STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **12**, July 2024, pp 1210–1230.
Published online in International Academic Press (www.IAPress.org)

In-depth Analysis of von Mises Distribution Models: Theory, Applications, and Future Directions

Said Benlakhdar [1,*], Mohammed Rziza [1], Rachid Oulad Haj Thami [2]

[1] *LRIT URAC 29, Faculty of Sciences, Mohammed V University in Rabat, Morocco*
[2] *RIITM, ENSIAS, Mohammed V University in Rabat, Morocco*

Abstract Multimodal and asymmetric circular data manifest in diverse disciplines, underscoring the significance of fitting suitable distributions for the analysis of such data. This study undertakes a comprehensive comparative assessment, encompassing diverse extensions of the von Mises distribution and the associated statistical methodologies, spanning from Richard von Mises' seminal work in 1918 to contemporary applications, with a particular focus on the field of wind energy. The primary objective is to discern the strengths and limitations inherent in each method. To illustrate the practical implications, three authentic datasets and a simulation study are incorporated to showcase the performance of the proposed models. Furthermore, this paper provides an exhaustive list of references related to von Mises distribution models.

Keywords Von Mises distribution, Mixture distributions, Hierarchical Model, Generalization of von Mises distributions, Specific EM algorithm, Wind Energy

DOI: 10.19139/soic-2310-5070-1919

1. Introduction

In various scientific fields, we find data measured in angles, relating to the orientation or direction of certain phenomena. Modeling angular variables is crucial in various fields such as biology [1, 2], astronomy[3], meteorology [4], earth sciences, and others [5]. For example, oceanographers may be interested in the direction of ocean currents; meteorologists to that of the winds and geologists to the orientation of the crystals of igneous rocks [6]. Biologists can consider the axis on which the bee deploys its dance according to light stimuli, or the deviation of the path of carrier pigeons from their destination [7]. However, direction measurements are not the only circular data: periods of cyclical phenomena can be represented as 360 degree circles. For example, one could determine that for a biological rhythm with a period of 24 hours and phase 0 h (or 24 h), an event X occurring at 3h corresponds to $45°$. Similarly, a circular representation can be used to express the frequency of earthquakes as a function of the month, over a period of one year. Psychological research also uses circular data, most often calculated from the geographic system.

The von Mises distribution (vM) and its mixed variants are primarily utilized as the parametric distribution for modeling the Wind Direction Probability Distribution (WDPD). For instance, Boente et al. [8] (2014) employed the von Mises distribution to estimate WDPDs for two Spanish locations, demonstrating moderate fitting performance. Another significant contribution is the introduction of a hierarchical von Mises distribution (HMvM) by Benlakhdar et al. [9] (2022), showcasing its accuracy in characterizing the probabilistic nature of wind direction. The mixture von Mises distributions (MvM) have gained popularity due to their adaptive adjustment of component numbers, providing a reasonable fit to WDPD. Carta et al. [10] (2008) illustrated this by enhancing goodness-of-fit through

*Correspondence to: Said Benlakhdar (Email: said-benlakhdar@um5s.net.ma).Laboratory of Research in Information Technology and Telecommunication (LRIT), Rabat IT Center, Faculty of Sciences, Mohammed V University in Rabat, Rabat B.P. 1014 RP, Morocco

ISSN 2310-5070 (online) ISSN 2311-004X (print)

an increased number of mixture components. Ovgor et al. [11] (2012) employed MvM to fit the WDPD of a Korean site, reporting a high coefficient of determination. Furthermore, Masseran et al. [12] (2013) utilized MvM to fit WDPDs across nine Malaysian locations, observing a small mean absolute percentage error. Soukissian et al [13](2017) extended the application to six sites, finding that a MvM with three to six components demonstrated a satisfactory goodness-of-fit. These diverse applications underscore the adaptability and efficacy of these models in capturing wind direction variability across different geographical locations.

The origins of the vM can betraced back to the 1918s, and it remains a popular tool in many fields today. The scientist and mathematician Richard Edler von Mises introduced this distribution, using it initially The exploration of discrepancies in atomic weights from integer values initiated a significant scientific inquiry. Subsequently, numerous researchers have drawn inspiration from von Mises' pioneering investigations to delve into and harness this phenomenon, laying the groundwork for circular distributions. The von Mises distribution stands out as the most extensive among univariate circular distributions, with its probability density function expressed as $f(\omega|\mu,k) = \frac{1}{2\pi I_0(k)} \exp k\cos(\omega-\mu)$, where $I_r(z) = (2\pi)^{-1}\int_0^{2\pi} \cos r\omega \times \exp\{z\cos\omega\}d\omega$, $z \in C$. For $\mu \in [0, 2\pi]$ and $k \geq 0$, denotes the concentration parameter μ , represents the mean angle, and $I_r(z)$ stands for the modified Bessel function of order r. It is possible to obtain a circular distribution that is flexible by using a finite mixture of simple distributions, such as a MvM. Nevertheless, this flexibility has some drawbacks. Mixture models require more intricate calculations and involve greater inference complexity because they lack sufficiency, invariance, and other factors.They often lead to irregular maximum likelihood problems. Furthermore, MvM lacks the essential theoretical properties inherent in the generalized von Mises distribution (GvMK), which demonstrates more flexibility compared to the standard von Mises distribution. Another significant category of continuous distributions for circular data is the envelope stable α (WαS) class, derived from the characteristic function of α-stable distributions in the real line. The density WαS can be represented as a Fourier series; $g(\theta) = \frac{1}{2\pi} + \frac{1}{\pi}\sum_{j=1}^{\infty} \exp\{-\tau^\alpha j^\alpha\} cos\{j(\theta-\mu) - \tau^\alpha j^\alpha \beta \tan\frac{\alpha\pi}{2}\}$, for $\theta \in [0, 2\pi), \tau > 0$ and the unimodality, tail behavior, and circular symmetry of WαS densities depend on the value of α and β. Interested readers can find more information on WαS distributions and inference in Gatto (2008). However, unlike GvM densities, WαS densities cannot exhibit bimodality and do not possess the theoretical characteristics outlined in Sections 2 to 4. As previously noted, GvM densities can display symmetry, asymmetry, unimodality, or bimodality.

In 2022, Benlakhdar et al [9](2022) introduced a novel hierarchical model of the von Mises distribution. The model was utilized to explore computational efficiency and generalization in directional problems at a large scale. Its main aim is to tackle the inflexibility issues of single models by exploiting the overall flexibility of the hierarchical structure and optimizing the model for maximum effectiveness.

This research offers a comparative study of the different alternatives to the von Mises distribution, examining the conditions, potential applications in analysis, and the limitations. The review further investigates how this knowledge has been applied to directional data problems, including classification. Additionally, the article emphasizes the potential for further improvements in developing more user-friendly models. To conduct this comparison, we have conducted bibliographic research using various scientific databases, such as Springer, Elsevier, Taylor and Francis, and Wiley, to compile a comprehensive list of cognitive explanatory models based on the von Mises distribution discussed in scientific literature. The forthcoming sections of this paper are structured as follows: Section 2 provides an extensive overview of the current literature. In Section 3, we introduce and delineate various models rooted in the von Mises distribution. Section 4 delineates the experimental methodology, as well as the presentation and discussion of the findings. Lastly, Section 5 serves to draw conclusions from the study.

2. Literature review

Circular data can be modeled using various circular distributions such as the wrapped Cauchy distribution, uniform circular distribution, Cardioid Distribution, Bingham Distribution and von Mises distribution models. Among these distributions, the wrapped Cauchy distribution stands out as a widely employed univariate symmetric circular distribution.Initially introduced by Lévy [18] and further explored by Wintner [19], the wrapped Cauchy distribution was developed by mapping Cauchy distribution onto the circle [20]. Building upon this groundwork, Kato and Pewsey [21] introduced a five-parameter bivariate wrapped Cauchy distribution tailored for toroidal data, ensuring consistency with univariate wrapped Cauchy distributions in both marginals and conditionals. This family of distributions maintains closure under conditioning and marginalization [31]. Additionally, Leguey et al. [31] proposed a tree-structured Bayesian network model for circular data based on the wrapped Cauchy distribution. Another distribution widely used in directional statistics is the Bingham distribution, introduced by Bingham in 1974 [22], is a versatile probability distribution applicable to data distributed across any-dimensional sphere[24]. It finds widespread use across diverse fields such as directional statistics [23] , computer graphics[25] , and neuroscience[26], to model the uncertainty in a parametric form[27] .Parameters of the Bingham distribution include the mean direction, specifying the distribution's central direction, and concentration parameters, regulating the distribution's dispersion around the mean direction and orthogonal directions. Its flexibility allows modeling of both unimodal and multimodal distributions, offering rotational symmetry around the mean direction. This distribution is instrumental in tasks like simulating surface normal distributions in computer graphics and studying neuronal process orientations in neuroscience. Additionally, the Bingham distribution can be extended to higher dimensions to model data on higher-dimensional spheres or hyperspheres.

The Cardioid (C) distribution [28] stands out as a significant model in circular data modeling. Although some of its structural attributes are defined, this distribution might not fully capture asymmetry and multimodal behaviors on the circle, thus requiring additional enhancements. Numerous general approaches are available for creating circular distributions, such as expansions of the C distribution using beta, Kumaraswamy, gamma, and Marshall–Olkin generators [29, 30].

The von Mises distribution, functioning as the circular analog of the univariate Gaussian distribution, stands out as a leading model in circular statistics. Mardia introduced the bivariate von Mises distribution [32] and later expanded it to the multivariate realm [74], revealing that the conditional distributions also adhere to von Mises distributions. However, the marginal distributions might exhibit either unimodal or bimodal characteristics. It's worth noting that the unimodal scenario approximates a von Mises distribution effectively only under conditions of a large concentration parameter. Moreover, contributions by Gatoo and Jammalamadaka [34](2007) are also discussed, providing insights into the theory and advanced topics in directional statistics. An extension of the von Mises distribution encompasses both unimodal and multimodal patterns, and it accommodates both symmetric and asymmetric characteristics within circular datasets. This generalization offers a versatile framework capable of capturing the diverse nature of circular data, whether it exhibits single or multiple peaks, and whether its distribution is symmetric or asymmetric.

3. Von Mises Distribution Models: Four Core Families

We classify techniques relying on the von Mises distribution into four main categories: the von Mises distribution model, the mixture of von Mises distributions model, the generalization of von Mises distributions model, and hierarchical von Mises mixture distributions model (refer to Table 1).

Table 1. List of references for alternative von Mises models

References	The model used	Application fields
(Swanson et al., 2023) [15]	Mixture von Mises distributions model	Molecules and Conformational Variability
(Campos-Aranda, 2023)[17]	von Mises distribution model	The administration of water resources in a river and the development of non-structural strategies for mitigating flood damage.
(Salvador & Gatto, 2022)[35]	Generalized von Mises distribution model	Bayesian inference

References	The model used	Application fields
(Benlakhdar et al., 2022)[9]	Hierarchical von Mises mixture distributions model	Wind energy
(Johnson, 2022)[36]	von Mises distribution model	Transformation of Directional Data to Approximate von Mises Distribution
(Thien et al., 2022)[37]	Mixture von Mises distributions model	Sign Indetermination in Phase Reconstruction: A DNN Approach with von Mises Mixture Model
(Marrelec & Giron, 2021)[38]	von Mises distribution model	The analysis of circular data
(Rønning et al., 2021)[39]	von Mises distributions model	Implementation of the Skewed Sine von Mises Distribution in Pyro and NumPyro for Protein Bioinformatics
(Mokhtar et al., 2021)[40]	von Mises distribution model	Modeling Wind Direction Relationships during the Southwest Monsoon in Langkawi Island
(Beh, 2021)[41]	Mixture von Mises distributions model	Classification of sound sources in the time-frequency
(Fan & Bouguila, 2020)[42]	Mixture von Mises distributions model	Non parametric Bayesian framework
(Mulder et al., 2020)[43]	Mixture von Mises distribution	Application to the music listening data
(Ye et al., 2019)[44]	von Mises distribution model	Statistical analysis of the wind speed and wind direction
(Takamichi et al., 2018)[45]	von Mises distribution model	Audio signal and speech processing
(Margon & Virtanen, 2018) [46]	Von Mises distribution model	Statistical model applied in audio signals
(Prokudin et al., 2018)[53]	Mixture von Mises distributions model	Image processing
(Karavasilis et al., 2017)[47]	Mixture von Mises distributions model	Modeling Circular Data in Visual Object Tracking
(Chinellato et al., 2017)[50]	Mixture von Mises distributions model	Modeling the Time of Occurrence of Events in Data Streams with Circular von Mises Distributions
(Qin et al., 2013)[55]	Generalized von Mises distribution model	Wind study
(Kim & SenGupta, 2013)[57]	Generalized von Mises distribution (k=3) model	A real environmental data
(Gabarda & Cristóbal, 2012)[58]	von Mises distribution model	Calculating the von Mises distribution of image entropy
(Marković & Petrović, 2012)[60]	Mixture von Mises distributions model	Bayesian Tracking of Bearings in Directional Scenarios using Mixture von Mises Distributions: Modeling, Inference, and Assessment.
(Calderara et al., 2011)[61]	Mixture von Mises distributions model	Classifying people trajectories in video surveillance

References	The model used	Application fields
(Ueaoki et al., 2011)[62]	von Mises distribution model	Scale-Invariant Shape Descriptors Using von Mises Distributions for Unordered Shapes in Image Matching and Retrieval
(Razavian et al., 2012)[63]	von Mises distribution model	Protein structures study
(Marković & Petrović, 2010)[60]	von Mises distribution model	Evaluation and Comparison of an Algorithm for Speaker Tracking: A Test on Synthetic Data Set and Real-world Recordings Against Particle Filter Representation
(Chen et al., 2008) [67]	Mixture von Mises distributions model	Data on movements of turtles are used as an illustration
(Vo & Oraintara, 2011)[68]	von Mises distribution model	Modeling the nature images in the transform domain
(Fu & Li, 2008)[69]	Mixture von Mises distributions model	Mixture Models with Application to Circadian Gene Expression
(Carta et al., 2008)[10]	Mixture von Mises distributions model	Representation the distribution of directional wind speed
(Gatto, 2008) [14]	Generalized von Mises distribution model	Model Comparison in Meteorological Data: GvM2, vM, and vM2 Models Applied to Arctic Wind Directions
(Muralidha & Parikh, 2007) [73]	von Mises distribution model	Analyzing Sea Star Movement: Bayesian Inference of Directional Data with Exponential Priors
(Gao et al., 2006) [75]	von Mises distributions model	Analyzing Disease Onset Patterns
(Mooney et al., 2003)[76]	Mixture von Mises distributions model	Sudden infant death syndrome
(Kent, 1983)[78]	Mixture von Mises distribution model	Directional data
(Stephens, 1982)[79]	von Mises distribution model	The one-way analysis-of-variance technique
(Damien & Walker, 1999)[80]	von Mises distribution model	Bayesian analysis of circular data

3.1. Von Mises Distribution

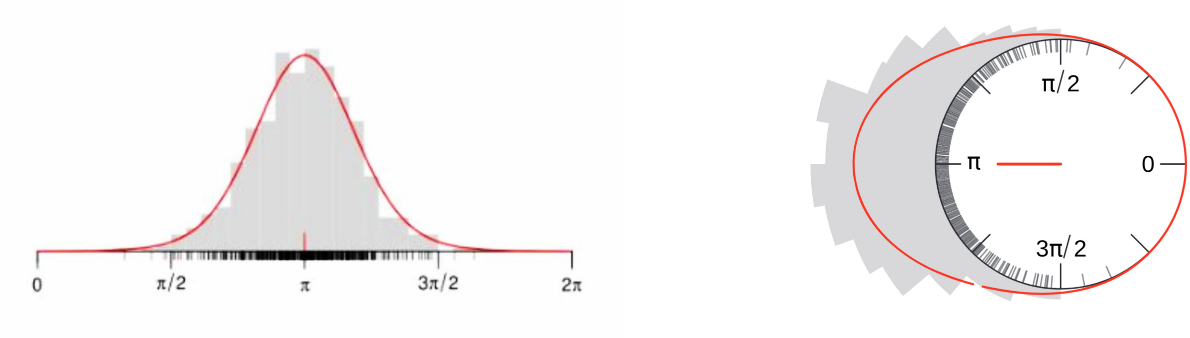
- **Description and Purpose**

The distribution function of a vM random variable X on the circle as described in [81], with mean direction μ_0 and a concentration parameter $k \geq 0$ is formulated as follows

$$G(\theta;\mu_0,k) = P(0 < X \leq \theta) = \int_0^\theta \frac{1}{2\pi_0(k)} \exp\{k\cos(X-\mu_0)\}\,dX \tag{1}$$

The procedure evaluates this function when $\mu_0 = 0$, the domain of definition of G is extended to the whole real line by integrating of the function $G(\theta + 2\pi) - G(\theta) = 1, -\infty < \theta < \infty$. When $\mu_0 \neq 0$, the vM function can be obtained from knowledge of the function by means$G(\theta, 0, k)$ of the equation

$$G(\theta;\mu_0,k) = G(\theta-\mu_0;0,k) - G(-\mu_0,0,k) \tag{2}$$



(a) Linear presentation (b) Circular presentation

Figure 1. Presentations of the von Mises distribution

Similary, the von Mises integral over any interval on the circle $(\theta_1, \theta_2), 0 \leqslant \theta_1 \leqslant \theta_2 \leqslant 2\pi$ is given by

$$(\theta_1 < X \leqslant \theta_2) = G(\theta_2 - \mu_0; 0, k) - G(\theta_1 - \mu_0; 0, k) \tag{3}$$

The vM was first introduced by von Mises in 1918 and further details on its theory and applications can be found in Mardia's (1972) work [82]. Kendall's (1974) publication [83] includes various useful charts that illustrate the behavior of the distribution.In 1967, Maksimov clarified several absolutely continuous circular distributions, each characterized by continous densities as described in the following expression

$$g(\theta)\alpha \exp\{\sum_{j=1}^{k} a_j \cos j\theta + b_j \sin j\theta\} \tag{4}$$

For $\theta \in [0, 2\pi)$ and for certain constants $a_1, a_2, \ldots, b_1, b_2, \ldots \in \mathbb{R}$. the renowned von Mises probability density function (pdf) is derived by specifically considering the sum in the exponent of Equation (4) for $k = 1$, yielding the subsequent equation

$$f(\theta|\mu, k) = \frac{1}{2\pi I_0(k)} \exp\{k \cos(\theta - \mu)\} \tag{5}$$

for $\theta \in [0, 2\pi)$,$\mu \in [0, 2\pi)$, $k > 0$ and where $I_r(z) = (2\pi)^{-1} \int_0^{\theta} \cos r\theta \exp\{z \cos\theta\} d\theta, z \in C$, is the modified Bessel function I of integer order r (e.g. Equation (4). If we use a unimodal vM for describing a set of n.i.i.d. Samples of angular features $\theta = \langle \theta_1, ..., \theta_n \rangle$, the parameters can be inferred using the maximum likelihood (ML), with the following equations

$$\theta_0^{ML} = \tan^{-1}\left\{\frac{\sum_{i=1}^{n} sin\theta_i}{\sum_{i=1}^{n} cos\theta_i}\right\} \tag{6}$$

$$A(m^{ML}) = \frac{I_1(m^{ML})}{I_0(ML)} = \frac{1}{n}\sum_{i=1}^{n} \cos(\theta_i - \theta_0^{ML}) \tag{7}$$

By inverting Equation (7) numerically, the solution for m^{ML} can be found. The von Mises distribution showed in Figure 5 must be viewed as a reasonable assumption for the a priori distribution of $X(G(\theta - \mu_0; 0, k))$ because Equation (5) it wraps around a circle.

In this section, we offer a concise summary of specific distributional properties linked with the von Mises distribution. To maintain generality, we assume that the parameter μ is zero in Equation (7). Consequently, we refer to the random variable X as having a standard von Mises distribution, represented as $X \sim vM(\theta, k)$, with its probability density function defined by

$$f(\theta) = \frac{e^{\cos\theta}}{2\pi I_0(k)}, -\pi < \theta < \pi, k \geq 0 \tag{8}$$

Where $I_0(k)$ is the modified Bessel function of the first kind and order 0 since

$$e^{k\cos\theta} = (I_0(k) + 2\sum_{j=1}^{\infty} I_j(k)\cos j\theta) \tag{9}$$

see Abramowitz and Stegun (1970), we can wite the pdf $f(\theta)$ and the cdf $F(\theta)$ as follows

$$f(\theta) = \frac{1}{2\pi}(1 + \frac{2}{I_0(k)} + \sum_{j=1}^{\infty} I_j(k)\cos j\theta), -\pi < \theta < \pi, k \geq 0 \tag{10}$$

and

$$f(\theta) = \frac{1}{2\pi}((\theta + \pi) + \frac{2}{I_0(k)} + \sum_{j=1}^{\infty} \frac{I_j(k)}{j} + \sin j\theta), -\pi < \theta < \pi, k \geq 0 \tag{11}$$

Where $I_j(k)$ is the modified Bessel function of the first kind and order j given by

$$I_j = (\frac{k}{2})^j \sum_{i=0}^{\infty} (\frac{k}{2})^{2j} \frac{1}{i!\Gamma(j+i+1)} \tag{12}$$

In what follows, we will consider several distributional properties of $X \sim vM(\theta, k)$ Based on these distributional properties, some characterizations of $X \sim vM(\theta, k)$ will be given. The standard vM distribution $X \sim vM(\theta, k)$ with the pdf as given in Equation(1) has the following properties.

I. It is symmetric around
II. For $k = 0$ in Equation (9), X has the uniform distribution on $(-\pi, \pi)$.
III. If $k \longrightarrow \infty$ in Equation (9), then X has the normal distribution on $(-\infty, \infty)$, with the pdf given by $f(\theta) = \frac{1}{\sigma\sqrt{2\pi}} \frac{-\theta^2}{2\sigma^2}$,$\sigma^2 = \frac{1}{k}$
IV. The mode is at $\theta = 0$,and it is $\frac{e^k}{2\pi I_0(k)}$
V. $k = \ln(\frac{f(0)}{f(\frac{\pi}{2})})$

- **Characterization by Truncated First Moment**

This section introduces characterizations of the von Mises distribution through truncated first moments, detailed in Theorem 1 below. Without loss of generality, we will focus on the standard vM distribution, $X \sim vM(\theta, k)$ denoted as $f(\theta)$, expressed as a series of modified Bessel functions.

$$f(\theta) = \frac{e^{k\cos\theta}}{2\pi I_0(k)} = \frac{1}{2\pi}(1 + \frac{2}{I_0(k)} \sum_{j=1}^{\infty} I_j(k)\cos j\theta), -\pi < \theta <, k \geq 0 \tag{13}$$

as provided in Equations(8) and (9) above.

Theorem 1
Suppose that X is absolutely continuous bounded random variable with cdf F(x) such that $F(-\pi) = 0$ and $F(\pi) = 1$ then $E(X|X < \theta) = g(\theta)\tau(\theta)$, where $\tau(\theta) = \frac{f(\theta)}{F(\theta)}$ and $g(\theta)$ is a continuous differentiable function of θ given by

$$g(\theta) = e^{-k\cos}(\frac{\theta^2 - \pi^2}{2} I_0(k) + 2\sum_{j=1}^{\infty} \frac{I_j(k)}{j^2}(j\theta\sin j\theta + \cos j\theta + \cos j\theta - \cos j\pi)) \tag{14}$$

if and only if $f(\theta) = \frac{1}{2\pi I_0(k)} e^{k\cos\theta}$,$-\pi < \theta < \pi, k \geq 0$.

Proof
If $f(\theta)=\frac{1}{2\pi I_0(k)}e^{k\cos\theta}$,$-\pi<\theta<\pi,k\geq 0$, Subsequently, when expressed as a series of modified Bessel functions, it becomes evident that, after integration and simplification,

$$g(\theta)=\frac{\int_{-u}^{\theta}uf(u)du}{f(\theta)}=\frac{\int_{-u}^{\theta}u(I_0(k)+2\sum_{j=1}^{\infty}I_j(k)\cos ju)du}{e^{k\cos\theta}} \tag{15}$$

$$=e^{-k\cos\theta}[\{(\theta^2-\pi^2)I_0(k)+2\sum_{j=1}^{\infty}\frac{I_j(k)}{j}\theta\sin j\theta\}-\{\frac{(\theta^2-\pi^2)}{2}I_0(k)-2\sum_{j=1}^{\infty}\frac{I_j(k)}{j^2}(\cos j\theta-\cos j\pi)\}] \tag{16}$$

$$=e^{-k\cos\theta}\theta(\frac{(\theta^2-\pi^2)}{2}I_0(k)+2\sum_{j=1}^{\infty}\frac{I_j(k)}{j^2}(j\theta\sin j\theta+\cos j\theta-\cos j\pi)) \tag{17}$$

Suppose that

$$g(\theta)=e^{-k\cos}(\frac{(\theta^2-\pi^2)}{2}I_0(k)+2\sum_{j=1}^{\infty}\frac{I_j(k)}{j^2}(j\theta\sin j\theta+\cos j\theta-\cos j\pi)) \tag{18}$$

Thus $\frac{\theta-g^{'}(\theta)}{g(\theta)}=-k\sin\theta$
from which, on using Lemma 1, we have
$\frac{f^{'}(\theta)}{f(\theta)}=\frac{\theta-g'(\theta)}{g(\theta)}=-k\sin\theta$
On integrating the above equation with respect to θ , we obtain $f(\theta)=ce^{k\cos\theta}$, where c is a constant to be determined.
Using the condition $\int_{-\pi}^{\pi}f(u)=1$, and recalling the integral representation of the modified Bessel function of the first kind and order 0, we easily obtain $c=\frac{1}{2\pi I_0(k)}$ and thus $=\frac{1}{2\pi_0(k)}e^{k\cos\theta},-\pi<\theta<\pi,k\geq 0$, which is the pdf of the standard von Mises distribution. □

3.2. Mixture of von Mises distribution

The mixture of von Mises distribution (MvM) is a statistical model that combines multiple von Mises distributions to represent complex directional data. In this model, each component distribution represents a distinct mode or cluster within the data, characterized by its own mean direction and concentration parameter.

The probability density function of MvM is a weighted sum of the individual von Mises density functions. Determining the weights and parameters of each component distribution within the model necessitates estimating them using methods like maximum likelihood estimation or Bayesian inference [84]. Moreover, the MvM is widely used in various applications, including modeling biological rhythms, analyzing directional data in environmental sciences, and clustering circular data in machine learning [85, 86, 43, 78]. By capturing the underlying patterns and structure in directional datasets, the MvM provides a flexible and powerful tool for understanding and modeling complex directional phenomena.

The attention given to finite mixture models over the years testifies to their usefulness as an extremely flexible modeling method, and their importance both theoretically and practically [89, 90, 91]. Often, we propose a MvM to describe a set of angles distribution because a unimodal von Mises model is not significant enough to represent dataset. It can be defined as

$$MvM(\theta|\theta_0,m)=\sum_{k=1}^{K}f(\theta|\theta_{0,k},m_k) \tag{19}$$

Where π_k is the weight of the k^{th} component (with k the number of mixture's components).

$$f(x,\theta)=\sum_{i=1}^{k}\frac{\pi_i}{2\pi I_0(k_i)}\exp\{k_i\cos(x-\mu_i)\},0<x\leqslant 2\pi \tag{20}$$

For $0 < x \leqslant 2\pi$,$-\infty < \mu_i < \infty$ are measures of locations, $k_i > 0$ are measures of concentrations and π_i are non-negative and sum to 1, where $I_\nu(.)$ denote the modified Bessel function of the first kind of order ν defined by

$$I_\nu(x) = \sum_{k=0}^{\infty} \frac{1}{\Gamma(k+v+1)k!}(\frac{x}{2})^{2k+\nu} \tag{21}$$

The n^{th} raw moment corresponding to MvM distribution is

$$m_n = \sum_{i=1}^{k} \alpha_i \frac{I_{|n|}(k_i)}{I_0(k_i)} \exp(in\mu_i) \tag{22}$$

Where $i = \sqrt{-1}$. Hence, the corresponding mean angle, mean resultant, circular variance, circular skewness and circular kurtosis are

$$\mu = \arcsin \frac{\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i)}{I_0(k_i)} sin(\mu_i)}{\sqrt{[\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i)}{I_0(k_i)} cos(\mu_i)]^2 + [\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i)}{I_0(k_i)} sin(\mu_i)]^2}} \tag{23}$$

$$\rho = \sqrt{[\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i)}{I_0(k_i)} cos(\mu_i)]^2 + [\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i)}{I_0(k_i)} sin(\mu_i)]^2} \tag{24}$$

$$v = 1 - \sqrt{[\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i)}{I_0(k_i)} \cos(\mu_i)]^2 + [\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i}{I_0(k_i)} \sin(\mu_i)]^2} \tag{25}$$

$$\gamma_1 = \frac{exp(-2\mu)}{v^{\frac{3}{2}}}[\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i)}{I_0(k_i)} sin(2\mu_i)] \tag{26}$$

and

$$\gamma_2 = \frac{exp(-2\mu)}{1-\rho^2}[\sum_{i=1}^{k} \alpha_i \frac{I_1(k_i)}{I_0(k_i)} cos(2\mu_i)] - \frac{\rho^4}{1-\rho^2} \tag{27}$$

The EM algorithm is a flexible method used to compute maximum likelihood estimates of mixture parameters. In mixture models, latent variables that are not directly observed define the ”responsibilities” of individual samples with respect to specific components within the mixture. The assumption is that all variables are independent, and the data stem from k joint distributions. However, the maximum likelihood framework described in Equation (8) is unsuitable for mixture distributions because it cannot handle singularities.

- **Simulation scenarios**

To assess the effectiveness of diverse methods across various scenarios, we employed four distinct data generation processes. These scenarios are illustrated in Figure 2 and include: Situation 1; two von Mises components where$\mu_1 = 0$, $\mu_2 = \pi$ and $k_1 = k_2 = 10$, Situation 2; two von Mises components where $\mu_1 = -\frac{\pi}{6}$, $\mu_2 = \frac{\pi}{6}$ and $k_1 = k_2 = 10$, Situation (3); two von Mises components where $\mu_1 = \frac{-\pi}{6}$ $\mu_2 = 0$ and $k_1 = k_2 = k_3 = 10$ and Situation 4; a uniform von Mises component where $\mu_1 = -\frac{\pi}{6}$, $\mu_2 = \frac{\pi}{6}$ and $k_1 = k_2 = 0$. Each scenario is simulated with 100 observations.

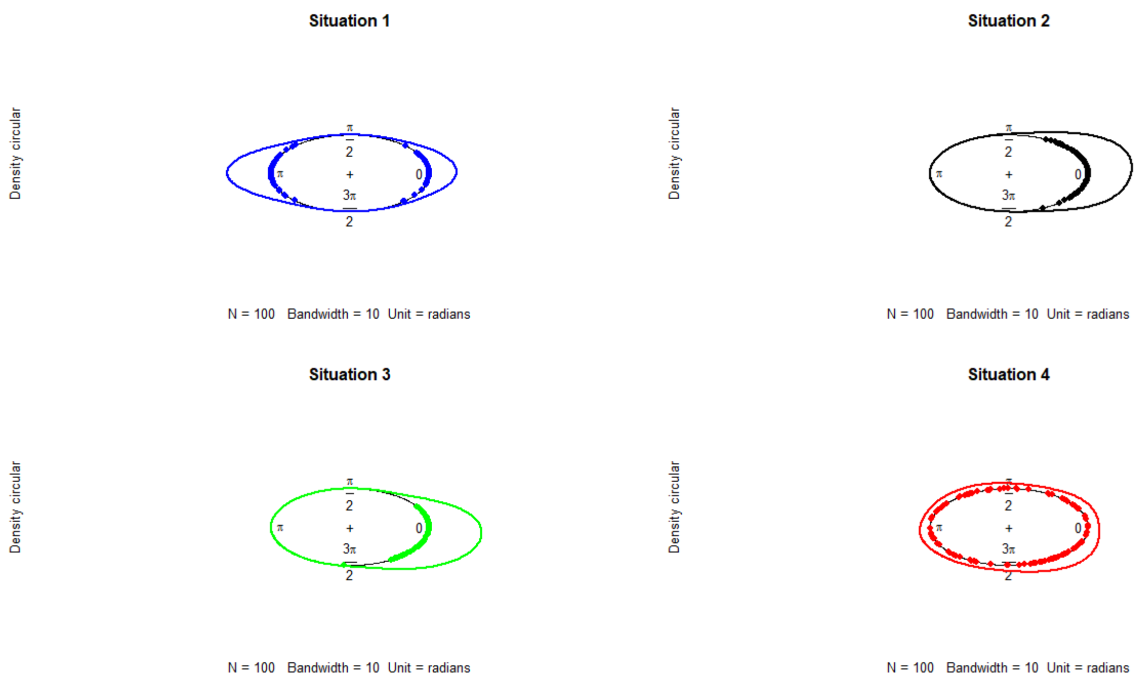


Figure 2. Visualization of simulation scenarios used for investigating sampler performance

3.3. *The generalization of von Mises distribution*

The GvMk of order k is a generalization of von Mises distribution that is able to model both unimodal and multimodal data, as well as symmetric and asymmetric data [92, 93, 94].The GvMk distribution offers greater flexibility than the von Mises distribution, which maintains circular symmetry and unimodality, featuring density that exponentially decreases on both sides of its center. Unlike the von Mises distribution, the GvMk distribution can effectively represent bimodality.

- **Model selection based on EMD**

The density function of the GvMk distribution is defined as

$$f(\theta, \mu_1, ..., \mu_k, k_1, ..., k_k) = \frac{1}{2\pi G_0^{(k)}(\delta_1, ..., \delta_{k-1}, k_1, ..., k_k)} \exp\{\sum_{j=1}^{k} k_j \cos j(\theta - \mu_j)\} \tag{28}$$

Where,
$k_1, ..., k_k > 0$, $\theta \in [0, 2\pi)$, $\mu_2 \in [0, \pi)$, $\mu_k \in [0, 2\pi/k)$, $\delta_1 = (\mu_1 - \mu_2) \mod \pi$ and $\delta_{k-1} = (\mu_1 - \mu_k) \mod (2\pi/k)$. The normalizing constant $G_0^{(k)}$ is given by

$$G_0^{(k)}(\delta_1, ..., \delta_{k-1}, ..., k_k) = \frac{1}{2\pi} \int_0^{2\pi} \exp\{k_1 \cos\theta + k_2 \cos 2(\theta + \delta_1) + ... + k_k \cos k(\theta + \delta_{k-1})\} d\theta \tag{29}$$

$$= \sum_{j=1}^{k} k_j \cos j(\theta - \mu_j) - \log[2\pi G_0^{(k)}(\delta_1, ..., \delta_{k-1}, k_1, ..., k_k] \tag{30}$$

In Equation (28), $\log f(\theta; \mu_1, ..., \mu_k, k_1, ..., k_k)$ can be perceived as the summation of a constant $-\log[2\pi G_0(k)]$ and multiple cosine functions of distinct frequencies j. The expression on the right side of Equation (28) bears resemblance to the outcome of empirical mode decomposition (EMD) [95]. In this paper, we will examine this category with a specific focus on the scenario where $k = 2$, which offers a significant expansion of vM and can be restated as

$$f(\theta|\mu_1, \mu_2, k_1, k_2) = \frac{1}{2\pi}\int_0^{2\pi} \exp\{k_1 \cos(\theta - \mu_1) + k_2 \cos 2(\theta - \mu_2)\} \tag{31}$$

for $\theta \in [0, 2\pi], \mu_1 \in [0, 2\pi), \mu_2 \in [0, \pi), \delta = (\mu_1 - \mu_2) \mod \pi, k_1, k_2 > 0$ and where the normalizing constant is given by

$$G_0(\delta, k_1, k_2) = \frac{1}{2\pi}\int_0^{2\pi} \exp\{k_1 \cos\theta + k_2 \cos 2(\theta + \delta)\} d\theta \tag{32}$$

The density outlined in Equation (12) will be termed the GvM density, and any circular random variable characterized by this density will be denoted as $\theta \sim GvM(\mu_1, \mu_2, k_1, k_2)$. Apart from Maksimov [96](1967), brief references to this distribution have been made in Yfantis and Borgman [97] (1982), particularly focusing on numerical considerations [98].

As mentioned earlier, GvM densities exhibit various shapes including symmetric, asymmetric, unimodal, or bimodal distributions. In this context, we present fundamental findings regarding the potential forms of these densities. Initially, we consider the hypothesis H_0: $\mu_2 = \mu_1 \mod \pi$, or equivalently, $H_0 : \delta = 0$, assuming k_1 and k_2 are both positive. Under this assumption, the density exhibits circular symmetry around μ_1, which, without loss of generality, can be set to 0. Through the process of differentiation, we ascertain that the critical points of the density satisfy the equation

$$\frac{k_1}{4k_2}\sin(\theta) + \sin(\theta)\cos(\theta) = 0 \tag{33}$$

Table 2. Critical points of the GvM density under $H_0 : \mu_2 = \mu_1 \mod \pi$ and for $k_1 < 4k_2$

Argument values	Type	Density values
$\pi_1 - \pi$	Maximum	$\{2\pi G_0(0, k_1, k_2)\}^{-1} \exp\{-k_1 + k_2\}$
$\mu_1 - \arccos(-\frac{k_1}{4k_2})$	Minimum	$\{2\pi G_0(0, K_1, k_2)\}^{-1} \exp\{-k_2 - \frac{k_1^2}{8k_2}\}$
μ_1	Maximum	$\{2\pi G_0(0, k_1, k_2)\}^{-1} \exp\{k_1 + k_2\}$
$\mu_1 + \arccos(-\frac{k_1}{4k_2})$	Minimum	$\{2\pi G_0(0, k_1, k_2)\}^{-1} \exp\{-k_2 - \frac{k_1^2}{8k_2}\}$

When $k_1 < 4k_2$, the interval $[-\pi, \pi)$ encompasses two trivial and two non-trivial critical points. Table 2 presents these critical points for a general μ_1, along with their characteristics and the corresponding values of the GvM density.

- **Member of the exponential family**

It's worth mentioning that the GvMk distribution can be expressed in the form of the canonical exponential family. We investigate a re-parametrization method for the GvMk density, where $\lambda_1 = k_1 \cos\mu_1$, $\lambda_2 = k_1 \sin\mu_1$, $\lambda_3 = \lambda_2 \sin 2\mu_2, \ldots, \lambda_{2K-1} = k_k \cos k\mu_k$. This involves expanding the cosines in Equation 28 and defining

$\lambda = (\lambda_1, \ldots, \lambda_{2K})^T \in \mathbb{R}^{2k}$ and $T(\theta) = (\cos\theta, \sin(\theta), \cos(2\theta), \sin(2\theta), \ldots, \cos k\theta, \sin k\theta)^T$, we can represent the GvMk density as

$$f^*(\theta|\lambda) = \exp \lambda^T T(\theta) - k(\lambda) \tag{34}$$

This transformation of the GvMk density conforms to the canonical exponential family with 2k parameters. The statistic $T(\theta)$ functions as a sufficient and complete statistic for λ, while the normalization constant is expressed as

$$K(\lambda) = \log(2\pi) + \log G_0^{(k)}(\delta_1, \ldots, \delta_{k-1}, \|\lambda^{(1)}\|, \ldots, \|\lambda^{(k)}\|), \tag{35}$$

where $\|.\|$ Euclidean norm is denoted by, $\lambda^{(1)} = (\lambda_1, \lambda_2)^T$, $\lambda^{(2)} = (\lambda_3, \lambda_4)^T, \ldots, \lambda^{(k)} = (\lambda_{2k-1}, \lambda_{2k})^T$, and $\delta_1 = (\arg\lambda^{(1)} - \arg\lambda^{(2)}/2) \mod \pi$, $\delta_2 = (\arg\lambda^{(1)} - \arg\lambda^{(3)}/3) \mod (2\pi/3), \ldots, \lambda_{k-1} = \arg\lambda^{(1)} - \arg\lambda^{(k)}/k) \mod (2\pi/k)$. In the case of $k = 2$, this constant can be evaluated using the following equation: $G_0(\delta, k_1, k_2) = I_0(k_1)I_0(k_2) + 2\sum_{j=1}^{\infty} I_{2j}(k_1)I_j(k_2)\cos(2j\delta)$. Furthermore, a referee has observed that the original Maksimov distribution with k summands is part of the 2k-parameter canonical exponential family, represented by

$$\lambda = (\lambda_1, \ldots, \lambda_k, \lambda_{k+1}, \ldots, \lambda_{2k}) = (a_1, \ldots, a_k, b_1, \ldots, b_k) \tag{36}$$

and

$$T(\theta) = (\cos(\theta), \ldots, \cos(k\theta), \sin(\theta), \ldots, \sin(k\theta)). \tag{37}$$

However, it's essential to note that the canonical re-parameterization lacks the intuitive clarity of the GvM form. For instance, the hypothesis $H_0 : \delta = 0$ is more straightforwardly apparent in the GvM parameterization, illustrating circular symmetry in both the first and second frequency components.

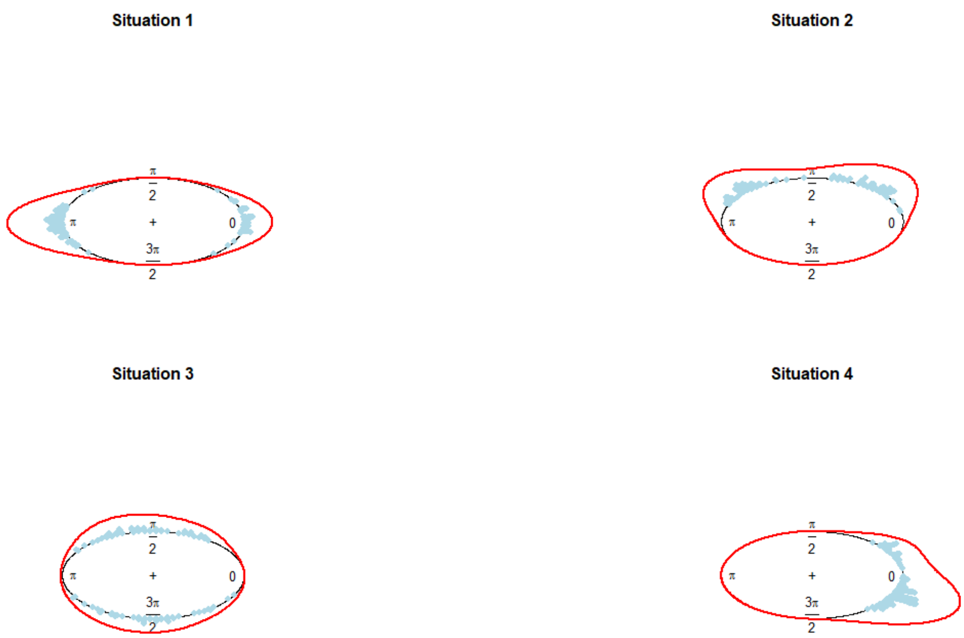


Figure 3. Some asymmetric GvM densities

3.4. *Hierarchical Von Mises Mixture Distributions*

Benlakhdar et al [9](2022) introduced the hierarchical von Mises mixture distribution model (HMvM-pdf) as a powerful tool for modeling complex, asymmetric, and multimodal datasets. The model is particularly suitable for addressing high-dimensional challenges, demonstrating enhanced performance concerning computational complexity, efficiency, and generalization.

In comparison to the von Mises mixture model, the HMvM exhibits various advantages. Its fundamental concept revolves around harnessing the comprehensive flexibility inherent in the complete hierarchy to offset the constrained flexibility observed in individual and mixture models. By using a simple hierarchy, the model enables straightforward interpretation of complex directional data and provides analytical and computational simplification benefits.

$$f(x,\theta) = \sum_{m+1}^{M} \sum_{i=l}^{L} \pi_{m,l} f_{m,l}(x,\theta_{m,l}) \tag{38}$$

$$f(x,\theta) = \sum_{m+1}^{M} \sum_{i=l}^{L} \pi_{m,l} \frac{1}{2\pi I_0(k_{m,l})} \exp[k_{m,l}\cos(x-\mu_{m,l})] \tag{39}$$

Where,
$\theta = \{\mu_{1,1}, \pi_{1,2}, ..., \pi_{m,l}, \theta_{1,1}, ..., \theta_{m,l}, \theta_{m,l} = \{(\mu_{1,1}, k_{1,1}), ..., (\mu_{M,L}, k_{M,L})\}$ and $\sum_{m=1}^{M}\sum_{l=1}^{L}\pi_{m,l} = 1$
The field of mixture model optimization has made significant progress over the past 45 years, largely due to the success of the maximum likelihood method. However, in the unsupervised case, maximizing the log-likelihood often results in equations without analytical solutions. To address this issue, one of the most commonly used methods is the Expectation-Maximization (EM) algorithm, proposed by Dumpsters et al [99] (1977). The EM algorithm stands as a widely recognized method for parameter estimation of mixture distributions, achieved by maximizing the log-likelihood of the completed data $L(x, y; \theta)$. A key advantage of the EM algorithm is that it always increases the observed likelihood, which is an important property (See Theorem 2).

Theorem 2
With each iteration of the EM algorithm, the observed likelihood L increases, namely

$$\varphi(\theta^{k+1}) > \varphi(\theta)^k \quad \forall k \geq 1$$

The EM algorithm is a commonly employed iterative method for estimating parameters in density models that treat observations as "incomplete data". It is particularly effective for mixture. In this framework, observations constitute the incomplete-data set, while each element of the complete-data set consists of an observation and an indicator specifying the mixture component. Unlike traditional methods, EM utilizes properties of the complete-data density, often resulting in more tractable estimation problems and accurate parameter estimates, especially with small sample sizes [99].
Benlakhdar et al [9](2022) developed an alternative to the EM algorithm which they named Specific EM to find parameter estimates of their HMvM model. A pseudo language version of the algorithm is presented (Algorithm 1). For more details see [9, 100].

- **The SEM algorithm in pseudo language**

Drawing upon the foundational principles of the EM algorithm, we have devised an algorithm termed SEM, as detailed in our aforementioned article "Statistical modeling of directional data using a robust hierarchical von mises distribution model: perspectives for wind energy" published in 2022 in the international journal of Computational Statistics.

In the SEM framework, the E-step entails computing the conditional expectations of the complete data log-likelihood conditioned on the observed data, while the M-step entails iteratively updating the parameter estimates to maximize this expectation. The iterative process continues until convergence is achieved, with parameter estimates refined at each iteration.

Algorithm 1 Specific EM algorithm

Require: Set χ of data points
Ensure: Estimation of unknown parameters
$\theta = \{\mu_{m,l}, \pi_{m,l}, k_{m,l}\}_{m,l=1}^{ML}$
$\ell = 0$
repeatStep E (Expectation) of the EM algorithm
  for $n = 1$ to N **do**
    for $m = 1$ to M **do**
      for $l = 1$ to L **do**
        $Q(\theta|Q^{(\ell)}) \leftarrow \sum_{n=1}^{N} \sum_{k=1}^{k} \sum_{l=1}^{L} \zeta_n^{(\ell)}(m,l) \ln(\pi_{m,l} f_{m,l}(x_n, \theta_{m,l}))$
      end for
    end for
  end for
until Convergence
End of step E
Step M (Maximization) of the EM algorithm
for $m = 1$ to M **do**
  for $l = 1$ to L **do**
    $\theta^{(m+\ell)} \leftarrow \arg\max_{\theta} Q(\theta|\theta^{(\ell)})$
  end for
end for
End of stage M
$\theta = \{\mu_{m,l}, \pi_{m,l}, k_{m,l}\}_{m,l=1}^{ML} \leftarrow \arg\{\mu_{m,l}^{*}, \pi_{m,l}^{*}, k_{m,l}^{*}\}_{m,l=1}^{ML}$

Based on the principles of the EM algorithm, maximizing results in the formulation of the following estimation equations:

$$\mu_{m,l}^{\ell+1} = \arctan\left(\frac{\sum_{n=1}^{N} \sin(x_n)\pi_{m,l}^{\ell} f_{m,l}(x_n, \theta_{m,l}^{\ell})}{\sum_{n=1}^{N} \cos(x_n)\pi_{m,l}^{\ell} f_{m,l}(x_n, \theta_{m,l}^{\ell})}\right) \tag{40}$$

$$= \arctan\left(\frac{\sum_{n=1}^{N} \zeta_n^{(\ell)}(m,l)\sin(x_n)}{\sum_{n=1}^{N} \zeta_n^{(\ell)}(m,l)\cos(x_n)}\right) \tag{41}$$

Mixing ratio

$$\pi_{m,l}^{(\ell+1)} = \frac{1}{N}\sum_{n=1}^{N} \zeta_n^{(\ell)}(m,l) f_{m,l}(x_n, \theta_{m,l}^{(\ell)}) \tag{42}$$

Concentration

$$A(K_{m,l}^{(\ell+1)}) = \frac{\sum_{i=1}^{N} \pi_{m,l}(\ell) f_{m,l}(x_n, \theta_{m,l}^{(\ell)}) \cos(x_n - \mu_{m,l}^{\ell+1})}{\sum_{i=1}^{N} \pi_{m,l}^{(\ell)} f_{m,l}(x_n, \theta_{m,l}^{(\ell)})} \tag{43}$$

$$K_{m,l}^{(\ell+1)} = A^{-1}\frac{\sum_{i=1}^{N} \zeta_n^{(\ell)} \cos(x_n - \mu_{m,l}^{(\ell+1)})}{\sum_{i=1}^{N} \zeta_n^{(\ell)}(m,l)} \tag{44}$$

To overcome the difficulty of converging to a local maximum, the EM algorithm is repeatedly run with a range of initial values. This approach increases the likelihood of identifying the global maximum. In Figure 4, the histogram depicts a comparison between the MvM and HMvM models based on 1000 pseudo-random observations. The contrast between the two outcomes is notably evident. Consequently, the accuracy of the Von Mises hierarchical

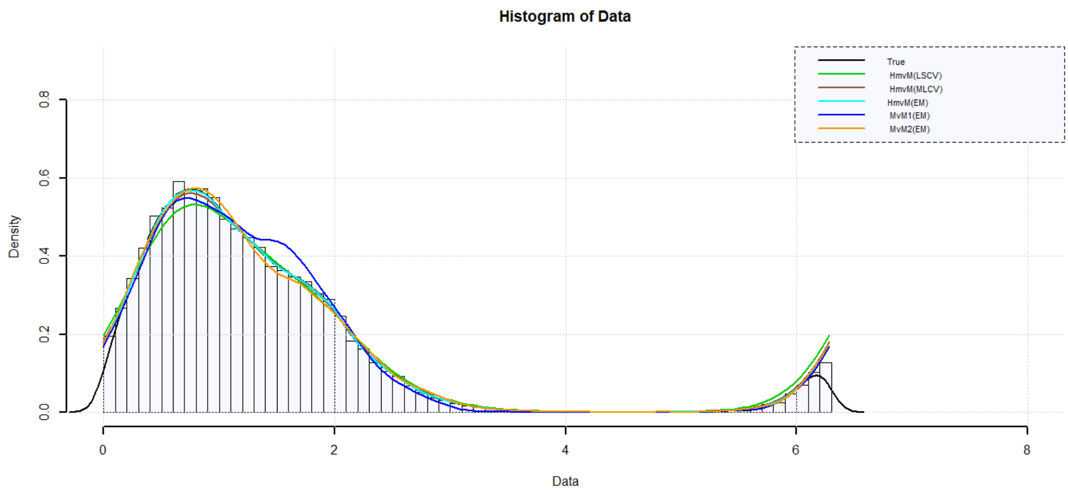


Figure 4. Representing the linear fit og HMvM-pdf and MvM-pdf using both parametric (EM) and (LSCV-MLCV)

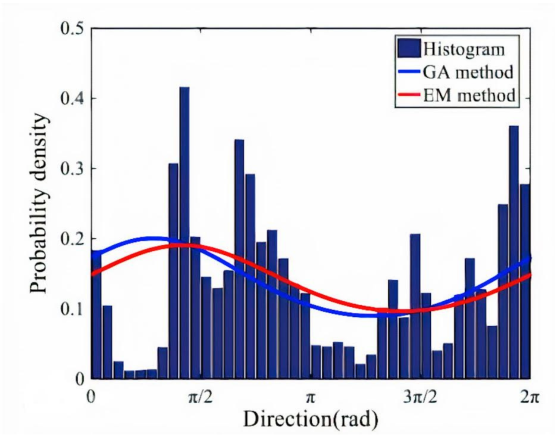
mixture distributions model, achieved through the modified EM algorithm, surpasses that of both the Von Mises mixture model and Kernel estimation [101, 102, 103].

- **Computational efficiency**

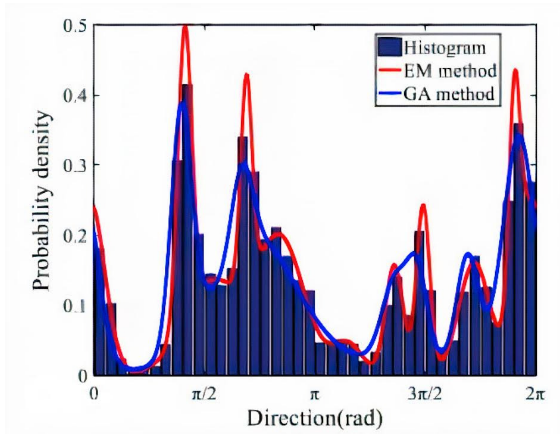
As the sample size increases, the computational burden of fitting models and calculating densities for von Mises distributions escalates, making efficient data structures and algorithms essential for handling large-scale data. Estimating parameters such as the mean direction and concentration can be particularly resource-intensive, with methods like Maximum Likelihood Estimation [104] commonly used for this purpose [105]. Optimization algorithms, including the Newton-Raphson method and Expectation-Maximization (EM) for mixture models, play a critical role in enhancing computational efficiency by reducing iterations and improving convergence rates [105]. A specialized algorithm for the hierarchical von Mises model, known as Specific EM (SEM)[106], is particularly tailored for large datasets. This algorithm demonstrates high efficiency in fitting model parameters; however, it is relatively time-consuming compared to the standard Expectation-Maximization (EM) algorithm. Employing parallel computing techniques can significantly boost efficiency by distributing the computational workload across multiple processors or cores [107], expediting tasks such as parameter estimation and density calculation. For instance, in R, parallel computing frameworks like the parallel package can be utilized to implement these techniques effectively, ensuring that large datasets are processed more rapidly and efficiently[108]. This integration of parallel computing and optimization not only minimizes computation time but also enhances the scalability of von Mises models fitting.

4. Results and Discussion

In the comparative scrutiny of models, it is imperative to recognize the challenges faced by the von Mises distribution model in faithfully encapsulating the observed data. As depicted in Figure 5a, the von Mises model contends with difficulties in accurately representing certain fluctuations or distinctive attributes of wind direction, while neglecting the diverse modalities inherent in the data. Instead, its emphasis lies on conforming to the most prominent mode within the dataset. This observation underscores the importance of exploring alternative methodologies, exemplified by the MvM distribution model. Such an approach offers heightened flexibility and an enhanced ability to capture the inherent intricacies within the dataset (refer to Figures 5b). This advanced model



(a) Single von Mises distribution



(b) Finite Mixture von Mises distribution

Figure 5. Single and Mixture von Mises distribution

affords the incorporation of distinct subpopulations, spatiotemporal variations, and inter-observation dynamics, thereby facilitating a more precise and realistic modeling of wind direction.
The Figures 5a and 5b depict the optimal vM model and the MvM model, respectively, attained through the EM method and the genetic algorithm (GA). It is noteworthy to mention that the EM method demonstrates superior efficacy compared to the GA method in both the construction of the vM model and the MvM model.
Throughout the entirety of January to December 2021, wind directions were systematically logged on a daily basis at four distinct Pan Arctic sites: the Pan Arctic, Greenland, Europe and North America basins. These datasets were acquired from ArcticRIMS (A Regional, Integrated Hydrological Monitoring System for the Pan Arctic Land Mass, http://rims.unh.edu), consist of n = 365 observations each. Employing statistical modeling, we applied the GvM2 and MvM2 distributions to these datasets, estimating their respective Maximum Likelihood parameters. Furthermore, We employed both the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) to evaluate the performance of these models.

Table 3. Evaluating AIC and BIC Values for GvM2 and MvM2 Distributions

Distribution tested	**Data**	**Observation Number**	**AIC**	**EM**
GvM2	Plan Arctic	365	765.177	613.692
	Europe	365	1007.608	934.028
	Greenland	365	1139.566	1023.140
	North America	365	944.590	902.517
MvM2	Plan Arctic	365	905.047	806.741
	Europe	365	1024.863	967.007
	Greenland	365	1135.567	1019.698
	North America	365	960.559	918.904

Table 3 exhibits the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for both the GvM2 model and the MvM2 model. Across all basins, the GvM2 model demonstrates superior performance in terms of AIC, with the exception of the Greenland Basin, where the MvM2 model displays a marginal advantage. Similarly, with regard to BIC, the GvM2 model consistently exhibits stronger performance, albeit with the narrowest discrepancy observed between GvM2 and MvM2 in the Greenland Basin.

However, accurately capturing certain nuances of the data proves challenging. As demonstrated in Table 4, expanding the dataset size leads to a notable reduction in the Bayesian Information Criterion (BIC) values across all models. Consequently, with an ample number of observations, both the GvMk and HMvM models exhibit enhanced capability in accurately modeling the data, capturing its intricate features more effectively. Our findings underscore the superiority of the HMvM model over both the MvM (see Figure 6) and vM models.

Table 4. Evaluating AIC and BIC Values for vM, MvM, GvM and HMvM Distributions

Distribution tested	Data	Observation Number	AIC	BIC
vM	Data 1	500	201.315	189.503
	Data 2	5000	180.629	173.702
	Data 3	50000	47.201	42.034
MvM	Data 1	500	179.382	166.642
	Data 2	5000	157.231	145.170
	Data 3	50000	33.712	29.314
GvM	Data 1	500	179.039	154.130
	Data 2	5000	120.887	116.983
	Data 3	50000	31.401	24.374
HMvM	Data 1	500	178.041	152.672
	Data 2	5000	121.230	117.103
	Data 3	50000	30.0185	22.526

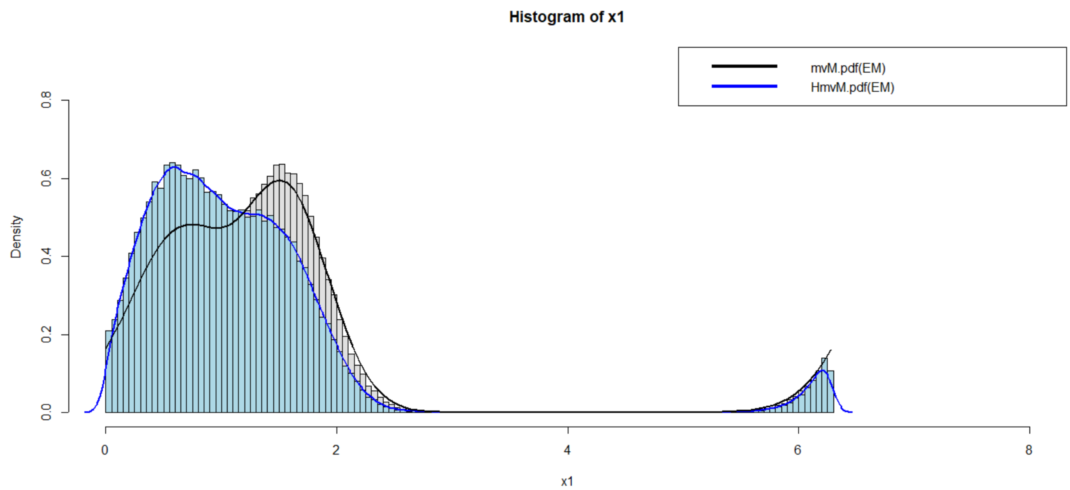


Figure 6. Concurrent Presentation of Models vM,MvM, GvM and HMvM Estimated by EM and SEM Algorithms

In conclusion, while the von Mises distribution model may offer a reasonable initial approximation of the wind direction distribution, its capacity to precisely conform to observed data, particularly in the presence of complex variations or specific characteristics, is constrained. In such contexts, opting for more advanced models such as GvMk or the HMvM model is advisable, as they offer a more refined and suitable representation of wind direction patterns.

5. Conclusion and future work

This paper provides an extensive examination of various substitutes for the von Mises distribution, coupled with a thorough discussion on utilizing circular statistics for the modeling of directional data and its numerical resolution. Through rigorous testing in challenging scenarios, we showcased the outstanding accuracy of the

GvMk and HMvM models in classifying directional data.The distributions under consideration present notably enhanced versatility in comparison to the presently utilized vM, MvM, and Circular Normal distributions. They maintain essential characteristics, including membership in the exponential family, association with the normal distribution, and maximized entropy. Our analysis has scrutinized various pivotal facets of these distributions, unveiling significant properties and characterizations. Furthermore, HMvM proves to be more precise and suitable for rapid implementations. This model can be broadly applied to represent directional patterns in regions with multiple dominant directions. To improve the quality of these models in future studies, a promising approach involves integrating the Ward algorithm with a mixture of Dirichlet processes. This concept is based on the fact that the Ward algorithm has the ability to group observations into homogeneous clusters based on their similarities. Subsequently, once the clusters are obtained, the Dirichlet process will be used to estimate the proportions of the clusters. Moreover, a Mixture von Mises distribution model will be employed to estimate the means and dispersions of the clusters. Finaly, proper management of overfitting heterogeneity, which occurs in high-dimensional mixture models when component-specific parameters of different components are identical, is essential. Frühwirth-Schnatter [109](2011)recommends the application of sparse priors for the component-specific location parameters to effectively address this issue. This approach will offer flexibility, robustness, and computational feasibility to capture the complex structure and variability of the data. Such refinements hold the potential to advance the field by providing more nuanced and accurate insights into directional data modeling.This model can be broadly applied to represent directional patterns in regions with multiple dominant directions.

Conflicts of interest

The author declares that there is no conflict of interest.

REFERENCES

1. S. Kato, C. Ley, and S. Loizidou, *The trivariate wrapped Cauchy copula–a multi-purpose model for angular data*, arXiv preprint arXiv:2401.10824, 2024.
2. J. J. Fernández-Durán and M. M. Gregorio-Domínguez, *Modeling angles in proteins and circular genomes using multivariate angular distributions based on multiple nonnegative trigonometric sums*, Statistical Applications in Genetics and Molecular Biology, vol. 13, no. 1, pp. 1–18, 2014.
3. M. Rice, K. Gerbig, and A. Vanderburg, *The Orbital Geometries and Stellar Obliquities of Exoplanet-hosting Multistar Systems*, The Astronomical Journal, vol. 167, no. 3, p. 126, 2024.
4. F. Hassanzadeh, *A smoothing spline model for multimodal and skewed circular responses: Applications in meteorology and oceanography*, Environmetrics, vol. 32, no. 2, p. e2655, 2021.
5. Mardia, K. V., & Jupp, P. E. (1999). Directional Statistics. Wiley. Chichester, UK.
6. Fisher, N. I., Lewis, T., & Embleton, B. J. (1993). Statistical analysis of spherical data. Cambridge university press.
7. Batschelet, E. (1981). Circular statistics in biology. ACADEMIC PRESS, 111 FIFTH AVE., NEW YORK, NY 10003, 1981, 388.
8. Boente, G., Rodriguez, D., & Manteiga, W. G. (2014). Goodness-of-fit Test for Directional Data. Scandinavian Journal of Statistics, 41(1), 259-275.
9. Benlakhdar, S., Rziza, M., & Thami, R. O. H. (2022). Statistical modeling of directional data using a robust hierarchical von Mises distribution model: perspectives for wind energy. Computational Statistics, 1-21.
10. Carta, J. A., Bueno, C., & Ramírez, P. (2008). Statistical modelling of directional wind speeds using mixtures of von Mises distributions: Case study. Energy conversion and management, 49(5), 897-907.
11. Ovgor, B., Lee, S. K., & Lee, S. (2012). A method of micrositing of wind turbine on building rooftop by using joint distribution of wind speed and direction, and computational fluid dynamics. Journal of mechanical science and technology, 26, 3981-3988.
12. Masseran, N., Razali, A. M., Ibrahim, K., & Latif, M. T. (2013). Fitting a mixture of von Mises distributions in order to model data on wind direction in Peninsular Malaysia. Energy Conversion and Management, 72, 94-102.
13. Soukissian, T. H., & Karathanasi, F. E. (2017). On the selection of bivariate parametric models for wind data. Applied energy, 188, 280-304.
14. Gatto, R. (2008). Some computational aspects of the generalized von Mises distribution. Statistics and computing, 18, 321-331.
15. Swanson, K., Williams, J. L., & Jonas, E. M. (2023, July). Von mises mixture distributions for molecular conformation generation. In International Conference on Machine Learning (pp. 33319-33342). PMLR.
16. Chakraborty, S., & Wong, S. W. (2023). On the circular correlation coefficients for bivariate von Mises distributions on a torus. Statistical Papers, 64(2), 643-675.
17. Campos-Aranda, D. F. (2023). Probabilistic characterization of the dates of occurrence of annual floods using the von Mises distribution. Tecnología y Ciencias del Agua, 14(2), 204-259.
18. P. Lévy, *L'addition des variables aléatoires définies sur une circonférence*, Bull. Soc. Math. France, vol. 67, pp. 1–41, 1939.
19. A. Wintner, *On the shape of the angular case of Cauchy's distribution curves*, The Annals of Mathematical Statistics, vol. 18, no. 4, pp. 589–593, 1947.

20. P. McCullagh, *Möbius transformation and Cauchy parameter estimation*, Ann. Statist., vol. 24, no. 2, pp. 787–808, 1996.
21. S. Kato and A. Pewsey, *A Möbius transformation-induced distribution on the torus*, Biometrika, vol. 102, no. 2, pp. 359–370, 2015.
22. C. Bingham, *An antipodally symmetric distribution on the sphere*, The Annals of Statistics, pp. 1201-1225, 1974.
23. C. J. Fallaize and T. Kypraios, *Bayesian Model Choice for Directional Data*, Journal of Computational and Graphical Statistics, vol. 33, no. 1, pp. 25-34, 2024.
24. S. Bhowmick, S. K. Biswas, and T. K. Mondal, *Quantifying fluid pressure events using shallow crustal veins*, International Journal of Earth Sciences, vol. 113, no. 1, pp. 145–160, 2024.
25. A. Goodyear, *A Nonlinear Bingham Filter for Pose Using Dual-Quaternions*, 2023.
26. H. Sato, T. Ikeda, and K. Nishiwaki, *A Probabilistic Rotation Representation for Symmetric Shapes With an Efficiently Computable Bingham Loss Function*, in 2023 IEEE International Conference on Robotics and Automation (ICRA), pp. 6923-6929, IEEE, May 2023.
27. S. N. Sotiropoulos, T. E. Behrens, and S. Jbabdi, *Ball and rackets: inferring fiber fanning from diffusion-weighted MRI*, Neuroimage, vol. 60, no. 2, pp. 1412-1425, 2012.
28. F. V. Paula, A. D. C. Nascimento, G. J. A. Amaral, and G. M. Cordeiro, *Generalized Cardioid Distributions for Circular Data Analysis*, Stats, vol. 4, pp. 634–649, 2021.
29. F. V. Paula, A. D. Nascimento, and G. J. Amaral, *A new extended Cardioid model: an application to wind data*, arXiv preprint arXiv:1712.01824, 2017.
30. T. Abe, A. Pewsey, and K. Shimizu, *On Papakonstantinou's extension of the cardioid distribution*, Statistics & probability letters, vol. 79, no. 20, pp. 2138-2147, 2009.
31. I. Leguey, C. Bielza, and P. Larrañaga, *Tree-structured Bayesian networks for wrapped Cauchy directional distributions*, in Advances in Artificial Intelligence, Vol. 9868, Springer, 2016, pp. 207–216.
32. K. V. Mardia, *Bayesian analysis for bivariate von Mises distributions*, Journal of Applied Statistics, 37(3), pp. 515–528, 2010.
33. K. V. Mardia, G. Hughes, C. C. Taylor, and H. Singh, *A multivariate von Mises distribution with applications to bioinformatics*, Canadian Journal of Statistics, 36(1), pp. 99–109, 2008.
34. R. Gatoo and S. R. Jammalamadaka, *The generalized von Mises distribution*, Statistical Methodology, 4(3), pp. 341–353, 2007.
35. Salvador, S., & Gatto, R. (2022). An algebraic analysis of the bimodality of the generalized von Mises distribution. Communications in Statistics-Theory and Methods, 1-17.
36. Johnson, R. A. (2022). Transformations to Improve the Approximation by a von Mises Distribution. In Directional Statistics for Innovative Applications: A Bicentennial Tribute to Florence Nightingale (pp. 147-158). Singapore: Springer Nature Singapore.
37. Thien, N. B., Wakabayashi, Y., Yuting, G., Iwai, K., & Nishiura, T. (2022, November). Von Mises Mixture Model-based DNN for Sign Indetermination Problem in Phase Reconstruction. In 2022 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC) (pp. 957-961). IEEE.
38. Marrelec, G., & Giron, A. (2021). Estimating the concentration parameter of a von Mises distribution: a systematic simulation benchmark. Communications in Statistics-Simulation and Computation, 1-13.
39. Rønning, O., Ley, C., Mardia, K. V., & Hamelryck, T. (2021, September). Time-efficient Bayesian Inference for a (Skewed) Von Mises Distribution on the Torus in a Deep Probabilistic Programming Language. In 2021 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI) (pp. 1-8). IEEE.
40. Mokhtar, N. A., Badyalina, B., Chang, K. L., Yaa'cob, F., Ghazali, A., & Shamala, P. (2021). Error-in-Variables Model of Malacca Wind Direction Data with the von Mises Distribution in Southwest Monsoon. Applied Mathematical Sciences, 15(9), 471-479.
41. Beh, J. (2021). MAP Adaptation with Mixtures of von Mises Distributions and Its Application to Underdetermined Convolutive Blind Source Separation. Journal of the Korea Academia-Industrial cooperation Society, 22(10), 752-759.
42. Fan, W., & Bouguila, N. (2020). Spherical data clustering and feature selection through nonparametric Bayesian mixture models with von Mises distributions. Engineering Applications of Artificial Intelligence, 94, 103781.
43. Mulder, K., Jongsma, P., & Klugkist, I. (2020). Bayesian inference for mixtures of von Mises distributions using reversible jump MCMC sampler. Journal of Statistical Computation and Simulation, 90(9), 1539-1556.
44. Ye, X. W., Xi, P. S., & Nagode, M. (2019). Extension of REBMIX algorithm to von Mises parametric family for modeling joint distribution of wind speed and direction. Engineering Structures, 183, 1134-1145.
45. S. Takamichi, Y. Saito, N. Takamune, D. Kitamura, and H. Saruwatari, *Phase reconstruction from amplitude spectrograms based on von-Mises-distribution deep neural network*, in 2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC), IEEE, pp. 286–290, 2018.
46. Magron, P., & Virtanen, T. *On modeling the STFT phase of audio signals with the von Mises distribution*, in 2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC), IEEE, September 2018, pp. 550-554.
47. Karavasilis, V., Nikou, C., & Likas, A. (2017). Real-time visual tracking using a spatially weighted von Mises mixture model. Pattern Recognition Letters, 90, 50-57.
48. Fernandez-Gonzalez, P., Bielza, C., & Larranaga, P. (2017). Univariate and bivariate truncated von Mises distributions. Progress in Artificial Intelligence, 6, 171-180.
49. Wu Navarro, A. K. (2019). Probabilistic Machine Learning for Circular Statistics: Models and inference using the Multivariate Generalised von Mises distribution (Doctoral dissertation, University of Cambridge).
50. Chinellato, E., Mardia, K. V., Hogg, D. C., & Cohn, A. G. (2017). An incremental von Mises mixture framework for modelling human activity streaming data. Proceedings ITISE, 2017.
51. Magron, P., & Virtanen, T. (2018, September). On modeling the STFT phase of audio signals with the von Mises distribution. In 2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC) (pp. 550-554). IEEE.
52. Holzapfel, G. A., Niestrawska, J. A., Ogden, R. W., Reinisch, A. J., & Schriefl, A. J. (2015). Modelling non-symmetric collagen fibre dispersion in arterial walls. Journal of the Royal Society Interface, 12(106), 20150188.
53. Prokudin, S., Gehler, P., & Nowozin, S. (2018). Deep directional statistics: Pose estimation with uncertainty quantification. In Proceedings of the European Conference on Computer Vision (ECCV) (pp. 534-551).

54. Mardia, K. V., & Voss, J. (2014). Some fundamental properties of a multivariate von Mises distribution. Communications in Statistics-Theory and Methods, 43(6), 1132-1144.
55. Qin, X., Zhang, J. S., & Yan, X. D. (2013). Model selection of the generalized von Mises distribution based on empirical mode decomposition with data analyses. Statistical Methodology, 14, 15-25.
56. Lark, R. M., Clifford, D., & Waters, C. N. (2013). Modelling complex geological angular data with the Projected Normal distribution and mixtures of von Mises distributions. Solid Earth Discussions, 5(2).
57. Kim, S., & SenGupta, A. (2013). A three-parameter generalized von Mises distribution. Statistical Papers, 54, 685-693.
58. Gabarda, S., & Cristóbal, G. (2012). No-reference image quality assessment through the von Mises distribution. JOSA A, 29(10), 2058-2066.
59. Hung, W. L., Chang-Chien, S. J., & Yang, M. S. (2012). Self-updating clustering algorithm for estimating the parameters in mixtures of von Mises distributions. Journal of Applied Statistics, 39(10), 2259-2274.
60. Marković, I., & Petrović, I. (2012, October). Bearing-only tracking with a mixture of von Mises distributions. In 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems (pp. 707-712). IEEE.
61. Calderara, S., Prati, A., & Cucchiara, R. (2011). Mixtures of von Mises distributions for people trajectory shape analysis. IEEE Transactions on Circuits and Systems for Video Technology, 21(4), 457-471.
62. Ueaoki, K., Iwata, K., Suematsu, N., & Hayashi, A. (2011). Matching handwritten line drawings with von Mises distributions. IEICE TRANSACTIONS on Information and Systems, 94(12), 2487-2494.
63. Razavian, N. S., Kamisetty, H., & Langmead, C. J. (2012, December). Learning generative models of molecular dynamics. In BMC genomics (Vol. 13, No. 1, pp. 1-13). BioMed Central.
64. T. F. Chan, S. Esedoglu, F. Park, and A. Yip, *Total variation image restoration: Overview and recent developments*, in Handbook of Mathematical Models in Computer Vision, edited by N. Paragios, Y. Chen, and O. Faugeras, Springer-Verlag, New York, pp. 17–31, 2006.
65. I. Marković and I. Petrović, *Applying von Mises distribution to microphone array probabilistic sensor modelling*, in ISR 2010 (41st International Symposium on Robotics) and ROBOTIK 2010 (6th German Conference on Robotics), pp. 1-7, VDE, 2010.
66. I. Marković and I. Petrović, *Speaker localization and tracking with a microphone array on a mobile robot using von Mises distribution and particle filtering*, Robotics and Autonomous Systems, vol. 58, no. 11, pp. 1185-1196, 2010.
67. J. Chen, P. Li, and Y. Fu, *Testing homogeneity in a mixture of von Mises distributions with a structural parameter*, Canadian Journal of Statistics, vol. 36, no. 1, pp. 129-142, 2008.
68. A. P. Vo, S. Oraintara, and T. T. Nguyen, *Statistical image modeling using von Mises distribution in the complex directional wavelet domain*, in 2008 IEEE International Symposium on Circuits and Systems, pp. 2885-2888, IEEE, 2008.
69. Y. Fu, J. Chen, and P. Li, *Modified likelihood ratio test for homogeneity in a mixture of von Mises distributions*, Journal of Statistical Planning and Inference, vol. 138, no. 3, pp. 667-681, 2008.
70. K. V. Mardia, C. C. Taylor, and G. K. Subramaniam, *Protein bioinformatics and mixtures of bivariate von Mises distributions for angular data*, Biometrics, vol. 63, no. 2, pp. 505-512, 2007.
71. J. A. Carta, C. Bueno, and P. Ramírez, *Statistical modelling of directional wind speeds using mixtures of von Mises distributions: Case study*, Energy conversion and management, vol. 49, no. 5, pp. 897-907, 2008.
72. R. Gatto and S. R. Jammalamadaka, *The generalized von Mises distribution*, Statistical Methodology, vol. 4, no. 3, pp. 341-353, 2007.
73. K. Muralidharan and R. Parikh, *Some Bayesian inferences for von Mises distribution*, American Journal of Mathematical and Management Sciences, vol. 27, no. 1-2, pp. 123-137, 2007.
74. K. V. Mardia, G. Hughes, C. C. Taylor, and H. Singh, *A multivariate von Mises distribution with applications to bioinformatics*, Canadian Journal of Statistics, vol. 36, no. 1, pp. 99-109, 2008.
75. F. Gao, K. S. Chia, I. Krantz, P. Nordin, and D. Machin, *On the application of the von Mises distribution and angular regression methods to investigate the seasonality of disease onset*, Statistics in medicine, vol. 25, no. 9, pp. 1593-1618, 2006.
76. J. A. Mooney, P. J. Helms, and I. T. Jolliffe, *Fitting mixtures of von Mises distributions: a case study involving sudden infant death syndrome*, Computational Statistics & Data Analysis, vol. 41, no. 3-4, pp. 505-513, 2003.
77. S. D. Grimshaw, D. G. Whiting, and T. H. Morris, *Likelihood ratio tests for a mixture of two von Mises distributions*, Biometrics, vol. 57, no. 1, pp. 260-265, 2001.
78. J. T. Kent, *Identifiability of finite mixtures for directional data*, The Annals of Statistics, pp. 984-988, 1983.
79. M. A. Stephens, *Use of the von Mises distribution to analyse continuous proportions*, Biometrika, vol. 69, no. 1, pp. 197-203, 1982.
80. P. Damien and S. Walker, *A full Bayesian analysis of circular data using the von Mises distribution*, The Canadian Journal of Statistics/La Revue Canadienne de Statistique, pp. 291-298, 1999.
81. K. V. Mardia and P. J. Zemroch, *Algorithm AS 86: The von Mises distribution function*, Journal of the Royal Statistical Society. Series C (Applied Statistics), vol. 24, no. 2, pp. 268-272, 1975.
82. K. V. Mardia, *A multi-sample uniform scores test on a circle and its parametric competitor*, Journal of the Royal Statistical Society: Series B (Methodological), vol. 34, no. 1, pp. 102-113, 1972.
83. S. B. Kendall, *Preference for Intermittent Reinforcement*, in Journal of the Experimental Analysis of Behavior, vol. 21, no. 3, pp. 463–473, 1974.
84. J. Heckenbergerová, P. Kromer, and M. Hasal, *Modelling daily traffic flow distributions by finite mixtures of simple von Mises distributions*, in 19th Conference on Applied Mathematics, APLIMAT 2020 Proceedings, Slovenská technická univezita v Bratislave, 2020.
85. J. Ren and R. G. Vaughan, *Spaced antenna design in directional scenarios using the von mises distribution*, in 2009 IEEE 70th Vehicular Technology Conference Fall, pp. 1-5, IEEE, 2009.
86. A. Chabani, C. Mehl, I. Cojan, R. Alais, and D. Bruel, *Semi-automated component identification of a complex fracture network using a mixture of von Mises distributions: Application to the Ardeche margin (South-East France)*, Computers & Geosciences, vol. 137, p. 104435, 2020.

87. K. Mulder, P. Jongsma, and I. Klugkist, *Bayesian inference for mixtures of von Mises distributions using reversible jump MCMC sampler*, Journal of Statistical Computation and Simulation, vol. 90, no. 9, pp. 1539-1556, 2020.
88. J. T. Kent, *The circle is the von Mises distribution. However, when modeling more complicated features such as multimodality there are two frequently used approaches. First, one can use a mixture of von Mises distributions*, The Annals of Statistics, vol. 11, no. 3, pp. 984-988, 1983.
89. G. J. McLachlan and D. Peel, *Robust cluster analysis via mixtures of multivariate t-distributions*, in Advances in Pattern Recognition: Joint IAPR International Workshops SSPR'98 and SPR'98 Sydney, Australia, August 11–13, 1998 Proceedings, pp. 658-666, Springer Berlin Heidelberg, 1998.
90. D. Peel and G. J. McLachlan, *Robust mixture modelling using the t distribution*, Statistics and computing, vol. 10, pp. 339-348, 2000.
91. Y. Wang, S. Son, S. M. Swartz, and N. C. Goulbourne, *A mixed von Mises distribution for modeling soft biological tissues with two distributed fiber properties*, International Journal of Solids and Structures, vol. 49, no. 21, pp. 2914-2923, 2012.
92. S. Pfyffer and R. Gatto, *An efficient simulation algorithm for the generalized von Mises distribution of order two*, Computational Statistics, vol. 28, no. 1, pp. 255-268, 2013.
93. T. Dietrich and W. D. Richter, *Classes of geometrically generalized von Mises distributions*, Sankhya B, vol. 79, pp. 21-59, 2017.
94. X. Qin, J. S. Zhang, and X. D. Yan, *Model selection of the generalized von Mises distribution based on empirical mode decomposition with data analyses*, Statistical Methodology, vol. 14, pp. 15-25, 2013.
95. S. R. Jammalamadaka and A. SenGupta, *Topics in circular statistics*, vol. 5, World Scientific, 2001.
96. V. M. Maksimov, *Necessary and sufficient statistics for the family of shifts of probability distributions on continuous bicompact groups*, Theory of Probability & Its Applications, vol. 12, no. 2, pp. 267-280, 1967.
97. E. A. Yfantis and L. E. Borgman, *An extension of the von Mises distribution*, Communications in Statistics-Theory and Methods, vol. 11, no. 15, pp. 1695-1706, 1982.
98. A. P. Dempster, N. M. Laird, and D. B. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the royal statistical society: series B (methodological), vol. 39, no. 1, pp. 1-22, 1977.
99. A. Dumpster, N. Laird, & D. Rubin, *Maximum likelihood from incomplete data via the EM algorithm*, Journal of the Royal Statistical Society, 39(1), 1-38, 1977.
100. J. N. Rouder and J. Lu, *An introduction to Bayesian hierarchical models with an application in the theory of signal detection*, Psychonomic bulletin & review, vol. 12, no. 4, pp. 573-604, 2005.
101. H. G. Muller and U. Stadtmuller, *Variable bandwidth kernel estimators of regression curves*, The Annals of Statistics, pp. 182-201, 1987.
102. S. Weglarczyk, *Kernel density estimation and its application*, In ITM Web of Conferences, vol. 23, p. 00037, EDP Sciences, 2018.
103. H. Peter D, *Kernel estimation of a distribution function*, Communications in Statistics-Theory and Methods, vol. 14, no. 3, pp. 605-620, 1985.
104. J. X. Pan, K. T. Fang, *Maximum likelihood estimation*, in Growth Curve Models and Statistical Diagnostics, pp. 77-158, 2002.
105. B. Barazandeh, M. Razaviyayn, *On the behavior of the expectation-maximization algorithm for mixture models*, in 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), pp. 61-65, IEEE, November 2018.
106. S. Benlakhdar, M. Rziza, and R. O. H. Thami, *Directional Data Classification Using a Hierarchical Model of Von Mises Distribution*, dans Proceedings of the 2nd International Conference on Big Data, Cloud and Applications, pp. 1–6, mars 2017.
bibitemBarazandeh18 B. Barazandeh and M. Razaviyayn, *On the behavior of the expectation-maximization algorithm for mixture models*, in 2018 IEEE Global Conference on Signal and Information Processing (GlobalSIP), IEEE, pp. 61–65, November 2018.
107. W. Castro, J. Oblitas, R. Santa-Cruz, H. Avila-George, *Multilayer perceptron architecture optimization using parallel computing techniques*, PLoS One, vol. 12, no. 12, e0189369, 2017.
108. A. J. Rossini, L. Tierney, N. Li, *Simple parallel statistical computing in R*, Journal of Computational and Graphical Statistics, vol. 16, no. 2, pp. 399-420, 2007.
109. S. Frühwirth-Schnatter, *Panel data analysis: a survey on model-based clustering of time series*, Advances in Data Analysis and Classification, 5, 251-280, 2011.