



# Modeling directional distributions of wind data in the United Arab Emirates at different elevations

Aishah Al Yammahi<sup>1</sup> · Prashanth R. Marpu<sup>2</sup> · Taha B. M. J. Ouarda<sup>3</sup>

Received: 14 March 2020 / Accepted: 1 March 2021

© The Author(s) 2021

## Abstract

Modeling wind speed and direction are crucial in several applications such as the estimation of wind energy potential and the study of the long-term effects on engineering structures. While there have been several studies on modeling wind speed, studies on modeling wind direction are limited. In this work, we use a mixture of von Mises distributions to model wind direction. Finite mixtures of von Mises (FMVM) distributions are used to model wind directions at two sites in the United Arab Emirates. The parameters of the FMVM distribution are estimated using the least square method. The results of the research show that the FMVM is the best suited distribution model to fit wind direction at these two sites, compared to other distributions commonly used to model wind direction.

**Keywords** Circular statistics · Directional data · Finite mixture of von Mises distributions · Wind direction

## Introduction

Wind energy is an important source of renewable energy, and the global installation capacity has been growing every year. Understanding the wind energy potential is very important for the site selection for installations of wind turbines. The key to analyzing the potential is to estimate the probability density function of wind speed. Several methods can be used to estimate the probability distribution using statistical models such as Weibull, Rayleigh, lognormal, normal, and gamma (Ouarda et al., 2015; Qin et al., 2011; Ouarda et al., n.d.; Shin et al., 2016; Wang et al., 2016; Safari, 2011). Wind energy is not only affected by wind speed but also by many

other factors such as terrain, climate, temperature, friction, wind direction, and air density (Zhang et al., 2014; Zhang et al., 2013; Naizghi & Ouarda, 2017). While modeling wind speed is enough for turbines with a rotating nacelle, modeling wind direction becomes important for turbines with fixed nacelle. Moreover, modeling wind direction has several other applications in aerosol transport, power transmission, and engineering structure design (Jammalamadaka & Lund, 2006; Petersen et al., 2005; McLean Sloughter et al., 2013).

Wind direction plays a vital role in transporting pollutants, which can have serious implications on human health (Jammalamadaka & Lund, 2006). The wind parameters have a significant effect on the thermal rating of transmission lines, which influences the transmission capacity of the lines (Davis, 1977). Atmospheric flow in the vicinity of a large elliptical mountain is affected by upstream wind direction and orographic features. In mountain areas, it is important to understand the interaction between Coriolis force and the upstream wind direction, because this will affect atmospheric flow in the vicinity of a large elliptical mountain (Petersen et al., 2005).

Wind direction modeling is also important in applications related to navigation and safety (McLean Sloughter et al., 2013) and civil engineering projects such as towers and bridges (Masseran et al., 2013; Abohela et al., 2013). The analysis of circular data, such as wind direction, is also common in several other fields. For example, in biology, circular data are present in animal navigation and movement directions (Mardia & Jupp, 2009). Spherical data is evident in the earth

---

Responsible Editor: Zhihua Zhang

✉ Aishah Al Yammahi  
aishah.alymahi@ku.ac.ae

Prashanth R. Marpu  
prashanth.marpu@g42.ai

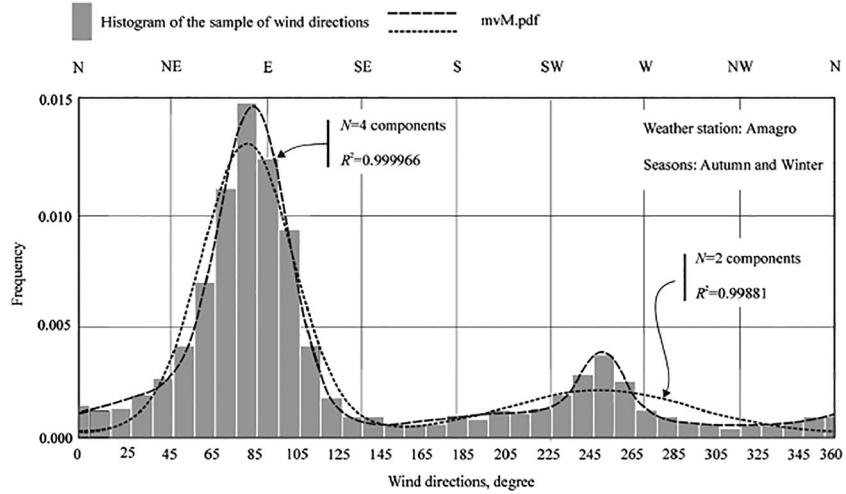
Taha B. M. J. Ouarda  
Taha.Ouarda@ete.inrs.ca

<sup>1</sup> Khalifa University, Abu Dhabi, United Arab Emirates

<sup>2</sup> Group 42, Abu Dhabi, United Arab Emirates

<sup>3</sup> Institut National de la Recherche Scientifique, INRS-ETE, Québec, Québec, Canada

**Fig. 1** Flow chart of hypothesis testing coefficient of determination R<sup>2</sup> for mixture of von Mises distribution

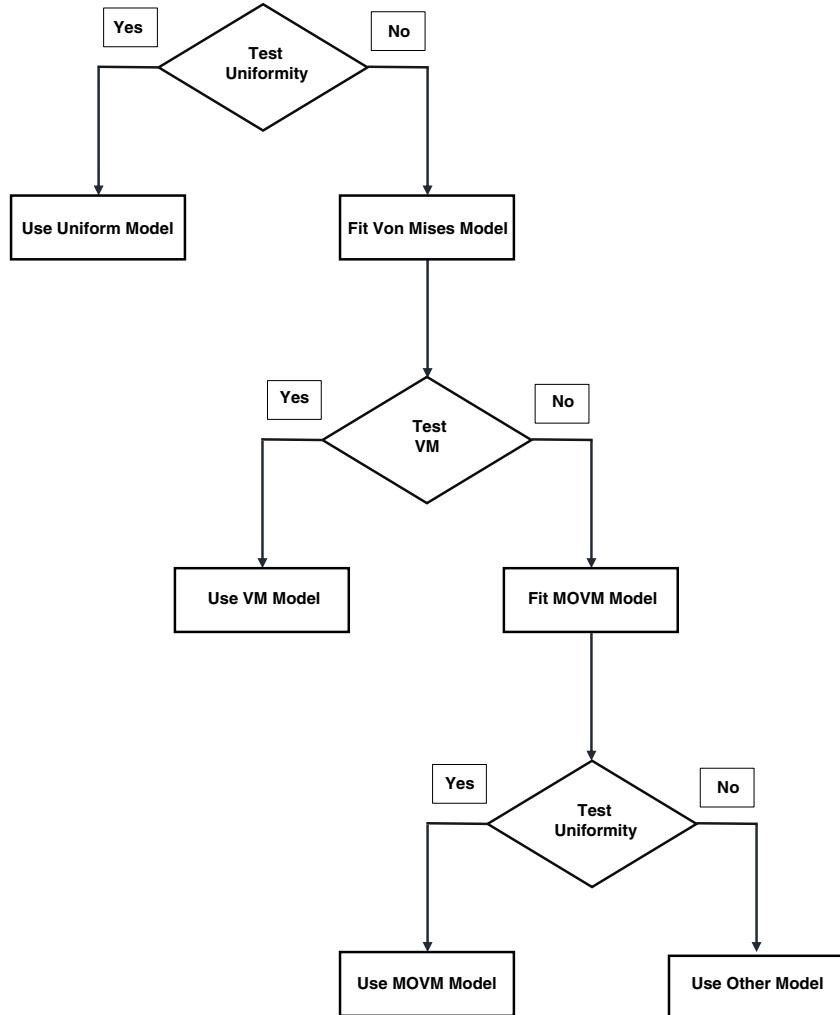


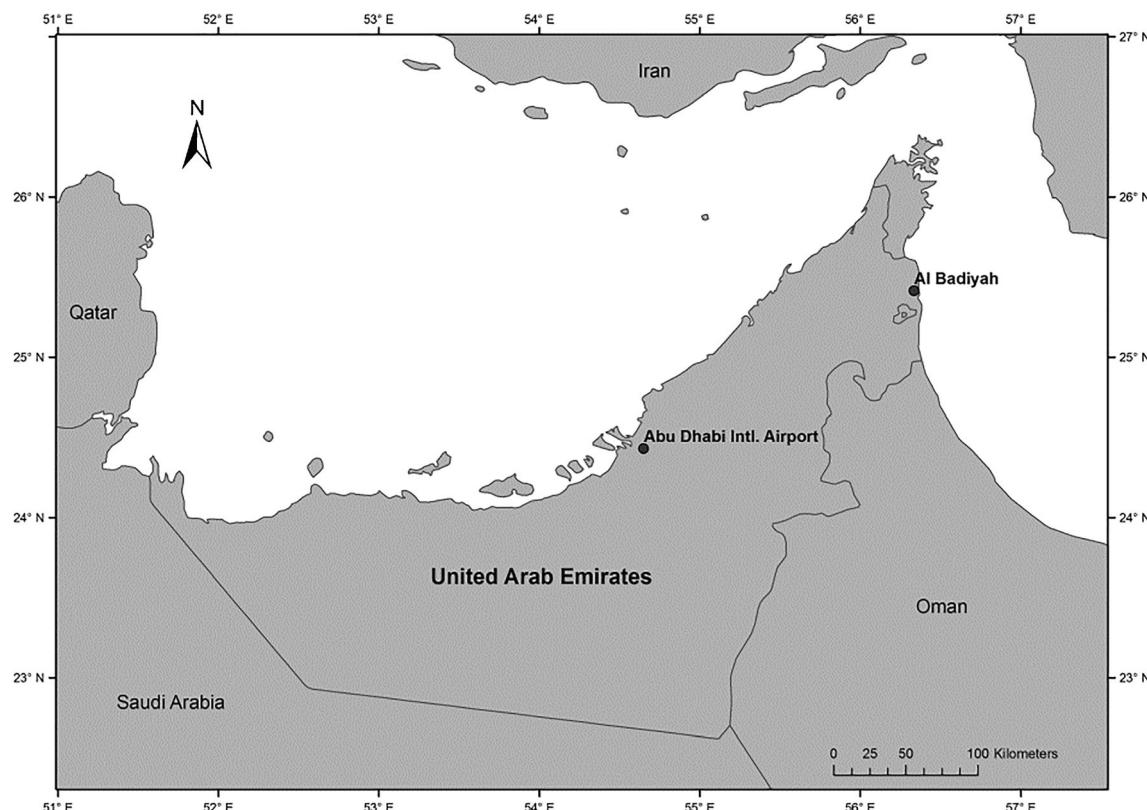
sciences field in the assessment of the relative rotation of tectonic plates, and directional data exists in studying the orientations of cross-bedding structures (Mardia & Jupp, 2009).

The circular random variable is measured in degrees from 0 to 360° or radians from zero to  $2\pi$ . These variables cannot be

modeled using standard statistical distributions. There are several periodical and directional statistical distribution models that are used to describe circular distributions such as the circular uniform distribution, von Mises distribution, wrapped normal distribution, wrapped Cauchy distribution, and the

**Fig. 2** Geographical location of the sites of the case study. Flow chart of hypothesis testing





**Fig. 3** (a) Wind rose diagrams for Al Badiyah station which gathered at 10 m height and (b) wind rose of ADIA station and wind blow up to 22.623 m. The scale represents the height in meter. Geographical location of the sites of the case study

Cardioid distribution (Jammalamadaka & Sengupta, 2001; Fisher, 1995).

A number of earlier research activities have modeled the wind direction using different methods. A model was developed for the prediction of surface wind direction over the Pacific Northwest using the von Mises distribution (Bao et al., 2010). Another research project forecasted wind directions over the North American Pacific Northwest using the applied Bayesian model averaging (BMA) method (McLean Sloughter et al., 2013). In (Masseran et al., 2013), a finite mixture of von Mises distributions is used to fit wind direction in nine stations in Malaysia. (Carta et al., 2008a) used the least squares method to estimate the parameters of a finite mixture of von Mises distributions to model wind direction in the Canary Islands. A comparison of four types of circular distribution, i.e., circular uniform distribution, von Mises

distribution, wrapped normal, distribution and wrapped Cauchy distribution, was carried out by (Kamisan et al., 2010). The study used graphical and numerical approaches to find the best fit to model southwesterly monsoon wind direction data recorded in 2005 at four locations in Malaysia. The results were compared graphically using the Q-Q plot. In addition, the mean circular distance and the chord length are also used to find the most adequate distributions. It was concluded that the von Mises distribution is the most appropriate model for the data in all four stations.

In the present work, the focus is on modeling the wind direction over the United Arab Emirates (UAE) using a finite mixture of von Mises distributions. The parameters are estimated at two different stations for different heights to look for the consistency in the distribution models. The paper is organized as follows: Section 2.1 presents the theoretical

**Table 1** Descriptive statistics of wind speeds for different elevations at Al Badiyah and ADIA

Stations	Al Badiyah					ADIA ADIA (27 m)
	Heights	10 m	50 m	80 m	100 m	
Maximum speed (m/s)	22.623	26.979	27.852	28.31	28.295	18
Minimum speed (m/s)	0.0425	0.1250	0.1318	0.1249	0.1182	0
Average speed (m/s)	4.0423	4.8159	4.8635	4.9049	4.8884	3.6837

**Table 2** Results of Rayleigh test for circular uniformity when vector mean is unknown

Stations	Al Badiyah					ADIA (27 m)
	Test when vector mean is unknown	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	
S*	$2.39 \times 10^4$	$2.51 \times 10^4$	$2.676 \times 10^4$	$2.752 \times 10^4$	$2.963 \times 10^4$	$2.055 \times 10^4$
Test Cutoff	5.991	5.991	5.991	5.991	5.991	5.991
P value	0.00	0.00	0.00	0.00	0.00	0.00
$H_0$	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected

**Table 3** Results of Rayleigh test for circular uniformity when vector mean is known

Stations	Al Badiyah					ADIA (27 m)
	Test when vector mean is known	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	
$\mu$	288°	90°	290°	280°	288°	338°
$\bar{C}$	0.3632	-0.3477	0.3822	0.3699	0.3946	0.2001
Test cutoff	0.004681	0.004681	0.004681	0.004681	0.004681	0.002749
P value	0.00	0.00	0.00	0.00	0.00	0.00
$H_0$	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected

background on circular statistics. Section 2.2 details the mathematical formulation of the von Mises distribution. Section 2.3 discusses the modeling of the mixture of two von Mises distributions. Hypothesis testing is discussed in section 3 of the paper. The case study in presented in section 4, and the results are illustrated in section 5. Finally, the discussion and conclusions are given in section 6.

## Analyzed models

### Circular statistics

Circular data is normally presented as points on the circumference of the unit circle which represent the direction's position. Any point in the plane can be denoted using rectangular coordinates  $(x, y)$  or polar coordinates  $(r, \alpha)$ , where  $r$  is the distance from the origin to the point and  $\alpha$  is the direction (angle). The trigonometric functions such as sine and cosine are used to transfer the data from rectangular coordinates to

polar coordinates and vice versa from the following equations (Jammalamadaka & Sengupta, 2001):

$$x = r \cos(\alpha), \text{ and } y = r \sin(\alpha) \quad (1)$$

The features of directional data are identified by measures of central tendency and dispersion such as mean direction, mean resultant length, circular variance, and circular deviation (Fisher, 1995; Trauth et al., 2007; Jones, 2006a). Finding the directional mean is different from finding the arithmetic mean. The proper way to combine vectors is to use vector addition:

$$C = \sum_{i=1}^n \cos \theta_i \quad (2)$$

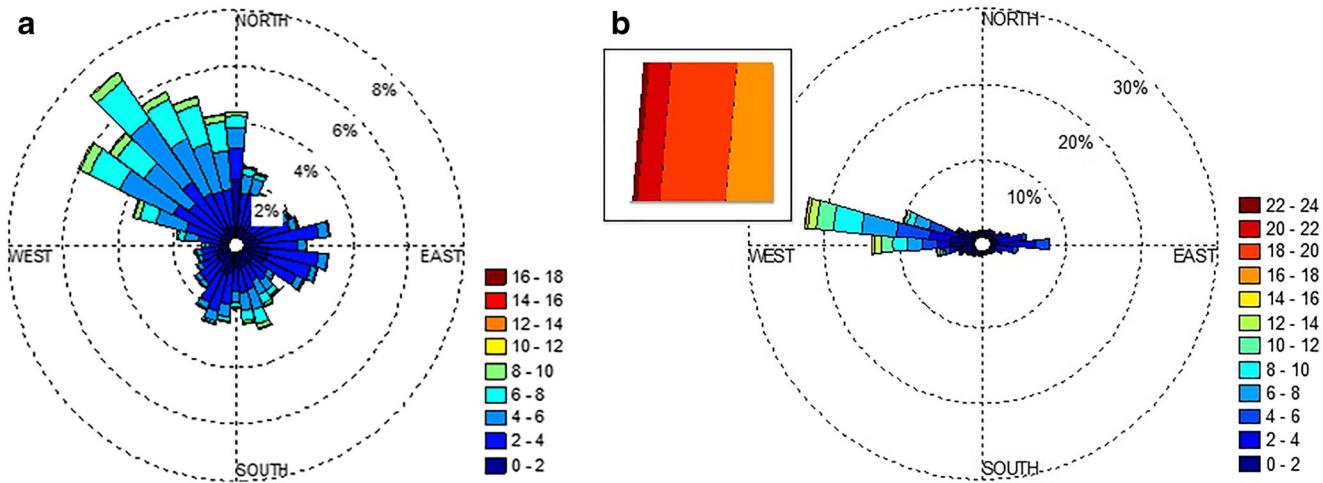
$$S = \sum_{i=1}^n \sin \theta_i \quad (3)$$

$$R^2 = C^2 + S^2 \quad (4)$$

The mean direction  $\bar{\theta}$  of the vector resultant of  $\theta_1, \dots, \theta_n$  can be computed from

**Table 4** Results of Watson's  $U^2$  test for circular uniformity

Stations	Al Badiyah					ADIA (27 m)
	$U^{*2}$ test statistic	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	
$U^{*2}$	1221	1120	1127	1145	1192	769.4
Test Cutoff	0.187	0.187	0.187	0.187	0.187	0.187
$H_0$	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected



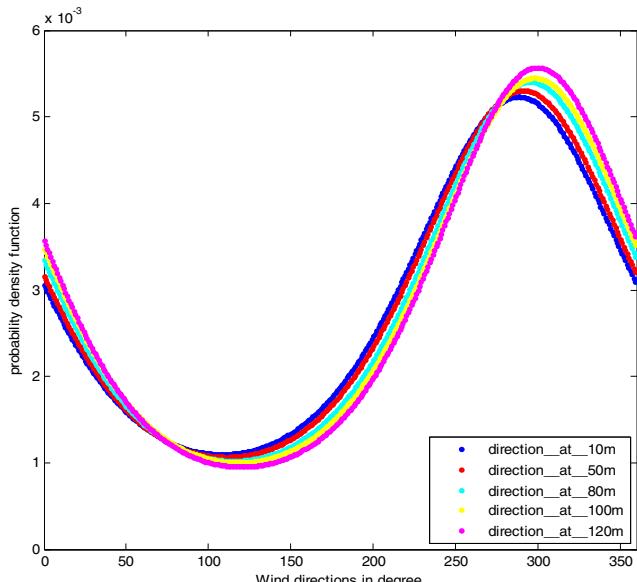
**Fig. 4** Probability density function of VM distribution for all elevations at Al Badiyah. (a) Wind rose diagrams for Al Badiyah station which gathered at 10 m height and (b) wind rose of ADIA station and wind blow up to 22.623 m. The scale represents the height in meter

$$\cos\bar{\theta} = C/R \quad (5)$$

$$\sin\bar{\theta} = S/R \quad (6)$$

$$\bar{\theta} = \begin{cases} \tan^{-1}\left(\frac{S}{C}\right) & S > 0, C > 0 \\ \tan^{-1}\left(\frac{S}{C}\right) + \pi & C < 0 \\ \tan^{-1}\left(\frac{S}{C}\right) + 2\pi & S < 0, C > 0 \end{cases} \quad (7)$$

The mean resultant length  $\bar{R}$  is the length of the mean vector and lies in the range (0, 1). It is defined in (Fisher, 1995; Bowers et al., 2000) as:



**Fig. 5** Probability density function of VM distribution at ADIA. Probability density function of VM distribution for all elevations at Al Badiyah

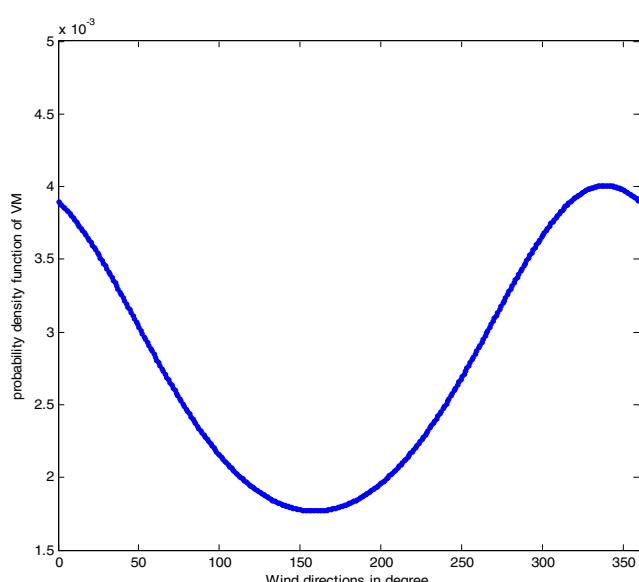
$$\bar{R} = R/n \quad (8)$$

The distribution is more concentrated when the circular variance is smaller. The sample circular variance  $V$  lies between 0 and 1 and is computed as:

$$V = 1 - \bar{R} \quad (9)$$

The corresponding circular deviation,  $S$ , is calculated as:

$$S = \sqrt{-2\log(1-V)} \quad (10)$$



**Fig. 6** Comparison of the empirical cumulative probability with the VM cumulative distribution of wind direction for all elevations at Al Badiyah. Probability density function of VM distribution at ADIA

**Table 5** Watson's  $U^2$  test results for VM when both vector mean and concentration parameters are known

Stations	Al Badiyah					ADIA
	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	Directions in 120 m	
$U^{*2}$ test statistic						ADIA (27 m)
$\mu$	288°	90°	290°	280°	288°	338°
$k$	0.7	0.7	0.9	0.8	0.8	0.4
$U^{*2}$	7306	1822	521.8	600.2	577	269.2
Test Cutoff	0.187	0.187	0.187	0.187	0.187	0.187
$H_0$	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected

## von Mises distribution

The von Mises distribution was introduced by von Mises in 1918, and it has some similarities with the normal distribution. This was highlighted by (Masseran et al., 2013; Gumbel et al., 1953; Masseran, 2015). The von Mises distribution  $VM(\mu, k)$  is a symmetric common unimodal model for circular data. The probability density function of the von Mises distribution is given as follows:

$$f(\theta) = \frac{1}{2\pi I_0(k)} e^{k\cos(\theta-\mu)}, \quad \theta \in [0, 2\pi] \quad (11)$$

where  $\mu$  is a parameter of the mean direction,  $k \in [0, \infty)$  is the concentration parameter around the center (Fisher, 1995; Costa et al., 2014), and  $I_0(k) = \frac{1}{2\pi} \int_0^{2\pi} e^{k\cos(\phi-\mu)} d\phi$  is the modified Bessel function of order zero.  $A_p(k)$  of order  $p$  is a ratio of two modified Bessel functions, where  $p$  is the order:

$$A_p(k) = A_{p-2}(k) - \frac{2(p-1)}{k} A_{p-1}(k) \quad (12)$$

$$A_p(k) = \frac{I_p(k)}{I_0(k)} \quad (13)$$

The distribution function from (Fisher, 1995) is:

$$F(\theta) = \frac{1}{2\pi I_0(k)} \int_0^\theta e^{k\cos(\phi-\mu)} d\phi \quad (14)$$

The concentration parameter can be estimated by using the maximum likelihood method, and it is a solution of the following equation [18]:

**Table 6** Watson's  $U^2$  test results for VM when both vector mean and concentration parameters are unknown

Stations	Al Badiyah					ADIA
	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	Directions in 120 m	
$U^2$ test statistic						ADIA (27 m)
$U^2$	7306	591.9	559.3	563.1	560	267.2
Test Cutoff	0.07267	0.07331	0.07419	0.07459	0.07568	0.0647
$H_0$	Rejected	Rejected	Rejected	Rejected	Rejected	Rejected

$$A_1 = \left( \hat{k}_{ML} \right) = \frac{R}{n} = \bar{R} \quad (15)$$

where  $\bar{R}$  is the mean resultant length and  $A_1(x) = \frac{I_1(x)}{I_0(x)}$ . The approximation for  $\hat{k}_{ML}$  from (Fisher, 1995; Bowers et al., 2000) is:

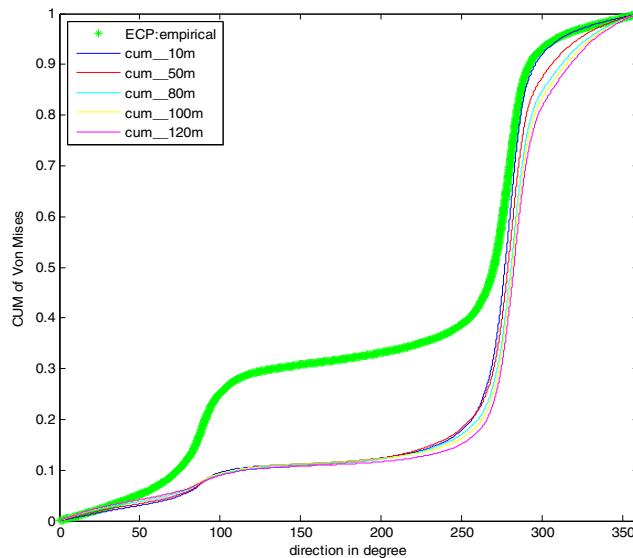
$$\hat{k}_{ML} = \begin{cases} 2\bar{R} + \bar{R}^3 + \frac{5\bar{R}^5}{6} & \bar{R} < 0.53 \\ -0.4 + 1.39\bar{R} + \frac{0.43}{1-\bar{R}} & 0.53 \leq \bar{R} < 0.85 \\ \frac{1}{\bar{R}^3 - 4\bar{R}^2 + 3\bar{R}} & \bar{R} \geq 0.85 \end{cases} \quad (16)$$

$\hat{k}_{ML}$  can be used to estimate the true value of  $k$  when the sample size is small and  $\bar{R} < 0.7$  (in other words, if the  $\bar{R} < 0.7$  and specifically  $\bar{R} < 0.45$ , then  $\hat{k}_{ML}$  is true and there is no need to go to the next step). If the sample size  $n \leq 15$ , then the following equation is preferred to estimate  $k$  after finding  $\hat{k}_{ML}$  (Fisher, 1995; Hakim et al., 2013).

$$\hat{k} = \begin{cases} \max \left( \hat{k}_{ML} - \frac{2}{\hat{k}_{ML}}, 0 \right) & \hat{k}_{ML} < 2 \\ \frac{(n-1)^3 \hat{k}_{ML}}{n^3 + n} & \hat{k}_{ML} \geq 2 \end{cases} \quad (17)$$

## Modeling mixture of two von Mises distributions

Generally, the mixture of von Mises distributions is derived from the following equation [29]:



**Fig. 7** Comparison of the empirical cumulative probability with the VM cumulative distribution of wind direction at ADIA. Comparison of the empirical cumulative probability with the VM cumulative distribution of wind direction for all elevations at Al Badiyah

$$f(\theta_i; \mu_i) = \sum_{i=1}^k \frac{p_i}{2\pi I_0(k)} e^{k_i \cos(\theta - \mu_i)} \quad (18)$$

The probability density function of a mixture of two von Mises distributions is written as (Fisher, 1995; Jones, 2006a; Mooney et al., 2003; Grimshaw et al., 2001):

$$f(\theta; \mu_1, \mu_2, k_1, k_2, p) = p_1 \frac{1}{2\pi I_0(k)} e^{k_1 \cos(\theta - \mu_1)} + (1-p_1) \frac{1}{2\pi I_0(k)} e^{k_2 \cos(\theta - \mu_2)}, \quad (19)$$

where  $\mu_1$  and  $k_1$  represent the vector mean and concentration parameters of the first component;  $\mu_2$  and  $k_2$ , represent the vector mean and concentration parameters of the second component; and  $\sum_{j=1}^n p_j = 1$  (Fisher, 1995; Carta et al., 2008a).

The five parameters are estimated using the least square method with numerical optimization. Least square method had been used in the literature to derive the equations from (20) up to (25) which had been used to compute the parameters as mentioned in (Carta et al., 2008a). The initial estimation of the parameters plays a significant role in convergence to the correct estimate. In (Fisher, 1995; Carta et al., 2008a), the method of moments was described, and it uses the following equations

**Table 7**  $R^2$  coefficients obtained for VM distribution fit

Stations	Al Badiyah					ADIA (27 m)
	Different heights	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	
$R^2$	0.858	0.8604	0.8604	0.8605	0.8602	0.9318

for the estimation of the above five parameters. The equations are as follows (Jones, 2006a):

$$pA_1(k_1)\cos(\mu_1) + (1-p)A_1(k_2)\cos(\mu_2) = \bar{C}_1, \quad (20)$$

$$pA_2(k_1)\cos(2\mu_1) + (1-p)A_2(k_2)\cos(2\mu_2) = \bar{C}_2, \quad (21)$$

$$pA_3(k_1)\cos(3\mu_1) + (1-p)A_3(k_2)\cos(3\mu_2) = \bar{C}_3, \quad (22)$$

$$pA_1(k_1)\sin(\mu_1) + (1-p)A_1(k_2)\sin(\mu_2) = \bar{S}_1, \quad (23)$$

$$pA_2(k_1)\sin(2\mu_1) + (1-p)A_2(k_2)\sin(2\mu_2) = \bar{S}_2, \quad (24)$$

$$pA_3(k_1)\sin(3\mu_1) + (1-p)A_3(k_2)\sin(3\mu_2) = \bar{S}_3, \quad (25)$$

where  $\bar{C}_p = (\frac{1}{n}) \sum_{i=1}^n \cos p\theta_i$  and  $\bar{S}_p = (\frac{1}{n}) \sum_{i=1}^n \sin p\theta_i$ .

If we denote the set of the parameters as  $\tilde{\mu}_1, \tilde{k}_1, \tilde{\mu}_2, \tilde{k}_2$ , and  $\tilde{p}$ , the aim is to find the residual difference between the sample moments  $\Delta C_1, \Delta C_2, \Delta C_3, \Delta S_1, \Delta S_2$ , and  $\Delta S_3$  which will be used in the least square criterion method to find the parameters of the two components of von Mises distributions.  $\Delta C_1$  is calculated from the following equation, and all residual differences can be obtained in the same manner (Jones, 2006a).

$$\Delta C_1 = \tilde{p}A_1(\tilde{k}_1)\cos(\tilde{\mu}_1) + (1-\tilde{p})A_1(\tilde{k}_2)\cos(\tilde{\mu}_2) - \bar{C}_1 \quad (26)$$

Then, the sum of the square criterion is:

$$r^2(\tilde{\mu}_1, \tilde{k}_1, \tilde{\mu}_2, \tilde{k}_2, \tilde{p}) = \Delta C_1^2 + \Delta C_2^2 + \Delta C_3^2 + \Delta S_1^2 + \Delta S_2^2 + \Delta S_3^2 \quad (27)$$

This method can be extended to more than two components of the von Mises distribution.

## Hypothesis testing

Three types of distributions, i.e., circular distribution, von Mises distribution, and mixture of von Mises distributions, are tested to find the best fitting model to the data as illustrated in the flow chart in Fig. 1.

Graphical assessment and formal assessment using 95% confidence intervals for the distributions are used to verify the hypothesis and results. First, we check the uniformity of

**Table 8** Estimates of the parameters of FMVM using the least square method

Components	Al Badiyah					ADIA (27 m)
	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	Directions in 120 m	
$k_1$	3.461	2.7693	2.3829	2.2828	2.2377	0.22424
$\mu_1$	81.8	81.9	82.1	82.3	82.4	107.7
$k_2$	7.8722	6.3049	6.2748	6.4047	6.3367	13.309
$\mu_2$	278.1	280.7	283.0	284.2	285.3	324.1
$p$	0.331	0.330	0.334	0.337	0.331	0.729

the data, and then a graphical assessment is carried out as in (Mardia & Jupp, 2009). Various methods of formal testing are used to check the uniformity, such as the test of randomness against any alternative (Kuiper's test for ungrouped data) (Fisher, 1995). Rayleigh test is used to test the hypothesis of uniformity of randomness against a unimodal alternative (Mardia & Jupp, 2009; Jones, 2006a). Watson's  $U^2$  test is also employed. It uses a corrected mean square deviation.  $U^2$  is well defined in this test and is invariant under rotation and reflection.  $U^2$  is based on comparing the theoretical cumulative distribution function with the practical distribution function (Jones, 2006a).

$$U^2 = \sum_{i=1}^n \left[ U_i - \bar{U} - \frac{i-\frac{1}{2}}{n} + \frac{1}{2} \right]^2 + \frac{1}{12n} \quad (28)$$

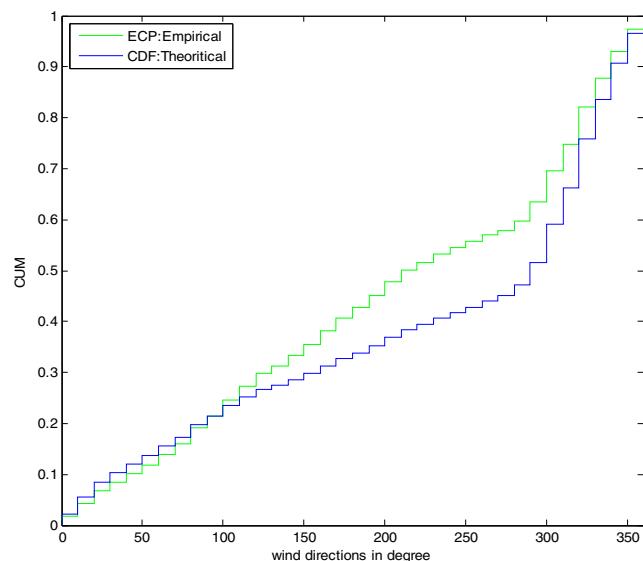
where  $U_i$  is the sorted cumulative distribution function,  $\bar{U}$  is the average of CDF, and  $N$  is the total number of records.

The second step is to test the hypothesis of von Mises distribution. The goodness of fit is measured using the  $R^2$

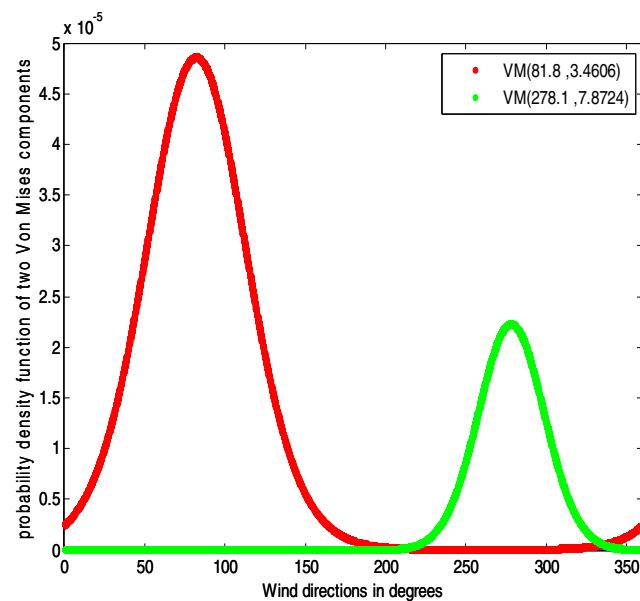
coefficient as it is mentioned in (Masseran et al., 2013; Carta et al., 2008a; Carta et al., 2008b). Watson's  $U^2$  test is used also to verify whether the distribution is a von Mises distribution or not. In this case, Watson's  $U^2$  test may take two different forms depending on whether the vector mean and concentration parameters are known (predicted) or not (both known, one of them known or both unknown) (Jones, 2006a). The last step is to check the goodness of fit for the finite mixture of two von Mises distributions using the  $R^2$  coefficient as it is mentioned in (Masseran et al., 2013; Carta et al., 2008a; Carta et al., 2008b). The mixture of von Mises is tested using the chi-square to check if the mixture of two von Mises is the best fit to the data (Heckenbergerová et al., 2013).

## Case study

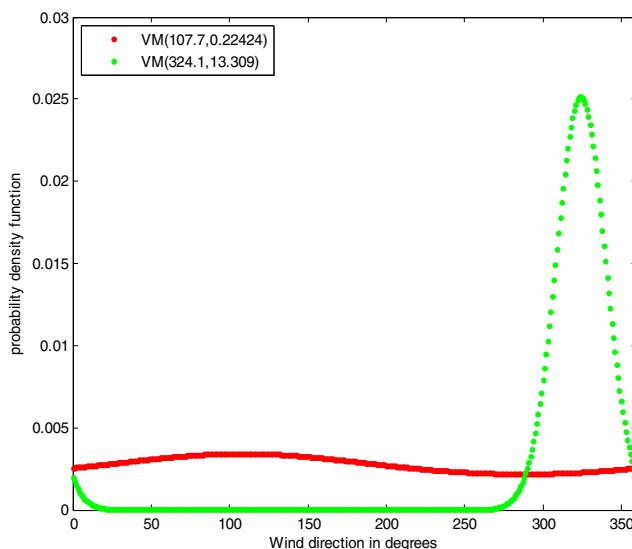
In this study, we use two data sets from two different locations in the UAE. The UAE is located in the southeastern part of



**Fig. 8** Probability density functions of the separated components of the FMVM distribution at 10 m height at Al Badiyah. Comparison of the empirical cumulative probability with the VM cumulative distribution of wind direction at ADIA



**Fig. 9** Probability density functions of the separated components of the FMVM distribution at ADIA. Probability density functions of the separated components of the FMVM distribution at 10 m height at Al Badiyah



**Fig. 10** Comparison of the empirical cumulative probability of wind direction at 10 m height with the FMVM cumulative distributions of wind direction for all elevations at Al Badiyah. Probability density functions of the separated components of the FMVM distribution at ADIA

the Arabian Peninsula between 22.5°N and 26°N and between 51°E and 56.5°E. The UAE is bordered by the Persian Gulf in the north, the Arabian Sea and Oman in the east, and Saudi Arabia in the south and west. It is composed of three different ecological areas: the marine coastal area, the northeastern mountainous area, and the desert inland area. The region is characterized by an arid climate with high temperatures. The coastal region is very humid with a summer relative humidity reaching 100%. In this region, temperatures range between 14 and 23 °C in the winter and can reach 46 °C in the summer (Ouarda et al., 2014). Both stations studied in this paper are located in the coastal region of the UAE.

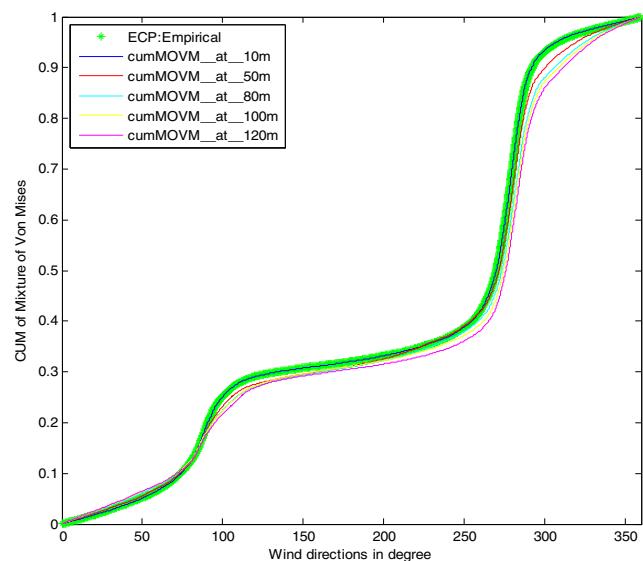
For the first site, the data is recorded at Abu Dhabi International Airport (ADIA) which is located at 24°26'N and 54°39'E. The data is measured at 27 m height from 1982 to 2010. The second site is located in the Al Badiyah region, which is near the coast and the mountains, at 25°24'N and 56°20'E. The wind direction and speed were simulated 24 h daily from 2003 to 2012 at five different elevations (10 m, 50 m, 80 m, 100 m, and 120 m) (Mooney et al., 2003). The locations of the two selected sites are illustrated in the map of Fig. 2. ADIA is located near the coast bordering the Arabian Gulf, while Al Badiyah is located near the coast bordering the Gulf of Oman and surrounded by a mountainous region. Figure 3 presents the wind rose diagrams at 10 m height for Al Badiyah and ADIA stations. The wind rose diagrams corresponding to other elevations at Al Badiyah are very similar to the one corresponding to the 10 m height. Table 1 provides an overview of the descriptive statistics of the data for the different elevations at Al Badiyah and at ADIA.

## Results

The objective of this work is to model wind direction distributions at two different locations in the UAE. For this, we explore three distribution models, i.e., uniform circular, von Mises, and mixture of two von Mises distributions. An appropriate model is selected based on the tests mentioned in section 5. The data proceed using the CircStat toolbox (Jones, 2006a; Jones, 2006b) applied in MATLAB R2014.

### Test of uniformity

The uniformity of the Al Badiyah and ADIA data was first verified based on graphical assessment using wind rose diagrams to represent wind direction as shown in Fig. 3. Wind directions are pointing to the east and to the west/northwest for ADIA and Al Badiyah, respectively. It is evident from the plots that the uniform distribution is not a good fit for the data. The hypothesis is also formally tested using the Rayleigh test with two assumptions, i.e. when the vector mean is known and when it is unknown. The hypothesis is also tested using Watson's  $U^2$  test. The results of formal assessment of the hypothesis under 0.05 confidence level are shown in Table 2, Table 3, and Table 4 for the Rayleigh test when vector means are unknown, Rayleigh test when vector means are known, and Watson's  $U^2$  test, respectively. It can be seen that the uniform circular distribution is not a good fit for the data at both locations. The test results indicate that there is a significant variation in the wind direction.



**Fig. 11** Comparison of the empirical cumulative probability with the FMVM cumulative distribution of wind direction at ADIA. Comparison of the empirical cumulative probability of wind direction at 10 m height with the FMVM cumulative distributions of wind direction for all elevations at Al Badiyah

**Table 9**  $R^2$  coefficients obtained for FMVM

Stations	Al Badiyah					ADIA (27 m)
	Different heights	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	
$R^2$	1	1	1	0.9999	0.9999	1

## Test of von Mises distribution

The graphical assessment of Al Badiyah data is illustrated in Fig. 4 which illustrates that there is a shift in the direction to the right (north) with an increase in the height. It can be observed that the von Mises distribution seems to be a good fit at 10 m height but deviates for the other heights. This indicates that as the height increases, the direction changes also and the highest speed is recorded at the altitude of 120 m Fig. 5. The graphical assessment of ADIA is shown in Fig. 6. The hypothesis was tested using Watson's  $U^2$  test (Tables 5 and Table 6). The goodness of fit was measured using the  $R^2$  coefficient as shown in Fig. 6 for Al Badiyah and in Fig. 7 for ADIA (Table 7). The tests show that the von Mises distribution is not a good fit for both data sets.

## Test of mixture of von Mises

The least square estimates of the parameters of the two components von Mises distribution are shown in Table 8. For the Al Badiyah data set, the table shows that there are two concentration parameters and two directional means. The concentration parameter  $k_2$  is higher than  $k_1$  which means that the wind direction below is highly concentrated in the western direction. The concentration parameter around the directional mean is higher at the 10 m elevation compared to the other elevations. It can be observed that as the elevation increases, the concentration parameters decrease. The reason for that might be because of the wind direction overlap in the height. The ADIA data set shows also that the concentration parameter  $k_2$  is higher than  $k_1$  and it concentrates in the northwest direction. The result in Table 8 is confirmed with the separation of two components of von Mises distributions in Fig. 8 (Al Badiyah) and Fig. 9 (ADIA). The plots at four different elevations are approximately similar to Fig. 8. The goodness of fit of the FMVM was then measured using the  $R^2$

coefficient as shown in Fig. 10, Fig. 11, and Table 9. These results illustrate that the FMVM and the empirical cumulative probability are approximately identical, and the  $R^2$  coefficient is equal to one for the heights 10 m, 50 m, 80 m, and at ADIA station. For the directions at the 100 m and 120 m heights, the  $R^2$  coefficient takes the value 0.9999. The result is comparable with (Carta et al., 2008a) as presented in Fig. 12 below.

The chi-square test was carried out to verify that the best fitting model is the FMVM as presented in Table 10. It is clear from the results corresponding to both stations that the finite mixture of two von Mises distributions is the best fitting model for wind direction.

$$X^2 = \sum_{i=1}^k \frac{(O_i - E_i)^2}{E_i} \quad (29)$$

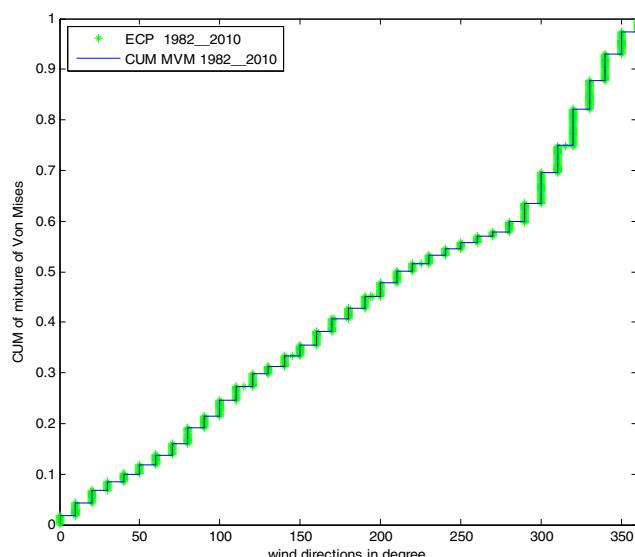
## Discussion and conclusions

Since the values of wind speed and direction is changing from one place to another, the data planned to be gathered in various locations. If the circular uniform model is accepted, this means that the wind below in all directions (360 degrees) are the same, so no much effort will be done to align the wind turbine since it will gather the same amount of energy. Whereas if the there is a particular direction for wind to blow, then that should be considered once planned to establish an electricity field of turbine. If the wind blows on two or more directions, this must be noted to place the wind turbine in a way to gather the most energy. This means that the best fitting models and number of Von Mises combination depend on the nature of the data.

To conclude, the first case study, which deals with modeling wind directions in Al Badiyah station at five different heights, shows that the mixture of two von Mises models fits the data better than the uniform circular and von Mises distributions. The mixture of von Mises model indicates that there

**Table 10** Chi-square values obtained for FMVM

Stations	Al Badiyah					ADIA (27 m)
	Different heights	Directions in 10 m	Directions in 50 m	Directions in 80 m	Directions in 100 m	
$\chi^2$	1.25	1.2501	1.8961	1.9011	1.9080	1.9080



**Fig. 12** Coefficient of determination  $R^2$  for mixture of von Mises distribution. Comparison of the empirical cumulative probability with the FMVM cumulative distribution of wind direction at ADIA

are two mean (average) wind directions which are different in concentration. All data sets are highly concentrated in interval range (278.1, 285.3) which points to the northwestern direction more than the second mean (average) direction in interval range (81.8, 82.4) which points to the northeastern direction for the five data sets. It is interesting to note that all heights (10 m, 50 m, 80 m, 100 m, and 120 m) have similar wind direction distributions. All chi-square values lie below the value of 3.841 when  $\alpha = 0.05$ . Consequently, the hypothesis that wind has two main directions cannot be rejected.

The second case study models wind direction at Abu Dhabi International Airport during the period 1982 to 2010. Wind direction at ADIA points to two main directions. The dominant wind direction is concentrated in the northwest (324.1), and the second direction points to the southeastern (107.7). This shows that wind direction characteristics at the ADIA station are significantly different from Al Badiyah station. This points to the spatial variability of wind direction in the UAE and the need for a more detailed assessment of wind directional distributions at any specific site of interest. The chi-square value is lower than 3.841 when the significant interval is  $\alpha = 0.05$ , and the hypothesis that wind has two main directions cannot be rejected.

Future research efforts may focus on the association between wind direction and speed, e.g., using linear-circular association methods (Fisher, 1995) in order to accurately estimate the wind power and better plan the development of wind farms. Future research may also explore the use of bivariate (wind speed and direction) and multivariate (wind speed and direction and air density) models using multivariate kernel density estimates (Zhang et al., 2013).

## Declarations

**Conflict of interest** The author(s) declare that they have no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Abohela I, Hamza N, Dudek S (2013) Effect of roof shape, wind direction, building height and urban configuration on the energy yield and positioning of roof mounted wind turbines. *Renew Energy* 50: 1106–1118
- Bao L, Gneiting T, Grimit EP, Guttorp P, Raftery AE (2010) Bias correction and Bayesian model averaging for ensemble forecasts of surface wind direction. *Mon Weather Rev* 138:1811–1821
- Bowers J, Morton I, Mould G (2000) Directional statistics of the wind and waves. *Appl Ocean Res* 22:13–30
- Carta JA, Bueno C, Ramírez P (2008a) Statistical modelling of directional wind speeds using mixtures of von Mises distributions: case study. *Energy Convers Manag* 49:897–907
- Carta JA, Ramírez P, Bueno C (2008b) A joint probability density function of wind speed and direction for wind energy analysis. *Energy Convers Manag* 49:1309–1320
- Costa M, Koivunen V, Poor HV (2014) Estimating directional statistics using wavefield modeling and mixtures of von-Mises distributions. *Signal Processing Letters IEEE* 21:1496–1500
- Davis MW (1977) A new thermal rating approach: the real time thermal rating system for strategic overhead conductor transmission lines—part I: general description and justification of the real time thermal rating system. *IEEE Transactions on Power Apparatus and Systems* 96:803–809
- N. I. Fisher, *Statistical analysis of circular data*: Cambridge University press, 1995
- Grimshaw SD, Whiting DG, Morris TH (2001) Likelihood ratio tests for a mixture of two von Mises distributions. *Biometrics* 57:260–265
- Gumbel E, Greenwood JA, Durand D (1953) The circular normal distribution: theory and tables. *J Am Stat Assoc* 48:131–152
- Hakim A, Marsland S, Guesgen HW (2013) Statistical Modelling of Complex Emotions using Mixture of von Mises Distributions. In: *Affective Computing and Intelligent Interaction (ACII), 2013 Humaine Association Conference on*, pp 517–522
- Heckenbergerová J, Musilek P, Mejznar J, Vančura M (2013) Estimation of wind direction distribution with genetic algorithms. In: *Electrical and Computer Engineering (CCECE), 2013 26th Annual IEEE Canadian Conference on*, pp 1–4
- Jammalamadaka SR, Lund UJ (2006) The effect of wind direction on ozone levels: a case study. *Environ Ecol Stat* 13:287–298
- Jammalamadaka SR, Sengupta A (2001) *Topics in circular statistics*, vol 5. World Scientific

- Jones TA (2006a) MATLAB functions to analyze directional (azimuthal) data—I: single-sample inference. *Comput Geosci* 32:166–175
- Jones TA (2006b) MATLAB functions to analyze directional (azimuthal) data—II: correlation. *Comput Geosci* 32:176–183
- Kamisan NAB, Hussin AG, Zubairi YZ (2010) Finding the best circular distribution for southwesterly monsoon wind direction in Malaysia. *Sains Malaysiana* 39:387–393
- Mardia KV, Jupp PE (2009) *Directional statistics*, vol 494. John Wiley & Sons
- Masseran N (2015) Modeling a multi-modal distribution of wind direction data in Kudat, Malaysia. *Pakistan Journal of Statistics and Operation Research* 11
- Masseran N, Razali A, Ibrahim K, Latif M (2013) Fitting a mixture of von Mises distributions in order to model data on wind direction in Peninsular Malaysia. *Energy Convers Manag* 72:94–102
- McLean Sloughter J, Gneiting T, Raftery AE (2013) Probabilistic wind vector forecasting using ensembles and Bayesian model averaging. *Mon Weather Rev* 141:2107–2119
- Mooney JA, Helms PJ, Jolliffe IT (2003) Fitting mixtures of von Mises distributions: a case study involving sudden infant death syndrome. *Computational Statistics & Data Analysis* 41:505–513
- Naizghi MS, Ouarda TBMJ (2017) Teleconnections and analysis of long-term wind speed variability in the UAE. *Int J Climatol* 37:230–248
- Ouarda T, Charron C, Shin J-Y, Marpu P, Al-Mandoos A, Al-Tamimi M et al (2015) Probability distributions of wind speed in the UAE. *Energy Convers Manag* 93:414–434
- Ouarda TBMJ, Charron C, Chebana F Review of criteria for the selection of probability distributions for wind speed data and introduction of the moment and L-moment ratio diagram methods, with a case study. *Energy Conversion and Management* 124:247–265 9/15/2016
- Ouarda TBMJ, Charron C, Nirajan Kumar K, Marpu PR, Ghedira H, Molini A, Khayal I (2014) Evolution of the rainfall regime in the United Arab Emirates. *J Hydrol* 514:258–270
- Petersen GN, Kristjansson JE, Ólafsson H (2005) The effect of upstream wind direction on atmospheric flow in the vicinity of a large mountain. *Q J R Meteorol Soc* 131:1113–1128
- Qin Z, Li W, Xiong X (2011) Estimating wind speed probability distribution using kernel density method. *Electr Power Syst Res* 81: 2139–2146
- Safari B (2011) Modeling wind speed and wind power distributions in Rwanda. *Renew Sust Energ Rev* 15:925–935
- Shin J-Y, Ouarda TBMJ, Lee T (2016) Heterogeneous mixture distributions for modeling wind speed, application to the UAE. *Renew Energy* 91:40–52
- Trauth MH, Gebbers R, Marwan N, Sillmann E (2007) *MATLAB recipes for earth sciences*, vol 34. Springer
- Wang J, Hu J, Ma K (2016) Wind speed probability distribution estimation and wind energy assessment. *Renew Sust Energ Rev* 60:881–899
- Zhang J, Chowdhury S, Messac A, Castillo L (2013) A multivariate and multimodal wind distribution model. *Renew Energy* 51:436–447
- Zhang Y, Wang J, Wang X (2014) Review on probabilistic forecasting of wind power generation. *Renew Sust Energ Rev* 32:255–270