\mu, \tag{8}$$

where $\Phi(u)$ is the cumulative distribution function for standardized Gaussian distribution with mean and variance being 0 and 1 respectively. The parameter μ is the angle for mean wind direction in radians. The inverse of the link function presented in Eq. (8) is expressed as follows,

$$g^{-1}(x) = \Phi^{-1}\left((x-\mu)/2\pi+\frac{1}{2}\right), \tag{9}$$

where Φ^{-1} is the probit function (i.e., the inverse of the standardized cumulative distribution function).

The circular variable (i.e., the wind direction) is converted using the inverse link function expressed in the above equation. Using the inverse link functions, the wind direction is converted to the linear variable. ARMA model is employed to forecast the linear variable. After obtaining the linear variable, using the link function, it is converted back to the circular variable, and forecasts on wind direction are obtained.

2.4. Vector ARMA approach

The Vector ARMA (VARMA) is a multivariate version of the ARMA model. In this research, we combine the wind attributes (speed and direction) to forecast the future values of wind attributes. For this purpose, we use the Vector Auto Regressive (i.e., VAR approach), where we perform regression on the previous values of wind attributes to predict the future values. The general equation for the VARMA model can be expressed as follows [34],

$$\Phi(B)Y_t = \delta + \Theta(B)\varepsilon_t, \tag{10}$$

where B denotes the backward shift operator, $\Phi(B) = I - \Phi_1 B^1 - \Phi_2 B^2 - \cdots \Phi_p B^p$ for the autoregressive term of order p, $\Theta(B) = I - \Theta_1 B^1 - \Theta_2 B^2 - \cdots \Theta_q B^q$ for the moving average term of order q, Y_t is column vector of the realization of the stochastic process at time t, ε_t is the column vector of disturbance (i.e. white noise) pertaining to period t, and δ is the column vector of the stochastic process of size n by 1. In addition, Φ_p is the autoregressive coefficient matrix with a size of n by n that belongs to lag p, where n is the number of variables included in the model, Θ_q is the moving average coefficient matrix of lag q of the same size, and I is the identity matrix. Based on the general equation for VARMA model, VAR model can be expressed in a simplified form,

$$\Phi(B)Y_t = \delta + \varepsilon_t. \tag{11}$$

VAR model is chosen in this study for two reasons. First, the persistence models that forecast the wind direction on the previous values are generally used as benchmarking purposes, and generally these models can provide satisfactory results for short-term forecasting. VAR models are the extension of persistence models where the wind directions are also included. Another reason is that for both the traditional/linked and component models, the specified ARMA models obtained (as discussed in a later section) involve the autoregressive terms. In order to delve into exploring the power of autoregressive models for prediction purposes, an approach employing the multivariate version of autoregressive model is thus developed.

A variation of VAR model is based on the restricting some parameters that are not statistically significant to 0. As previously discussed, this approach is suggested to overcome the potential problem of over-parameterization. This approach is called restricted VAR model, which is also included in the analysis and the accompanying results are discussed.

2.5. Forecast performance measure

In this study, we consider mean absolute error (MAE) for assessing the quality of the forecast. MAE can be considered as the mean of absolute values of the forecast errors and can be expressed with the following formula,

$$\mathrm{MAE} = \frac{\sum_{t=1}^{N}|y_t - y'_t|}{N}, \tag{12}$$

where y_t is the realized value at time period t, y'_t is the forecast pertaining to that period, and N is the number of data points.

The above approach is valid for the linear variables. However, in the case of the circular variables, a slightly different approach should be taken to calculate the corresponding MAE value for the circular variable. The formula that can provide the MAE values for the circular variables can be expressed as follows [44]

$$\mathrm{MAE} = \left\{ \begin{array}{ll} \frac{\sum_{t=1}^{N}(|\theta_t-\theta'_t|)}{N} & 0 < |\theta_t - \theta'_t| < \pi, \\ \frac{\sum_{t=1}^{N}(2\pi-(|\theta_t-\theta'_t|))}{N} & \pi < |\theta_t - \theta'_t| < 2\pi \end{array} \right\}. \tag{13}$$

3. Model building and parameter estimation

3.1. Data source

By employing anemometers and wind direction sensors, the wind speed and direction data were measured and captured at a wind observation site in North Dakota, an upper Midwest state in the United States with abundant wind resources. The data were recorded continuously and averaged over every hour to obtain the wind attributes. In the study, the hourly wind dataset of May 1–October 21, 2002 is adopted. Basically, the forecasting models are constructed by using the first 3936 hourly records of wind dataset, and 1-h ahead forecasts are made for the remaining 240 hourly measurements. Fig. 1 depicts the run plot for the wind speed for the period of May 1–October 11, 2002 for this particular site.

3.2. Component model

In order to obtain the parameters for the component model, the first step is to find the prevailing wind direction. The prevailing wind direction is calculated based on Eq. (3). It turns out that the wind site has a prevailing wind direction of 210.66° based on the dataset collected. Based on this information, the lateral and longitudinal components are obtained. Table 1 provides the correlation coefficients between the lateral and longitudinal components based on the axis formed at different angles with respect to the north axis. Based on the different angles formed with respect to the north axis, it can be seen that forming the axis at the prevailing wind direction yields the lowest correlation coefficient for the wind site.

It should be noted that although forming longitudinal and lateral component based on the prevailing wind direction could generate the minimum correlation coefficient in absolute figures for this case, this property may not be universal for all cases. As such, if any other direction would provide the lowest correlation coefficient, it might be worthy to build the lateral and longitudinal components based on that particular direction as well as the prevailing

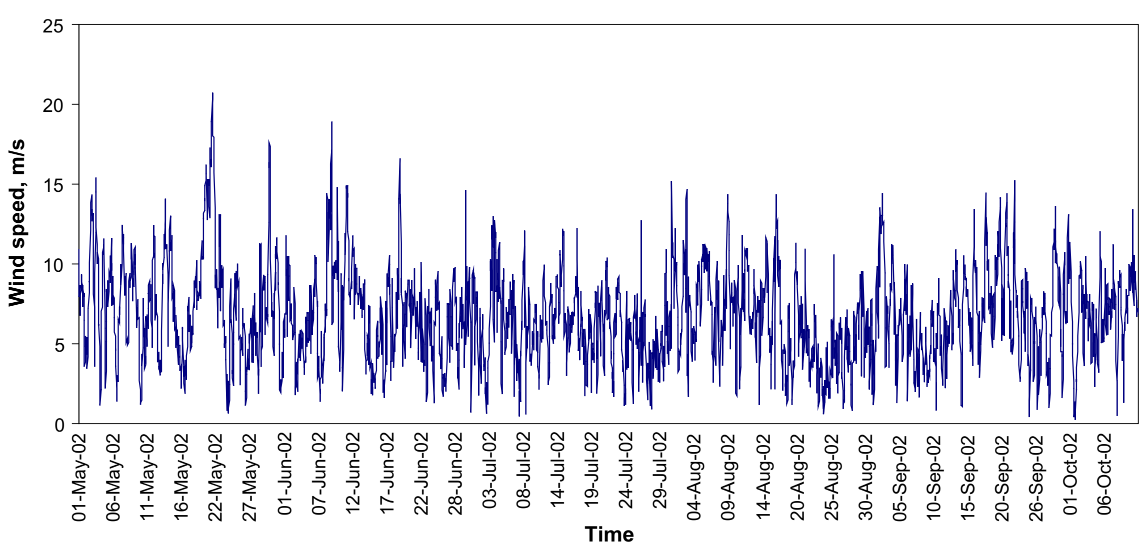

Fig. 1. Run plot for wind speed during May 1–October 11, 2002.

Table 1
Correlation coefficients between lateral and longitudinal components.

Angle of axis (°)	0	60	120	180	240	300	Prevailing direction
Correlation coefficient	−0.229	0.208	0.023	−0.229	0.208	0.096	−0.017

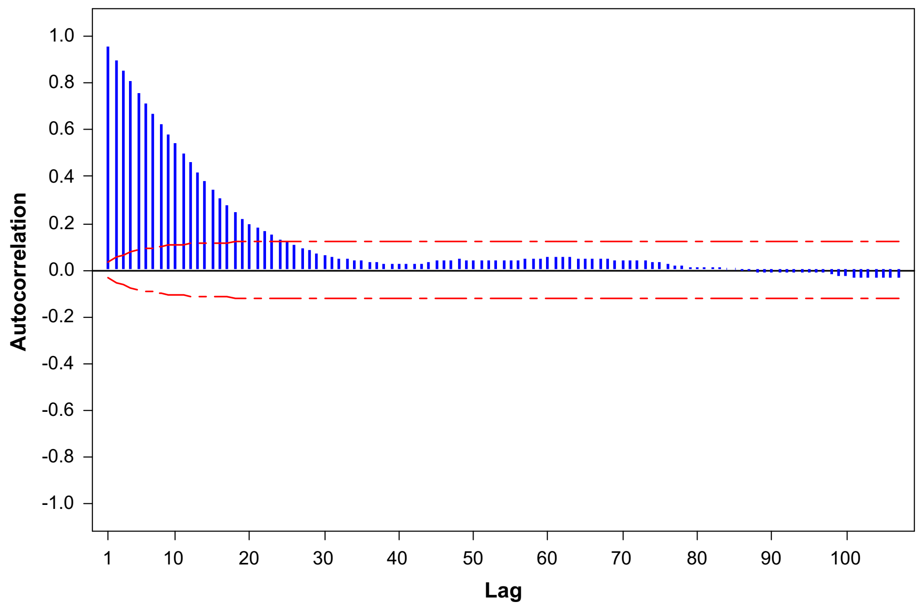

Fig. 2. ACF graph for the lateral wind component.

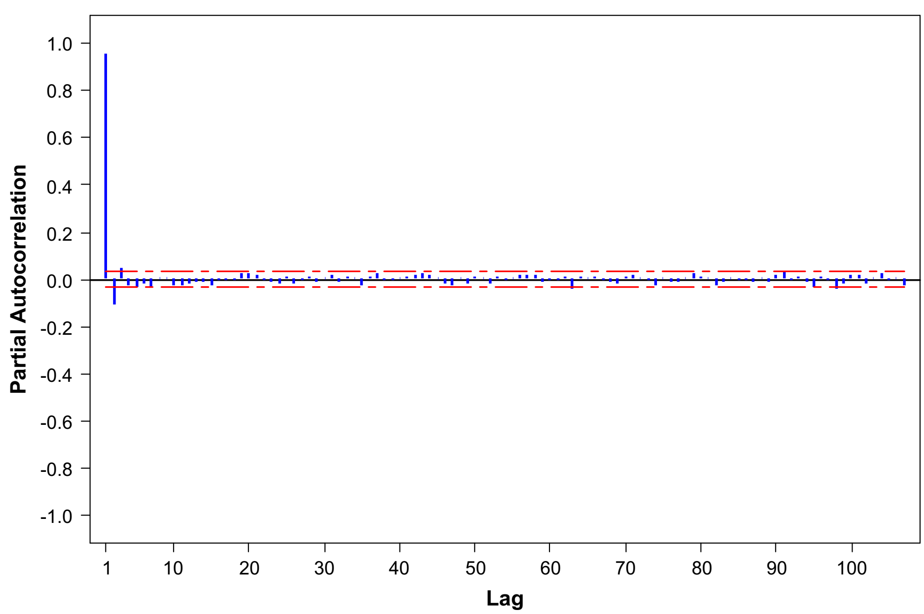

Fig. 3. PACF graph for the lateral wind component.

wind direction, and compare these two models in terms of the forecasting power.

The ACF and PACF graphs for the lateral component of wind vector are provided in Figs. 2 and 3. The ACF graph shows an exponential decaying pattern, while the PACF graph shows a cutoff value after lag 2 or probably lag 3. The exponential decay pattern is an indication of the autoregressive process with stationarity assumption provided. A visual inspection of the run order plot in Fig. 1 indicates that the data is stationary. Based on this assumption, ARMA(3, 0) model is considered for the lateral component of the wind speed data. A similar analysis is performed for the longitudinal component. For the purpose of brevity, the corresponding ACF and PACF graphs are not provided here. According to the patterns on ACF and PACF graphs, the autoregressive function with order 3 (i.e., ARMA(3, 0) model) is considered for the lateral components, and the ARMA(2, 0) model is considered to be valid for the longitudinal components.

For specification of the model, we employ the MINIC method in the PROC ARIMA module in SAS™ software. The MINIC method makes use of the Bayesian Information Criterion (BIC) to compute the MINIC table to specify the tentative orders for the autoregressive and moving average [45]. The combination yielding the smallest BIC values is selected. In this method, BIC values can be computed by,

$$\mathrm{BIC} = \ln(\hat{\sigma}_e^2) + \frac{2k \ln n}{n}, \tag{14}$$

where n is the number of points in the specified model, k is the number of estimated parameters, $\hat{\sigma}_e^2$ is the term for the error variance, which can be defined as,

$$\hat{\sigma}_e^2 = \frac{1}{n} \sum_{t=t_0}^{n} (x_i - \hat{x}_i)^2, \tag{15}$$

where x_i is the actual observation and $\hat{x}_i$ is the value obtained by the specified ARMA model, t is the data index. The parameter t_0 consists of two components. The first component is the order of the underlying autoregressive model for approximating error terms and the second component is the maximum of the autoregressive and moving average component of the specified ARMA model whose BIC value is calculated.

Table 2 illustrates the corresponding BIC values for the various ARMA models for the lateral component with autoregressive and moving average orders varying from 0 to 5. It can be seen that the minimum BIC value is obtained with ARMA(3, 0) model. This also supports our conclusion that a model with the autoregressive order of 3 with no moving average component fits our data for explaining the underlying stochastic pattern.

The model parameters are provided in Table 3. The ARMA model parameters are obtained through the non-linear least squares estimation. The procedure aims to minimize the sum of square of the residuals using a backward approach, where the values in time series are reversed and parameters are obtained based on the residuals of the fitted model [34]. The Ljung–Box test statistics can verify the obtained models provide a fairly good fit for the selected data set. For instance, the Ljung–Box test statistics for the lateral component are provided in Table 4. It can be seen that the Ljung–Box test statistics are smaller than corresponding critical χ^2 values at the significance level of 0.05, and thus we can conclude that residuals of the fitted model are not auto correlated based on the selected number of the lags.

Table 2
BIC values for various ARMA models for lateral components.

Model order	MA(0)	MA(1)	MA(2)	MA(3)	MA(4)	MA(5)
AR(0)	3.5228	3.4188	3.3250	3.2302	3.1277	3.0233
AR(1)	1.1143	1.1038	1.1038	1.1058	1.1064	1.1082
AR(2)	1.1045	1.1050	1.1055	1.1067	1.1078	1.1098
AR(3)	1.1037	1.1058	1.1074	1.1085	1.1098	1.1119
AR(4)	1.1052	1.1061	1.1081	1.1098	1.1118	1.1138
AR(5)	1.1061	1.1078	1.1097	1.1118	1.1139	1.1159

Table 3
ARMA model parameters for component model.

	Model	φ_1	φ_2	φ_3
Lateral	ARMA(3, 0)	1.0608	−0.1612	0.0523
Longitudinal	ARMA(2, 0)	1.0302	−0.112	

Table 4
Ljung–Box test statistics for the fitted ARMA(3, 0) model of the lateral component.

Lag	12	24	36	48
df	9	21	33	45
χ^2	11.48	32.16	44.89	55.37
p-value	0.2445	0.0564	0.0811	0.1384

3.3. Traditional/linked ARMA models

The second approach uses the traditional ARMA model for forecasting wind speed, and the linked ARMA model for forecasting wind direction. Our analysis starts with checking the independence assumption between the wind speed and wind direction. Rather than employing the traditional formula for calculating correlation between the two variables, we adopt the approach utilized by Mardia [46]. For calculating the correlation coefficient, we employ the following equation,

$$r^2 = \frac{r_{vc}^2 + r_{vs}^2 - 2r_{vc}r_{vs}r_{cs}}{1 - r_{cs}^2}, \tag{16}$$

where $r_{vc} = corr(v, cos\theta)$, $r_{vs} = corr(v, sin\theta)$, and $r_{cs} = corr(cos\theta, sin\theta)$ respectively. For this particular site, the obtained correlation coefficient is 0.0946, which is relatively small and indicates only slight statistical dependence between the two variables.

Similar to the procedure taken in the component model, we employ the traditional ARMA model for forecasting the wind speed. However, unlike the previous case, rather than using the component model to decompose the lateral and longitudinal components, we directly employ the ARMA approach to come up with the forecasts of wind speed. By following the same procedure (i.e., the ACF and PACF plots, and BIC tables), ARMA(3, 0) is obtained as a valid model. For the linked ARMA model, the inverse of a link function is used so that the circular variable can be converted to the linear variable. Then the traditional ARMA method is implemented to obtain the forecasted values on the linear variable. Using the link function in Eq. (8), it is possible to convert the linear variable back to the circular variable again and obtain the forecasts. Again, by following the same procedure, we conclude that ARMA(3, 0) model is suitable for forecasting the wind direction. The parameters of the ARMA models for the linearized wind direction based on inverse link functions and wind speed are provided in the Table 5.

3.4. VAR model

As mentioned above, the correlation coefficient between the wind speed and direction is 0.0946. This value indicates there might be slight correlation between the two attributes. As such, it is worthwhile to combine these two variables under the same model for forecasting of the tuples of wind speed and direction. For this reason, VAR based models can be developed which assume statistical dependency between the selected attributes. In a sense, VAR models share similar properties with the component models where the wind direction and wind speed information are used in conjunction for forecasting purposes. However, the main difference is that, the wind direction is incorporated implicitly in the

Table 5
ARMA model parameters for wind speed and direction.

Attribute	Model	δ	φ_1	φ_2	φ_3
Wind speed	ARMA(3, 0)	0.6877	1.0276	−0.2031	0.0756
Wind direction	ARMA(3, 0)		0.4655	0.1491	0.1607

component models, whereas in the VAR based models, the inverse link function is used to convert the circular wind direction variable to a linear variable. The linearized variable is then used explicitly for establishing the VAR model.

One major aspect of building the VAR models is to decide on the value of p (i.e., the order of the autoregressive terms). For this purpose, we mainly employ different p values associated with the VAR model, and compare the BIC values similar to the procedure conducted in Section 3.2. After building the model, various multivariate and univariate diagnostic tests, such as the portmanteau and Durbin–Watson tests, are employed for model validation.

We establish five different values for the number of autoregressive terms and compare those VAR models respectively (i.e., VAR(1), VAR(2), VAR(3), VAR(4), and VAR(5)) based on the BIC values. We set the limit for the number of autoregressive terms to five, based on the discussion provided in previous section that models involving autoregressive orders not exceeding three are employed for wind speed and direction separately. Therefore, leaving a cushion for safety, we limit the number of autoregressive term as five. Table 6 provides the corresponding BIC values for the different VAR models. It can be seen that VAR(3) model has the lowest BIC values, and thus it is chosen.

Using SAS 9™ software, we obtain the parameter values as shown in Table 7 for the autoregressive coefficients and constant terms. We can also express the model parameters in the above table as the elements of the Φ_1, Φ_2, Φ_3, and δ matrices that are provided in Eq. (11) where Φ_1, Φ_2, Φ_3 are 2×2 matrices that provide the model coefficients for the autocorrelation terms and δ is a 2×1 matrix representing the constant terms in our case. To cite an instance, the coefficient -0.0439 (the coefficient indicating the effect of wind direction observed two periods before on the current wind speed) corresponds to the element located at the second row and first column of the Φ_2 matrix (i.e., $\varphi_2(2,1) = -0.0439$). The model coefficient parameters and corresponding significance levels can be located in Table 7.

Table 6
Corresponding BIC values of VAR models.

VAR model	BIC
VAR(1)	−0.24566
VAR(2)	−0.30854
VAR(3)	−0.33331
VAR(4)	−0.32728
VAR(5)	−0.31987

Table 7
Coefficient values for VAR(3) model.

Lag	Wind attribute	Speed	Direction
1	Speed	1.0255 (<0.0001)	0.0026 (0.761)
	Direction	−0.0776 (0.0001)	0.4645 (<0.0001)
2	Speed	−0.2057 (<0.0001)	0.001 (0.9355)
	Direction	−0.0439 (0.1714)	0.1495 (<0.0001)
3	Speed	0.0792 (0.0001)	0.0049 (0.5649)
	Direction	0.0175 (0.5505)	0.1629 (<0.0001)
Constant term		0.68254 (<0.0001)	−0.0604 (0.0314)

Table 8
Coefficient values for restricted VAR(3) model.

Lag	Wind attribute	Speed	Direction
1	Speed	1.0263 (<0.0001)	0
	Direction	−0.0964 (<0.0001)	0.4646 (<0.0001)
2	Speed	−0.2061 (<0.0001)	0.0078 (0.0319)
	Direction	0	0.1491 (<0.0001)
3	Speed	0.0791 (<0.0001)	0
	Direction	0	0.1631 (<0.0001)
Constant term		0.6814 (<0.0001)	−0.0556 (0.0383)

The portmanteau test is also conducted to test the statistical significance of the cross correlations of the residuals. The portmanteau test indicates that the corresponding significance level for the residuals up to lags 4 and 12 are 0.105 and 0.05 respectively which indicates that there is no severe cross correlation problem among residuals and the stated model would be able to identify the underlying stochastic process. Therefore, we cannot reject the hypothesis that the residuals are not cross-correlated. Similarly, the Durbin–Watson test as a univariate test for checking the autocorrelation of the residuals among individual wind attributes indicates that the residuals are not also correlated. For wind speed, the corresponding Durbin–Watson test statistic is 1.999, and for wind direction, this value is 2.001. The fact that those two values are close to 2.0 leads us to conclude that the residuals are not correlated with each other for individual wind attributes [47].

Examining Table 7 reveals that for some of the parameter values, the corresponding significance level (i.e., the numbers in brackets) are greater than 0.05. For example, the term, $\varphi_2(2,1)$ has a value of -0.0439, with a significance level of 0.1714 showing that that the effect of wind direction observed two periods ago on the current wind direction might not play a significant role and statistically insignificant at 5% significance level. The non-significant terms is especially encountered for the terms describing the effect of a particular wind attribute on the other one (i.e., the effect of wind speed on wind direction or vice versa). The non-significant terms deteriorate the parsimoniousness of the model and lead to the over-parameterization. In order to handle this problem, we propose a variation of the VAR model, where the coefficients of the statistically non-significant parameters in the VAR model are restricted to be equal to 0. It is an iterative process where the non-significant terms are removed from the model and significant terms are added with the selected significance level of 5%. In a sense, this iterative process is a similar to the stepwise regression with the difference that more than one parameter might be included or excluded in the next iteration. Subsequent iterations are made until all the significant terms are included in the model, and all non-significant ones are excluded. The parameter values and their corresponding significance levels are obtained and summarized in Table 8.

4. Performance analysis and discussion

This section provides a general analysis on the forecasting performance of four approaches based on the results obtained for the wind observation site. To further analyze the observations, a summary of forecasting results for a second dataset is provided. In addition, the possible ways to extend the current research are discussed.

4.1. Forecasting performance

After obtaining the model parameters, the next step is to forecast the wind speed and wind direction based on the four different approaches. Fig. 4 provides the absolute difference between the observed values and forecasts based on the component, traditional-linked ARMA, VAR(3), and restricted VAR models for wind speed. Fig. 5 provides the absolute difference between the forecasted and observed values for those models for wind direction. The MAE figures for the wind speed and wind direction are pre-

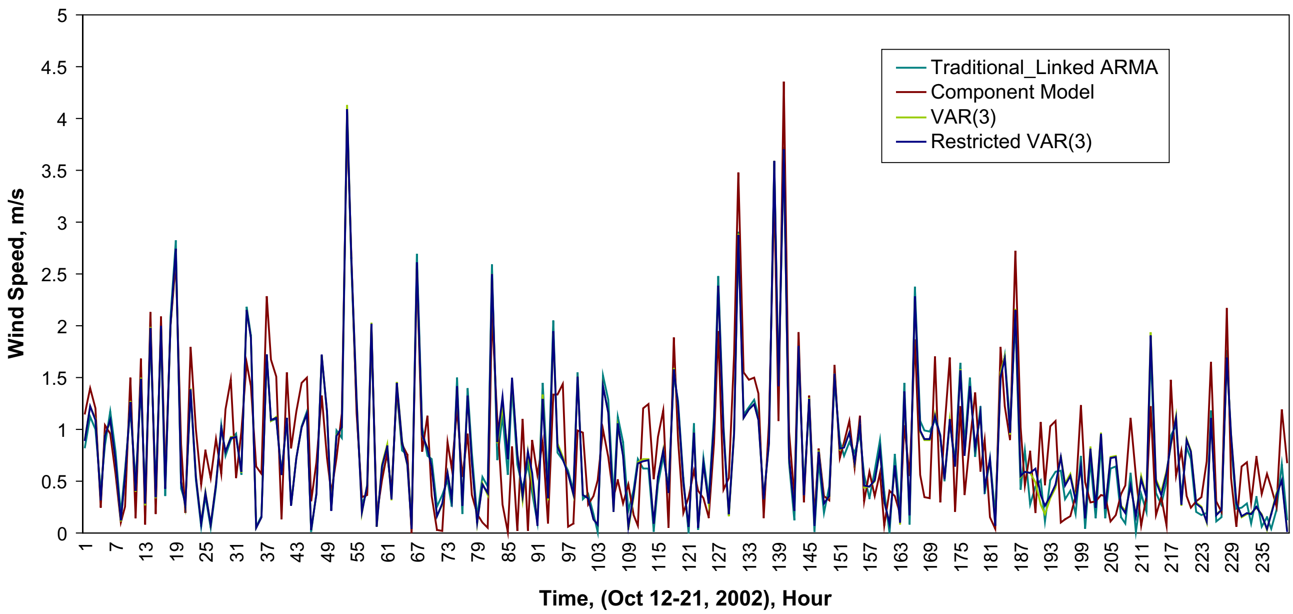

Fig. 4. Absolute forecast errors for wind speed.

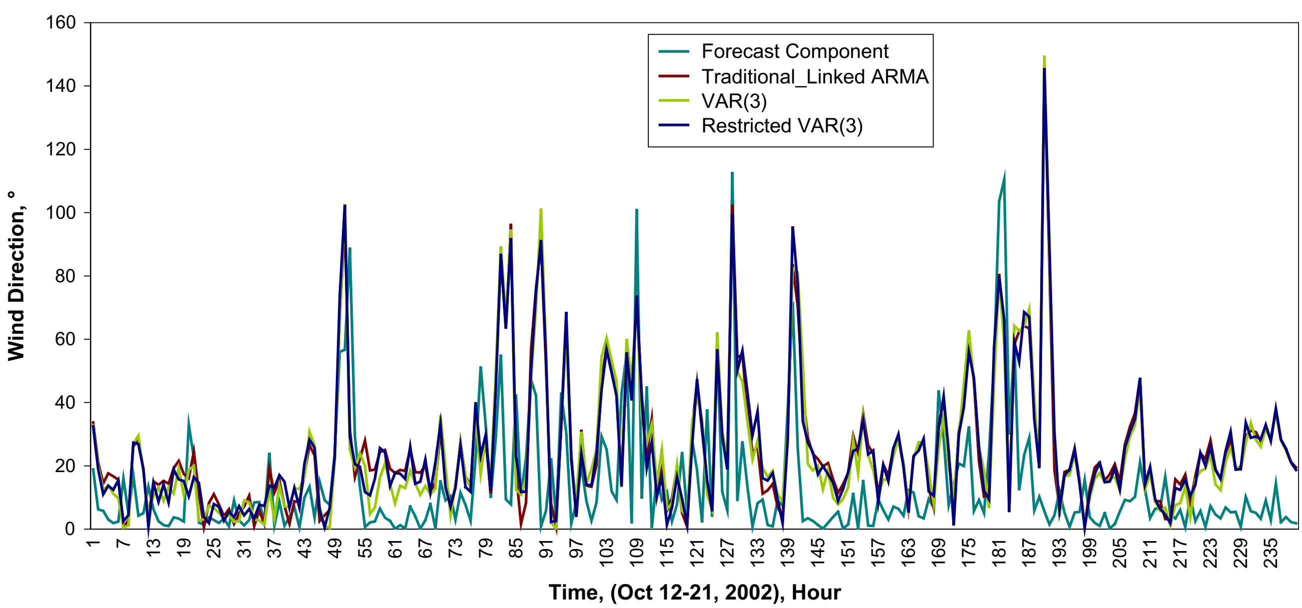

Fig. 5. Absolute forecast errors for wind direction.

sented in Table 8. By examining Figs. 4 and 5, and Table 9, it can be seen that with regard to the wind speed, the traditional-linked ARMA, VAR model, and restricted VAR model have almost identical forecasting performance. Their MAE values are very close, and their absolute error profiles as shown in Fig. 4 overlap in the entire forecast range. All three models clearly outperform the component model. Nevertheless, with respect to the wind direction, the component model appears to be better compared with the other three models, with its corresponding MAE value being about 50% of those of the other models, among which the traditional-linked ARMA model performs the worst. The results indicate that the models based on the link functions (i.e., traditional-linked ARMA, VAR, and restricted VAR models) may have difficulty in capturing the rapid changing wind attributes to a certain extent. This might be partly due to the fact that the employed link function might not be able to map the directional wind variable to the linear scale in an accurate and unbiased manner.

Table 9
MAE values of four ARMA based approaches.

Wind attribute	Component model	Traditional-linked ARMA	VAR(3)	Restricted VAR(3)
Speed (m/s)	0.8311	0.8045	0.8098	0.8094
Direction (°)	13.4702	25.6622	24.1575	25.0185

Meanwhile, it can be seen that the traditional-linked ARMA model performs negligibly (i.e. less than 1% in terms of MAE) better as compared to the VAR and restricted VAR models for forecasting the wind speed. This might be partly explained with the low correlation coefficient between the wind attributes, and the same underlying model specification (i.e., the autoregressive form) specified for both the traditional-linked ARMA and VAR based models.

Table 10
MAE values of four ARMA based approaches for the second site.

Wind attribute	Component model		Traditional-linked ARMA		VAR(3)	Restricted VAR(3)
	Lateral: ARMA(3, 0)	Longitudinal: ARMA(4, 1)	Speed: ARMA(2, 0)	Direction: ARMA(3, 1)		
Speed (m/s)	2.302		1.149		1.159	1.154
Direction (°)	28.309		37.861		19.149	21.081

Another reason for the similar performance might be associated with employing the same link function for both cases. In terms of forecasting wind direction, the VAR based models are better as compared to the univariate counterparts with the difference not exceeding 7%. This might be due to the fact that the same link function being employed for converting circular variable to the linear one, and similar model structure employed for multivariate and univariate cases as in the case of forecasting the wind speed. In addition, by comparing the VAR and restricted VAR models, it is seen that their forecasting performances are nearly identical in terms of the wind speed, while the VAR model performs slightly better (less than 3.5% in terms of MAE) in terms of the wind direction. However, the restricted VAR approach might be still preferred in that it leads to more parsimonious models with less number of parameters.

4.2. Further discussion and analysis

In an attempt to generalize the findings, we also conduct a similar analysis on another dataset from a different wind site. Due to the limited number of available observations, the dataset only covers the period of March 1, 2002–May 31, 2002, and forecasts are performed to predict the wind speed for 4 days (i.e., 96 data points). It spans a shorter time horizon for model building, but the results would be helpful for providing more insightful information. Basically, the same procedure is followed for developing the models for this new wind dataset. For the sake of brevity, the details on the model development, and graphs on forecast errors are not presented here. It should be noted that the correlation coefficient in Eq. (15) is computed to be 0.179 for this site, and the prevailing direction for component model is 301.29°. Also, for the VAR approaches, VAR(3) model is chosen. Table 10 presents the model structures and MAE figures obtained from the analysis of this wind site.

From Table 10, a similar phenomenon can be observed about the VAR and restricted VAR models. Both approaches have almost identical MAE values in predicting wind speed, while the VAR model appears to be slightly better than the restricted VAR model in predicting wind direction. Basically, this is consistent with the observation for the first site in presented in Table 9. More importantly, it can be found that the component model is not as impressive as in the previous case. Although its performance in wind direction is still significantly better than that of the traditional-linked ARMA model, it is clearly outperformed by the VAR based models for both wind speed and direction. The deterioration of performance might be attributed to the different volatility nature of the wind at this site and the fact that less number of data points is used for building the component models. In addition, the performance of the component model is worse as compared to other models for forecasting the wind speed. This is consistent with the findings obtained by analyzing the first case.

Another interesting observation is that the performance gap between traditional-linked ARMA and multivariate counterparts widens for predicting wind direction. The traditional-linked ARMA model is the best in terms of the forecasting wind speed, as in the first case as well. On the other hand, its MAE value of wind direction forecasting is nearly twice as much as those of VAR models. This might be associated with two different factors. The first factor is that, the correlation coefficient between the wind speed and wind direction for this new site during the period of observation is higher than that of the first case (i.e., 17.9% versus 9.46%). Employing the multivariate version instead of the univariate one might help capture the relation between these two wind attributes, thus lead to more accurate forecasts. Another reason might be the different model structures underlying the univariate and multivariate versions. As the name implies, the multivariate version is based on autoregressive process, however, for univariate version it is a mixture of autoregressive and moving average processes (e.g., ARMA(3, 1)). This argument could be also supported by the similar performance in forecasting the wind speed between the traditional-linked ARMA model and the VAR models. In this case, the underlying process for the wind speed in univariate model is ARMA(2, 0), a similar structure to the multivariate counterpart.

In the future, to improve the forecast quality of multivariate models, the general VARMA based models that consider both autoregression and moving average terms might be taken into account. Meanwhile, VAR and traditional-linked ARMA based models rely on the link function to convert the circular wind direction variable into a linear one. To improve the performance of these two models, other link functions can be tested. Also, the forecasting accuracy of the approaches based on the circular autoregressive models might be compared against the models using the link functions. In addition, 1-h ahead forecast might be appropriate for the systems that retrieve and analyze the information on hourly basis. However, in order to bring a more comprehensive approach for assessing the forecasting power of those models, the comparison based on longer forecasting horizons (e.g., 6 h) might be conducted.

Another potential research direction is to develop suitable forecasting approaches using the autoregressive processes for directional data. Rather than relying on the link functions to make the conversion between the circular and linear variables, an autoregressive process that depends on a suitable circular distribution can also be introduced for forecasting the wind direction. Autoregressive processes based on wrapped normal and von Mises distributions are proposed in the literature for forecasting the wind direction [48]. Apart from those two distributions, any circular distribution symmetric around it mean can be used, and the moving average terms can also be incorporated in the model [43]. Employing the autoregressive process around the circle might lead to powerful time series models for forecasting the wind direction [48]. Those approaches might be compared against the models developed in the current study.

5. Concluding remarks

In this study, four different approaches for forecasting wind speed and direction are proposed, and the performance of forecasts is evaluated using the metric of MAE. By analyzing the prediction results for two wind observation sites, the following points can be summarized. First, the component model is a better choice than the traditional-linked ARMA model for forecasting the wind direction, whereas the opposite holds between the two models for forecasting the wind speed. Second, compared with the traditional-linked ARMA model, the VAR models offer higher forecasting accuracy in wind direction and close performance in wind speed. Depending on the correlation between wind speed and direction, the VAR models can outperform the component model when the correlation is modestly significant. Thirdly, using restricted version of the VAR models would be a key for more parsimonious models but a suitable approach must be taken for model building for retaining the forecasting power compared to the non-restricted counterparts. Those findings could be significant for the research on short-term wind forecasting, and it will stimulate follow-up studies in the future.

References

[1] World Wind Energy Association. World Wind Energy Report; 2009. <http://www.wwindea.org/home/images/stories/worldwindenergyreport2009_s.pdf> [retrieved April 2010].
[2] Dale L, Milborrow D, Slark R, Strbac G. A shift to wind is not unfeasible (total cost estimates for large-scale wind scenarios in UK). Power UK 2003;109:17–25.
[3] Lei M, Jiang C, Liu H, Zhang Y. A review on forecasting wind data and wind output. Renew Sustain Energy Rev 2009;13(4):915–20.
[4] Giebel G, Brownsword R, Kariniotakis G. The state-of-the-art in short-term prediction of wind power: a literature overview. Project ANEMOS, Deliverable Report D1.1; 2003. <http://anemos.cma.fr/download/ANEMOS_D1.1_StateOfTheArt_v1.1.pdf> [retrieved April 2010].
[5] Landberg L. Short-term prediction of local wind conditions. J Wind Eng Ind Aerodyn 2001;89(3/4):235–45.
[6] Landberg L. Short-term prediction of the power production from wind farms. J Wind Eng Ind Aerodyn 1999;80:207–22.
[7] Landberg L, Watson SJ. Short-term prediction of local wind conditions. Bound Layer Meteorol 1994;70:171–95.
[8] Box GEP, Jenkins GM, Reinsel GC. Time series analysis, forecasting and control. 3rd ed. Englewood Cliffs, USA: Prentice-Hall; 1994.
[9] Lalarukh K, Yasmin ZJ. Time series models to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. Sol Energy 1997;61(1):23–32.
[10] Brown BG, Katz RW, Murphy AH. Time series models to simulate and forecast wind speed and power. J Climate Appl Meteorol 1984;23(11):1184–95.
[11] Torres JL, García A, De Blas M, De Francisco A. Forecast of hourly average wind speed with ARMA models in Navarre. Sol Energy 2005;79(1):65–77.
[12] Gong L, Shi J. On comparing three artificial neural networks for wind speed forecasting. Appl Energy 2010;87:2313–20.
[13] Mabel MC, Fernandez E. Analysis of wind power generation and prediction using ANN: a case study. Renew Energy 2008;33(5):986–92.
[14] Damousis IG, Alexiadis MC, Theocharis JB, Dokopoulos PS. A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation. IEEE Trans Energy Convers 2004;19(2):352–61.
[15] Thiaw L, Sow G, Fall SS, Kasse M, Sylla E, Thioye S. A neural network based approach for wind resource and wind generators production assessment. Appl Energy 2010;87(5):1744–8.
[16] Barbounis TG, Theocharis JB. A locally recurrent fuzzy neural network with application to the wind speed prediction using spatial correlation. Neurocomputing 2007;70(7/9):1525–42.
[17] Morales JM, Minguez R, Conejo AJ. A methodology to generate statistically dependent wind speed scenarios. Appl Energy 2010;87(3):843–55.
[18] Xydis G, Koroneos C, Loizidou M. Exergy analysis in a wind speed prognostic model as a wind farm sitting selection tool: a case study in Southern Greece. Appl Energy 2009;86(11):2411–20.
[19] Zhang J, Pu J. A Bayesian approach for short-term transmission line thermal overload risk assessment. IEEE Trans Power Deliv 2002;17:770–8.
[20] Daniele P, Lirer L, Petrosino P, Spinelli N, Peterson R. Applications of the PUFF Model to forecasts of volcanic clouds dispersal from Etna and Vesuvio. Comput Geosci 2009;35(5):1035–49.
[21] McMillan N, Bortnick SM, Irwin ME, Berliner LM. A hierarchical Bayesian model to estimate and forecast ozone through space and time. Atmos Environ 2005;39(8):1373–82.
[22] Finzi G, Fronza G, Rinaldi S. Stochastic modelling and forecast of the dosage area product. Atmos Environ 1978;12(4):831–8.
[23] Burlando MAP, Villa L, Ratto CF, Cassulo G. Preliminary estimate of the large-scale wind energy resource with few measurements available: the case of Montenegro. J Wind Eng Ind Aerodyn 2009;97(11–12):497–511.
[24] Bao L, Gneiting J, Grimit P. Bias correction and Bayesian model averaging for ensemble forecasts of surface wind direction. MonWeather Rev 2010;138(5):1811–21.
[25] Kalusner Z, Kaplan H, Fattal E. The similar days method for predicting near surface wind vectors. Meteorol Appl 2009;16(4):569–79.
[26] Potter CW. Negnevitsky M very short term wind forecasting for Tasmanian power generation. IEEE Trans Power Syst 2006;21(2):965–72.
[27] Carta JA, Ramirez P, Bueno C. A joint probability function of wind speed and direction for wind energy analysis. Energy Convers Manage 2008;49:1309–20.
[28] Erdem E, Shi J. Comparison of bivariate distribution construction approaches for analyzing wind speed and direction data. Wind Energy, in press. doi:10.1002/we.400.
[29] Fisher NI, Lee AJ. Time series analysis of circular data. J Roy Stat Soc Ser B (Methodology) 1994;56(2):327–39.
[30] Vereda L, Lopes H, Fukuda R. Estimating VAR models for the term structure of interest rates. Insur Math Econ 2008;42(2):548–59.
[31] Buckle RA, Kim K, Kirkham H, McLellan N, Sharma J. A structural VAR business cycle model for a volatile small open economy. Econ Model 2007;24(6):990–1017.
[32] Bruggemann R. Model reduction methods for vector autoregressive processes. 1st ed. Berlin–Heidelberg, Germany: Springer-Verlag; 2004.
[33] Robertson JC, Tallman EW. Improving federal-funds rate forecasts in VAR models used for policy analysis. J Bus Econ Stat 2001;19(3):324–30.
[34] Montgomery DC, Jennings CL, Kulahci M. Introduction to tine series analysis and forecasting. Hoboken, New Jersey, USA: John Wiley & Sons Inc.; 2008.
[35] Iqbal N, Baksh K, Maqbool A, Ahmad AS. Use of the ARIMA model for forecasting wheat area and production in Pakistan. J Agric Social Sci 2005;1(2):120–2.
[36] Kamal L, Jafri YZ. Time series model to simulate and forecast hourly averaged wind speed in Quetta, Pakistan. Sol Energy 1997;61(1):23–32.
[37] Asma M. Mean reversion of stock returns of an emerging market: empirical evidence from Botswana stock exchange. SSRN: working paper. <http://ssrn.com/abstract=1462088> [accessed April 2010].
[38] Fraser P, Taylor MP. An empirical examination of long-run purchasing power parity as theory of international commodity arbitrage. Appl Econ 1991;23(11):1749–59.
[39] Bohn H. The behavior of US public debt and deficits. Quart J Econ 1998;113(3):949–63.
[40] Ljung GM, Box GEP. On a measure of a lack of fit in time series models. Biometrika 1978;65:297–303.
[41] Carta JA, Ramírez P, Bueno C. A joint probability density function of wind speed and direction for wind energy analysis. Energy Convers Manage 2008;49(6):1309–20.
[42] McWilliams B, Newmann MM, Sprevak D. The probability distribution of wind velocity and direction. Wind Eng 1979;3:269–73.
[43] Fisher NI. Statistical analysis of circular data. 1st ed. Cambridge, UK: Cambridge University Press; 1993.
[44] Grimit EP, Gneiting T, Berrocal V, Johnson NA. The continuous rank probability score for circular variables. Technical Report No. 493. Department of Statistics, University of Washington; 2006.
[45] SAS Institute. SAS/ETS™ 9.2 user's guide. The MINIC method. <http://support.sas.com/documentation/cdl/en/etsug/60372/HTML/default/etsug_arima_sect030.htm> [retrieved August 2010].
[46] Mardia KV. Linear–circular correlation coefficients and rhythmometry. Biometrika 1976;63:403–5.
[47] Gujarati DN. Basic econometrics. 4th ed. Boston, USA: McGraw-Hill; 2003.
[48] Breckling J. The analysis of directional time series: applications to wind speed and direction. In: Beger J, Fienberg S, Gani J, Krickeberg K, Olkin I, Singer B, editors. Lecture notes in statistics. Berlin, Germany: Springer-Verlag; 1989.