



OPEN

Enhancing wind direction prediction of South Africa wind energy hotspots with Bayesian mixture modeling

Najmeh Nakhaei Rad^{1,2,✉}, Andriette Bekker¹ & Mohammad Arashi^{1,3}

Wind energy production depends not only on wind speed but also on wind direction. Thus, predicting and estimating the wind direction for sites accurately will enhance measuring the wind energy potential. The uncertain nature of wind direction can be presented through probability distributions and Bayesian analysis can improve the modeling of the wind direction using the contribution of the prior knowledge to update the empirical shreds of evidence. This must align with the nature of the empirical evidence as to whether the data are skew or multimodal or not. So far mixtures of von Mises within the directional statistics domain, are used for modeling wind direction to capture the multimodality nature present in the data. In this paper, due to the skewed and multimodal patterns of wind direction on different sites of the locations understudy, a mixture of multimodal skewed von Mises is proposed for wind direction. Furthermore, a Bayesian analysis is presented to take into account the uncertainty inherent in the proposed wind direction model. A simulation study is conducted to evaluate the performance of the proposed Bayesian model. This proposed model is fitted to datasets of wind direction of Marion island and two wind farms in South Africa and show the superiority of the approach. The posterior predictive distribution is applied to forecast the wind direction on a wind farm. It is concluded that the proposed model offers an accurate prediction by means of credible intervals. The mean wind direction of Marion island in 2017 obtained from 1079 observations was 5.0242 (in radian) while using our proposed method the predicted mean wind direction and its corresponding 95% credible interval based on 100 generated samples from the posterior predictive distribution are obtained 5.0171 and (4.7442, 5.2900). Therefore, our results open a new approach for accurate prediction of wind direction implementing a Bayesian approach via mixture of skew circular distributions.

The future of the energy industry lies in clean power that minimizes or entirely removes pollutants from the process of power generation. The perfect clean energy mix occurs where green energy, derived from natural sources, meets renewable energy from sources that are constantly being replenished. Wind energy is one of the most important sustainable forms of this ideal clean energy and one of the fastest-growing energy sources. A sophisticated knowledge, based on statistical analysis, of wind characteristics is crucial for the future harnessing of this important renewable energy resource. Wind power is developing as a renewable energy source in a number of countries and it will be increasingly important to find an effective and predictable way of integrating this intermittent but environmentally friendly power source into the existing electrical grid system.

In South Africa, there is an increasing transition towards an environmentally sustainable, climate-change resilient, low-carbon economy. In October 2020, the South African Wind Energy Association (SAWEA) reported that wind technology has already attracted R209.7 billion in investment for the development of projects in South Africa. In fact wind power comprises a larger share of the planned renewable energy investments to date. It is estimated by 2030 that 22.7% of the required electricity in South Africa, namely 17742 MW, will be generated from wind energy. In terms of job creation, the 22 wind Independent Power Producers (IPPs) that have successfully reached commercial operations to date, have created 2723 jobs for South African citizens.

¹Department of Statistics, University of Pretoria, Pretoria 0002, South Africa. ²DSI-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS), Pretoria, South Africa. ³Department of Statistics, Faculty of Mathematical Sciences, Ferdowsi University of Mashhad, Mashhad, Iran. ✉email: najmeh.nakhaeirad@up.ac.za

Wind as an energy source is only practical in areas that have strong and steady winds. South Africa's climatology allows for significant wind energy production especially along the coastal areas of the Eastern and Western Capes. The first large-scale wind farm in South Africa became operational in 2014 and based on the SAWEA report, there are 33 wind farms: 22 fully operational and 11 in construction. In this paper, we will study the wind direction of two operational wind farms in South Africa: (1) Jeffreys Bay (Humansdorp), located in the Eastern Cape; (2) Noupoot located in the Northern Cape. In addition, we will investigate the wind direction data from Marion Island, part of the Western Cape Province which possesses excellent potential for wind studies.

Unlike conventional energy resources that are available at any time, wind speed and wind direction need to be forecasted in advance in order to estimate production and plan its contribution to a nation's grid system. As the use of wind power increases, accurate forecasts are essential to maximize output from the wind farms. This includes the most important decision of all, the location of a wind farm and the placement of its turbines¹.

The location of an industrial-scale wind farm, defined as a cluster of wind turbines used to produce electricity, is of paramount importance. Measuring the farm-specific wind characteristics including mean wind speed, wind speed distribution (diurnal, seasonal, annual patterns), distribution of wind direction, short-term fluctuations, long-term fluctuations and wind shear profile are essential for determining the location of farm and turbines. This can strongly influence the performance of the wind turbines and thus the power generated by the wind farms². Moreover, interactions among multiple turbines change the power generation efficiency of turbines. Specifically, the wakes from upwind turbines can greatly affect the power production of downstream turbines, and this effect depends strongly on the wind direction³. Generally, downstream turbines produce less power compared to upwind turbines, but changes in wind direction can cause heterogeneity in the power curve of each turbine such that some upstream turbines can become downstream turbines⁴. Porté-Agel et al.⁵ presented a study about the effects of wind direction on turbine wakes and power losses at a large wind farm. Castellani et al.⁶ showed how the alignment of wind turbines to wind direction affects efficiency (see also Kazacoks et al.⁷ and Gomez and Lundquist⁸).

Predicting wind speed and wind direction are crucial to choose the location of wind farm and the placement of its turbines and also to estimate wind power production. To the best of the authors' knowledge none of the existing literature follows a directional statistics approach for prediction of the wind direction. The interested reader is referred to some contributions in which several approaches have been proposed for forecasting wind direction. El-Fouly et al.²⁰ suggested a linear time-series-based model for prediction of wind speed and direction. Garcia-Planas and Gongadze²¹ constructed a predictive model for wind speed and direction based on linear Markov chains under linear algebra point of view (see also Zeng et al.²², Fan et al.²³, Zheng et al.²⁴, Chen et al.²⁵, Liu et al.²⁶, Giangregorio et al.²⁷, Wang et al.²⁸). Note however that this paper approaches skew directional models from the Bayesian statistical angle.

Circular statistics can be applied to obtain the distribution of wind direction while Weibull, gamma, normal, Rayleigh, log-normal, inverse Gaussian, logistic distributions are some common models for the wind speed (see Deep et al.⁹ and Gugliani et al.¹⁰). For example, mixtures of von Mises (VM) distributions have been widely applied to model wind direction for different locations^{10–17}. Gugliani et al.¹⁸ have applied Kato and Jones circular distribution¹⁹ to model wind direction.

However, wind datasets usually exhibit skew and multimodal patterns while most of the well-known circular distributions are symmetric such as the von Mises. Therefore in this paper, the application of skewed multimodal distributions is investigated for modeling the wind direction of South Africa hotspots from Bayesian viewpoint. The k sine-skewed von Mises (SSVM) distribution²⁹ and mixtures of SSVM are ideal candidates to model wind direction data exhibiting both skewness and multimodality behaviour. Due to the fact that the likelihood-based inference and also the expectation maximization (EM) algorithm techniques for mixture models can be computationally complicated, a Bayesian approach can overcome such computational difficulties. It provides more accurate results for small datasets. Bayesian inference is conditional on the data and is exact, without reliance on asymptotic approximations. The Bayesian predictive posterior function can be used to forecast the wind direction.

Two important contributions of a Bayesian stochastic model are as follows:

(1) Inclusion of uncertainty about the parameters of the wind direction distribution results in using a more practical predictive distribution for the wind direction. This implies the predictive distribution is more disperse than the probability distributions when the uncertainty about the parameters is neglected. (2) The prior distributions of the parameters can represent the heterogeneity of the distributions of the wind direction over a wind farm. The wind direction distributions for various turbines on a farm may belong to the same family, such as the skew-von Mises, but the model parameters of each turbine may be different randomly according to some probability distributions. The Bayesian predictive distribution aggregates the non-homogeneous distributions into a single distribution that captures the variation among the probability distributions of the wind directions at the turbines' locations on a wind farm.

There is a vast literature on the Bayesian approach for symmetric directional data specifically, Bayesian analysis using the symmetric von Mises and von Mises-Fisher distributions^{30–39}. The von Mises-Fisher mixture model is implemented by Taghia et al.⁴⁰ and Roge et al.⁴¹. Mulder et al.⁴² provided a Bayesian inference for mixtures of von Mises distributions using the reversible jump Markov chain Monte Carlo (MCMC) sampler and focused on noninformative priors. From the preceding it follows there is a gap in the literature that inspired us to propose novel Bayesian analysis of *skew directional wind data*. Recently Nakhaei Rad et al.^{43,44} provided Bayesian analysis for skew von Mises-Fisher distribution and skew Wrapped Cauchy mixture model.

In “[Site location and wind data](#)”, we provide details of the datasets that are analyzed in this paper. “[Materials and methods](#)” revisited the k sine-skewed von Mises distribution and the maximum likelihood estimates (MLEs) of the mixture of SSVM parameters. The Bayesian inference of the mixture of SSVM is also presented, followed by the posterior predictive distribution to forecast the wind direction. In “[Evaluation and results](#)”, a simulation study is conducted to show the performance of the proposed Bayesian approach. Finally, SSVM and mixture of



Figure 1. Marion island (created by the University of Pretoria) and meteorological research station on the island (provided by Antarctic Legacy of South Africa <http://www.antarcticlegacy.org> and <https://blogs.sun.ac.za>).

SSVM are fitted to these datasets for different values of k together with their competitor, namely the mixture of von Mises distributions.

Site location and wind data

The first dataset (A) shows the wind direction of Marion island which is recorded daily at 08:00, 14:00 and 20:00 South Africa standard time (SAST) (relates to the main synoptic hours). Marion Island is part of South Africa with a climate that is highly oceanic in nature, coupled with the influence of passing frontal weather systems. In fact, the geographic location of Marion Island, lying directly in the path of eastward moving depressions all year round make it an excellent location for meteorological studies. Powerful regional winds, colloquially known as the ‘Roaring Forties’, so called as they have found between the latitudes of 40° and 50° in the Southern Hemisphere, blow almost every day in a north-westerly direction. The exceptional research potential of Marion Island for wind studies, as well the rate and impacts of climate change, is demonstrated by the presence of a permanent meteorological research station on the island. This station was established as early as 1948, and run by the South African National Antarctic Programme (see Fig. 1).

The second dataset (B) reflects the wind direction of Jeffreys Bay wind farm, recorded every 10 min at 60 m height. Jeffreys Bay is one of the biggest wind farms in South Africa spanning 3700 hectares with a 138 MW capacity. This site’s optimal wind conditions, relatively flat topography, minimal environmental constraints and its close proximity to the Eskom (electricity supply commission of South Africa) grid line, make it an ideal wind energy resource (see Fig. 2, left).

The last dataset (C) shows wind direction of Noupoort wind farm comprising 7500 hectares and providing a 80 MW capacity, recorded every 10 min at 20 m height. This site is significant because of the excellent wind conditions, its proximity to national roads for wind turbine transportation, the favourable construction conditions, municipality and local stakeholder support and the straightforward electrical connection into the Eskom grid (see Fig. 2, right). Figure 3 shows the map of South Africa with the locations of Marion island, Jeffreys Bay and Noupoort wind farms and rose plots of the wind direction in these regions.

Table 1 shows the descriptive information about the datasets. The results in Table 1, confirm skewness presence in these datasets. Also the Boxplots and kernel density plots of these datasets in Fig. 4. The Boxplots emphasize that these wind direction datasets reveal skew patterns and the kernel density plots confirm multimodal patterns. kernel density estimate is a smoothed version of the histogram which is a useful alternative to the histogram for continuous data. Unlike the histogram, the kernel technique produces a smooth estimate of the density function, uses all sample points’ locations and more convincingly suggests multimodality.

Materials and methods

Sine-skewed von Mises distribution. Most of the distributions on the unit circle share the common feature of being symmetric about their location $\mu \in [-\pi, \pi)$. However, since the assumption that data is symmetric is often rejected, Ref.²⁹ introduced the k sine-skewed von Mises distribution with density function

$$f_{SSVM}(\theta; \mu, \tau, \lambda) = \frac{1}{2\pi I_0(\tau)} \exp(\tau \cos(\theta - \mu))(1 + \lambda \sin(k(\theta - \mu))), \quad (1)$$

where $I_0(\cdot)$ is the modified Bessel function of the first kind of order 0, $\mu \in [-\pi, \pi)$ is the location parameter, $\tau > 0$ is the concentration parameter, $-1 \leq \lambda \leq 1$ is the skewness parameter and k is a positive integer. $\lambda > 0$



Figure 2. Jeffreys Bay (Humansdorp) wind farm <https://jeffreysbaywindfarm.co.za> (left) and Noupoot wind farm <https://noupootwind.co.za> (right).

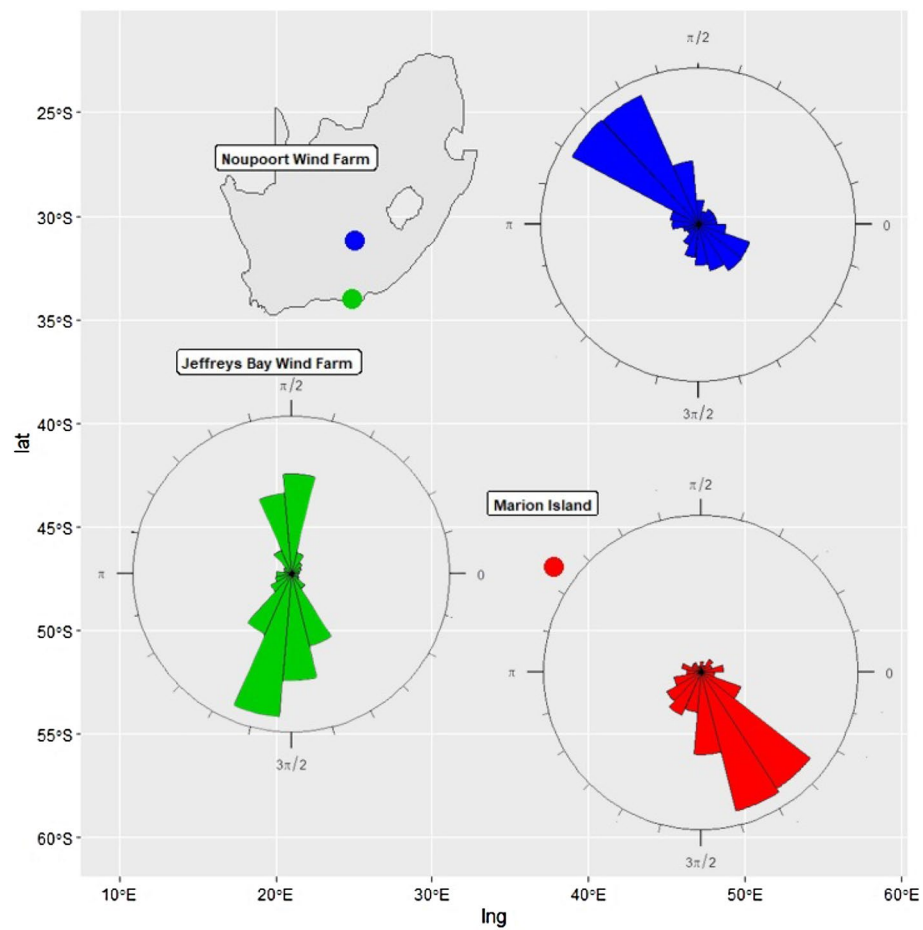


Figure 3. Map of South Africa with the locations of Marion island, Jeffreys Bay and Noupoot wind farms and rose plots of the wind direction (created by R programming language version 4.1.3 <https://www.r-project.org>).

Id	Location	Begin	End	Duration (days)	n	Mean	Variance	Mean resultant length	Skewness	Kurtosis
A	Marion	01-Jan-2017	31-Dec-2017	365	1079	5.0242	0.4376	0.5624	0.4039	0.9686
B	Jeffreys Bay	01-Jan-2019	31-Jan-2019	31	4464	4.3498	0.7720	0.2279	0.5051	0.8084
C	Noupoort	01-Feb-2019	29-Feb-2019	29	4032	2.3351	0.7923	0.2076	- 0.1160	0.7220

Table 1. Descriptive statistics for the wind direction data.

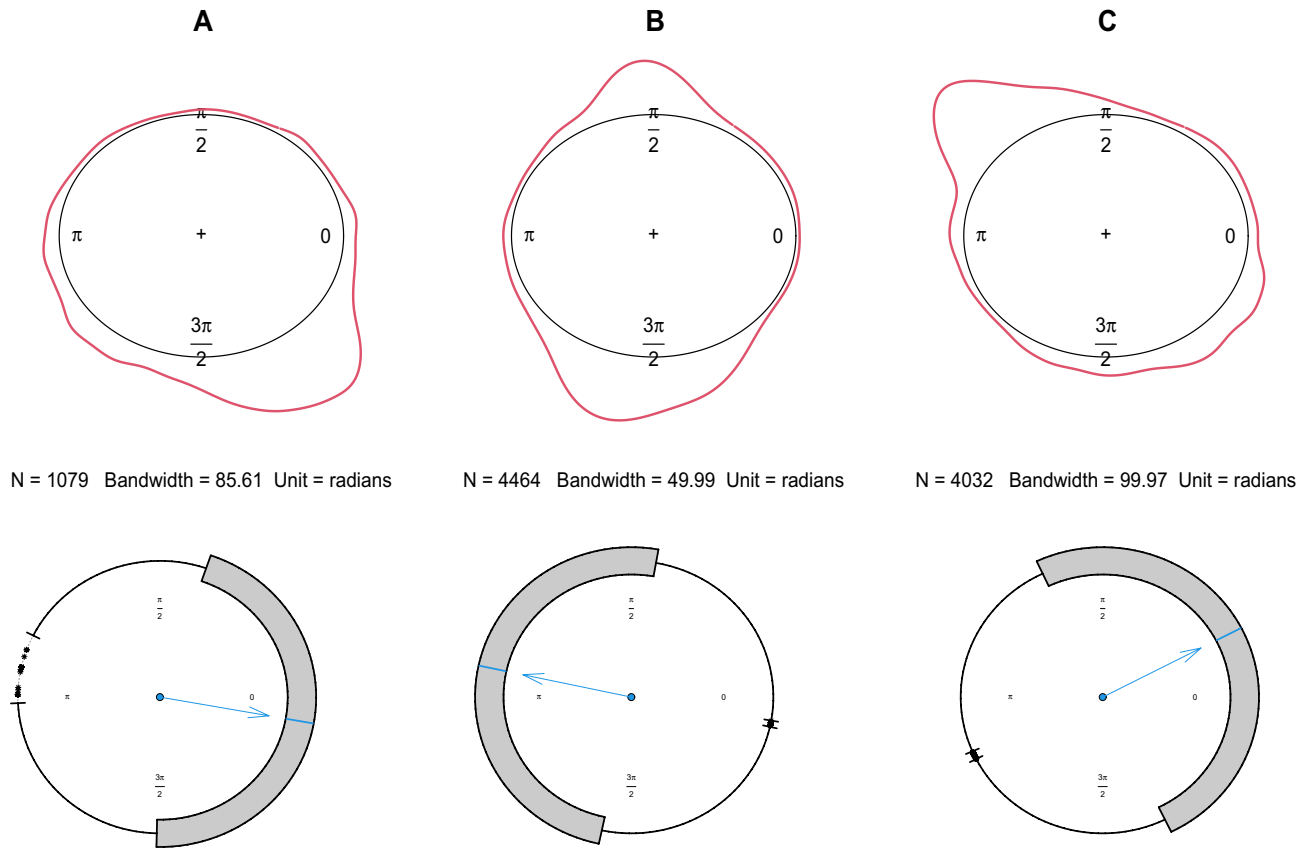


Figure 4. Boxplots and kernel density plots of the wind direction datasets A-C from Marion island, Jeffreys Bay and Noupoort wind farms.

leads to left skewed distributions and $\lambda < 0$ provides right skewed distributions. The symmetric von Mises distribution is retrieved if $\lambda = 0$. For $k \geq 2$, (1) has a multimodal form but for $k = 1$ it can be both unimodal and bimodal. Figure 5 shows plots of SSVM density functions (see (1)) for $\mu = 0$, $\tau = 0.5$, $\lambda = -0.8, -0.2, 0.5, 1$ and $k = 1, 2$. As can be seen with $k = 2$ bimodal distributions follows. A mixture of SSVM distributions with $M \in \mathbb{Z}^+$ components is expressed as

$$f_M(\theta; \mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda}) = \sum_{j=1}^M w_j f_{SSVM}(\theta; \mu_j, \tau_j, \lambda_j), \tag{2}$$

where $\boldsymbol{\mu} = (\mu_1, \dots, \mu_M)$, $\boldsymbol{\tau} = (\tau_1, \dots, \tau_M)$ and $\boldsymbol{\lambda} = (\lambda_1, \dots, \lambda_M)$ are vectors of parameters, $\tau_j > 0$, $\mu_j \in [-\pi, \pi]$ and $\lambda_j \in [-1, 1]$. $\mathbf{w} = (w_1, \dots, w_M)$ is a vector of the weights containing the relative size of each component in the total sample satisfy the constraints $0 \leq w_j \leq 1$ and $\sum_{j=1}^M w_j = 1$.

Algorithm 1⁴⁵ can be used to generate a sample from the SSVM distribution in (1).

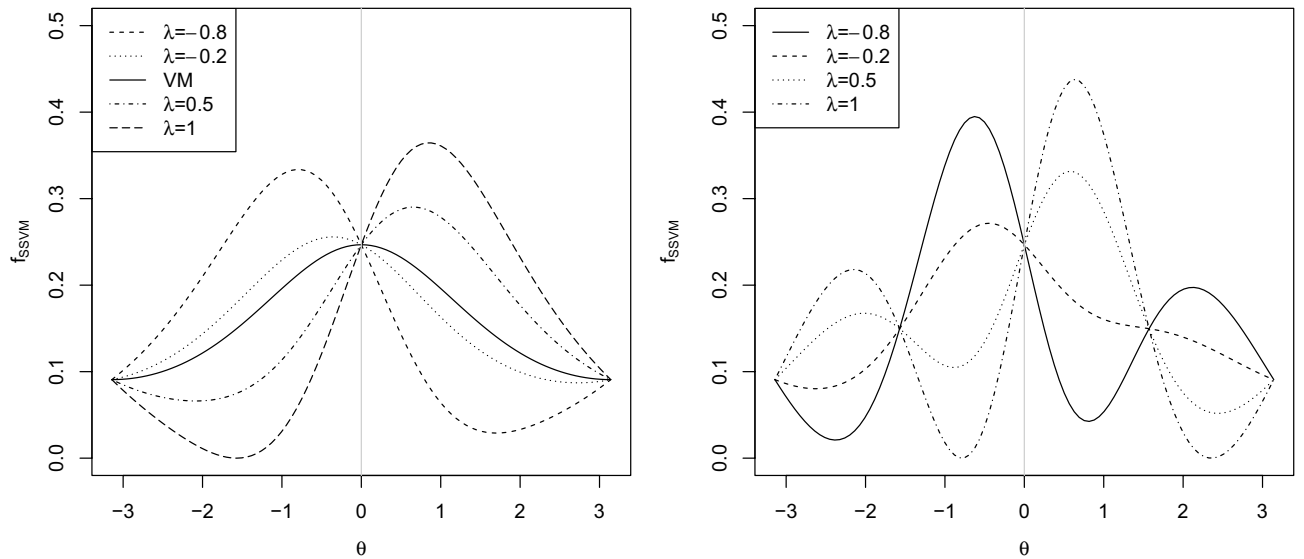


Figure 5. Density functions of the SSVM for $\tau = 0.5, \mu = 0, \lambda = -0.8, -0.2, 0.5, 1$ and $k = 1$ (left) and $k = 2$ (right).

Algorithm 1 Steps to generate samples from SSVM.

- 1: Generate Θ_{VM} from von Mises distribution with parameter μ and κ .
- 2: Generate U from $U(0,1)$.
- 3: If $U \leq (1 + \lambda \sin(\Theta - \mu))/2$, then $\Theta_{SSVM} = \Theta_{VM}$. Otherwise $\Theta_{SSVM} = -\Theta_{VM}$.

Parameter estimation. In this section, first, the MLEs of parameters for a mixture of SSVM is presented, followed by a Bayesian inference when all the weight, location, concentration and skewness parameters (w, μ, τ, λ) are unknown.

Maximum likelihood estimation. The log-likelihood function of a mixture of SSVM in (2), can be represented as follows:

$$l(w, \mu, \tau, \lambda | \theta) = \sum_{i=1}^n \log \left(\sum_{j=1}^M w_j f_{SSVM}(\theta_i; \mu_j, \tau_j, \lambda_j) \right). \tag{3}$$

By setting the partial derivatives of (3) with respect to (w, μ, τ, λ) to zero, the MLEs of (w, μ, τ, λ) can be obtained. Since no closed-form expressions exist, numerical methods should be used to obtain the estimates. The DEoptim package⁴⁶ in R software which is based on the Differential Evolution (DE) algorithm⁴⁷ is used to obtain the MLEs. Differential evolution is a heuristic evolutionary method for global optimization that is effective in many problems of interest in science and technology and its significant performance as a global optimization algorithm on continuous numerical minimization problems has been extensively studied⁴⁸. DEoptim has made this algorithm possible to easily apply in the R language and environment. DEoptim relies on repeated evaluation of the objective function in order to move the population toward a global minimum⁴⁶.

Bayes estimation. Let $\theta = (\theta_1, \theta_2, \dots, \theta_n)$ be a random sample of size n from a mixture of SSVM (see (2)). It should be noted that the number of components M is considered as a known parameter. Suppose the latent variable $d = (d_1, \dots, d_n)$ allocates the component that θ is sampled from. The probability of being attributed to component j is given by

$$P(d_i = j | w) = w_j.$$

Therefore, for $i = 1, \dots, n$ and $j = 1, \dots, M$

$$f(\theta_i | d_i = j) = f_{SSVM}(\theta_i; \mu_j, \tau_j, \lambda_j).$$

It implies that conditional on d_i , θ_i is an independent observation from its respective component j that makes the inference easier because the problem reduces to inference for a single SSVM component. Therefore, conditional on \mathbf{d} , the likelihood function can be expressed as

$$L(\boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda} | \boldsymbol{\theta}, \mathbf{d}) = \prod_{i=1}^n f_{SSVM}(\theta_i; \mu_{d_i}, \tau_{d_i}, \lambda_{d_i}). \tag{4}$$

Subsequently, we measure the uncertainty in the parameters with the following prior distributions for $(\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda})$. If the sample size is small, or available data provides only indirect information about the parameters of interest, the prior distribution becomes more important⁴⁹. Ghaderinezhad et al.⁵⁰ implemented the Wasserstein impact measure (WIM) as a measure of quantifying prior impact. It helps us to choose between two or more given priors. Nakhaei Rad et al.⁴⁴ by using the WIM measure demonstrated that the combination of the von Mises, gamma and truncated normal distributions decreases the execution time in the Gibbs sampling algorithm. Thus, providing accurate parameter estimates for the skew Fisher-von Mises distribution⁵¹ as well.

Therefore, consider independent von Mises and gamma distributions with parameters $(\boldsymbol{\mu}_0, \boldsymbol{\tau}_0)$ and $(\boldsymbol{\alpha}, \boldsymbol{\beta})$ as priors for $\boldsymbol{\mu}$ and $\boldsymbol{\tau}$, respectively:

$$\pi(\mu_j, \tau_j; \mu_{0j}, \tau_{0j}, \alpha_j, \beta_j) \propto \exp(\tau_{0j} \cos(\mu_j - \mu_{0j})) \tau_j^{\alpha_j - 1} \exp(-\beta_j \tau_j), \tag{5}$$

where $\tau_{0j}, \alpha_j, \beta_j > 0, \mu_{0j} \in [-\pi, \pi]$ and $j = 1, 2, \dots, M$.

For the skewness parameter $\boldsymbol{\lambda}$, the truncated normal distribution on $[-1, 1]$ is proposed with parameters $\boldsymbol{\xi}$ and $\boldsymbol{\sigma}^2$:

$$\pi(\lambda_j; \xi_j, \sigma_j) = \frac{1}{\sigma_j} \frac{\phi\left(\frac{\lambda_j - \xi_j}{\sigma_j}\right)}{\Phi\left(\frac{1 - \xi_j}{\sigma_j}\right) - \Phi\left(\frac{-1 - \xi_j}{\sigma_j}\right)}, \quad \lambda_j \in [-1, 1]. \tag{6}$$

where $\xi_j \in \mathbb{R}, \sigma_j > 0, j = 1, 2, \dots, M$, $\phi(\cdot)$ is the density function of standard normal distribution and $\Phi(\cdot)$ is its cumulative distribution function.

For the weight parameter \boldsymbol{w} , the Dirichlet distribution with parameter \boldsymbol{c} is considered as prior:

$$\pi(\boldsymbol{w}; \boldsymbol{c}) = \frac{1}{B(\boldsymbol{c})} \prod_{j=1}^M w_j^{c_j - 1}, \tag{7}$$

where $c_j > 0$ for $j = 1, \dots, M$ and $B(\boldsymbol{c}) = \frac{\prod_{j=1}^M \Gamma(c_j)}{\Gamma(\sum_{j=1}^M c_j)}$. Thus the marginal distribution of w_j is $Beta(c_j, \sum_{i=1}^M c_i - c_j)$ ⁵².

Subsequently, the posterior distribution is:

$$\pi(\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda} | \boldsymbol{\theta}) \propto \pi(\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda}) L(\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda} | \boldsymbol{\theta}), \tag{8}$$

with $\pi(\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda})$ from (5), (6) and (7). The full conditionals of parameters $(\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda}, \mathbf{d})$ for using in the Gibbs algorithm follow from (8). Therefore the Gibbs sampler is as follows (see Algorithm 2):

Algorithm 2 Steps to generate samples from posterior distribution.

1: Update \boldsymbol{w} via $\pi(\boldsymbol{w} | \dots) = \frac{1}{B(\boldsymbol{c}')} \prod_{j=1}^M w_j^{c'_j - 1}$ where $\boldsymbol{c}' = \boldsymbol{c} + \sum_{i=1}^n \mathbf{1}_{d_i=j}$.

2: Update $(\boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda})$ via

$$\begin{aligned} \pi(\mu_j | \dots) &\propto \pi(\mu_j) \prod_{d_i=j} f_{SSVM}(\theta_i; \mu_j, \tau_j, \lambda_j), \\ \pi(\tau_j | \dots) &\propto \pi(\tau_j) \prod_{d_i=j} f_{SSVM}(\theta_i; \mu_j, \tau_j, \lambda_j), \\ \pi(\lambda_j | \dots) &\propto \pi(\lambda_j) \prod_{d_i=j} f_{SSVM}(\theta_i; \mu_j, \tau_j, \lambda_j). \end{aligned}$$

3: Update the latent variable \mathbf{d} via $\pi(d_i = j | \dots) = \frac{w_j f_{SSVM}(\theta_i | \mu_j, \tau_j, \lambda_j)}{\sum_{s=1}^M w_s f_{SSVM}(\theta_i | \mu_s, \tau_s, \lambda_s)}$.

For $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots, \theta_n)$, a set of observations and $\boldsymbol{\omega} = (\boldsymbol{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda})$, the posterior predictive distribution for a new data point θ_{new} and d_{new} (the corresponding latent switch variable associated with θ_{new}) is:

$$\pi(\theta_{new} | \boldsymbol{\theta}) = \sum_{d_{new}} \int_{\boldsymbol{\omega}} f(\theta_{new} | d_{new}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda}) p(d_{new} | \boldsymbol{w}) \pi(\boldsymbol{\omega} | \boldsymbol{\theta}) d\boldsymbol{\omega},$$

	Parameter	Actual value	Mean	SD	Q ₁	Median	Q ₃
$n = 500$	μ	3.00	2.9634	0.2267	2.7158	2.9990	3.2771
	τ	2.00	1.9826	0.4958	1.1513	1.9568	3.0635
	λ	0.50	0.4915	0.0084	0.4858	0.4919	0.5012
$n = 100$	μ	3.00	3.1094	0.0443	2.9982	3.1180	3.1753
	τ	2.00	2.0177	0.3556	1.3373	2.0351	2.6634
	λ	0.50	0.4836	0.0240	0.4592	0.4712	0.5342
$n = 50$	μ	3.00	3.1925	0.0380	3.0926	3.1954	3.2491
	τ	2.00	1.9310	0.2712	1.3405	1.9381	2.4485
	λ	0.50	0.5214	0.0220	0.4669	0.5306	0.5390

Table 2. Bayes estimates of parameters of SSVM with prior parameters, $\mu_0 = 0$, $\tau_0 = 0.01$, $\alpha = 4$, $\beta = 2$, $\xi = 0.5$ and $\sigma = 0.01$.

where θ_{new} is independent of the sample data θ . Sometimes the form of $\pi(\theta_{new}|\theta)$ can be derived directly, but it is often easier to sample from $\pi(\theta_{new}|\theta)$ using Monte Carlo methods. For generating an iid sample $(\theta_{new}^{(1)}, \theta_{new}^{(2)}, \dots, \theta_{new}^{(n)})$ from $\pi(\theta_{new}|\theta)$ Algorithm 3 is followed:

Algorithm 3 Steps to generate samples from posterior predictive distribution.

- 1: Sample $(\mathbf{w}^{(t)}, \boldsymbol{\mu}^{(t)}, \boldsymbol{\tau}^{(t)}, \boldsymbol{\lambda}^{(t)})$ from posterior $\pi(\mathbf{w}, \boldsymbol{\mu}, \boldsymbol{\tau}, \boldsymbol{\lambda}|\theta)$ in (8) using Algorithm 2.
- 2: Sample $\theta_{new}^{(t)}$ from $\sum_{j=1}^M w_j^{(t)} f(\theta|\mu_j^{(t)}, \tau_j^{(t)}, \lambda_j^{(t)})$, $t = 1, 2, \dots, n$.

Model selection criteria. Model selection is an important part of any statistical analysis and many tools for selecting the “best model” have been suggested in the literature. Here, three different criteria are applied to evaluate the models. Suppose $\boldsymbol{\omega}$ is the vector of parameters with k elements, $l(\boldsymbol{\omega}|\theta)$ is the log-likelihood function and n is the sample size. The Akaike information criterion (AIC)⁵³ and the Bayesian information criterion (BIC)⁵⁴ as penalized-likelihood criteria are given by

$$\begin{aligned} AIC &= -2l(\boldsymbol{\omega}|\theta) + 2k, \\ BIC &= -2l(\boldsymbol{\omega}|\theta) + k \log n. \end{aligned}$$

As can be seen, BIC penalizes parameters more heavily than AIC. Spiegelhalter et al.⁵⁵ proposed the deviance information criterion (DIC), as

$$DIC = 2\bar{D}(\boldsymbol{\omega}) - D(\bar{\boldsymbol{\omega}}),$$

where $D(\boldsymbol{\omega}) = -2l(\boldsymbol{\omega}|\theta)$, $\bar{\boldsymbol{\omega}}$ is the posterior mean of $\boldsymbol{\omega}$ and $\bar{D}(\cdot)$ is the average of $D(\cdot)$ over the samples of $\boldsymbol{\omega}$. DIC is usually applied in Bayesian model selection problems where the posterior distribution has been obtained by MCMC simulation.

Evaluation and results

Simulation. In this section, to assess the performance of the proposed Bayesian approach a simulation study was conducted to estimate the parameters of SSVM in (1) and mixture of SSVM in (2). SSVM with parameters $\mu = 3$, $\tau = 2$, $\lambda = 0.5$ and prior parameters $\mu_0 = 0$, $\tau_0 = 0.01$, $\alpha = 4$, $\beta = 2$, $\xi = 0.5$, $\sigma = 0.01$ and a mixture of SSVM with two components ($M = 2$) with parameters $w = 0.8$, $\mu_1 = 3$, $\tau_1 = 0.2$, $\lambda_1 = 0.75$, $\mu_2 = 3.14$, $\tau_2 = 0.6$, $\lambda_2 = -0.3$ and prior parameters $\mu_{01} = 3$, $\tau_{01} = 0.1$, $\alpha_1 = 4$, $\beta_1 = 2$, $\xi_1 = 0.9$, $\sigma_1 = 0.15$, $c_1 = 1$ and $\mu_{02} = 0$, $\tau_{02} = 0.1$, $\alpha_2 = 6$, $\beta_2 = 2$, $\xi_2 = -1$, $\sigma_2 = 1.0$, $c_2 = 1$ were considered. Samples of sizes $n = 50, 100, 500$ were generated from the posterior distribution in (8) for each model, using Gibbs sampling in Algorithm 2. The Bayes estimates of parameters were obtained based on the squared error and absolute error loss functions. The posterior mean and the posterior median are the Bayes estimators under the squared error and absolute error loss functions, respectively. In order to obtain the Bayes estimates of the parameters, the mean and median of the generated samples from the posterior distribution (8) were calculated along with some other descriptive statistics. The results, including the sample mean, standard deviation (sd) and quartiles (Q₁, median, and Q₃) of the posterior distribution are summarized in Tables 2 and 3. As can be seen the differences between true values of the parameters and the posterior sample mean and the posterior sample median are minimal. Therefore, the proposed Bayesian approach provides accurate estimates for the parameters. The traceplots of the generated samples from the posteriors and the compare-partial plots⁵⁶ are shown in Fig. 6 for the mixture of SSVM. A traceplot is used for evaluating convergence which shows the time series of the sampling process from the posterior distribution. It is expected to get a traceplot that looks completely random. A compare-partial plot provides overlapped kernel density plots related to the last part

	Parameter	Actual value	Mean	SD	Q ₁	Median	Q ₃
<i>n</i> = 500	<i>w</i>	0.80	0.8135	0.0172	0.7820	0.8132	0.8453
	μ_1	3.00	3.0803	0.0685	2.9469	3.0810	3.2094
	τ_1	0.20	0.2357	0.0710	0.0996	0.2352	0.3673
	λ_1	0.75	0.7817	0.0005	0.7806	0.7817	0.7829
	μ_2	3.14	3.1413	0.0100	3.1223	3.1412	3.1654
	τ_2	0.60	0.5925	0.1354	0.3143	0.5969	0.8621
	λ_2	-0.30	-0.3017	0.0022	-0.3067	-0.3017	-0.2965
<i>n</i> = 100	<i>w</i>	0.80	0.8419	0.0334	0.7775	0.8395	0.9014
	μ_1	3.00	3.1114	0.0516	3.0224	3.1101	3.2243
	τ_1	0.20	0.1945	0.0554	0.0977	0.1871	0.3186
	λ_1	0.75	0.7316	0.0028	0.7269	0.7314	0.7373
	μ_2	3.14	3.1413	0.0058	3.1322	3.1427	3.1579
	τ_2	0.60	0.5964	0.1206	0.3761	0.5998	0.8152
	λ_2	-0.30	-0.3326	0.0038	-0.3410	-0.3320	-0.3266
<i>n</i> = 50	<i>w</i>	0.80	0.8351	0.0487	0.7383	0.8360	0.9212
	μ_1	3.00	3.2101	0.0789	3.1226	3.2164	3.3998
	τ_1	0.20	0.1903	0.0665	0.0847	0.1912	0.3147
	λ_1	0.75	0.7320	0.0032	0.7270	0.7314	0.7378
	μ_2	3.14	3.1420	0.0033	3.1342	3.1418	3.1489
	τ_2	0.60	0.6164	0.1145	0.3946	0.6158	0.7955
	λ_2	-0.30	-0.3321	0.0033	-0.3390	-0.3321	-0.3274

Table 3. Bayes estimates of parameters of a mixture of SSVM with prior parameters, $\mu_{01} = 3$, $\tau_{01} = 0.1$, $\alpha_1 = 4$, $\beta_1 = 2$, $\xi_1 = 0.9$, $\sigma_1 = 0.15$, $c_1 = 1$ and $\mu_{02} = 0$, $\tau_{02} = 0.1$, $\alpha_2 = 6$, $\beta_2 = 2$, $\xi_2 = -1$, $\sigma_2 = 1.0$, $c_2 = 1$.

of the chain (the last 10 values, in green) and the whole chain (in black). The overlapped kernel densities are expected to be similar. It means the initial and final parts of the chain should be sampling in the same target posterior distribution. These plots in Fig. 6 confirm the convergence of the chains and show that the Gibbs sampler recovers the values that actually generate the dataset.

To evaluate the accuracy of the obtained Bayes estimates, the mean squared errors (MSE) of the estimates under squared error and absolute error loss functions for the mixture of SSVM with two components ($M = 2$) with parameters which are mentioned above were obtained for different sample sizes $n = 10, 25, 50, 100, 200, 300, 500$ with 100 repetitions. The results in Fig. 7 show that by increasing n , MSE decreases and also, the MSEs of the estimates for absolute error loss function are less than squared error loss function because outliers have a smaller effect on the median.

Real data. To demonstrate the performance of the SSVM for the wind direction data for South African hot-spots, three real skewed datasets as discussed in “Site location and wind data” (see Table 1) were analyzed. Due to the multimodal pattern of the datasets observed in Fig. 4, the following distributions were assumed:

- mixtures of von Mises distributions with $M = 2, 3, 4$ components,
- SSVM with $k = 2$,
- mixtures of SSVM with $k = 1$ and $M = 2$ components,
- mixtures of SSVM with $k = 2$ and $M = 2$ components.

The MLEs of parameters (μ, τ, λ, p) were obtained by using the DEoptim package in R. The results including MLEs and corresponding log-likelihood, AIC and BIC are reported in Table 4. A model with the maximum log-likelihood and minimum values of AIC and BIC provides better fit for the data. Therefore, for dataset A, the mixture of SSVM with $k = 1$ provides the best fit. Mixture of SSVM with $k = 2$ and the mixture of von Mises with $M = 2$ are the second and third best models, respectively. For datasets B and C, the mixture of SSVM with $k = 2$ provides the best fit and the mixture of von Mises with $M = 4$ is the second best model. In all of these datasets, the difference in the AIC and BIC values of the mixture of SSVM in comparison to the mixture of von Mises are remarkable. Furthermore, the mixture of SSVM with smaller value of M , outperformed the mixture of von Mises. The kernel density plots of the datasets and the fitted curves consisting of the best mixture of von Mises and mixture of SSVM for $k = 1, 2$ are shown in Fig. 8.

To demonstrate the performance of the proposed Bayesian approach, a mixture of two SSVM distributions is fitted to dataset A for $k = 1$, and to dataset B and C with $k = 2$. A sample of size $n = 500$ was generated from the posterior distribution in (8) for each model, using the Gibbs sampling outlined in Algorithm 2. The Bayes estimates of the parameters were obtained based on the squared error, absolute error and zero-one loss functions. For our purpose, the posterior mean, posterior median and posterior mode were calculated from the generated

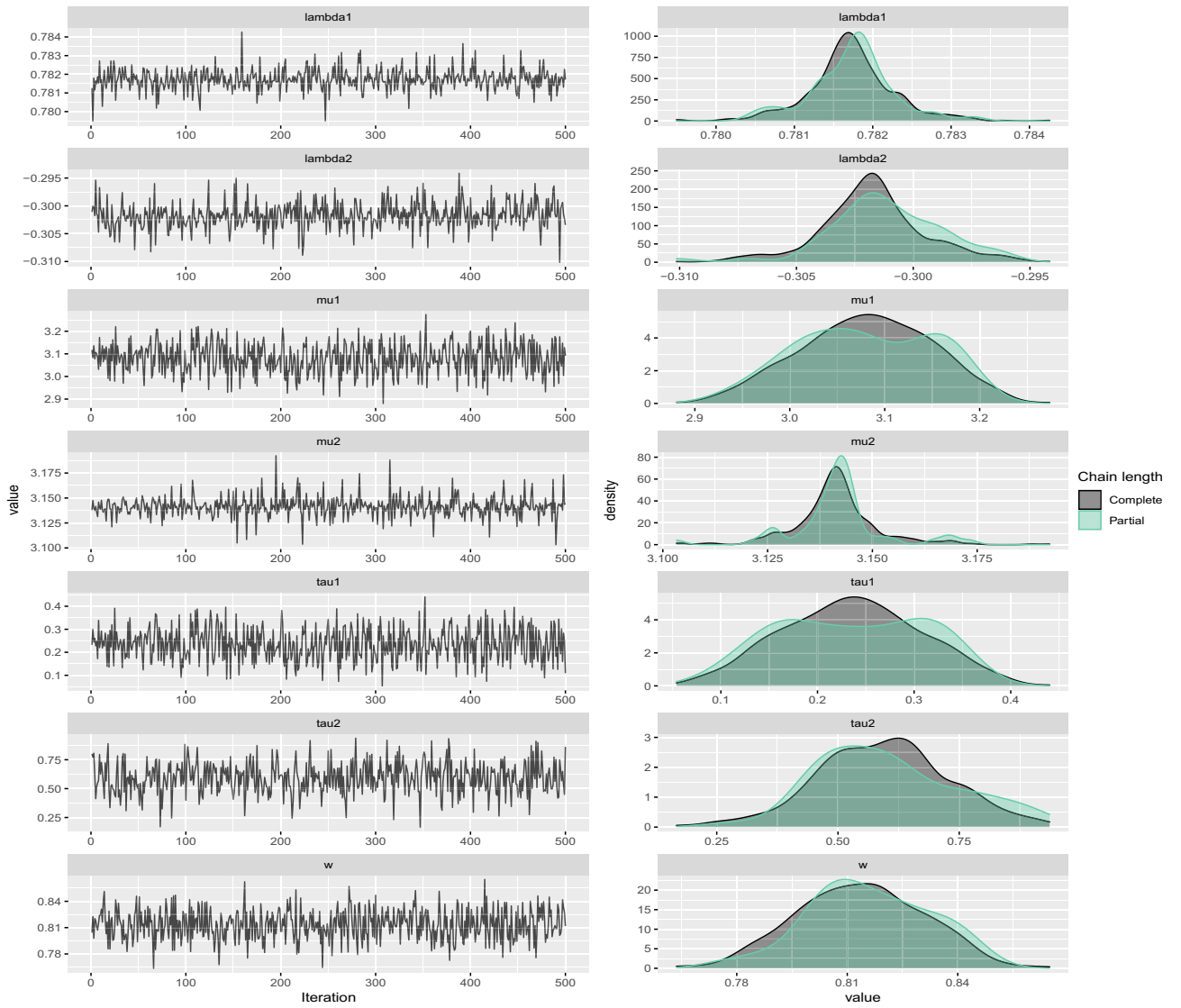


Figure 6. Traceplots and estimated posterior density plots of generated samples for $(w, \mu_1, \tau_1, \lambda_1, \mu_2, \tau_2, \lambda_2)$ in Table 3 for $n = 500$.

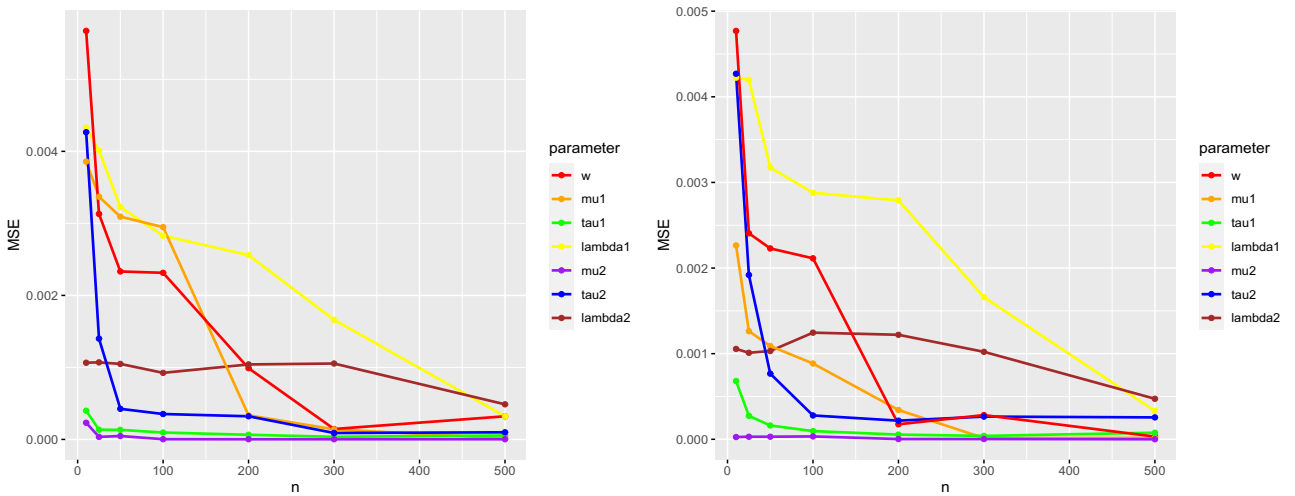


Figure 7. MSE of Bayes estimates under the squared error (left) and absolute error (right) loss functions, for $n = 10, 25, 50, 100, 200, 300, 500$.

Data	Model	$\hat{\tau}$	$\hat{\mu}$	$\hat{\lambda}$	\hat{w}	Log-likelihood	AIC	BIC
A	Mixture of VM ($M = 2$)	0.8264	4.6437	-	0.6388	-1524.7750	3059.5490	3084.4680
		13.4279	5.2866	-	0.3612			
	Mixture of VM ($M = 3$)	1.6861	4.0380	-	0.2852	-2241.4770	4498.9530	4538.8230
		11.7538	2.2718	-	0.4200			
		0.6421	5.6606	-	0.2948			
	Mixture of VM ($M = 4$)	1.6430	4.0014	-	0.3179	-1522.9620	3067.9240	3122.7460
		9.2575	5.1609	-	0.1863			
		0.7727	6.0288	-	0.2365			
	Mixture of SSVM ($k = 1, M = 2$)	0.5490	3.4434	0.8831	0.4291	-1248.6560	2511.3130	2546.2000
		5.9863	5.2451	0.0447	0.5709			
	SSVM ($k = 2$)	1.3283	4.8196	0.4113	-	-1575.1960	3156.3930	3171.3440
	Mixture of SSVM ($k = 2, M = 2$)	0.7644	4.4362	0.5208	0.5974	-1437.3610	2888.7220	2923.6090
11.8642		5.2842	0.1428	0.4026				
B	Mixture of VM ($M = 2$)	3.9011	4.5829	-	0.6284	-6392.3200	12794.6400	12826.6600
		4.1262	1.6053	-	0.3716			
	Mixture of VM ($M = 3$)	0.6536	1.8472	-	0.2602	-6066.5200	12149.0400	12200.2700
		6.8578	4.6102	-	0.5356			
		37.5722	1.6121	-	0.2042			
	Mixture of VM ($M = 4$)	1.2487	1.5608	-	0.2000	-6060.9610	12143.9200	12214.3600
		39.7624	1.6653	-	0.1962			
		1.5231	4.1872	-	0.1187			
	Mixture of SSVM ($k = 1, M = 2$)	3.7053	1.6860	0.4816	0.3799	-6295.287	12604.5700	12649.4000
		4.2219	4.5949	-0.7337	0.6201			
		0.4141	3.8738	0.6329	-			
	Mixture of SSVM ($k = 2, M = 2$)	1.2525	2.1711	-0.8901	0.4731	-5372.1610	10758.3200	10803.1500
7.3277		4.6315	-0.2355	0.5269				
C	Mixture of VM ($M = 2$)	0.9550	5.3272	-	0.5384	-6238.2750	12486.5500	12518.0600
		10.1064	2.2563	-	0.4616			
	Mixture of VM ($M = 3$)	2.4565	5.3344	-	0.2757	-6203.316	12422.6300	12473.0500
		0.1095	2.3723	-	0.3075			
		12.3062	2.2591	-	0.4168			
	Mixture of VM ($M = 4$)	1.8131	5.3286	-	0.4136	-6187.3030	12396.6100	12465.9300
		1.3339	2.2543	-	0.1532			
		24.8131	2.2987	-	0.2757			
	Mixture of SSVM ($k = 1, M = 2$)	0.8520	5.0994	-0.2553	0.5582	-6220.1690	12454.3400	12498.4500
		10.9951	2.2543	0.7743	0.4418			
		0.3378	2.9753	-0.7547	-			
	Mixture of SSVM ($k = 2, M = 2$)	0.4357	4.6249	0.7508	0.6137	-5584.4470	11182.8900	1127.0100
14.8628		2.2538	0.0835	0.3863				

Table 4. Maximum likelihood estimates and corresponding log-likelihood, AIC and BIC for datasets. The best model is indicated in bold.

samples as the Bayes estimates of parameters under the different mentioned loss functions. The results including the Bayes estimates of the parameters and corresponding DIC are reported in Table 5. A model with minimum value of DIC has better fit for the data. The mentioned models above with parameters estimated based on the absolute error loss function provide more accurate fit for the datasets. The kernel density plots of the datasets and the fitted curves are shown in Fig. 9.

In Table 6, using Algorithm 3, the predicted means of wind direction were obtained, based on absolute error loss function, for $n = 20, 50, 100$. Also, 95% credible intervals are derived. We focused on the assumption of absolute error loss function as a result of the performance observed in Table 5. As can be seen, by increasing n , the mean value of the predictive wind direction distributions are getting closer to the mean value of the datasets.

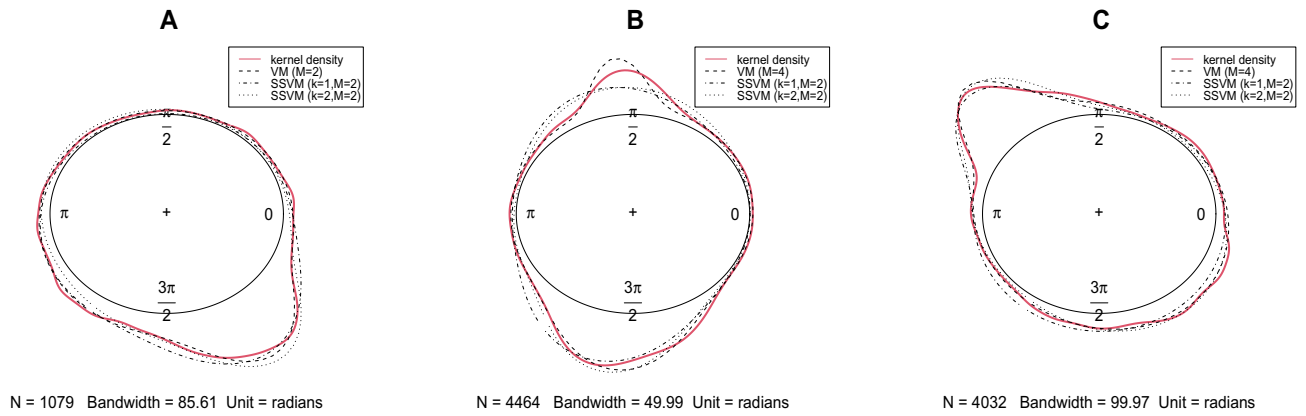


Figure 8. Kernel density plots of datasets and fitted curves based on MLEs.

Data	Model	Loss function	$\hat{\tau}_1$	$\hat{\mu}_1$	$\hat{\lambda}_1$	$\hat{\tau}_2$	$\hat{\mu}_2$	$\hat{\lambda}_2$	\hat{w}	DIC
A	Mixture of SSVM ($k = 1, M = 2$)	Squared error	0.4609	3.4528	0.7013	6.1235	5.2314	0.3395	0.5002	3086.42
		Absolute error	0.4368	3.4715	0.6973	6.0937	5.2244	0.3398	0.4997	3086.15
		Zero-one	0.3274	3.4560	0.5831	5.8139	5.1360	0.3334	0.4881	3087.48
B	Mixture of SSVM ($k = 2, M = 2$)	Squared error	1.5323	2.0121	-0.8772	7.4175	4.6898	-0.2405	0.4997	12839.28
		Absolute error	1.5743	2.0463	-0.8969	7.3480	4.6784	-0.2303	0.4996	12837.10
		Zero-one	1.5048	1.9889	-0.9049	7.3433	4.6634	-0.1981	0.5046	12885.87
C	Mixture of SSVM ($k = 2, M = 2$)	Squared error	0.4832	4.6294	0.7985	14.8946	2.2955	0.0995	0.6087	12796.50
		Absolute error	0.4038	4.6262	0.7794	14.6273	2.3348	0.0898	0.6122	12795.83
		Zero-one	0.4014	4.6227	0.7811	15.5572	2.3375	0.0829	0.6044	12812.50

Table 5. Bayes estimates of parameters under different loss functions and corresponding DIC for datasets. The best model is indicated in bold.

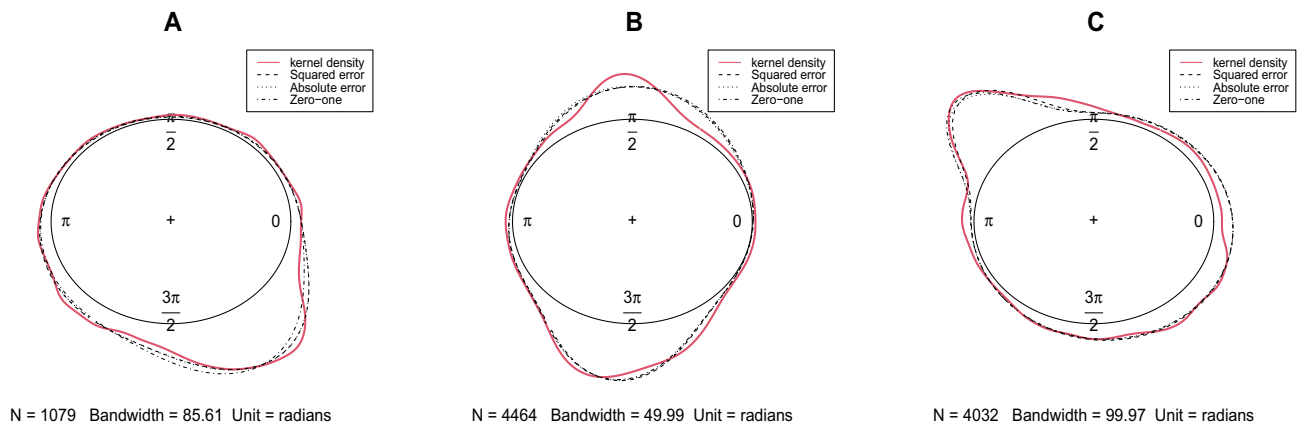


Figure 9. Kernel density plots of datasets and fitted curves based on Bayes estimates.

In addition, the length of the credible intervals is short. Therefore, our approach provides accurate prediction of wind direction.

Conclusion

In this paper, due to the skew and multimodal patterns of wind direction datasets from South Africa, a skew and multimodal mixture model, namely mixture of sine-skewed von Mises distributions is proposed for modeling wind direction. Our proposed model outperforms mixtures of von Mises distributions (with larger number of components) which is extensively used in literature to model wind direction. Due to the difficulties in estimating parameters for mixture models using maximum likelihood method, a Bayesian approach is implemented for estimating the parameters of a mixture of sine-skewed von Mises distributions using a Gibbs sampler. The results show this approach provides accurate estimates for parameters. In addition the posterior predictive distribution can be applied for wind direction prediction (see Table 6) which provides accurate forecasts. Future work

Data	Mean	Model	n	Predicted mean	95% Credible interval
A	5.0242	SSVM ($k = 1, M = 2$)	20	4.8754	4.4275,5.3233)
			50	5.1249	4.4611,5.3887)
			100	5.0171	(4.7442,5.2900)
B	4.3498	SSVM ($k = 2, M = 2$)	20	4.4918	(3.6417,5.3419)
			50	4.4652	(3.9834,4.9470)
			100	4.3580	(3.8963,4.8198)
C	2.3351	SSVM ($k = 2, M = 2$)	20	2.5216	(1.7277,3.3154)
			50	2.2784	(1.7142,2.8426)
			100	2.3726	(1.9737,2.7714)

Table 6. Predicted wind direction based on absolute error loss function for different values of n .

may consist of implementing the models of Bekker et al.⁵⁷ and Kato and Jones¹⁹ and investigating the impact of other prior choices⁵⁰. One can use our proposal to improve the wind energy potential as described and detailed in Arashi et al.⁵⁸.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 20 February 2022; Accepted: 6 June 2022

Published online: 06 July 2022

References

- Lahouar, A. & Ben Hadj Slama, J. Wind speed and direction prediction for wind farms using support vector regression. In *2014 5th International Renewable Energy Congress (IREC)*, 1–6. <https://doi.org/10.1109/IREC.2014.6826932> (2014).
- Huang, S. *et al.* Hierarchical optimal control for synthetic inertial response of wind farm based on alternating direction method of multipliers. *IEEE Trans. Sustain. Energy* **12**, 25–35. <https://doi.org/10.1109/TSTE.2019.2963549> (2021).
- Lerner, J., Grundmeyer, M. & Garvert, M. The importance of wind forecasting. *Renew. Energy Focus* **10**, 64–66. [https://doi.org/10.1016/S1755-0084\(09\)70092-4](https://doi.org/10.1016/S1755-0084(09)70092-4) (2009).
- You, M., Liu, B., Byon, E., Huang, S. & Jin, J. Direction-dependent power curve modeling for multiple interacting wind turbines. *IEEE Trans. Power Syst.* **33**, 1725–1733. <https://doi.org/10.1109/TPWRS.2017.2737529> (2018).
- Porté-Agel, F., Ting Wu, Y. & Chen, C. A numerical study of the effects of wind direction on turbine wakes and power losses in a large wind farm. *Energies* **6**, 5297–5313. <https://doi.org/10.3390/EN6105297> (2013).
- Castellani, F. *et al.* How wind turbines alignment to wind direction affects efficiency? A case study through scada data mining. *Energy Proced.* **75**, 697–703. <https://doi.org/10.1016/j.egypro.2015.07.495> (2015).
- Kazacoks, R., Amos, L. & Leithead, W. Effect of wind flow direction on the loads at wind farm. *Energy Proced.* <https://doi.org/10.1088/1742-6596/1356/1/012005> (2019).
- Gomez, M. S. & Lundquist, J. The effect of wind direction shear on turbine performance in a wind farm in central IOWA. *Wind Energy Sci. Discuss.* <https://doi.org/10.5194/wes-2019-22> (2019).
- Deep, S., Sarkar, A., Ghawar, M. & Rajak, M. K. Estimation of the wind energy potential for coastal locations in India using the Weibull model. *Renew. Energy* **161**, 319–339. <https://doi.org/10.1016/j.renene.2020.07.054> (2020).
- Gugliani, G., Sarkar, A., Mandal, S. & Agrawal, V. Location wise comparison of mixture distributions for assessment of wind power potential: A parametric study. *Int. J. Green Energy* **14**, 737–753. <https://doi.org/10.1080/15435075.2017.1327865> (2017).
- Carta, J., Ramirez, P. & Velázquez, S. A review of wind speed probability distributions used in wind energy analysis: Case studies in the canary islands. *Renew. Sustain. Energy Rev.* **13**, 933–955. <https://doi.org/10.1016/j.rser.2008.05.005> (2009).
- Carta, J. A., Bueno, C. & Ramirez, P. Statistical modelling of directional wind speeds using mixtures of von mises distributions: Case study. *Energy Convers. Manage.* **49**, 897–907. <https://doi.org/10.1016/j.enconman.2007.10.017> (2008).
- Qin, X., Zhang, J. & Yan, X. A new circular distribution and its application to wind data. *J. Math. Res.* **2**, 12. <https://doi.org/10.5539/jmr.v2n3p12> (2010).
- Gatto, R. & Jammalamadaka, S. R. The generalized von mises distribution. *Stat. Methodol.* **4**, 341–353. <https://doi.org/10.1016/j.stamet.2006.11.003> (2007).
- Masseran, N., Razali, A., Ibrahim, K. & Latif, M. Fitting a mixture of von mises distributions in order to model data on wind direction in peninsular Malaysia. *Energy Convers. Manage.* **72**, 94–102. <https://doi.org/10.1016/j.enconman.2012.11.025> (2013).
- Belu, R. & Koracin, D. Statistical and spectral analysis of wind characteristics relevant to wind energy assessment using tower measurements in complex terrain. *J. Wind Energy* <https://doi.org/10.1155/2013/739162> (2013).
- Quill, R., Sharples, J. J., Wagenbrenner, N. S., Sidhu, L. A. & Forthofer, J. M. Modeling wind direction distributions using a diagnostic model in the context of probabilistic fire spread prediction. *Front. Mech. Eng.* **5**, 5. <https://doi.org/10.3389/fmech.2019.00005> (2019).
- Gugliani, G., Sarkar, A., Ley, C. & Mandal, S. New methods to assess wind resources in terms of wind speed, load, power and direction. *Renew. Energy* **129**, 168–182. <https://doi.org/10.1016/j.renene.2018.05.088> (2018).
- Kato, S. & Jones, M. A tractable and interpretable four-parameter family of unimodal distributions on the circle. *Biometrika* **102**, 181–190. <https://doi.org/10.1093/biomet/asu059> (2015).
- El-Fouly, T. H. M., El-Saadany, E. F. & Salama, M. M. A. One day ahead prediction of wind speed and direction. *IEEE Trans. Energy Convers.* **23**, 191–201. <https://doi.org/10.1109/TEC.2007.905069> (2008).
- García-Planas, M. I. & Gongadze, T. Wind profile prediction using linear markov chains: A linear algebra approach. *IEEE Latin Am. Trans.* **16**, 536–541. <https://doi.org/10.1109/TLA.2018.8327410> (2018).
- Zeng, Y., Zhou, H., Lai, Y. & Wen, B. Wind-direction mapping with a modified wind spreading function by broad-beam high-frequency radar. *IEEE Geosci. Remote Sens. Lett.* **15**, 679–683. <https://doi.org/10.1109/LGRS.2018.2809558> (2018).

23. Fan, S. *et al.* Estimation of wind direction in tropical cyclones using c-band dual-polarization synthetic aperture radar. *IEEE Trans. Geosci. Remote Sens.* **58**, 1450–1462. <https://doi.org/10.1109/TGRS.2019.2946885> (2020).
24. Zheng, G. *et al.* Development of a gray-level co-occurrence matrix-based texture orientation estimation method and its application in sea surface wind direction retrieval from sar imagery. *IEEE Trans. Geosci. Remote Sens.* **56**, 5244–5260. <https://doi.org/10.1109/TGRS.2018.2812778> (2018).
25. Chen, X., Huang, W. & Haller, M. C. A novel scheme for extracting sea surface wind information from rain-contaminated x-band marine radar images. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **14**, 5220–5234. <https://doi.org/10.1109/JSTARS.2021.3078902> (2021).
26. Liu, X., Huang, W. & Gill, E. W. Wind direction estimation from rain-contaminated marine radar data using the ensemble empirical mode decomposition method. *IEEE Trans. Geosci. Remote Sens.* **55**, 1833–1841. <https://doi.org/10.1109/TGRS.2016.2635078> (2017).
27. Giangregorio, G., Galdi, C. & Bisceglie, M. D. Wind direction estimation by deconvolution of gnss delay-doppler maps: A simulation analysis. *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **13**, 2409–2418. <https://doi.org/10.1109/JSTARS.2020.2986448> (2020).
28. Wang, J., Xiong, X., Li, Z., Wang, W. & Zhu, J. Wind forecast-based probabilistic early warning method of wind swing discharge for ohtls. *IEEE Trans. Power Deliv.* **31**, 2169–2178. <https://doi.org/10.1109/TPWRD.2016.2519599> (2016).
29. Abe, T. & Pewsey, A. Sine-skewed circular distributions. *Stat. Pap.* **52**, 683–707. <https://doi.org/10.1007/s00362-009-0277-x> (2011).
30. Bagchi, P. Empirical Bayes estimation in directional data. *J. Appl. Stat.* **21**, 317–326. <https://doi.org/10.1080/757583874> (1994).
31. Bagchi, P. & Guttman, I. Theoretical considerations of the multivariate von mises-fisher distribution. *J. Appl. Stat.* **15**, 149–169. <https://doi.org/10.1080/02664768800000022> (1988).
32. Bagchi, P. & Kadane, J. Laplace approximations to posterior moments and marginal distributions on circles, spheres, and cylinders. *Can. J. Stat.* **19**, 67–77. <https://doi.org/10.2307/3315537> (1991).
33. Bangert, M., Hennig, P. & Oelfke, U. Using an infinite von mises-fisher mixture model to cluster treatment beam directions in external radiation therapy. In *2010 Ninth International Conference on Machine Learning and Applications*, 746–751. <https://doi.org/10.1109/ICMLA.2010.114> (2010).
34. Damien, P. & Walker, S. A full Bayesian analysis of circular data using the von Mises distribution. *Can. J. Stat.* **27**, 291–298. <https://doi.org/10.2307/3315639> (1999).
35. Dowe, D. L., Oliver, J. J., Baxter, R. A. & Wallace, C. S. Bayesian estimation of the von Mises concentration parameter. In *Maximum Entropy and Bayesian Methods* (eds Hanson, K. M. & Silver, R. N.) (Springer, 1996). <https://doi.org/10.1007/978-94-011-5430-7-6>.
36. Guttorp, P. & Lockhart, R. A. Finding the location of a signal: A Bayesian analysis. *J. Am. Stat. Assoc.* **83**, 322–330. <https://doi.org/10.2307/2288846> (1988).
37. Hornik, K. & Grün, B. On conjugate families and Jeffreys priors for von Mises-fisher distributions. *J. Stat. Plan. Inference* **143**, 992–999. <https://doi.org/10.1016/j.jspi.2012.11.003> (2013).
38. Nuñez-antonio, G. & Gutiérrez-peña, E. A Bayesian analysis of directional data using the von Mises-fisher distribution. *Commun. Stat. Simul. Comput.* **34**, 989–999. <https://doi.org/10.1080/03610910500308495> (2005).
39. Rodrigues, J., Galvão Leite, J. & Milan, L. A. Theory & methods: An empirical Bayes inference for the von Mises distribution. *Aust. N. Z. J. Stat.* **42**, 433–440. <https://doi.org/10.1111/1467-842X.00140> (2000).
40. Taghia, J., Ma, Z. & Leijon, A. Bayesian estimation of the von-mises fisher mixture model with variational inference. *IEEE Trans. Pattern Anal. Mach. Intell.* **36**, 1701–1715. <https://doi.org/10.1109/TPAMI.2014.2306426> (2014).
41. Roge, R., Madsen, K. H., Schmidt, M. N. & Mørup, M. Infinite von Mises-Fisher mixture modeling of whole brain fmri data. *Neural Comput.* **29**, 2712–2741. https://doi.org/10.1162/neco_a_01000 (2017).
42. Mulder, K., Jongmsa, P. & Klugkist, I. Bayesian inference for mixtures of von mises distributions using reversible jump mcmc sampler. *J. Stat. Comput. Simul.* **90**, 1539–1556. <https://doi.org/10.1080/00949655.2020.1740997> (2020).
43. Nakhaei Rad, N., Bekker, A. & Arashi, M. Bayesian inference for skew-wrapped Cauchy mixture model using a modified Gibbs sampler. In *2021 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI)*, 1–6. <https://doi.org/10.1109/MFI52462.2021.9591181> (IEEE, 2021).
44. Nakhaei Rad, N., Bekker, A., Arashi, M. & Ley, C. Coming together of Bayesian inference and skew spherical data. *Front. Big Data* <https://doi.org/10.3389/fdata.2021.769726> (2021).
45. Ley, C. & Verdebout, T. *Modern Directional Statistics* (Chapman and Hall, 2017).
46. Mullen, K., Ardia, D., Gil, D. L., Windover, D. & Cline, J. Deoptim: An r package for global optimization by differential evolution. *J. Stat. Softw.* **40**, 1–26. <https://doi.org/10.18637/jss.v040.i06> (2011).
47. Storn, R. & Price, K. Differential evolution: A simple and efficient heuristic for global optimization over continuous spaces. *J. Glob. Optim.* **11**, 341–359. <https://doi.org/10.1023/A:1008202821328> (1997).
48. Price, K., Storn, R. M. & Lampinen, J. A. *Differential Evolution: A Practical Approach to Global Optimization* (Springer, 2006).
49. Carlin, B. P. & Louis, T. A. *Bayesian Methods for Data Analysis* (CRC Press, 2008).
50. Ghaderinezhad, F., Ley, C. & Serrien, B. The wasserstein impact measure (wim): A practical tool for quantifying prior impact in Bayesian statistics. *Comput. Stat. Data Anal.* **1**, 107352. <https://doi.org/10.1016/j.csda.2021.107352> (2021).
51. Ley, C. & Verdebout, T. Skew-rotationally-symmetric distributions and related efficient inferential procedures. *J. Multivar. Anal.* **159**, 67–81. <https://doi.org/10.1016/j.jmva.2017.02.010> (2017).
52. Kotz, S., Balakrishnan, N. & Johnson, N. L. *Continuous Multivariate Distributions. Models and Applications* (Wiley, 2004).
53. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**, 716–723. <https://doi.org/10.1109/TAC.1974.1100705> (1974).
54. Schwarz, G. Estimating the dimension of a model. *Ann. Stat.* **6**, 461–464. <https://doi.org/10.2307/2958889> (1978).
55. Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Van Der Linde, A. Bayesian measures of model complexity and fit. *J. R. Stat. Soc. B* **64**, 583–639. <https://doi.org/10.1111/1467-9868.00353> (2002).
56. Fernández-i-Marín, X. Ggmcmc: Analysis of mcmc samples and Bayesian inference. *J. Stat. Softw.* **70**, 1–20. <https://doi.org/10.18637/jss.v070.i09> (2016).
57. Bekker, A., Nakhaei Rad, N., Arashi, M. & Ley, C. *Generalized skew-symmetric circular and toroidal distributions* (Springer, 2021).
58. Arashi, M., Nagar, P. & Bekker, A. Joint probabilistic modeling of wind speed and wind direction for wind energy analysis: A case study in humansdorp and noupoort. *Sustainability* <https://doi.org/10.3390/su12114371> (2020).

Acknowledgements

We would like to thank two anonymous reviewers for the constructive comments that significantly improved the presentation. Also, we would like to thank Mr. Lourens Snyman (Department of Geography, Geoinformatics and Meteorology at the University of Pretoria), Ms. Ria Olivier (Department of Botany and Zoology, Stellenbosch University) and Antarctic Legacy of South Africa for providing Fig. 1. This work was based upon research supported in part by the Visiting Professor Programme, University of Pretoria and the National Research Foundation (NRF) of South Africa, SARChI Research Chair UID: 71199; Ref.: IFR170227223754 Grant No. 109214; Ref.: SRUG190308422768 Grant No. 120839, the South African DST-NRF-MRC SARChI Research Chair in

Biostatistics (Grant No. 114613), STATOMET at the Department of Statistics at the University of Pretoria and DSI-NRF Centre of Excellence in Mathematical and Statistical Sciences (CoE-MaSS), South Africa. The research of the third author (M. Arashi) is supported by a grant from Ferdowsi University of Mashhad (N.2/56073). The opinions expressed and conclusions arrived at are those of the authors and are not necessarily to be attributed to the CoE-MaSS or the NRF.

Author contributions

All authors contributed equally to this work.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to N.N.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



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/>.

© The Author(s) 2022