Available online at www.sciencedirect.com
ScienceDirect
Energy Conversion and Management 49 (2008) 897–907
ENERGY CONVERSION & MANAGEMENT
www.elsevier.com/locate/enconman

Statistical modelling of directional wind speeds using mixtures of von Mises distributions: Case study

José A. Carta [a,*], Celia Bueno [b], Penélope Ramírez [b]

[a] *Department of Mechanical Engineering, University of Las Palmas de Gran Canaria, Campus de Tafira s/n, 35017 Las Palmas de Gran Canaria, Canary Islands, Spain*

[b] *Department of Renewable Energies, Technological Institute of the Canary Islands, Pozo Izquierdo Beach s/n, 35119 Santa Lucia, Gran Canaria, Canary Islands, Spain*

Received 24 March 2007; accepted 30 October 2007
Available online 19 December 2007

Abstract

A finite mixture model of continuous variable probability is used in this paper to represent the distribution of directional wind speed. The model is comprised of a finite mixture of von Mises (vM-pdf) distributions.

The parameters of the model are estimated using the least squares method. The range of integration to compute the mean angle and the standard deviation of wind direction is adjusted to minimum variance requirements. The suitability of the distributions is judged from the coefficient of determination R^2. The model is applied in this paper to wind direction data recorded at several weather stations located in the Canary Islands (Spain). The conclusion reached is that the mixture distribution used in this paper provides a very flexible model for wind direction studies and can be applied in a widespread manner to represent the wind direction regimes in zones with several modes or prevailing wind directions. In the case of the Canary Islands, mixtures of two vM-pdfs provide better fits in all the cases analysed than those obtained with the models proposed in the specialised literature on wind energy. The conclusion is also drawn that when the number of components N of the mixture distribution increases, the value of R^2 increases. However, the variations in R^2 are not significant for values of N greater than six.

Keywords: Mixture distributions; Wind direction; Circular normal distributions; von Mises distributions

1. Introduction

In wind energy analysis, the use of continuous probability density functions (pdf) is common [1–4]. However, discrete models are usually employed in the analysis of directional wind speed [5–9]. These models normally classify the wind speed directions into sectors or bins.[1]

In the specialised literature on wind energy and other renewable energy sources, the only continuous models to which we have reference are those proposed by Smith [10–12] and McWilliams et al. [13,14]. The horizontal wind direction θ is an angular variable with a period of 360°, so these authors use an offset normal distribution (or angular Gaussian or projected normal distribution). The offset normal models derive from a bivariate normal distribution, whose variables are the horizontal Cartesian components of the wind speed, to deduce a wind direction distribution law. Smith [10–12] takes into account the existing correlation between the two horizontal wind components. The model of McWilliams et al. [13,14] is derived from the assumption that the wind speed component along the prevailing wind direction is normally distributed with non-zero mean and a given variance, while the wind speed component along a direction orthogonal to it is independent and normally distributed with zero mean and the same variance. That is to say, as indicated by Weber [15], we are dealing with an isotropic Gaussian model.

* Corresponding author. Tel.: +34 928 45 14 83; fax: +34 928 45 14 84.
E-mail address: jcarta@dim.ulpgc.es (J.A. Carta).

[1] Wind directions are often recorded to the nearest 8 or 16 compass points.

0196-8904/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.enconman.2007.10.017

A large number of studies have been published that propose the use, in various fields of science, of a variety of pdfs to describe frequency distributions of circular or angular variables [16–18]. Nevertheless, most pdfs are symmetric and unimodal, with the von Mises (or circular normal) distribution being the most commonly used circular distribution.

We use in this paper a continuous variable probability model to represent distributions of directional wind speed, irrespective of the number of modes or prevailing wind speeds. The model is comprised of a finite mixture of von Mises distributions (vM-pdf). The parameters of the models are estimated using the least squares method [19], which is resolved in this paper using the Levenberg–Marquardt algorithm [20]. The range of integration to compute the mean angle and standard deviation of the wind direction is adjusted to minimum variance requirements. The suitability of the distributions is judged from the coefficient of determination R^2 [20]. A comparison is made between Smith's model [10–12] and the proposed mixture distributions. This comparison is based on an analysis of the level of fit to the cumulative frequencies of hourly wind directions recorded at two weather stations located in the Canary Islands (Spain).

2. Models

2.1. Smith's model

Smith [10,11] uses the bivariate Gaussian distribution $f(v_x, v_y)$ of the horizontal wind velocity components v_x and v_y. Eq. (1)

$$f(v_x, v_y) = A\exp\left[-\frac{1}{2(1-\rho^2)}\left(\frac{v_x-\bar{v}_x}{\sigma_x}\right)^2 -2\rho\left(\frac{v_x-\bar{v}_x}{\sigma_x}\right)\left(\frac{v_y-\bar{v}_y}{\sigma_y}\right)+\left(\frac{v_y-\bar{v}_y}{\sigma_y}\right)^2\right] \quad (1)$$

where

$$A = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \quad (2)$$

$\bar{v}_x$ and $\bar{v}_y$ are the sample mean components, Eq. (3); σ_x and σ_y are the standard deviations of v_x and v_y, respectively, Eq. (4) and ρ is the correlation coefficient, Eq. (5). n is the number of sample data.

$$\bar{v}_x = \frac{1}{n}\sum_{i=1}^{n}(v_x)_i = \frac{1}{n}\sum_{i=1}^{n}v_i\cos\theta_i; \quad \bar{v}_y = \frac{1}{n}\sum_{i=1}^{n}(v_y)_i = \frac{1}{n}\sum_{i=1}^{n}v_i\sin\theta_i \quad (3)$$

$$\sigma_x = \left\{\frac{1}{n-1}\sum_{i=1}^{n}[v_i\cos\theta_i-\bar{v}_x]^2\right\}^{\frac{1}{2}}; \quad \sigma_y = \left\{\frac{1}{n-1}\sum_{i=1}^{n}[v_i\sin\theta_i-\bar{v}_y]^2\right\}^{\frac{1}{2}} \quad (4)$$

$$\rho = \frac{n\sum_{i=1}^{n}v_i^2\cos\theta_i\sin\theta_i-\left(\sum_{i=1}^{n}v_i\cos\theta_i\right)\left(\sum_{i=1}^{n}v_i\sin\theta_i\right)}{\left[n\sum_{i=1}^{n}v_i^2\cos^2\theta_i-\left(\sum_{i=1}^{n}v_i\cos\theta_i\right)^2\right]^{\frac{1}{2}}\left[n\sum_{i=1}^{n}v_i^2\sin^2\theta_i-\left(\sum_{i=1}^{n}v_i\sin\theta_i\right)^2\right]^{\frac{1}{2}}} \quad (5)$$

To deduce the distribution law of the wind direction θ, Smith [10,11] converts (Eq. (1)) to polar coordinates, Eq. (6)

$$f(\theta) = \frac{A}{B}\left[1+C\sqrt{2\pi}\exp\left(\frac{C^2}{2}\right)\phi(C)\right]\exp\left(-\frac{D}{2}\right) \quad (6)$$

where

$$B = \frac{1}{1-\rho^2}\left(\frac{\cos^2\theta}{\sigma_x^2}-2\rho\frac{\cos\theta\sin\theta}{\sigma_x\sigma_y}+\frac{\sin^2\theta}{\sigma_y^2}\right) \quad (7)$$

$$C = \frac{1}{1-\rho^2}\left(\frac{\bar{v}_x\cos\theta}{\sigma_x^2}-\rho\frac{\bar{v}_x\sin\theta+\bar{v}_y\cos\theta}{\sigma_x\sigma_y}+\frac{\bar{v}_y\sin\theta}{\sigma_y^2}\right)\frac{1}{\sqrt{B}} \quad (8)$$

$$D = \frac{1}{1-\rho^2}\left(\frac{v_x^2}{\sigma_x^2}-2\rho\frac{v_xv_y}{\sigma_x\sigma_y}+\frac{v_y^2}{\sigma_y^2}\right) \quad (9)$$

$$\phi(C) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{C}\exp\left(-\frac{t^2}{2}\right)\mathrm{d}t \quad (10)$$

2.2. von Mises mixture model

The proposed continuous probability model mvM(θ) is comprised of a sum of N von Mises probability densities vM$_j(\theta)$, (Eq. (11))

$$\mathrm{mvM}(\theta) = \sum_{j=1}^{N}\omega_j\mathrm{vM}_j(\theta) \quad (11)$$

where the ω_j are nonnegative quantities that sum to one [21,22]; that is, (Eq. (12))

$$0 \leqslant \omega_j \leqslant 1 \quad (j=1,\ldots,N) \quad \text{and} \quad \sum_{j=1}^{N}\omega_j = 1 \quad (12)$$

A random variable θ has a von Mises distribution, vM-pdf, if its probability density function is defined by Eq. (13) [17]

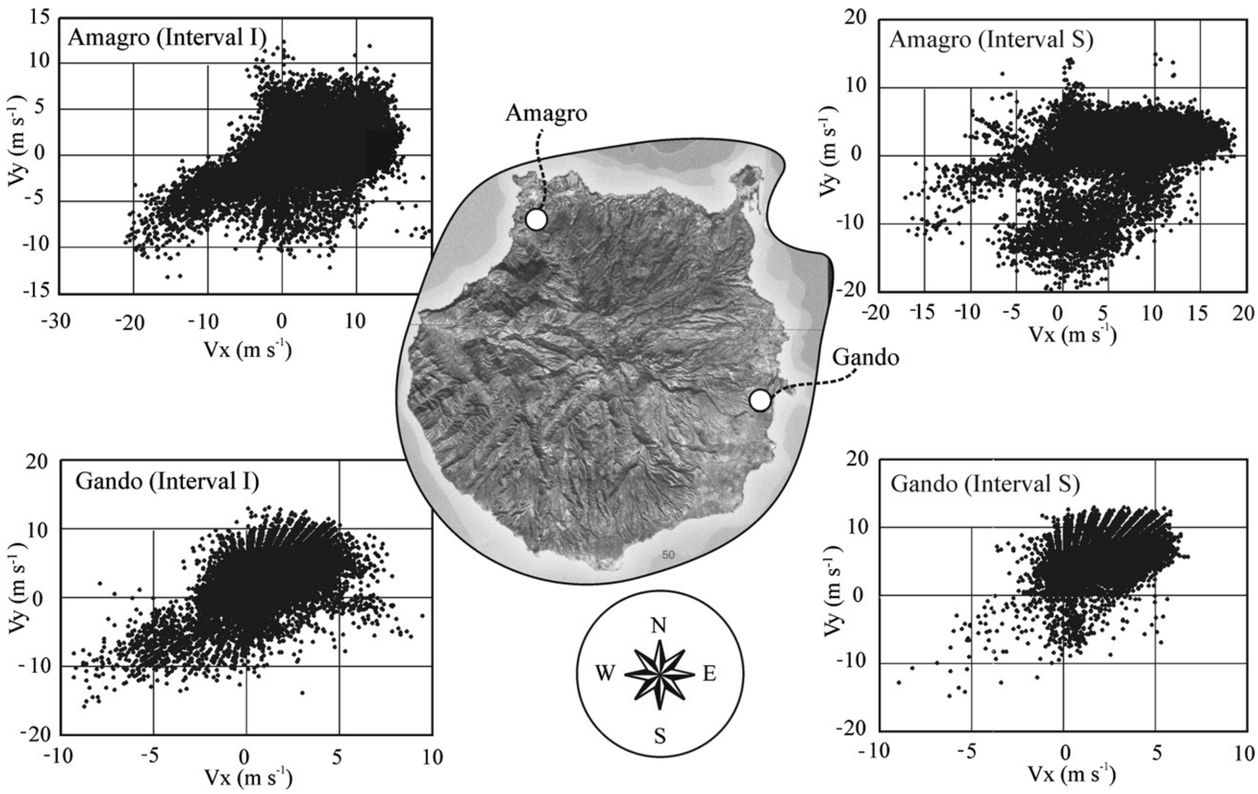

Fig. 1. Location of the weather stations used.

$$\mathrm{vM}_j(\theta;\kappa_j,\mu_j)=\frac{1}{2\pi I_0(\kappa_j)}\exp\left[\kappa_j\cos\left(\theta-\mu_j\right)\right],\quad 0\leqslant\theta<2\pi \tag{13}$$

where $\kappa_j \geqslant 0$ and $0 \leqslant \mu_j < 2\pi$ are parameters. The distribution is unimodal and is symmetrical about $\theta = \mu_j$. In this paper, the angle corresponding to the northerly direction[2] is taken as angle 0°. The parameter μ_j is the mean direction and the parameter κ_j is known as the concentration parameter. Here, $I_0(\kappa_j)$ is the modified Bessel function of the first kind and order zero [23] and is given by Eq. (14)

$$I_0(\kappa_j)=\frac{1}{\sqrt{2\pi}}\int_0^{2\pi}\exp\left[\kappa_j\cos\theta\right]\mathrm{d}\theta=\sum_{k=0}^{\infty}\frac{1}{(k!)^2}\left(\frac{\kappa_j}{2}\right)^{2k} \tag{14}$$

The distribution law mvM(θ), given by Eq. (11), can be numerically integrated [17] between two given values of θ to obtain the probability that the wind direction is found within a particular angular sector. The cumulative distribution function of the mixture (MvM-cdf) is given by Eq. (15), which has to be evaluated numerically [17]

$$\mathrm{MvM}(\theta;\kappa_j,\mu_j,\omega_j)=\sum_{j=1}^{N}\frac{\omega_j}{2\pi I_0(\kappa_j)}\int_0^{\theta}\exp\left[\kappa_j\cos\left(\theta-\mu_j\right)\right]\mathrm{d}\theta \tag{15}$$

3. Estimation of the parameters of the mixture models

Various methods can be used to estimate the $3N$ parameters on which the mixture of N vM-pdfs depends [24]. In this paper, the least squares (LS) method has been used. In the LS method, the $3N$ unknown values of the parameters can be estimated by looking for the numerical values of the parameters that minimise the sum of the squares of the deviations between the experimental data and those obtained with the model under linear inequality constraints, Eq. (16). In this paper, the LS method is applied to the cumulative distribution function Eq. (15) with μ_j, κ_j and ω_j ($j = 1,\ldots,N$) as the unknown parameters. P is a vector that contains the experimental cumulative relative frequencies function obtained from a sample of $n$observations, $\theta_i(i = 1,\ldots,n)$. In other words, if the observed wind direction values are grouped into $T(T \geqslant N)$ wind direction sectors[3] $0° - \theta_1,\ \theta_1 - \theta_2,\ldots,\theta_{T-1} - \theta_T$ ($\theta_T = 360°$) and to each sector is assigned its relative frequency of occurrence

[2] In meteorology, the angle is measured clockwise from the north.

[3] 37 sectors have been used in this paper ($T = 37$).

$fr_1, fr_2, \ldots, fr_T$, then the cumulative frequencies will be given by: $P_1 = fr_1, P_2 = P_1 + fr_2, \ldots, P_T = P_{T-1} + fr_T$

$$\mathrm{Min}S = \sum_{k=1}^{T} \left\{ P_k - \sum_{j=1}^{N} \omega_j \int_0^{\theta_k} \frac{1}{2\pi I_0(\kappa_j)} \times \exp\left[\kappa_j \cos(\theta - \mu_j)\right] \mathrm{d}\theta \right\}^2 \tag{16}$$

Inequality constraints $\kappa_j \geqslant 0$; $0 \leqslant \mu_j < 2\pi$; $0 \leqslant \omega_j \leqslant 1$; $\sum_{j=1}^{N} \omega_j = 1$.

The Levenberg–Marquardt algorithm[4] (LMA) [20] has been used to solve Eq. (16).

The nonlinear programming technique used requires a starting point or base point.[5] Good starting values will often allow an iterative technique to converge to a solution much faster than would otherwise be possible.[6] In order to determine the initial values of the $3N$ parameters, the observed wind direction values are grouped into N wind directional sectors $0° - \theta'_1$, $\theta'_1 - \theta'_2$, $\ldots$,$\theta'_{N-1} - \theta'_N$ ($\theta'_N = 360'$) and to each sector is assigned its relative frequency of occurrence $fr'_1, fr'_2, \ldots, fr'_N$. Then, the coefficients μ_j are estimated with Eq. (17)

$$\mu_j = \begin{cases} \arctan\left(\frac{\bar{s}_j}{\bar{c}_j}\right); & \bar{s}_j \geqslant 0,\ \bar{c}_j > 0 \\ \frac{\pi}{2}; & \bar{s}_j > 0,\ \bar{c}_j = 0 \\ \pi + \arctan\left[\frac{\bar{s}_j}{\bar{c}_j}\right]; & \bar{c}_j < 0 \end{cases}$$

$$\mu_j = \begin{cases} \pi; & \bar{s}_j = 0,\ \bar{c}_j = -1 \\ 2\pi + \arctan\left[\frac{\bar{s}_j}{\bar{c}_j}\right]; & \bar{s}_j < 0,\ \bar{c}_j > 0 \\ \frac{3\pi}{2}; & \bar{s}_j < 0,\ \bar{c}_j = 0 \end{cases} \tag{17}$$

where $\bar{s}_j$ and $\bar{c}_j$ are given by the expressions, Eq. (18))

$$\bar{s}_j = \frac{\sum_{i=1}^{n_j} \sin\theta_i}{n_j}; \quad \bar{c}_j = \frac{\sum_{i=1}^{n_j} \cos\theta_i}{n_j} \tag{18}$$

n_j is the number of directional wind speed data pertaining to sector j. The coefficients κ_j are obtained as the solution of Eq. (19) [16]

$$A(\kappa_j) = \frac{I_1(\kappa_j)}{I_0(\kappa_j)} = \left[\bar{s}_j^2 + \bar{c}_j^2\right]^{\frac{1}{2}} \tag{19}$$

$A(\kappa_j)$ is the ratio of two Bessel functions: $I_0(\kappa_j)$, Eq. (14), and $I_1(\kappa_j) \cdot I_1(\kappa_j)$ is the modified Bessel function of the first kind and order one, (Eq. (20))

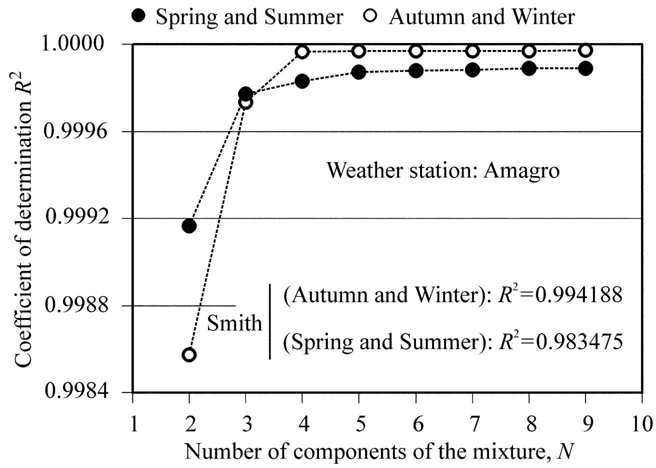

Fig. 2. Coefficient of determination as function of the number of components of the mixture distribution at Amagro station.

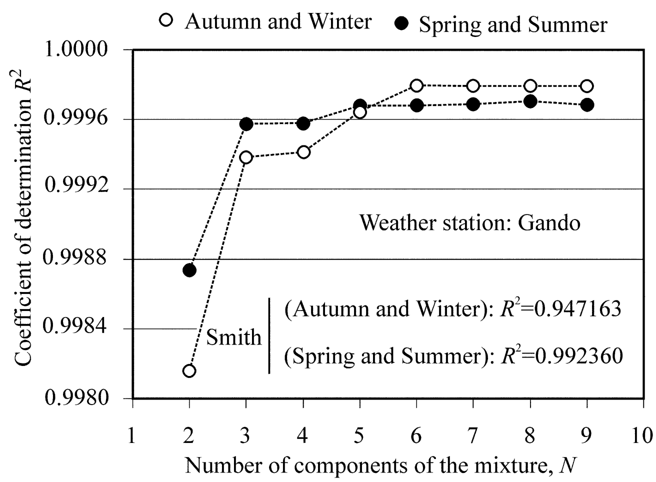

Fig. 3. Coefficient of determination as function of the number of components of the mixture distribution at Gando station.

$$I_1(\kappa_j) = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} \cos\theta \exp\left[\kappa_j \cos\theta\right] \mathrm{d}\theta = \sum_{k=0}^{\infty} \frac{1}{\Gamma(k+2)\Gamma(k+1)} \left(\frac{\kappa_j}{2}\right)^{2k+1} \tag{20}$$

where Γ is the gamma function [23]. $A(\kappa_j)$ is a strictly monotonic increasing function of κ_j, and so, to resolve Eq. (19), in this paper, we propose to resolve the following approximate equation[7], (Eq. (21))

$$\kappa_j = \left\{ 23.29041409 - 16.8617370\left(\bar{s}_j^2 + \bar{c}_j^2\right)^{0.25} - 17.4749884 \exp\left[-\left(\bar{s}_j^2 + \bar{c}_j^2\right)\right] \right\}^{-1} \tag{21}$$

4. Test of fit used

In this paper, we use the coefficient of determination (R^2) [16], Eq. (22), to estimate the degree of fit of the cumu-

[4] The Mathcad® Software 2001*i* programme of MathSoft Engineering and Education, Inc [25] is used to find the values of the $3N$ parameters.

[5] All the iterative procedures require initial values of the parameters to be estimated.

[6] However, in many cases, the LMA finds a solution even if it starts very far from the final minimum.

[7] This equation has a coefficient of determination of 0.999996.

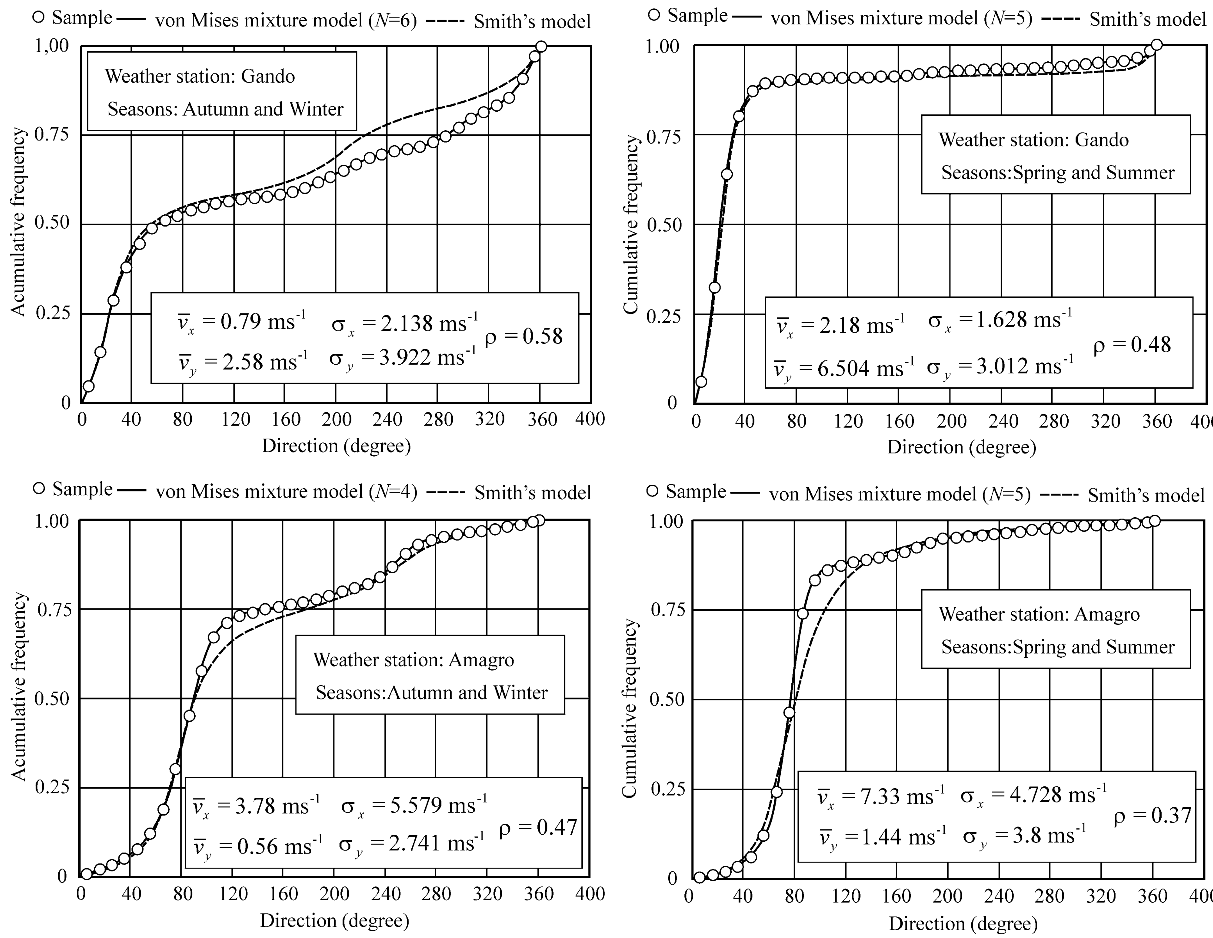

Fig. 4. Cumulative frequencies of Smith's model and of the model proposed in this paper for the different weather stations and monthly intervals considered.

lative sample distributions and the cumulative distribution functions

$$R^2 = 1 - \frac{\sum_{k=1}^{T}(P_k - MG_k)^2}{\sum_{k=1}^{T}(P_k - \overline{MG})^2} \tag{22}$$

where MG_k are the values of the theoretical cumulative relative frequencies and $\overline{MG}$ is the mean of the MG_k values. The value of R^2 varies between 0 and 1. The higher R^2 is, the better is the fit.

5. Estimation of mean and standard deviation of wind direction

Since the horizontal wind direction is a circular variable, its mean μ_θ and standard deviation σ_θ cannot be directly estimated by on line methods [15,18]. If we try to use expressions similar to those used for linear variables, the conclusion is reached that these expressions depend strongly on the position of the lower bound θ_l of the interval of integration [15], (Eqs. (23) and (24))

$$\mu_\theta(\theta_l) = \int_{\theta_l}^{\theta_l+2\pi} \theta g(\theta)\mathrm{d}\theta \tag{23}$$

$$\sigma_\theta(\theta_l) = \left\{\int_{\theta_l}^{\theta_l+2\pi} [\theta - \mu_\theta(\theta_l)]^2 g(\theta)\mathrm{d}\theta\right\}^{\frac{1}{2}} \tag{24}$$

where $g(\theta)$ is the pdf.

The value of θ_l can be obtained by establishing the condition that the standard deviation, defined by Eq. (24), is minimal. Then, the valid value of θ_l is that obtained when setting the derivative of the standard deviation with respect to θ_l, Eq. (25), equal to zero and requiring the second derivative of the standard deviation with respect to θ_l be positive

$$\frac{\partial\sigma_\theta(\theta_l)}{\partial\theta_l} = 0 \rightarrow [\pi + \theta_l - \mu_\theta(\theta_l)]g(\theta_l) = 0 \tag{25}$$

When θ_l has been determined, the mean and standard deviation can be calculated through Eqs. (23) and (24), respectively.

6. Meteorological data used

In order to determine the flexibility of the proposed model to represent different distributions of directional

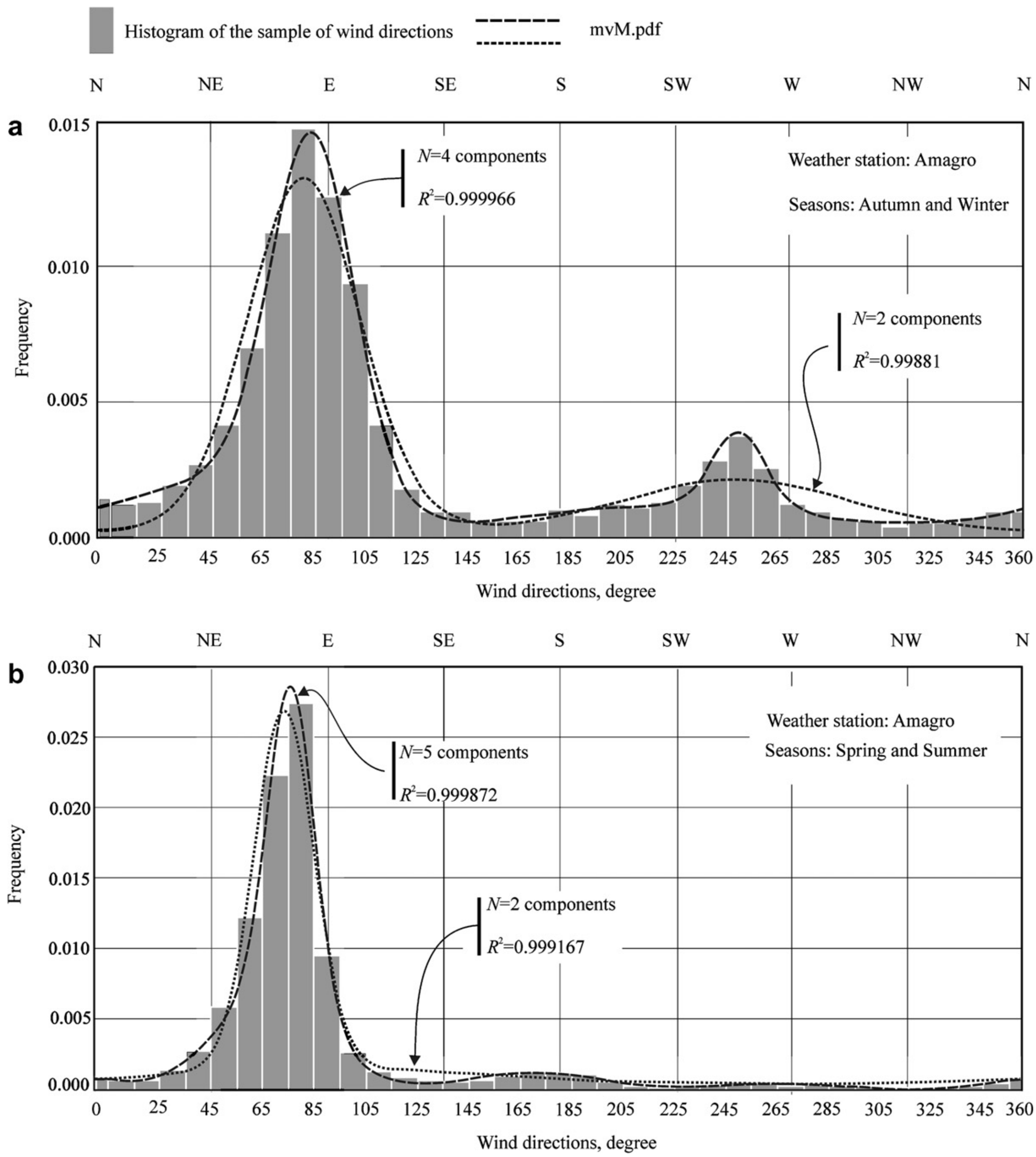

Fig. 5. Frequency histograms of wind directions at Amagro station: (a) months of interval I, (b) months of interval S.

wind speed, a study was conducted of samples recorded at different anemometer stations located in the Canarian Archipelago.[8] These stations have been installed in areas suitable for development of large scale exploitation of wind energy [26], and have, thus, avoided terrain of complex topography. However, the effect of the sea breezes in certain areas of the islands means that there are different wind speed distributions for the same island [3,24]. Even so, there are no significant differences between most of the stations with respect to the frequency histograms of the observed wind speed directions. Therefore, two stations have been selected that are representative of the different wind direction distributions in the island of Gran Canaria,[9] Fig. 1. For the station called Amagro [3], located in the

[8] The Canarian Archipelago is located northwest of the African continent, between latitude 27°37′ to 29°25′ north (subtropical position) and longitude 13°20′ to 18°10′ west of Greenwich. The Canarian Archipelago is approximately 1000 km from the Spanish mainland coast, and the closest and furthest distances from the African coast are 100 km and 500 km, respectively.

[9] The island of Gran Canaria, located in the centre of the Archipelago, is almost circular in shape, with a width of 47 km and a length of 55 km. It is a large rocky massif which peaks near its geographical centre at an approximate height of 2000 m. The surface area of the island is some 1532 km^2.

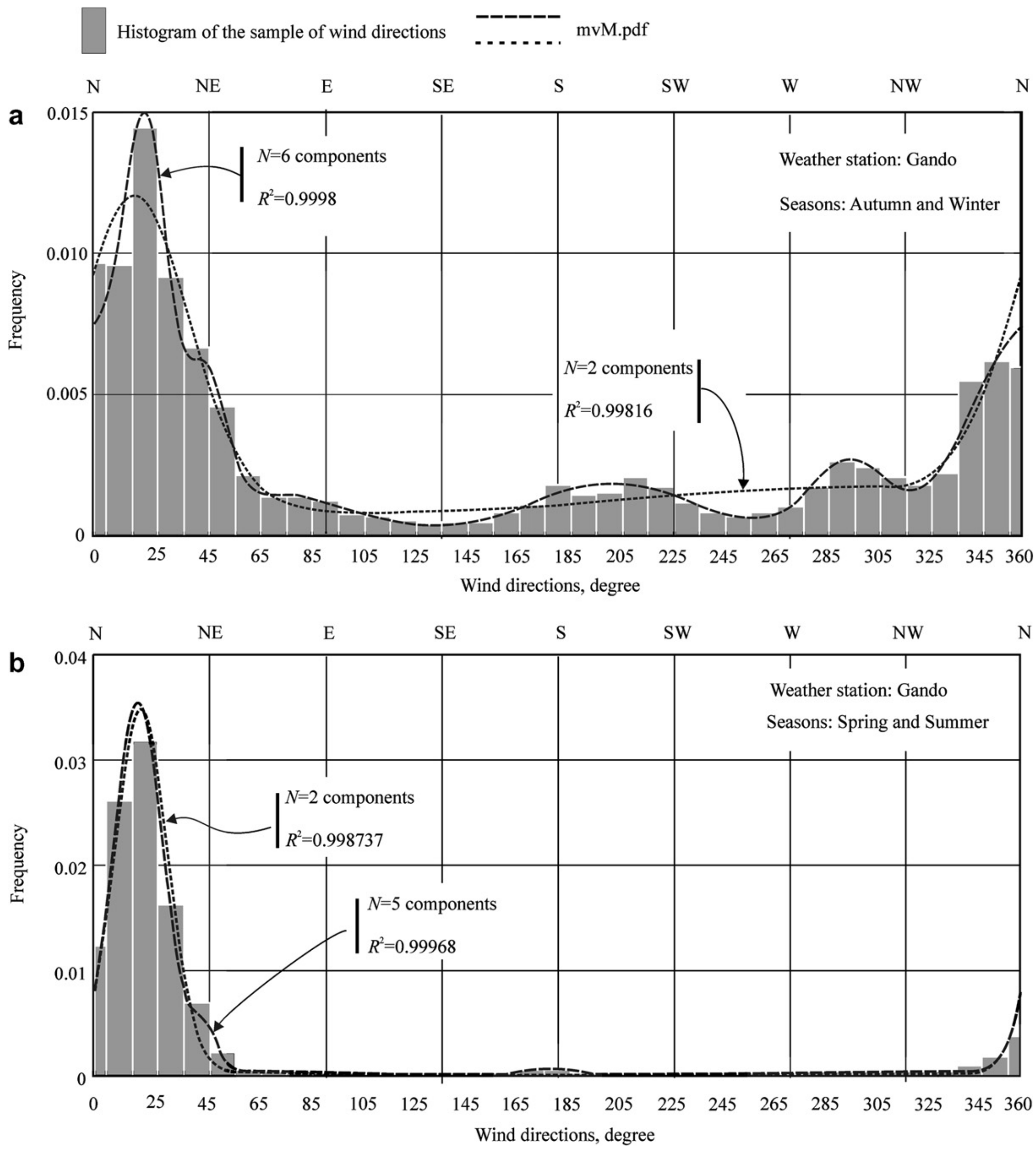

Fig. 6. Frequency histograms of wind directions at Gando station: (a) months of interval I, (b) months of interval S.

Table 1
Numerical values of the parameters of the mvM-pdfs at Amagro in different seasonal periods

j	Autumn and Winter						j	Spring and Summer					
	$N = 4$			$N = 2$				$N = 5$			$N = 2$		
	μ_j rad	κ_j –	ω_j –	μ_j rad	κ_j –	ω_j –		μ_j rad	κ_j –	ω_j –	μ_j rad	κ_j –	ω_j –
1	0.732	2.796	0.106	1.385	6.839	0.735	1	1.182	7.406	0.359	1.283	24.406	0.741
2	1.455	12.067	0.577	4.313	1.617	0.265	2	1.327	40.284	0.51	1.564	0.847	0.259
3	3.892	1.034	0.189				3	2.97	5.106	0.079			
4	4.352	30.464	0.074				4	4.644	5.269	0.029			
							5	6.271	16.656	0.024			

Table 2
Numerical values of the parameters of the mvM-pdfs at Gando in different seasonal periods

j	Autumn and Winter						j	Spring and Summer					
	$N = 6$			$N = 2$				$N = 5$			$N = 2$		
	μ_j rad	κ_j –	ω_j –	μ_j rad	κ_j –	ω_j –		μ_j rad	κ_j –	ω_j –	μ_j rad	κ_j –	ω_j –
1	0	7.817	0.366	0.29	7.512	0.579	1	0.299	34.416	0.86	0.312	30.533	0.899
2	0.357	51.813	0.209	5.021	0.381	0.421	2	0.747	144.385	0.045	0.631	0.578	0.101
3	0.758	53.947	0.086				3	1.004	3.002	0.027			
4	1.299	4.928	0.093				4	3.081	44.976	0.013			
5	3.512	3.386	0.149				5	5.673	1.163	0.055			
6	5.11	14.489	0.097										

north, mean hourly wind direction data from 7 years (1997–1999, 2001–2003, 2005) have been used . For the station called Gando [3], located in the east, data have been used from 4 years (2001–2002, 2004–2005). All the wind direction records were taken at a height of 10 m above ground level.

Because of its latitudinal situation and the proximity of the anticyclone of the Azores, the Canarian Archipelago is affected almost all year round by the northeast trade winds. During summer and much of spring, the frequency of the trade winds regime is very high, from 90% to 95% of the time during summer,[10] though in winter,[11] its frequency drops to about 50%.

In addition to the trade winds, other winds blow in the Archipelago, not constantly but with a certain local regularity. These include notably the Saharan winds of southerly and easterly components, which usually make their appearance in autumn and spring, and the tropical winds of southwesterly and westerly components, which blow in winter periods especially.

As the directional wind speed in the Canarian Archipelago has a seasonal behaviour, the model proposed will be applied to data recorded at the two previously mentioned stations and for the following two monthly intervals: winter and autumn months (January–March and October–December), which we will call Interval I, and spring and summer months (April–September), which we will call Interval S. In Fig. 1, we can see the horizontal components v_x and v_y of the wind speed at the two stations under study for both Interval I and Interval S.

7. Analysis of results

Fig. 2 shows, for the case of Amagro, the values obtained for the coefficient of determination R^2 as a function of the number N of vM-pdfs which comprises the mixture distribution. This relation is shown for the months of Interval I and Interval S. Fig. 3 is a similar representation for the Gando station.

It can be observed from Figs. 2 and 3 that, independently of the type of month interval considered, when increasing the number of components N of the mixture distribution, the value of the coefficient of determination R^2 increases. However, the variations in R^2 are not pronounced for values of N greater than six.

It can also be seen in these figures that mixtures of two vM-pdfs provide better fits in all the cases analysed than those obtained with the models proposed by Smith [10,11] in the specialised literature on wind energy. Despite this, the model proposed by Smith [10,11] can provide an acceptable visual fit in those periods where there is a single easily observable mode (Fig. 4). However, it should be pointed out that this model requires the use of wind speed and wind direction data, while the model proposed in this paper requires only the use of wind direction data.

Fig. 5 shows the wind direction frequency histograms for the Amagro station during the interval I months (Fig. 5a) and the interval S months (Fig. 5b). Fig. 6 shows the corresponding histograms for the Gando station.

For purposes of comparison, two mvM-pdfs are presented in both Figs. 5 and 6. The number of components N of one of the distributions has been chosen with the criterion that higher values of it do not entail appreciable increases of R^2. The number of components of the other distribution is the minimum to make a mixture distribution, that is $N = 2$.

It can be seen in Figs. 5b and 6b that in the interval S months, due to the existence of just one clearly prevailing wind direction (a single easily observable mode), the use of a mvM-pdf with $N = 2$ at Amagro (Table 1) and at Gando (Table 2) is sufficient to achieve a significant visual fit. With the use of a mvM-pdf with $N = 5$, values of R^2 are achieved very close to 0.9999.

However, in Figs. 5a and 6a, which represent the interval I months, due to the influence of various wind directions, mvM-pdfs with $N = 2$ present a poorer visual fit. To achieve values of R^2 close to 0.9999, mvM-pdfs with $N = 4$ are required at Amagro (Table 1) with two modes or prevailing wind directions, and with $N = 6$ at Gando (Table 2), which has more modes or prevailing wind

[10] In summer, the anticyclone is situated further from the Canary Islands, in the Azores, and, thus, the action of the trade winds is more intense.

[11] In winter, the nucleus of the anticlone is found very near to the Canary Islands, in Madeira, and so, the action of the trade winds is less important.

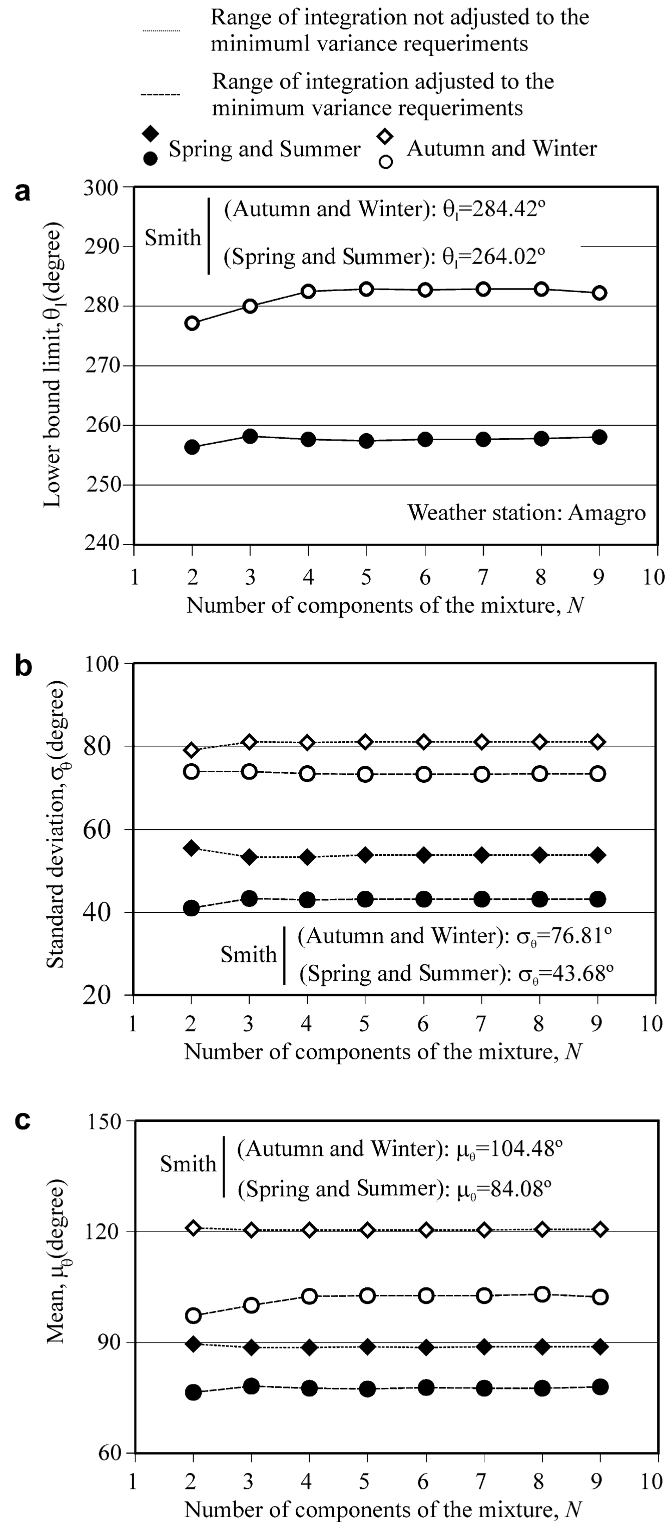

Fig. 7. Lower bound of the range of integration (a), angular standard deviation (b) and mean direction (c), as a function of the number of components of the mixture, at Amagro station.

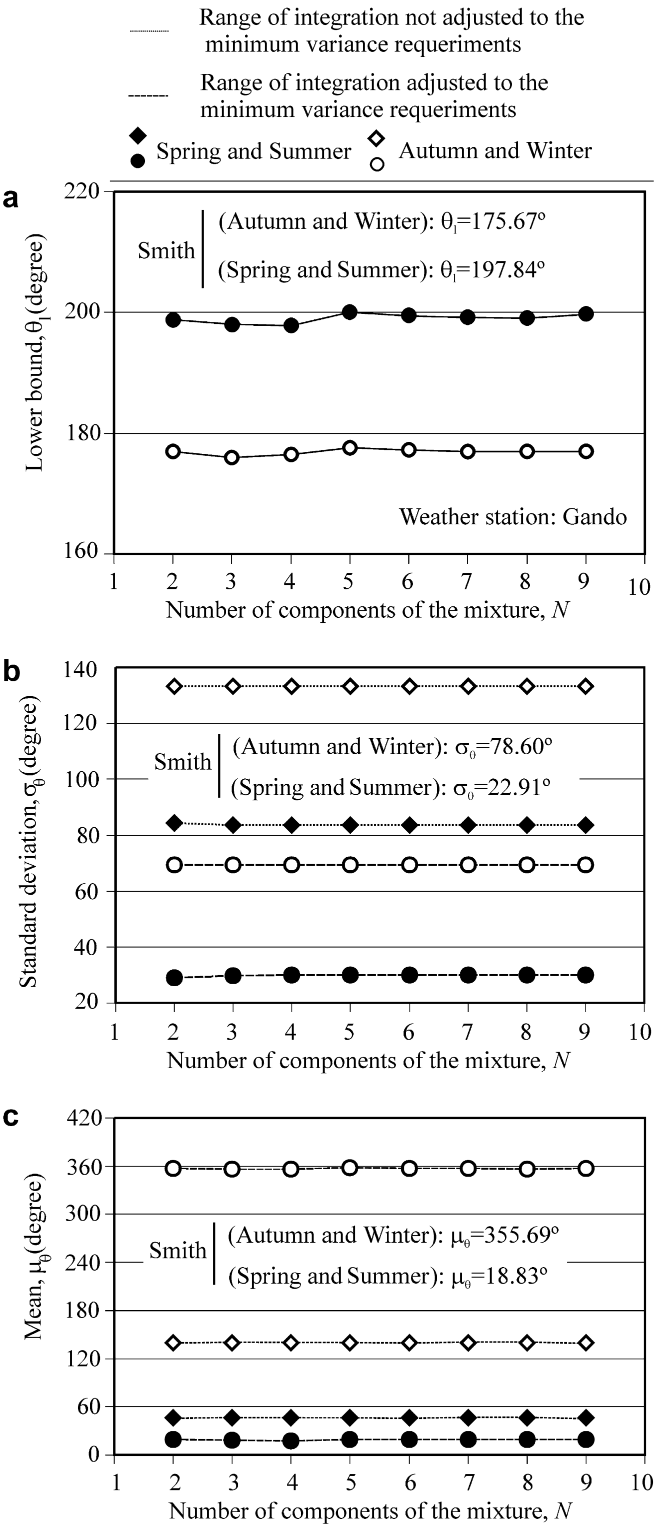

Fig. 8. Lower bound of the range of integrations (a), angular standard deviation (b) and mean direction (c), as a function of the number of components of the mixture, at Gando station.

directions. It should be pointed out that although these mixed distributions with several components enrich the modelling and enable high degrees of fits, they have the added complication of an increase in the number of parameters involved.

Fig. 7 shows, for Amagro, the relation between the number of components *N*of the mixture distribution and the lower angle of the range of integration (Fig. 7a), the standard deviation of the wind direction (Fig. 7b) and its mean angle (Fig. 7c). Fig. 8 is a similar representation for Gando station. The method explained in Section 5 was used to make these representations, and as an example (Fig. 9), the case of Gando is shown for the months of interval I and for $N = 6$. In all the cases analysed, we have found, as shown in Fig. 9, a sole lower bound value of integration that provides the minimum standard deviation and, consequently, a sole value of the mean angle of wind direction.

It can be seen from Figs. 7 and 8 that at the lower angles of the range of integration, the angular standard deviation and the mean angle vary slightly with N, mainly for values of N greater than three. However, for all values of N, there are notable discrepancies existing between the angular standard deviations and the mean angles estimated through the use of a range of integration that is adjusted to minimum variance requirements and those estimated without taking into account this range of integration. This shows that the approach made by McWilliams and Sprevak [14] was incorrect when they calculated the moments of the angles without considering the range of integration adjusted to minimum variance requirements.

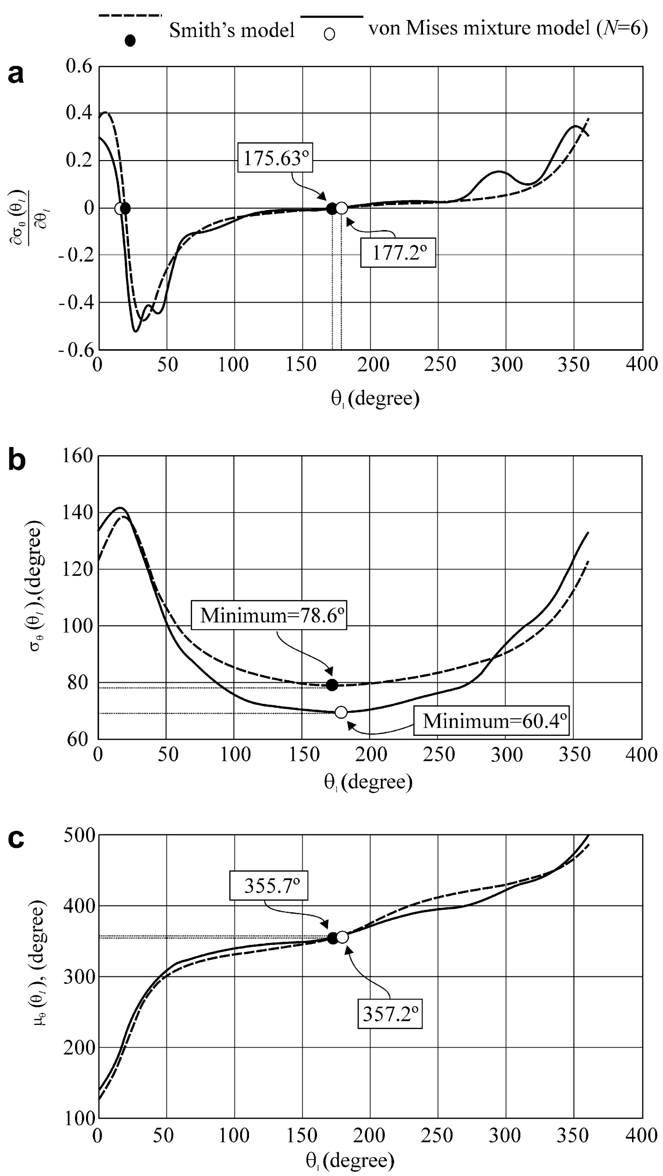

Fig. 9. Derivative of the standard deviation with respect to the lower bound of the range of integration (a), standard deviation as a function of the lower bound of the interval of integration (b) and mean direction as a function of the lower bound of the interval of integration (c), for Gando station in the months interval I.

8. Conclusions

The conclusion reached is that the mvM-pdfs presented in this paper provide a very flexible model for wind speed direction studies and can be applied in a widespread manner to represent the wind direction regimes in zones with one or several prevailing wind directions.

As a result of the application of the mixture model proposed in this paper to the data for wind direction recorded at two weather stations in the Canarian Archipelago, the following conclusions are drawn: (a) independently of the type of month interval considered, when increasing the number of components N of the mixture distribution, the value of the coefficient of determination R^2 increases. However, the variations in R^2 are not pronounced for values of N greater than six; (b) two vM-pdfs provide better fits in all the cases analysed than those obtained with the models proposed by Smith [10,11] in the specialised literature on wind energy; (c) there are notable discrepancies existing between the angular standard deviations and the mean angles estimated through the use of a range of integration that is adjusted to minimum variance requirements and those estimated without taking into account this range of integration.

References

[1] Ramírez P, Carta JA. Influence of the data sampling interval in the estimation of the parameters of the Weibull wind speed probability density distribution: a case study. Energy Convers Manage 2005;46:2419–38.

[2] Ramírez P, Carta JA. The use of wind probability distributions derived from the maximum entropy principle in the analysis of wind energy a case study. Energy Convers Manage 2006;47:2564–77.

[3] Carta JA, Ramírez P. Use of finite mixture distribution models in the analysis of wind energy in the Canarian Archipelago. Energy Convers Manage 2007;48:281–91.

[4] Chadee JC, Sharma C. Wind speed distributions: a new catalogue of defined models. Wind Eng 2001;25:319–37.

[5] Salmon JR, Walmsley JL. A two-site correlation model for wind speed, direction and energy estimates. J Wind Eng Ind Aerod 1999;79:233–68.

[6] Lange B, Højstrup J. Evaluation of wind-resources estimation program WAsP for offshore applications. J Wind Eng Ind Aerod 2001;89:271–91.

[7] García-Rojo R. Algorithm for the estimation of the long-term wind climate at a meteorological mast using a joint probabilistic approach. Wind Eng 2004;28:213–24.

[8] Matsu M, Ishihara T, Hibi K. Directional characteristics of probability distribution of extreme wind speeds by typhoon simulation. J Wind Eng Ind Aerod 2002;90:1541–53.

[9] Ettoumi FY, Sauvageot H, Adane AEH. Statistical bivariate modelling of wind using first-order Markov chain and Weibull distribution. Renew Energy 2003;28:1787–802.
[10] Smith OE. An application of distributions derived from the bivariate normal density function, In: Proceedings of the international symposium on probability and statistics in the atmospheric sciences; 1971;162–8.
[11] Justus CG. Winds and wind system performance. 1st ed. US: Franklin Institute Press; 1978.
[12] Koeppl GW. Putnam's power from the wind. 2nd ed. New York: Van Nostrand Reinhold; 1982.
[13] McWilliams B, Newmann MM, Sprevak D. The probability distribution of wind velocity and direction. Wind Eng 1979;3: 269–273.
[14] McWilliams B, Sprevak D. The estimation of the parameters of the distribution of wind speed and direction. Wind Eng 1980;4: 227–38.
[15] Weber R. Estimator for the standard deviation of wind direction based on moments of the Cartesian components. J Appl Meteorol 1991;30:1341–53.
[16] Rao S, SenGupta A. Topics in circular statistics. 1st ed. Singapore: World Scientific; 2001.
[17] Mardia KV, Jupp PE. Directional statistics. 2nd ed. New York: John Wiley and Sons; 2000.
[18] Fisher NI. Statistical analysis of circular data. 2nd ed. Cambridge: Cambridge University Press; 1995.
[19] Canavos G. Applied probability statistical methods. 1st ed. New York: Little Brown and Company; 1998.
[20] Draper NR, Smith H. Applied regression analysis. 1st ed. New York: John Wiley and Sons Inc; 1998.
[21] Titterington DM, Smith AFM, Makov UE. Statistical analysis of finite mixture distributions. 2nd ed. New York: John Wiley and Sons Inc; 1995.
[22] McLachlan G, Peel D. Finite mixture models. 1st ed. New York: Wiley; 2000.
[23] Zhang S, Jin J. Computation of special functions. 1st ed. New York: Wiley; 1996.
[24] Carta JA, Ramírez P. Analysis of two-component mixture Weibull statistics for estimation of wind speed distributions. Renew Energy 2007;32:518–31.
[25] Kiryanov D. The Mathcad 2001i handbook. 1st ed. USA: Delmar Thomson Learning; 2002.
[26] Calero R, Carta JA. Action plan for wind energy development in the Canary Islands. Energy Policy 2004;32:1185–97.