Prog Artif Intell (2017) 6:171–180
DOI 10.1007/s13748-016-0109-x

REGULAR PAPER

Univariate and bivariate truncated von Mises distributions

Pablo Fernandez-Gonzalez[1] · **Concha Bielza**[1] · **Pedro Larrañaga**[1]

Received: 23 November 2016 / Accepted: 20 December 2016 / Published online: 10 January 2017

Abstract In this article we study the univariate and bivariate truncated von Mises distribution, as a generalization of the von Mises distribution. This implies the addition of two or four new truncation parameters in the univariate and, bivariate cases, respectively. The results include the definition, properties of the distribution and maximum likelihood estimators for the univariate and bivariate cases. Additionally, the analysis of the bivariate case shows how the conditional distribution is a truncated von Mises distribution, whereas the marginal is a generalization of the non-truncated marginal distribution. From the viewpoint of applications, we test the distribution with data regarding leaf inclination angles. This research aims to assert this probability distribution as a potential option for modeling or simulating any kind of phenomena where circular distributions are applicable.

Keywords Angular probability distributions · Directional statistics · von Mises distribution · Truncated probability distributions

1 Introduction

The von Mises distribution has received undisputed attention in the field of directional statistics [4] and in other areas like supervised classification [5]. Thanks to desirable properties such as its symmetry, mathematical tractability and convergence to the wrapped normal distribution [6] for high concentrations, it is a viable option for many statistical analyses. However, angular phenomena may present constraints on the outcomes that are not properly accounted for by the density function of the von Mises probability distribution. Thus, a truncated distribution with the capabilities of the von Mises distribution is strongly suggested. Additionally, there is hardly any literature in this direction, and to the best of our knowledge, only one paper [2], proposes a definition of the truncated von Mises distribution.

In this article, we propose a new definition of truncated probability distribution for angular values whose parent distribution is the von Mises distribution. The univariate and bivariate cases of this distribution are explicitly developed.

Section 2 introduces the definition for the univariate case and derives some properties of the distribution, calculates the maximum likelihood estimators of the parameters and studies the distribution moments. Section 3 addresses the definition of the bivariate truncated von Mises, maximum likelihood estimation of the parameters and the definition and study of the conditional and marginal truncated distributions. Section 4 shows real data application where the distribution successfully models the data. Finally, Sect. 5 discusses the summary and conclusions.

The first and second sections of the supplementary material cover the simulations and additional experiments on real datasets (for dihedral angles in protein chains in the bivariate case). In the third section, the proofs of all results can be found.

2 Univariate truncated von Mises distribution

Definition 2.1 The truncated von Mises distribution is presented as a four-parameter generalization of the non-truncated case for truncation parameters a, b as

Electronic supplementary material The online version of this article (doi:10.1007/s13748-016-0109-x) contains supplementary material, which is available to authorized users.

✉ Pablo Fernandez-Gonzalez
pablo.fernandezgonz@fi.upm.es

[1] Department of Artificial Intelligence, Universidad Politécnica de Madrid, Madrid, Spain

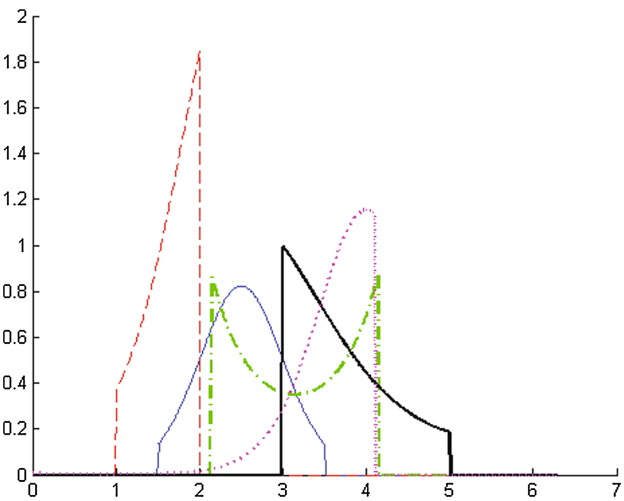

Fig. 1 Several truncated von Mises distributions. Symmetrical function with maxima not at the extrema (*thin continuous line*), strictly increasing function (*dashed line*), strictly decreasing function (*thick continuous line*), unique critical point that is a minimum (*dash-dot line*) and two critical points, a maximum and a minimum (*dotted line*)

$$f_{\text{tvM}}(\theta;\mu,\kappa,a,b)=\begin{cases}\frac{e^{\kappa\cos(\theta-\mu)}}{\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta} & \text{if } \theta\in\mathbb{O}_{a,b}\\ 0 & \text{if } \theta\in\mathbb{O}_{b,a}\end{cases} \qquad (1)$$

where $\mu\in\mathbb{O}$ is the location parameter, $\kappa>0$ the concentration parameter, $\mathbb{O}$ is the circular set of points ($\mathbb{O}:(x,y)$ such that $x^2+y^2=1$), $\mathbb{O}_{a,b}\subset\mathbb{O}$ is obtained by selecting the points in the circular path from $a\in\mathbb{O}$ to $b\in\mathbb{O}$ in the preferred direction (counterclockwise) and $\mathbb{O}_{b,a}$ is its counterpart w.r.t. $\mathbb{O}$.

Our proposed definition differs from Bistrian and Iakob [2] in the circular definition of the truncation parameters, not bounded to a linear definition involving the location parameter. The additional developments covered in this article can also be considered a novelty.

To illustrate the differences with the non-truncated case for these parameters, Fig. 1 represents multiple examples of truncated von Mises distributions.

It is a well-known result [1] that $2\pi I_0(\kappa)=\int_0^{2\pi}e^{\kappa\cos(\theta-\mu)}d\theta$, where $I_0(\kappa)$ is the modified Bessel function of the first kind and order 0, that is,

$$I_0(\kappa)=\sum_{m=0}^{\infty}\frac{x^{2m}}{(m!)^2 2^m}.$$

The above expression suffices for truncation parameters a,b such that $\mathbb{O}_{a,b}=\mathbb{O}$. However, it is necessary to calculate the general case for non-restricted truncation parameters. Taking $w=\lfloor\frac{n}{2}\rfloor+\operatorname{mod}\frac{n}{2}-1$, we have obtained:

Lemma 2.1 $\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta = I(b;\mu,\kappa)-I(a;\mu,\kappa)$, *where*

$$I(\theta;\mu,\kappa)=\sum_{n=0}^{\infty}\frac{\kappa^n}{n!}\left(\sin(\theta-\mu)\sum_{i=0}^{w}\left(\cos^{n-2i-1}(\theta-\mu)\prod_{j=0}^{2i}(n-j)^{-(-1)^j}\right)+\frac{((-1)^n+1)\prod_{j=0}^{w}(n-j)^{-(-1)^j}(\theta-\mu)}{2}\right). \qquad (2)$$

$I(\theta;\mu,\kappa)$ is the distribution function of the positive support of the truncated von Mises density. (Note then that while truncation parameters are circular quantities, the values for the integration coefficients are linear.)

Proof See supplementary material (Section 3.1). □

2.1 Maximum likelihood estimation

Provided we have a sample of observations $\theta_1,\theta_2,\ldots,\theta_n$ from a truncated von Mises distribution (1), we obtain:

$$\ln L(\mu,\kappa,a,b;\theta_1,\theta_2,\ldots,\theta_n)=\sum_{i=1}^{n}\ln\left(\frac{e^{\kappa\cos(\theta_i-\mu)}}{\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta}\right)=\sum_{i=1}^{n}\kappa\cos(\theta_i-\mu)-n\ln\left(\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta\right), \qquad (3)$$

where $\ln L(\mu,\kappa,a,b;\theta_1,\theta_2,\ldots,\theta_n)$ is the log-likelihood function for the truncated von Mises distribution.

We now seek to solve the system of four log-likelihood equations created by the four parameters of the distribution. For parameters μ,κ, we have

$$\frac{\partial\ln L}{\partial\mu}=0$$
$$\frac{\partial\ln L}{\partial\kappa}=0.$$

As parameters a,b, define the region of the greater-than-zero density, we find that all $\theta_1,\ldots,\theta_n$ observations necessarily lie within the subset $\mathbb{O}_{a,b}$. This, together with the $-n\ln(\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta)$ sub-term of (3), allows us to isolate the estimators

$$\mathbb{O}_{\hat{a},\hat{b}}=\underset{a,b}{\operatorname{argmax}}(\max(\{A(\mathbb{O}_{\theta'_1,\theta'_2}),\ldots,A(\mathbb{O}_{\theta'_{n-1},\theta'_n}),A(\mathbb{O}_{\theta'_n,\theta'_1})\})), \qquad (4)$$

where $A(\mathbb{O}_{\theta_1,\theta_2})$ is the angle between θ_1 and θ_2, and $\{\theta'_1,\ldots,\theta'_n\}$ is the sample sorted in ascending order. Intuitively, the truncation parameters are separated by the largest angle and are contiguous in a sorted finite circular sample.

From this result, we can say that the truncation parameters of the truncated von Mises distribution have population-only dependent maximum likelihood estimators. For parameters μ and κ, interdependency is a consequence of the possibly non-symmetrical shape of the distribution. If we observe the expressions

$$\frac{1}{n}\sum_{i=1}^{n}\sin(\theta_i-\mu)-\frac{e^{\kappa\cos(a-\mu)}-e^{\kappa\cos(b-\mu)}}{\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta}=0$$
$$\frac{1}{n}\sum_{i=1}^{n}\cos(\theta_i-\mu)-\frac{\int_a^b \cos(\theta-\mu)e^{\kappa\cos(\theta-\mu)}d\theta}{\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta}=0,$$

$e^{\kappa\cos(a-\mu)}-e^{\kappa\cos(b-\mu)}=0$ holds if a, b are symmetrical w.r.t. μ, reducing the location parameter estimator to that of the non-truncated case [6], the circular sample mean $\hat{\mu}$. As no population-only dependent expressions of the parameters μ and κ were found, optimization techniques to maximize the log-likelihood function for those parameters are needed.

2.2 Moments

The moments in circular statistics are particular values of the characteristic function. The rth moment about a direction d can be written as

$$m_{r_{\text{tvM}}}=\mathbb{E}[e^{ir(X-d)}].$$

The first moment about the 0 direction for the truncated von Mises is calculated as

$$m_{1_{\text{tvM}}}=\frac{\int_a^b \cos(\theta)e^{\kappa\cos(\theta-\mu)}d\theta}{\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta}+\frac{i\int_a^b \sin(\theta)e^{\kappa\cos(\theta-\mu)}d\theta}{\int_a^b e^{\kappa\cos(\theta-\mu)}d\theta}, \tag{5}$$

and we can relate (5) to the first moment about the μ direction, denoted as $m'_{1_{\text{tvM}}}$ as

$$m_{1_{\text{tvM}}}=e^{i\mu}m'_{1_{\text{tvM}}}. \tag{6}$$

Notice that if $\cos(a-\mu)=\cos(b-\mu)$, then $m'_{1_{\text{tvM}}}=\frac{\int_a^b \cos(x-\mu)e^{\kappa\cos(x-\mu)}d\theta}{\int_a^b e^{\kappa\cos(x-\mu)}d\theta}=R$, the mean resultant length of μ and thus $m_{1_{\text{tvM}}}=e^{i\mu}R$.

An alternative expression for $m_{1_{\text{tvM}}}$ can be found by considering equations $\mathbb{E}[\cos(x)]=R'\cos(\mu')$ and $\mathbb{E}[\sin(x)]=R'\sin(\mu')$, where R' and μ' are the sample mean resultant length and sample mean, respectively. We can then state

$$\begin{aligned}m_{1_{\text{tvM}}}&=\mathbb{E}[\cos(x)]+i\mathbb{E}[\sin(x)]\\&=R'\cos(\mu')+iR'\sin(\mu')=R'e^{i\mu'}.\end{aligned} \tag{7}$$

Thus, merging Eqs. (6) and (7), we obtain

$$e^{i(\mu'-\mu)}R'=m'_{1_{\text{tvM}}},$$

which can be seen as a valuable expression as it contains the sample mean (μ') and the location parameter of the distribution (μ).

3 Bivariate truncated von Mises distribution

The non-truncated bivariate von Mises distribution was first proposed by Singh [9] and extended and developed in Mardia et al. [7] and Mardia and Voss [8]. It is a unimodal/bi-modal function on the torus $f_{\text{btvM}}:\mathbb{O}\times\mathbb{O}\to\mathbb{R}$ obtained by replacing the quadratic and linear terms of the normal bivariate distribution with their circular analogues. This distribution is known as the "sin variant bivariate von Mises distribution" and is defined for dependent pairs of angular variables. It is expressed for variables θ_1 and θ_2, as

$$f(\theta_1,\theta_2)=Ce^{\kappa_1\cos(\theta_1-\mu_1)+\kappa_2\cos(\theta_2-\mu_2)+\lambda\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)},$$

where $\kappa_1,\kappa_2\geq 0$, $\lambda\in\mathbb{R}$, $\mu_1,\mu_2\in\mathbb{O}$ and C is the normalization constant. We propose the density function for the truncated case as a nine-parameter function with density defined as follows.

Definition 3.1 The density function for the truncated case is a nine-parameter function with density

$$f_{\text{btvM}}(\theta_1,\theta_2;\boldsymbol{W})=\begin{cases}\frac{f_{\text{ubvM}}(\theta_1,\theta_2;\boldsymbol{W})}{\int_{a_1}^{b_1}\int_{a_2}^{b_2}f_{\text{ubvM}}(\theta_1,\theta_2;\boldsymbol{W})d\theta_2 d\theta_1} & \text{if } \theta_1\in\mathbb{O}_{a_1,b_1},\theta_2\in\mathbb{O}_{a_2,b_2},\\ 0 & \text{otherwise}\end{cases} \tag{8}$$

where $\boldsymbol{W}=\{\lambda,\mu_1,\mu_2,\kappa_1,\kappa_2,a_1,b_1,a_2,b_2\}$ is the parameter vector and $f_{\text{ubvM}}(\theta_1,\theta_2;\boldsymbol{W})=e^{\kappa_1\cos(\theta_1-\mu_1)+\kappa_2\cos(\theta_2-\mu_2)+\lambda\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)}$ is the unnormalized bivariate von Mises distribution. Parameters μ_1,μ_2 and κ_1,κ_2 are analogous to parameters μ and κ, respectively, in the univariate truncated case. Truncation parameters a_1,b_1,a_2 and b_2 are similar to the univariate truncation parameters. The $\lambda\in\mathbb{R}$ parameter accounts for the dependency between the variable components (Fig. 2). If $\lambda=0$, then θ_1 and θ_2 are independent and each is distributed as a univariate von Mises distribution. Also, if θ_1,θ_2 are independent, then $\lambda=0$.

A desirable property of a joint distribution is having closed distributions under marginalization and conditioning, i.e., the marginal and conditional distributions should also follow the univariate distribution. Particularizing for the von Mises

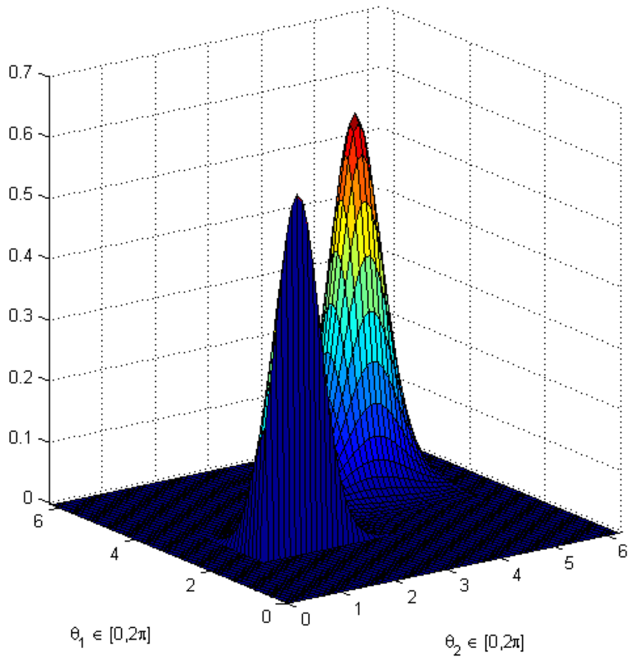

Fig. 2 Example of the bi-dimensional von Mises distribution showing truncated bi-modality

family, the bivariate von Mises distribution presents closed distributions only under conditioning as shown by Singh [9]. We want to find out whether this also holds for the truncated case.

3.1 Maximum likelihood estimation

The maximum likelihood estimator for the bivariate distribution takes data of the form $\{(\theta_{1i}, \theta_{2i})\}$ $i = 1, \ldots, n$. The resulting log-likelihood function is

$$\begin{aligned}
&\ln L(\boldsymbol{W}; (\theta_{11}, \theta_{21}), \ldots, (\theta_{1n}, \theta_{2n}))\\
&= \sum_{i=1}^{n} \ln \left(\frac{e^{\kappa_1 \cos(\theta_{1i}-\mu_1)+\kappa_2 \cos(\theta_{2i}-\mu_2)+\lambda \sin(\theta_{1i}-\mu_1)\sin(\theta_{2i}-\mu_2)}}{\int_{a_1}^{b_1}\int_{a_2}^{b_2} e^{\kappa_1 \cos(\theta_1-\mu_1)+\kappa_2 \cos(\theta_2-\mu_2)+\lambda \sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)} d\theta_2 d\theta_1} \right)\\
&= \sum_{i=1}^{n} (\kappa_1 \cos(\theta_{1i}-\mu_1) + \kappa_2 \cos(\theta_{2i}-\mu_2)\\
&\quad + \lambda \sin(\theta_{1i}-\mu_1)\sin(\theta_{2i}-\mu_2))\\
&\quad - n \ln \left(\int_{a_1}^{b_1}\int_{a_2}^{b_2} e^{\kappa_1 \cos(\theta_1-\mu_1)+\kappa_2 \cos(\theta_2-\mu_2)+\lambda \sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)} d\theta_2 d\theta_1 \right).
\end{aligned}$$

Thus we have

$$\frac{\partial}{\partial \mu_1} \ln L(\boldsymbol{W}; (\theta_{11}, \theta_{21}), \ldots, (\theta_{1n}, \theta_{2n})) = 0,$$

that is,

$$\begin{aligned}
&\sum_{i=1}^{n} \kappa_1 \sin(\theta_{1i}-\mu_1) - \lambda \cos(\theta_{1i}-\mu_1)\sin(\theta_{2i}-\mu_2)\\
&\quad - \frac{n\left(\int_{a_2}^{b_2} f_{\text{ubvM}}(a_1, \theta_2) - f_{\text{ubvM}}(b_1, \theta_2) d\theta_2\right)}{\int_{a_1}^{b_1}\int_{a_2}^{b_2} f_{\text{ubvM}}(\theta_1, \theta_2) d\theta_2 d\theta_1} = 0,
\end{aligned}$$

where $f_{\text{ubvM}}(\theta_1, \theta_2)$ is the following unnormalized bivariate truncated von Mises function

$$f_{\text{ubvM}}(\theta_1, \theta_2) = e^{\kappa_1 \cos(\theta_1-\mu_1)+\kappa_2 \cos(\theta_2-\mu_2)+\lambda \sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)}.$$

Similarly, the partial derivate w.r.t. μ_2 gives

$$\begin{aligned}
&\sum_{i=1}^{n} \kappa_2 \sin(\theta_{2i}-\mu_2) - \lambda \cos(\theta_{2i}-\mu_2)\sin(\theta_{1i}-\mu_1)\\
&\quad - \frac{n\left(\int_{a_1}^{b_1} f_{\text{ubvM}}(\theta_1, a_2) - f_{\text{ubvM}}(\theta_1, b_2) d\theta_1\right)}{\int_{a_1}^{b_1}\int_{a_2}^{b_2} f_{\text{ubvM}}(\theta_1, \theta_2) d\theta_2 d\theta_1} = 0.
\end{aligned}$$

For κ_1 we have

$$\frac{\partial}{\partial \kappa_1} \ln L(\boldsymbol{W}; (\theta_{11}, \theta_{21}), \ldots, (\theta_{1n}, \theta_{2n})) = 0,$$

that is,

$$\begin{aligned}
&\frac{1}{n}\sum_{i=1}^{n} \cos(\theta_{1i}-\mu_1)\\
&\quad - \frac{\int_{a_1}^{b_1}\int_{a_2}^{b_2} \cos(\theta_1-\mu_1) f_{\text{ubvM}}(\theta_1, \theta_2) d\theta_2 d\theta_1}{\int_{a_1}^{b_1}\int_{a_2}^{b_2} f_{\text{ubvM}}(\theta_1, \theta_2) d\theta_2 d\theta_1} = 0. \qquad (9)
\end{aligned}$$

Similarly, the partial derivate w.r.t. κ_2 gives

$$\frac{1}{n}\sum_{i=1}^{n}\cos(\theta_{2i}-\mu_2) \\ -\frac{\int_{a_1}^{b_1}\int_{a_2}^{b_2}\cos(\theta_2-\mu_2)f_{\text{ubvM}}(\theta_1,\theta_2)\mathrm{d}\theta_2\mathrm{d}\theta_1}{\int_{a_1}^{b_1}\int_{a_2}^{b_2}f_{\text{ubvM}}(\theta_1,\theta_2)\mathrm{d}\theta_2\mathrm{d}\theta_1}=0. \quad (10)$$

At this point, we can see that both equations (9) and (10), involving κ_1, κ_2 parameters, respectively, preserve their analogy with the univariate case. Their second addend corresponds to the definition of the estimators of $\mathbb{E}[\cos(\theta_1-\mu_1)]$ and $\mathbb{E}[\cos(\theta_2-\mu_2)]$, respectively.

For the parameter λ we obtain

$$\frac{\partial}{\partial\lambda}\ln L(\boldsymbol{W};(\theta_{11},\theta_{21}),\ldots,(\theta_{1n},\theta_{2n}))=0,$$

that is,

$$\frac{1}{n}\sum_{i=1}^{n}\sin(\theta_{1i}-\mu_1)\sin(\theta_{2i}-\mu_2) \\ -\frac{\int_{a_1}^{b_1}\int_{a_2}^{b_2}\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)f_{\text{ubvM}}(\theta_1,\theta_2)\mathrm{d}\theta_2\mathrm{d}\theta_1}{\int_{a_1}^{b_1}\int_{a_2}^{b_2}f_{\text{ubvM}}(\theta_1,\theta_2)\mathrm{d}\theta_2\mathrm{d}\theta_1}=0,$$

which analogously corresponds to the estimator of $\mathbb{E}[\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)]$. As in the univariate case, the truncation parameters has the following isolated estimators

$$\mathbb{O}_{\hat{a}_1,\hat{b}_1}=\underset{a_1,b_1}{\operatorname{argmax}}\left(\max\left(\left\{A\left(\mathbb{O}_{\theta'_{11},\theta'_{12}}\right),\ldots,A\left(\mathbb{O}_{\theta'_{1n-1},\theta'_{1n}}\right),\right.\right.\right. \\ \left.\left.\left.A\left(\mathbb{O}_{\theta'_{1n},\theta'_{11}}\right)\right\}\right)\right) \\ \mathbb{O}_{\hat{a}_2,\hat{b}_2}=\underset{a_2,b_2}{\operatorname{argmax}}\left(\max\left(\left\{A\left(\mathbb{O}_{\theta'_{21},\theta'_{22}}\right),\ldots,A\left(\mathbb{O}_{\theta'_{2n-1},\theta'_{2n}}\right),\right.\right.\right. \\ \left.\left.\left.A\left(\mathbb{O}_{\theta'_{2n},\theta'_{21}}\right)\right\}\right)\right),$$

while as yielded by the above calculations, the expressions regarding the non-truncation parameters exhibit interdependency.

$$f_{\text{mtvM}}(\theta_1;\boldsymbol{W})=\begin{cases}\dfrac{\int_{a_2}^{b_2}\mathrm{e}^{\kappa_1\cos(\theta_1-\mu_1)+\kappa_2\cos(\theta_2-\mu_2)+\lambda\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)}\mathrm{d}\theta_2}{\int_{a_1}^{b_1}\int_{a_2}^{b_2}\mathrm{e}^{\kappa_1\cos(\theta_1-\mu_1)+\kappa_2\cos(\theta_2-\mu_2)+\lambda\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)}\mathrm{d}\theta_2\mathrm{d}\theta_1} & \text{if } \theta_1\in\mathbb{O}_{a_1,b_1} \\ 0 & \text{otherwise}\end{cases} \quad (12)$$

3.2 Conditional truncated von Mises distribution

The density of the conditional truncated von Mises distribution is defined as:

Definition 3.2 The conditional truncated von Mises distribution has density

$$f_{\text{ctvM}}(\theta_2|\theta_1;\lambda,\mu_1,\mu_2,\kappa 2,a_2,b_2) \\ =\begin{cases}\dfrac{\mathrm{e}^{\kappa_2\cos(\theta_2-\mu_2)+\lambda\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)}}{\int_{a_2}^{b_2}\mathrm{e}^{\kappa_2\cos(\theta_2-\mu_2)+\lambda\sin(\theta_1-\mu_1)\sin(\theta_2-\mu_2)}\mathrm{d}\theta_2} & \text{if } \theta_2\in\mathbb{O}_{a_2,b_2}. \\ 0 & \text{otherwise}\end{cases} \quad (11)$$

It is a six-parameter distribution where the parameters hold the same meaning as in the bivariate case, with the simplification of parameters κ_1, a_1, b_1 for $f_{\text{ctvM}}(\theta_2|\theta_1)$ (or κ_2, a_2, b_2 for $f_{\text{ctvM}}(\theta_1|\theta_2)$). Worthy of note, however, is that $\theta_1\in\mathbb{O}_{a_1,b_1}$ in $f_{\text{ctvM}}(\theta_2|\theta_1)$ since otherwise, by the definition of the conditional distribution ($f_{\text{ctvM}}(\theta_2|\theta_1)=\frac{f_{\text{btvM}}(\theta_2,\theta_1)}{f_{\text{tvM}}(\theta_1)}$), $f_{\text{ctvM}}(\theta_2|\theta_1)$ is not defined.

Theorem 3.1 *A conditional truncated von Mises distribution corresponds to the univariate truncated von Mises distribution*

$$f_{\text{ctvM}}(\theta_2|\theta_1;\lambda,\mu_1,\mu_2,\kappa_2,a_2,b_2) \\ =f_{\text{tvM}}\left(\theta_2;\mu_2+\arctan\left(\frac{\lambda\sin(\theta_1-\mu_1)}{\kappa_2}\right),\right. \\ \left.\times\sqrt{\kappa_2^2+(\lambda\sin(\theta_1-\mu_1))^2},a_2,b_2\right),$$

which completely specifies the behavior and properties of the conditional distribution and is analogous to the non-truncated conditional case [9].

Proof See supplementary material (Section 3.2). □

3.3 Marginal truncated von Mises distribution

We can define the density function of the marginal truncated von Mises distribution as:

Definition 3.3 The density function of the marginal truncated von Mises distribution can be written as

It is a nine-parameter distribution that shares all the parameters with the bivariate truncated von Mises distribution. In the original publication, Singh [9] studied the distribution and reported the "frontiers" of bi-modality (for $\mu=0$) as

$$\frac{I_1(\kappa_2)}{I_0(\kappa_2)}=\frac{\kappa_1\kappa_2}{\lambda^2},$$

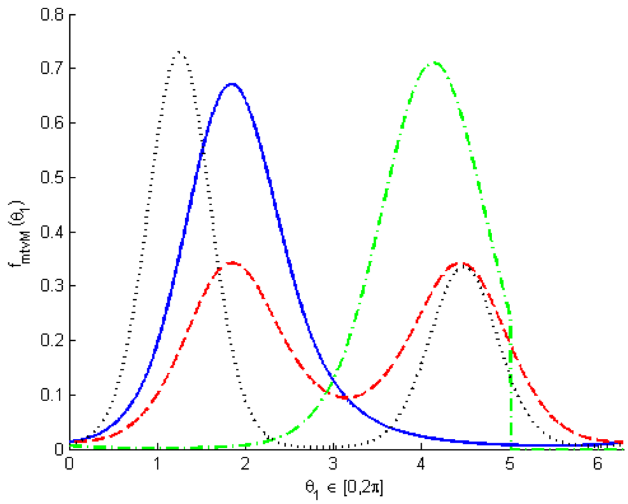

Fig. 3 Several truncated marginal distributions showing unimodality (*continuous line*), two equal maxima (*dashed line*), truncated unimodality (*dash-dot line*) and two distinct maxima (*dotted line*)

where the distribution is unimodal if $\frac{I_1(\kappa_2)}{I_0(\kappa_2)} \geq \frac{\kappa_1\kappa_2}{\lambda^2}$, and bimodal with two equal maxima otherwise. Additionally, the modes were calculated to be symmetrical w.r.t μ_1 and at the distance value θ_1^* that solves the equation (for $\mu_1 = 0$):

$$\frac{A\left(\sqrt{\kappa_2 + \lambda^2 \sin^2(\theta_1^*)}\right)}{\sqrt{\kappa_2 + \lambda^2 \sin^2(\theta_1^*)}} \cos(\theta_1^*) = \frac{\kappa_1}{\lambda^2},$$

where $A(x) = \frac{I_1(x)}{I_0(x)}$. In order to generalize this analysis to cover the truncated case in Eq. (12), we need to account for the contribution made by the parameters μ_2, a_2 and b_2 to the shape of the distribution. Contrary to the non-truncated case, a truncated marginal distribution that exhibits two maxima may have only one global maximum, and the distribution is not necessarily centered around the mean (Fig. 3). Therefore, our analysis determines the different parameter configurations that produce the whole range of behaviors, focusing on bi-modality/unimodality.

If, without loss of generality, we take $\theta_{1'} = \theta_1 - \mu_1$, we can postulate the following theorem:

Theorem 3.2 *All different behaviors w.r.t. the unimodality/bi-modality of the marginal truncated von Mises distribution can be accounted for as follows:*

1. $f_{\mathrm{mtvM}}(\theta_{1'})$ *is unimodal with mode (maximum) in* μ_1*, if and only if* $T(\lambda, \mu_2, \kappa_1, \kappa_2, a_2, b_2) < 0$ *and* $\cos(b_2 - \mu_2) = \cos(a_2 - \mu_2)$.
2. $f_{\mathrm{mtvM}}(\theta_{1'})$ *is bi-modal with equal maxima, if and only if* $T(\lambda, \mu_2, \kappa_1, \kappa_2, a_2, b_2) > 0$ *and* $\cos(b_2 - \mu_2) = \cos(a_2 - \mu_2)$*. Also in this case, a minimum is found at* $\theta_{1'} = 0$.
3. $f_{\mathrm{mtvM}}(\theta_{1'})$ *presents two differentiated maxima if and only if one of the two following cases applies*:
   (a) $\cos(b_2 - \mu_2) < \cos(a_2 - \mu_2)$ *and* $f'_{\mathrm{umtvM}}(\theta_{1'}; \lambda, \mu_1, \mu_2, \kappa_1, \kappa_2, \mu_2, a_2, b_2)$ *has exactly two zero points in* $\theta_{1'} \in [-\frac{\pi}{2}, 0]$.
   (b) $\cos(b_2 - \mu_2) > \cos(a_2 - \mu_2)$ *and* $f'_{\mathrm{umtvM}}(\theta_{1'}; \lambda, \mu_1, \mu_2, \kappa_1, \kappa_2, \mu_2, a_2, b_2)$ *has exactly two zero points in* $\theta_{1'} \in [0, \frac{\pi}{2}]$.
4. $f_{\mathrm{mtvM}}(\theta_{1'})$ *is unimodal with mode not at* μ_1 *if the parameters do not match any of the above cases*,

where $T(\lambda, \mu_2, \kappa_1, \kappa_2, a_2, b_2)$ *is the test function and is defined as*

$$T(\lambda, \mu_2, \kappa_1, \kappa_2, a_2, b_2) = -\frac{\kappa_1}{\lambda^2} + \frac{\int_{a_2}^{b_2} \sin^2(\theta_2 - \mu_2) e^{\kappa_2 \cos(\theta_2 - \mu_2)} d\theta_2}{\int_{a_2}^{b_2} e^{\kappa_2 \cos(\theta_2 - \mu_2)} d\theta_2}, \quad (13)$$

and $f'_{\mathrm{umtvM}}(\theta_{1'}; \lambda, \mu_1, \mu_2, \kappa_1, \kappa_2, \mu_2, a_2, b_2)$ *is the unnormalized truncated marginal von Mises derivative function.*

Proof See supplementary material (Section 3.3). □

Table 1 Parameter values obtained after conducting the first study

	μ	κ	a	b	No. samples
All data	1.0063	5.9602	0	1.5708	741

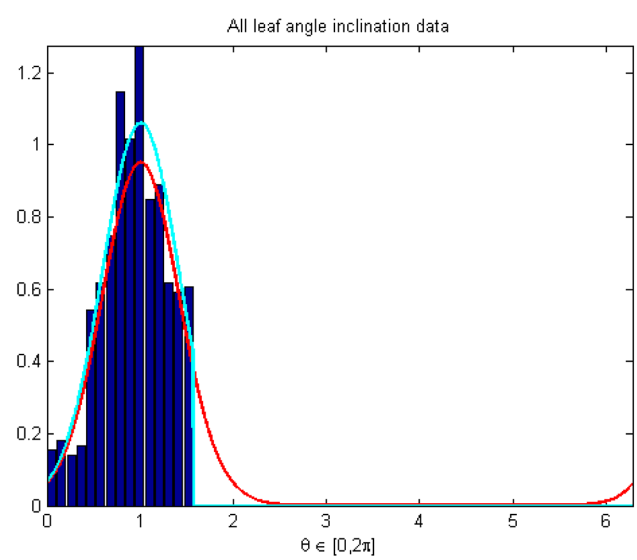

Fig. 4 The study distribution and data representation of the entire dataset. The estimated truncated von Mises distribution (*lighter line*) clearly has higher density values than its associated von Mises distribution (*darker line*). The data are grouped by value intervals in order to observe its relative frequency (*bars*)

Table 2 Modified Rayleigh statistic values for the second study

	Truncated von Mises S^*	Non-truncated von Mises S^*
A. erioloba	3.014	3.5534
Grewia flava	0.0038	20.6273
A. leuderitzii	2.6073	10.1990
A. mellifera	1.3157	7.3046

Table 3 Parameter values yielded after conducting the second study

	μ	κ	a	b	No. samples
A. erioloba	0.8516	11.1894	0	1.5359	100
Grewia flava	1.1261	5.2668	0	1.5708	254
A. leuderitzii	1.0706	5.5138	0	1.5708	184
A. mellifera	0.9125	5.7396	0	1.5708	203

4 Real data application

4.1 Leaf angle inclination

The data in Bowyer and Danson [3] was collected during a safari along the Kalahari Transect, southwest Botswana in 2001. It contains measurements of leaf inclination angles of four different woody plant species (*Acacia erioloba*, *Grewia flava*, *Acacia leuderitzii* and *Acacia mellifera*) across three different regions (Mabuasehube, Tsabong and Tshane). The measurements were taken using a clinometer.

In order to formally test the goodness-of-fit of the estimated distributions, we transform the data by means of the random variable $U = 2\pi \frac{[I(\theta,\mu,\kappa)-I(a,\mu,\kappa)]}{\int_a^b e^{\kappa\cos(\theta-\mu)}\mathrm{d}\theta} \mod 2\pi$ that is applied over the sorted sample $\theta_1, \ldots, \theta_n$. If the data distribute according to the truncated von Mises distribution, then the above random variable has a uniform distribution. As shown in Mardia and Jupp [6], the modified Rayleigh statistic $S^* = (1-\frac{1}{2n})2nR^2 + \frac{nR^4}{2}$, where n is the sample size and R the mean resultant length, distributes as a χ^2_2 distribution.

1. For the first study, the whole dataset containing a total of 741 samples was observed (Table 1; Fig. 4). A visual inspection of the plot clearly shows that the truncated von Mises distribution performs better. Formally, for the truncated case we have $S^* = 2.8887$, which corresponds to p-value $\in (0.2, 0.3)$. For the non-truncated case, $S^* = 25.5028$, with is a clear rejection p-value < 0.001. From these results we conclude that the truncated distribution is significantly better for these data. Truncation parameters conform the circular interval $\mathbb{O}_{0,\frac{\pi}{2}}$, which indicates no angle greater than 90° was measured in this study.

2. For the second study, we grouped the data by plant types without regard for region. This yielded four different distributions. A visual inspection shows that the univariate distributions are clearly better than the non-truncated von Mises distribution at describing the resulting data (Table 3; Fig. 5), except for the case of *A. erioloba*. The goodness-of-fit tests (Table 2) revealed that the

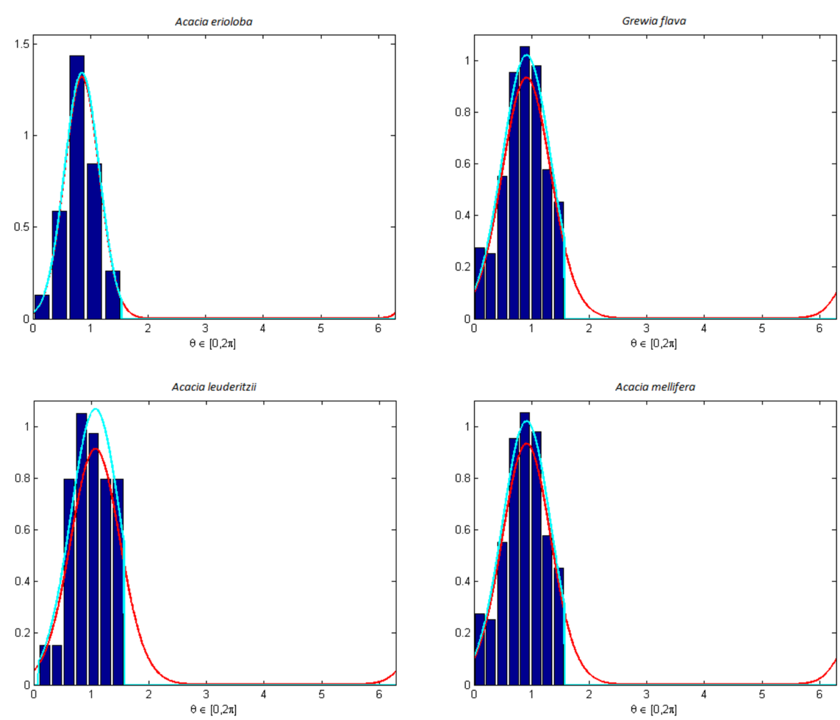

Fig. 5 Studies of each type of plant

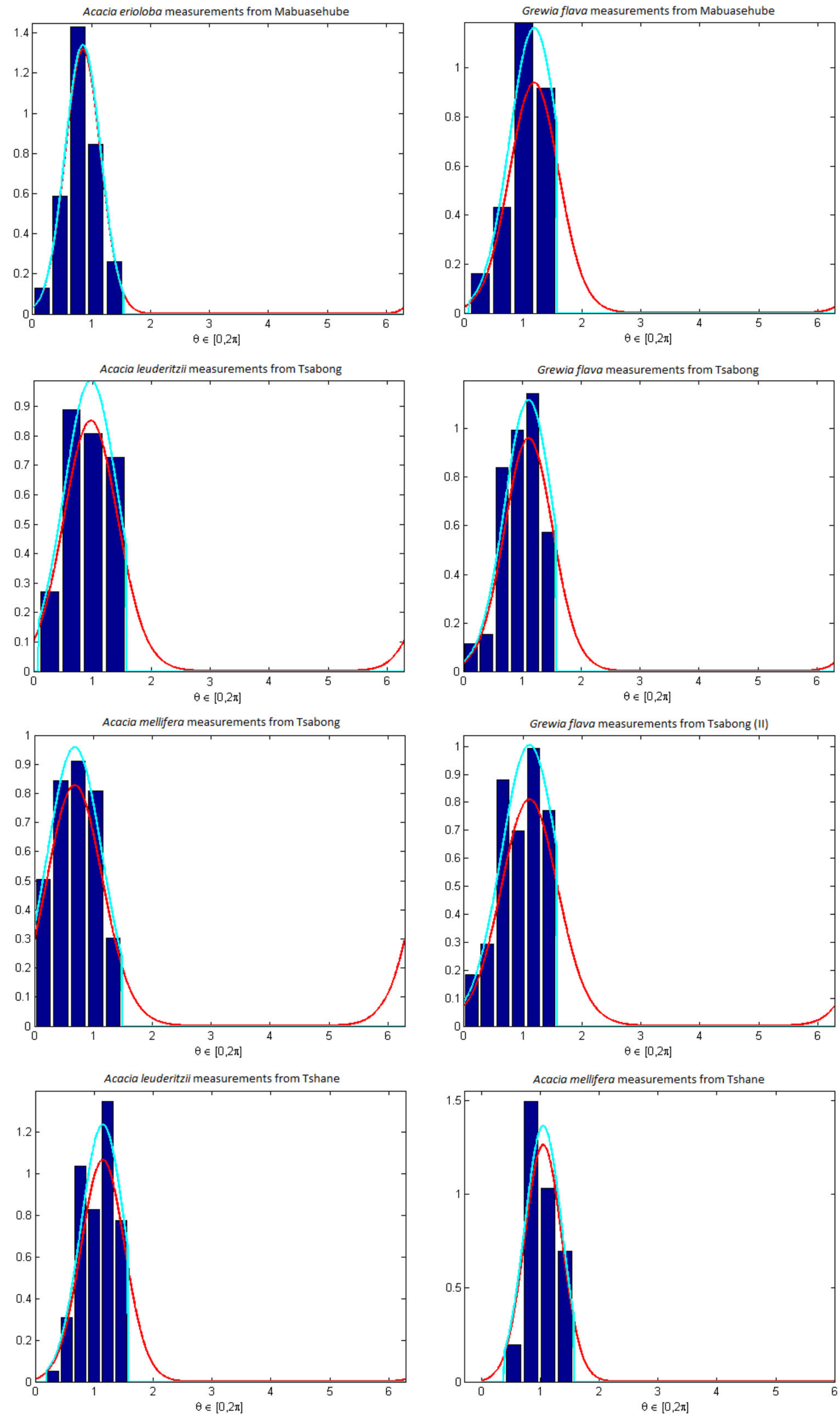

Fig. 6 Studies of each type of plant in each region

Table 4 Parameter values yielded after conducting the third study

	Truncated von Mises S^*	Non-truncated von Mises S^*
A. erioloba, Mabuasehube	3.014	3.5534
Grewia flava, Mabuasehube	1.1543	8.9599
A. leuderitzii, Tsabong	2.0981	7.3115
Grewia flava, Tsabong	0.2050	3.8702
A. mellifera, Tsabong	0.1199	4.2131
Grewia flava(2), Tsabong	0.1165	9.7290
A. leuderitzii, Tshane	0.7002	2.8717
A. mellifera, Tshane	1.0525	10.2656

Table 5 Parameter values yielded after conducting the third study

	μ	κ	a	b	No. samples
A. erioloba, Mabuasehube	0.8516	11.1894	0	1.5359	100
Grewia flava, Mabuasehube	1.1882	5.8142	0.0873	1.5708	50
A. leuderitzii, Tsabong	0.9712	4.8340	0.0873	1.5708	100
Grewia flava, Tsabong	1.1082	6.0832	0	1.5708	100
A. mellifera, Tsabong	0.6844	4.5884	0	1.4835	100
Grewia flava (2), Tsabong	1.1091	4.4078	0	1.5708	104
A. leuderitzii, Tshane	1.1474	7.4245	0.1920	1.5708	84
A. mellifera, Tshane	1.0525	10.2656	0.4014	1.5708	103

non-truncated distribution is rejected in all cases but in *A.erioloba*, whereas the truncated distribution hypothesis was more strongly accepted than that of the non-truncated distribution in all cases. Thus we can conclude that, for this study, the truncated distribution models the data better.

Truncation parameters were consistently found to be in $\mathbb{O}_{0,\frac{\pi}{2}}$ except for *A. erioloba*, which also presented a significantly higher concentration parameter than in any of the other estimations. The irregularities in *A. erioloba* could partially be explained by the small sample size, which causes the estimations to be less reliable. On the whole, the remaining studies show few variations in the location-concentration parameters, which closely resemble the ones obtained in the first study.

3. For the third study, fitted univariate truncated distributions for each plant in each region. Since not all plants were measured in all regions, this procedure produced eight different univariate truncated von Mises estimations. The distributions are generally observed to clearly differ from their associated non-truncated von Mises distribution, except in the first of the eight plots (Table 5; Fig. 6). The goodness-of-fit tests (Table 4) are also consistent with previous studies. All truncated von Mises hypotheses were accepted, while around half of the non-truncated distributions were rejected. Thus, there is a strong suggestion that the truncated von Mises distribution properly models the underlying behavior that yielded the data.

For this study, each distribution was estimated from a relatively small sample size ranging from 50 to 104 samples, which may have caused estimations to be less precise than desired. The concentration parameter shows the highest variability across the different cases (from 4.4078 to 11.1894 across the whole study or even from 4.8340 to 7.4245 in the case of *A. leuderitzii*). With more data it might be possible to distinguish if the variations in the concentration parameter are clearly influenced by the region of the plant species or the small sample size. Regarding the location parameter, there are few variations in the parameter value on the whole, *A. mellifera* being the species that experienced the highest variations w.r.t. one of the measurements in the first study. Truncation parameters remained consistently within the $\mathbb{O}_{0,\frac{\pi}{2}}$ interval.

5 Summary and conclusions

In this article, we developed the theoretical framework of the univariate and the bivariate truncated von Mises distribution. To do this, we gave

1. The definition of a truncated von Mises distribution in the circle $\mathbb{O}$. The circular distribution is defined by means of

the $\mathbb{O}$ subset, as the periodicity and properties of the circle have to be naturally acknowledged for.

2. The successfully determined expressions of the maximum likelihood estimators. For both univariate and bivariate cases, solely sample-dependent maximum likelihood estimators of the truncation parameters were found, while the other parameters showed interdependency.
3. The resulting moments of the univariate case and existing interrelationships.
4. The bivariate case and studies of the shape and behavior of marginal and conditional distributions. We determined that every conditional truncated von Mises distribution is a univariate truncated von Mises distribution. For the case of the marginal distribution, we concluded that only for parameter $\lambda = 0$ does the distribution behave like a truncated univariate von Mises distribution. When $\lambda \neq 0$, the resultant marginal distribution is potentially bimaximal and not a von Mises distribution. The modality behavior of this distribution has been accounted for in Theorem 3.2.

Acknowledgements This work has been partially supported by the Spanish Ministry of Economy and Competitiveness through the Cajal Blue Brain (C080020-09; the Spanish partner of the Blue Brain initiative from EPFL) and TIN2013-41592-P projects, by the Regional Government of Madrid through the S2013/ICE-2845-CASI-CAM-CM project.

References

1. Abramowitz, M., Stegun, I.: Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables. Applied Mathematics Series. Dover Publications, New York (1964)
2. Bistrian, D.A., Iakob, M.: One-dimensional truncated von Mises distribution in data modeling. Ann. Fac. Eng. Hunedoara tome **VI**, fascicule 3 (2008)
3. Bowyer, P., N.M.T., Danson, F.M.: SAFARI 2000 Canopy Structural Measurements, Kalahari Transect, Wet Season 2001. Data set (2005)
4. Jupp, P.E., Mardia, K.V.: A unified view of the theory of directional statistics, 1975–1988. Int. Stat. Rev. **57**(3), 261–294 (1989)
5. Lopez-Cruz, P., Bielza, C., Larrañaga, P.: Directional naive Bayes classifiers. Pattern Anal. Appl., pp. 1–22 (2013)
6. Mardia, K., Jupp, P.: Directional Statistics. Wiley Series in Probability and Statistics (2000)
7. Mardia, K.V., Hughes, G., Taylor, C.C., Singh, H.: A multivariate von Mises distribution with applications to bioinformatics. Can. J. Stat. **36**, 99–109 (2008)
8. Mardia, K.V., Voss, J.: Some fundamental properties of a multivariate von Mises distribution. Commun. Stat. Theory Methods **43**(6), 1132–1144 (2014)
9. Singh, H.: Probabilistic model for two dependent circular variables. Biometrika **89**(3), 719–723 (2002)