Bayesian tests of symmetry for the generalized Von Mises distribution

Sara Salvador¹ · Riccardo Gatto¹

Received: 2 July 2020 / Accepted: 30 August 2021 / Published online: 13 October 2021

© The Author(s) 2021

Abstract

Bayesian tests on the symmetry of the generalized von Mises model for planar directions (Gatto and Jammalamadaka in *Stat Methodol* 4(3):341–353, 2007) are introduced. The generalized von Mises distribution is a flexible model that can be axially symmetric or asymmetric, unimodal or bimodal. A characterization of axial symmetry is provided and taken as null hypothesis for one of the proposed Bayesian tests. The Bayesian tests are obtained by the technique of probability perturbation. The prior probability measure is perturbed so to give a positive prior probability to the null hypothesis, which would be null otherwise. This allows for the derivation of simple computational formulae for the Bayes factors. Numerical results reveal that, whenever the simulation scheme of the samples supports the null hypothesis, the null posterior probabilities appear systematically larger than their prior counterpart.

Keywords Axial symmetry · Bayes factor · Circular distribution · Probability perturbation · Uni- and bimodality

1 Introduction

In various scientific fields measurements can take the form of directions: the direction flight of a bird and the direction of earth's magnetic pole are two examples. These directions can be in the plane, namely in two dimensions, as in the first example, or they can be in the space, namely in three dimensions, as in the second example. These

The authors are grateful to two anonymous Referees and an Associate Editor for several suggestions and corrections.

✉ Sara Salvador
sarasalva91@gmail.com

Riccardo Gatto
gatto@stat.unibe.ch

¹ Department of Mathematics and Statistics, Institute of Mathematical Statistics and Actuarial Science, University of Bern, Alpeneggstrasse 22, 3012 Bern, Switzerland

measurements are called directional data and they appear in various scientific fields: in the analysis of protein structure, in machine learning, in forestry, in ornithology, in palaeomagnetism, in oceanography, in meteorology, in astronomy, etc. A two-dimensional direction is a point in $\mathbb{R}^2$ without magnitude, e.g. a unit vector. It can also be represented as a point on the circumference of the unit circle or as an angle, measured for example in radians and after fixing the null direction and the sense of rotation (clockwise or counter-clockwise). Because of this circular representation, observations on two-dimensional directional data are distinctively called circular data. During the last two or three decades, there has been a raise of interest for statistical methods for directional data. Recent applications can be found e.g. in Ley and Verdebout (2018). Some monographs on this topic are Mardia and Jupp (2000), Jammalamadaka and SenGupta (2001), Ley and Verdebout (2017) and also Pewsey et al. (2013). For a review article, see e.g. Gatto and Jammalamadaka (2014).

The popular probability distribution for circular data, or circular distribution, is the circular normal or von Mises (vM) distribution, whose density is given in (3) below. This distribution is symmetric around its unique mode. Until a couple decades ago, very few asymmetric circular distributions were available, two of these can be found in Sections 15.6 and 15.7 of Batschelet (1981). In recent years, various asymmetric or multimodal circular distributions have been introduced, for example: Umbach and Jammalamadaka (2009), Kato and Jones (2015), Abe et al. (2013), Gatto and Jammalamadaka (2003) and the generalized von Mises (GvM) of Gatto and Jammalamadaka (2007). This article proposes three Bayesian tests for the GvM distribution. This distribution has density given

$$f(\theta \mid \mu_1, \mu_2, \kappa_1, \kappa_2) = \frac{1}{2\pi G_0(\delta, \kappa_1, \kappa_2)} \exp\{\kappa_1 \cos(\theta - \mu_1) + \kappa_2 \cos 2(\theta - \mu_2)\}, \quad (1)$$

$\forall \theta \in [0, 2\pi)$, for given $\mu_1 \in [0, 2\pi)$, $\mu_2 \in [0, \pi)$, $\delta = (\mu_1 - \mu_2) \bmod \pi$, $\kappa_1, \kappa_2 > 0$, and where the normalizing constant is given by

$$G_0(\delta, \kappa_1, \kappa_2) = \frac{1}{2\pi} \int_0^{2\pi} \exp\{\kappa_1 \cos \theta + \kappa_2 \cos 2(\theta + \delta)\} d\theta. \quad (2)$$

We denote this distribution by $\text{GvM}(\mu_1, \mu_2, \kappa_1, \kappa_2)$. The well-known vM density is obtained by setting $\kappa_2 = 0$ in (1), giving

$$f(\theta \mid \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}, \quad (3)$$

$\forall \theta \in [0, 2\pi)$, for given $\mu \in [0, 2\pi)$, $\kappa > 0$ and where $I_\nu(z) = (2\pi)^{-1} \int_0^{2\pi} \cos \nu \theta \exp\{z \cos \theta\} d\theta$, $z \in \mathbb{C}$, is the modified Bessel function I of order ν , with $\Re \nu > -1/2$, cf. 9.6.18 at p. 376 of Abramowitz and Stegun (1972). We denote this distribution by $\text{vM}(\mu, \kappa)$.

Besides its greater flexibility in terms of asymmetry and bimodality, the GvM distribution possesses the following properties that other asymmetric or multimodal circular distributions do not have.

1. After a reparametrization, the GvM distribution belongs to the canonical exponential class. In this form, it admits a minimal sufficient and complete statistic; cf. Sect. 2.1 of Gatto and Jammalamadaka (2007).
2. The maximum likelihood estimator and the trigonometric method of moments estimator of the parameters are the same; cf. Sect. 2.1 of Gatto (2008). In this context, we should note that the computation of the maximum likelihood estimator is simpler with the GvM distribution than with the mixture of two vM distributions, as explained some lines below.
3. It is shown in Sect. 2.2 of Gatto and Jammalamadaka (2007) that for fixed trigonometric moments of orders one and two, the GvM distribution is the one with largest entropy. The entropy gives a principle for selecting a distribution on the basis of partial knowledge: one should always choose distributions having maximal entropy, within distributions satisfying the partial knowledge. In Bayesian statistics, whenever a prior distribution has to be selected and information on the first two trigonometric moments is available, then the GvM is the optimal prior. For other information theoretic properties of the GvM, see Gatto (2009).

The mixture of two vM distributions is perhaps a more popular bimodal or asymmetric model than the GvM. However, the mixture does not share the given properties 1–3 of the GvM. The mixture is not necessarily more practical. While the likelihood of the GvM distribution is bounded, the likelihood of the mixture of the $\text{vM}(\mu_1, \kappa_1)$ and the $\text{vM}(\mu_2, \kappa_2)$ distributions is unbounded. As $\kappa_1 \rightarrow \infty$, the likelihood when μ_1 is equal to any one of the sample values tends to infinity. This follows from $I_0(\kappa_1) \sim (2\pi\kappa_1)^{-1/2}e^{\kappa_1}$, as $\kappa_1 \rightarrow \infty$; cf. Abramowitz and Stegun (1972), 9.7.1 at p. 377. For alternative estimators to the maximum likelihood for vM mixtures, refer to Spurr and Koutbeiy (1991).

Some recent applications of the GvM distributions are: Zhang et al. (2018), in meteorology, Lin and Dong (2019), in oceanography, Astfalck et al. (2018), in offshore engineering, Christmas (2014), in signal processing, and Gatto (2021) in time series analysis.

The symmetry of a circular distribution is a fundamental question and, as previously mentioned, this topic has been studied in recent years. In the context of testing symmetry, one can mention: Pewsey (2002), who proposes a test of symmetry around an unknown axis based on the second sine sample moment, and Pewsey (2004), who considers the case where the symmetry is around the median axis. Both tests are frequentist and no Bayesian test of symmetry appears available in the literature. In fact, Bayesian analysis for circular data has remained underdeveloped, partly because of the lack of nice conjugate classes of distributions. Moreover, Bayesian analysis has focused on the vM model, which is symmetric. We refer to p. 278–279 of Jammalamadaka and SenGupta (2001) for a review on Bayesian analysis for circular data.

In this context, this article proposes Bayesian tests of symmetry for the GvM model (1). The first test proposed concerns the parameter δ . The null hypothesis is $\delta = 0$, that is, no shift between cosines of frequency one and two. In this case, the distribution

is symmetric around the axis passing through μ_1 . It is bimodal with one mode at μ_1 and the other one at $\mu_1 + \pi$, whenever $\kappa_1 < 4\kappa_2$. If $\kappa_1 \geq 4\kappa_2$, then it is unimodal with mode at μ_1 . We refer to Table 1 of Gatto and Jammalamadaka (2007). The second test is on the precise characterization of axial symmetry, i.e. on $\delta = 0$ or $\delta = \pi/2$. So far $\kappa_2 > 0$ is considered and the third test is for $\kappa_2 = 0$, so that the distribution is no longer GvM but vM, which is axially symmetric. The Bayesian tests rely on the method of probability perturbation, where the probability distribution of the null hypothesis is slightly perturbed, in order to give a positive prior probability to the null hypothesis, which would be null otherwise. It would be interesting to consider the above null hypotheses under the frequentist perspective, perhaps with the likelihood ratio approach. This topic is not studied in this article, in order to limit its length.

The remaining part of this article is organized as follows. Section 2 gives the derivation of these Bayesian tests and their Bayes factors: Section 2.1 presents the approach used for these tests, Sect. 2.2 considers the test of no shift between cosines, Sect. 2.3 considers the test of symmetry and Sect. 2.4 considers the test of vM axial symmetry. Numerical results are presented in Sect. 3: Section 3.1 presents a Monte Carlo study of the tests of Sect. 2.1 whereas Sect. 3.2 presents the application to some real data. Final remarks are given in Sect. 4.

2 Bayesian tests and perturbation method for the GvM model

The proposed tests rely on Bayes factors. The Bayes factor B_{01} indicates the evidence of the null hypothesis with respect to (w.r.t.) the general alternative. Let us denote by $\theta = (\theta_1, \dots, \theta_n)$ the sample. Then

$$B_{01} = \frac{P[\theta|H_0]}{P[\theta|H_1]} = \frac{P[H_0|\theta]}{P[H_1|\theta]} \cdot \frac{P[H_1]}{P[H_0]} = \frac{R_1}{R_0}, \quad (4)$$

where

$$R_0 = \frac{P[H_0]}{P[H_1]} = \frac{P[H_0]}{1 - P[H_0]} \quad \text{and} \quad R_1 = \frac{P[H_0|\theta]}{P[H_1|\theta]} = \frac{P[H_0|\theta]}{1 - P[H_0|\theta]},$$

are the prior and the posterior odds, respectively. The case $B_{01} > 1$ indicates evidence for H_0 . Interpretations of the values of the Bayes factor can be found in Jeffreys (1961) and Kass and Raftery (1995). Our synthesis of these interpretations is given in Table 1, which provides a qualitative scale for the Bayes factor.

The null hypotheses of this article are simple, in the sense that they concern only points of the parametric space. The fact that these points have probability null does not allow for the computation of Bayes factors. Therefore we use an approach with probability perturbation explained in the next section.

Table 1 Guidelines for the interpretation of Bayes factors

B_{01}	Evidence for H_0
< 1	Negative
1 to 1.5	Not worth more than a bare mention
1.5 to 5	Positive
5 to 10	Substantial
10 to 20	Strong
> 20	Decisive

2.1 Bayesian tests of simple hypotheses

The practical relevance of a simple null hypothesis, i.e. of the type $H_0 : \xi = \xi_0$, has been widely debated in the statistical literature.

According to Berger and Delampady: “it is rare, and perhaps impossible, to have a null hypothesis that can be exactly modelled as $\theta = \theta_0$ ”. They illustrate their claim by the following example. “More common precise hypotheses such as H_0 : *Vitamin C has no effect on the common cold* are clearly not meant to be thought of as exact point nulls; surely vitamin C has some effects, although perhaps a very miniscule effect.” A similar example involving forensic science can be found in Lindley (1977). When the parameter ξ is of continuous nature, it is usually more realistic to consider null hypotheses of the type $H_{0,\varepsilon} : |\xi - \xi_0| \leq \varepsilon/2$, for some small $\varepsilon > 0$. This solves also the problem of the vanishing prior probability of H_0 , namely $P[\xi = \xi_0] = 0$. This problem is sometimes addressed by giving a positive probability to $\{\xi = \xi_0\}$. However, Berger and Sellke (1987) explain that the two approaches should be related. “It is convenient to specify a prior distribution for the testing problem as follows: let $0 < \pi_0 < 1$ denote the prior probability of $H_0 : \theta = \theta_0 \dots$ One might question the assignment of a positive probability to H_0 , because it is rarely the case that it is thought possible for $\theta = \theta_0$ to hold exactly ... H_0 is to be understood as simply an approximation to the realistic hypothesis $H_0 : |\theta - \theta_0| \leq b$ and π_0 is to be interpreted as the prior probability that would be assigned to $\{\theta : |\theta - \theta_0| \leq b\}$.” Accordingly, we assign to the original simple hypothesis $H_0 : \xi = \xi_0$ the prior probability $p_0 > 0$ of $H_{0,\varepsilon} : \xi \in [\xi_0 - \varepsilon/2, \xi_0 + \varepsilon/2]$, for some $\varepsilon > 0$. Thus, we replace the prior probability measure P by its perturbation, obtained by the assignment of the probability $p_0 > 0$ to $\{\xi_0\}$. We denote by P_0 the probability measure P with the p_0 -perturbation. To summarize: the point null hypotheses is made relevant with $p_0 = P_0[\xi = \xi_0] = P[\delta \in [\xi_0 - \varepsilon/2, \xi_0 + \varepsilon/2]] > 0$.

The length ε of the neighbourhood of ξ_0 , which determines the prior probability p_0 of H_0 under the perturbed model, should not be too small. A significant value of p_0 for the null hypothesis is in fact coherent with the frequentist approach of hypotheses tests, where computations of rejection regions or P -values are carried over under the null hypothesis. Berger (1985, p. 149) states that ε has to be chosen such that any ξ in $(\xi_0 - \varepsilon/2, \xi_0 + \varepsilon/2)$ becomes “indistinguishable” from ξ_0 , while Berger and Sellke state that ε has to be “small enough” so that $H_{0,\varepsilon}$ can be “accurately approximated” by H_0 . A related reference is Berger and Delampady (1987), who studied this problem

with a Gaussian model, and Berger (1985, p. 149), who obtains an upper bound for the radius $\varepsilon/2$ under a simple Gaussian model. Two other references on the practical relevance of simple null hypotheses are Jeffreys (1961) and Zellner (1984).

We end this section with some comments regarding the choice of the prior distribution of ξ . This is a generally unsolved problem of Bayesian statistics and widely discussed in the literature, see e.g. Jeffreys (1961) and Kass and Wasserman (1996). According to Berger and Delampady (1987), there is “no choice of the prior that can claim to be objective”. In this article we follow Berger and Delampady (1987) and Berger and Sellke (1987), where various details on the choice of the prior are presented and some classes of priors are analysed. In absence of prior information, the prior should be symmetric about ξ_0 and non-increasing with respect to $|\xi - \xi_0|$. Our choices of priors are presented in Sect. 3: for each test of the study we compute Bayes factors under priors obtained by varying the concentration around the generic value ξ_0 .

2.2 Test of no shift between cosines of GvM

Consider the Bayesian test on the GvM model (1) of the null the hypothesis

$$H_0 : \delta = 0,$$

where $\delta = (\mu_1 - \mu_2) \bmod \pi$ and where the values of $\mu_1, \kappa_1, \kappa_2$ are assumed known and equal to $\mu_1^0, \kappa_1^0, \kappa_2^0$, respectively. Under the original probability measure P , the random parameter δ has an absolutely continuous prior distribution and so $P[\delta = 0] = 0$. According to Sect. 2.1 we define the perturbation of the probability measure P , denoted P_0 , for which $p_0 = P_0[\delta = 0] > 0$. This perturbation is the assignment to $\{\delta = 0\}$ of the probability mass that initially lies close to that P -null set. Let $\varepsilon > 0$ and consider the set

$$\mathcal{A}_\varepsilon = \left\{ \delta \in [0, \pi) \mid \delta \in \left[0, \frac{\varepsilon}{2}\right] \cup \left[\pi - \frac{\varepsilon}{2}, \pi\right) \right\}. \quad (5)$$

The complement is

$$\mathcal{A}_\varepsilon^c = \left\{ \delta \in [0, \pi) \mid \delta \in \left(\frac{\varepsilon}{2}, \pi - \frac{\varepsilon}{2}\right) \right\}.$$

Note that (5) refers to a neighbourhood of the origin of the circle of circumference π . We thus assign to p_0 the value

$$p_0 = P[\mathcal{A}_\varepsilon], \quad (6)$$

for some suitably small $\varepsilon > 0$. The prior distribution function (d.f.) under the perturbed probability measure P_0 at any $\delta' \in [0, \pi)$ is given by

$$p_0 \Delta(\delta') + (1 - p_0) G(\delta'). \quad (7)$$

where G denotes the prior d.f. of δ and where Δ is the Dirac d.f., which assigns mass one to the origin. Denote by g the density of G . If $0 \notin (\delta', \delta' + d\delta')$, for some $\delta' \in (0, \pi)$, where the relations $\in$ and $\notin$ are meant circularly over the circle of circumference π , then (7) implies

$$P_0[\delta \in (\delta', \delta' + d\delta')] = (1 - p_0)g(\delta')d\delta' = (1 - p_0)P[\delta \in (\delta', \delta' + d\delta')]. \quad (8)$$

Let $\theta_1, \dots, \theta_n$ be independent circular random variables that follow the GvM distribution (1). For simplicity, we denote the joint density of $\theta = (\theta_1, \dots, \theta_n)$, with the fixed values $\delta', \mu_1^0, \kappa_1^0$ and κ_2^0 , as

$$f(\theta|\delta') = \left\{ 2\pi G_0(\delta', \kappa_1^0, \kappa_2^0) \right\}^{-n} \exp \left\{ \kappa_1^0 \sum_{i=1}^n \cos(\theta_i - \mu_1^0) + \kappa_2^0 \sum_{i=1}^n \cos 2(\theta_i - \mu_1^0 + \delta') \right\}. \quad (9)$$

When considered as a function of δ' , (9) becomes the likelihood of δ . Then, by (8) the marginal density of $\theta = (\theta_1, \dots, \theta_n)$ under the perturbed probability is given by

$$\begin{aligned} m(\theta) &= \int_{[0, \pi)} f(\theta|\delta') P_0[\delta \in (\delta', \delta' + d\delta')] \\ &= \int_{\mathcal{A}_\varepsilon} f(\theta|\delta') P_0[\delta \in (\delta', \delta' + d\delta')] + (1 - p_0) \int_{\mathcal{A}_\varepsilon^c} f(\theta|\delta') g(\delta') d\delta' \\ &= p_0 f(\theta|0) + (1 - p_0) \int_{\mathcal{A}_\varepsilon} f(\theta|\delta') g(\delta') d\delta' + (1 - p_0) \int_{\mathcal{A}_\varepsilon^c} f(\theta|\delta') g(\delta') d\delta' \\ &\sim 2p_0 f(\theta|0) + (1 - p_0) \int_{\mathcal{A}_\varepsilon} f(\theta|\delta') g(\delta') d\delta', \quad \text{as } \varepsilon \rightarrow 0. \end{aligned} \quad (10)$$

The above asymptotic equivalence is due to

$$\begin{aligned} &(1 - p_0) \int_{\mathcal{A}_\varepsilon} f(\theta|\delta') g(\delta') d\delta' \\ &= (1 - p_0) p_0 \int_{\mathcal{A}_\varepsilon} f(\theta|\delta') \frac{g(\delta')}{p_0} d\delta' \sim p_0 f(\theta|0), \quad \text{as } \varepsilon \rightarrow 0. \end{aligned}$$

The posterior perturbed probability, namely the conditional perturbed probability of $\{\delta = 0\}$ given θ , can be approximated as follows,

$$\begin{aligned} P_0[\delta = 0|\theta] &\sim \frac{p_0 f(\theta|0)}{2p_0 f(\theta|0) + (1 - p_0) \int_{\mathcal{A}_\varepsilon} f(\theta|\delta') g(\delta') d\delta'} \\ &= \frac{1}{2} \left(1 + \frac{1 - p_0}{p_0} \frac{\int_{\mathcal{A}_\varepsilon} f(\theta|\delta') g(\delta') d\delta'}{2f(\theta|0)} \right)^{-1}, \quad \text{as } \varepsilon \rightarrow 0. \end{aligned} \quad (11)$$

In order to compute the Bayes factor for this test, we define the prior odds $R_0 = p_0/(1 - p_0)$ and the posterior odds $R_1 = P_0[\delta = 0|\theta]/(1 - P_0[\delta = 0|\theta])$. The Bayes factor is the posterior over the prior odds, namely $B_{01} = R_1/R_0$. Clearly $p_0 \leq P_0[\delta = 0|\theta]$ iff $B_{01} \geq 1$ and, the larger $P_0[\delta = 0|\theta] - p_0$ becomes, the larger B_{01} becomes: a large Bayes factor tells that the data support the null hypothesis. From the approximation

$$R_1 \sim \left[1 + \frac{1 - p_0}{p_0} \frac{\int_{\mathcal{A}_\varepsilon} f(\theta|\delta')g(\delta')d\delta'}{f(\theta|0)} \right]^{-1}$$

and from some simple algebraic manipulation, we obtain the computable approximation to the Bayes factor $B_{01} = R_1/R_0$ given by

$$B_{01} \sim \frac{f(\theta|0)}{\int_{\mathcal{A}_\varepsilon} f(\theta|\delta')g(\delta')d\delta'}, \quad \text{as } \varepsilon \rightarrow 0. \quad (12)$$

The representation of the Bayes factor (12) is asymptotically correct and we remind that, in the context where we approximate the null hypothesis with a neighbourhood by the point null hypothesis, the reasoning is always of asymptotic nature. A reference for this perturbation technique is Berger (1985, p. 148–150).

Regarding the large sample asymptotics of the proposed test, it is known that, for a sample of n independent random variables with common distribution with true parameter ξ_0 , the posterior distribution converges to the distribution with total mass over ξ_0 , as $n \rightarrow \infty$. This means that the posterior mode is a consistent estimator. We deduce that, under H_0 ,

$$P_0[\delta = 0|\theta] = P[\mathcal{A}_\varepsilon|\theta] \xrightarrow{P} 1, \quad \text{as } n \rightarrow \infty.$$

Consequently, $R_1 = P_0[\delta = 0|\theta]/(1 - P_0[\delta = 0|\theta]) \xrightarrow{P} \infty$ and $B_{01} = R_1/R_0 \xrightarrow{P} \infty$, as $n \rightarrow \infty$. The Bayesian test of $H_0 : \delta = 0$ is consistent in this sense.

We now give some computational remarks that are also valid for the tests of Sects. 2.3 and 2.4. The integral appearing in the denominator of (12) can be easily evaluated by Monte Carlo integration. For a given large integer s , we generate $\delta^{(i)}$, for $i = 1, \dots, s$, from the density g and then we compute the approximation

$$\begin{aligned} & \int_{\mathcal{A}_\varepsilon^c} f(\theta|\delta')g(\delta')d\delta' \\ &= \int_{\mathcal{A}_\varepsilon^c} f(\theta|\delta')P[\delta \in (\delta', \delta' + d\delta')] \simeq \frac{1}{s} \sum_{i=1}^s f(\theta|\delta^{(i)})I\{\delta^{(i)} \in \mathcal{A}_\varepsilon^c\}, \end{aligned} \quad (13)$$

where $I\{A\}$ denotes the indicator of statement or event A . For the computation normalizing constant of the GvM distribution given in (2) one can use the Fourier series

$$G_0(\delta, \kappa_1, \kappa_2) = I_0(\kappa_1)I_0(\kappa_2) + 2 \sum_{j=1}^{\infty} I_{2j}(\kappa_1)I_j(\kappa_2) \cos 2j\delta, \quad (14)$$

where $\delta \in [0, \pi)$ and $\kappa_1, \kappa_2 > 0$, see e.g. Gatto and Jammalamadaka (2007).

2.3 Test of axial symmetry of GvM

In this section we consider the Bayesian test of axial symmetry for the GvM model (1). A circular density g is symmetric around the angle $\alpha/2$, for some $\alpha \in [0, 2\pi)$, if $g(\theta) = g(\alpha - \theta)$, $\forall \theta \in [0, 2\pi)$. In this case we have also $g(\theta) = g((\alpha + 2\pi) - \theta)$, so that symmetry around $\alpha/2 + \pi$ holds as well: the symmetry is indeed an axial one.

Proposition 2.1 (Characterization of axial symmetry for the GvM distribution) *The GvM distribution (1) is axially symmetric iff*

$$\delta = 0 \text{ or } \delta = \frac{\pi}{2}.$$

In both cases, the axis of symmetry has angle μ_1 .

The proof of Proposition 2.1 is given in Appendix A.

Note that δ is defined modulo π and that for $\kappa_2 = 0$ or $\kappa_1 = 0$ the GvM reduces respectively to the vM or to the axial vM, defined later as vM₂ and given in (18). These two distributions are clearly symmetric, but Proposition 2.1 gives the characterization of symmetry in terms of δ only since we define the GvM distribution in (1) with concentration parameters $\kappa_1, \kappa_2 > 0$.

As mentioned at the beginning of the section, symmetry of a circular distribution around an angle is the symmetry around an axis. For the GvM density, this is made explicit in (21), where adding 2π to α would not have any influence. Figure 1 provides two numerical illustrations of the axial symmetry of the GvM distribution. The graph in Fig. 1a shows the density of the GvM($\pi, \pi, 0.1, 5.5$) distribution: $\delta = 0$ and the axis of symmetry is at angle $\mu_1 = \pi$. The graph in Fig. 1b shows the density of the GvM($\pi/2, 0, 5.5, 0.1$) distribution: $\delta = \pi/2$ and the axis of symmetry is at angle $\mu_1 = \pi/2$.

Thus, Proposition 2.1 allows us to write the null hypothesis of axial symmetry as

$$H_0 : \delta = 0 \text{ or } \delta = \frac{\pi}{2},$$

(a) GvM($\pi, 0, 0.1, 5.5$) density ($\delta = 0$).

(b) GvM($\pi/2, 0, 5.5, 0.1$) density ($\delta = \pi/2$).

Fig. 1 Two axially symmetric GvM densities over the interval $(-2\pi, 2\pi)$ and with their axis of symmetry at angle μ_1 shown by vertical dashed lines

where the values of $\mu_1, \kappa_1, \kappa_2$ are assumed known and equal to $\mu_1^0, \kappa_1^0, \kappa_2^0$, respectively. The Bayesian test is obtained by perturbation of the probability measure P , which is denoted P_0 . The probabilities

$$p_0 = P_0[\delta = 0] > 0 \quad \text{and} \quad p_{\frac{\pi}{2}} = P_0\left[\delta = \frac{\pi}{2}\right] > 0$$

are the probabilities masses of $\{\delta = 0\}$ and $\{\delta = \frac{\pi}{2}\}$ of the perturbed measure, respectively. They are obtained from

$$p_0 = P\left[\delta \in \left[0, \frac{\varepsilon}{2}\right] \cup \left[\pi - \frac{\varepsilon}{2}, \pi\right)\right] \text{ and } p_{\frac{\pi}{2}} = P\left[\delta \in \left[\frac{\pi}{2} - \frac{\varepsilon}{2}, \frac{\pi}{2} + \frac{\varepsilon}{2}\right]\right],$$

for suitably small $\varepsilon > 0$. As is Sect. 2.2, the prior d.f. of δ under the perturbed probability P_0 at any $\delta' \in [0, \pi)$ is given by

$$p_0 \Delta(\delta') + p_{\frac{\pi}{2}} \Delta\left(\delta' - \frac{\pi}{2}\right) + \left\{1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right\} G(\delta'), \quad (15)$$

where G is the prior d.f. of δ under P . It follows from (15) that for $0, \pi/2 \notin (\delta', \delta' + d\delta')$, for some $\delta' \in (0, \pi) \setminus \{\pi/2\}$,

$$\begin{aligned} P_0[\delta \in (\delta', \delta' + d\delta')] &= \left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] g(\delta') d\delta' \\ &= \left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] P[\delta \in (\delta', \delta' + d\delta')], \end{aligned}$$

where g is the density of G .

Let

$$\begin{aligned} \mathcal{B}_{\varepsilon,0} &= \left\{\delta \in [0, \pi) \mid \delta \in \left[0, \frac{\varepsilon}{2}\right] \cup \left[\pi - \frac{\varepsilon}{2}, \pi\right)\right\} \text{ and} \\ \mathcal{B}_{\varepsilon,\frac{\pi}{2}} &= \left\{\delta \in [0, \pi) \mid \delta \in \left[\frac{\pi}{2} - \frac{\varepsilon}{2}, \frac{\pi}{2} + \frac{\varepsilon}{2}\right]\right\}. \end{aligned}$$

Define

$$\mathcal{B}_{\varepsilon} = \left\{\delta \in [0, \pi) \mid \delta \in \left[0, \frac{\varepsilon}{2}\right] \cup \left[\pi - \frac{\varepsilon}{2}, \pi\right) \vee \delta \in \left[\frac{\pi}{2} - \frac{\varepsilon}{2}, \frac{\pi}{2} + \frac{\varepsilon}{2}\right]\right\} = \mathcal{B}_{\varepsilon,0} \cup \mathcal{B}_{\varepsilon,\frac{\pi}{2}}.$$

Its complement is given by

$$\mathcal{B}_{\varepsilon}^c = \left\{\delta \in [0, \pi) \mid \delta \in \left(\frac{\varepsilon}{2}, \frac{\pi}{2} - \frac{\varepsilon}{2}\right) \cup \left(\frac{\pi}{2} + \frac{\varepsilon}{2}, \pi - \frac{\varepsilon}{2}\right)\right\}.$$

The marginal density of $\theta = (\theta_1, \dots, \theta_n)$ with respect to the perturbed probability P_0 is given by

$$\begin{aligned} m(\theta) &= \int_{[0,\pi)} f(\theta|\delta') P_0[\delta \in (\delta', \delta' + d\delta')] \\ &= \int_{\mathcal{B}_{\varepsilon}} f(\theta|\delta') P_0[\delta \in (\delta', \delta' + d\delta')] + \left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] \int_{\mathcal{B}_{\varepsilon}^c} f(\theta|\delta') g(\delta') d\delta' \\ &= p_0 f(\theta|0) + p_{\frac{\pi}{2}} f\left(\theta \mid \frac{\pi}{2}\right) + \left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] \int_{\mathcal{B}_{\varepsilon}} f(\theta|\delta') g(\delta') d\delta' \\ &\quad + \left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] \int_{\mathcal{B}_{\varepsilon}^c} f(\theta|\delta') g(\delta') d\delta' \end{aligned}$$

$$\begin{aligned} &\sim 2p_0 f(\boldsymbol{\theta}|0) + 2p_{\frac{\pi}{2}} f\left(\boldsymbol{\theta}\left|\frac{\pi}{2}\right.\right) \\ &+ \left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] \int_{\mathcal{B}_\varepsilon^c} f(\boldsymbol{\theta}|\delta') g(\delta') d\delta', \text{ as } \varepsilon \rightarrow 0. \end{aligned}$$

In the asymptotic equivalence, as in Sect. 2.2, we note that

$$\begin{aligned} &\left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] \int_{\mathcal{B}_\varepsilon} f(\boldsymbol{\theta}|\delta') g(\delta') d\delta' \\ &= \left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] \left[\int_{\mathcal{B}_{\varepsilon,0}} f(\boldsymbol{\theta}|\delta') g(\delta') d\delta' + \int_{\mathcal{B}_{\varepsilon,\frac{\pi}{2}}} f(\boldsymbol{\theta}|\delta') g(\delta') d\delta' \right] \\ &= \left[1 - \left(p_0 + p_{\frac{\pi}{2}}\right)\right] \left[p_0 \int_{\mathcal{B}_{\varepsilon,0}} f(\boldsymbol{\theta}|\delta') \frac{g(\delta')}{p_0} d\delta' + p_{\frac{\pi}{2}} \int_{\mathcal{B}_{\varepsilon,\frac{\pi}{2}}} f(\boldsymbol{\theta}|\delta') \frac{g(\delta')}{p_{\frac{\pi}{2}}} d\delta' \right] \\ &\sim p_0 f(\boldsymbol{\theta}|0) + p_{\frac{\pi}{2}} f\left(\boldsymbol{\theta}\left|\frac{\pi}{2}\right.\right), \text{ as } \varepsilon \rightarrow 0. \end{aligned}$$

The posterior probability of $\{\delta = 0 \vee \delta = \pi/2\}$ under the perturbed probability measure is given by

$$\begin{aligned} P_0\left[\delta = 0 \vee \delta = \frac{\pi}{2} \middle| \boldsymbol{\theta}\right] &\sim \frac{p_0 f(\boldsymbol{\theta}|0) + p_{\frac{\pi}{2}} f(\boldsymbol{\theta}|\frac{\pi}{2})}{2p_0 f(\boldsymbol{\theta}|0) + 2p_{\frac{\pi}{2}} f(\boldsymbol{\theta}|\frac{\pi}{2}) + [1 - (p_0 + p_{\frac{\pi}{2}})]\mathcal{I}_1(\vec{\boldsymbol{\theta}})} \\ &= \frac{1}{2} \left[1 + \frac{[1 - (p_0 + p_{\frac{\pi}{2}})]\mathcal{I}_1(\boldsymbol{\theta})}{2p_0 f(\boldsymbol{\theta}|0) + 2p_{\frac{\pi}{2}} f(\boldsymbol{\theta}|\frac{\pi}{2})} \right]^{-1}, \text{ as } \varepsilon \rightarrow 0, \end{aligned}$$

where

$$\mathcal{I}_1(\boldsymbol{\theta}) = \int_{\mathcal{B}_\varepsilon^c} f(\boldsymbol{\theta}|\delta') g(\delta') d\delta'.$$

With this we obtain the following approximation to the posterior odds,

$$\begin{aligned} R_1 &= \frac{P_0[\delta = 0 \vee \delta = \frac{\pi}{2} | \boldsymbol{\theta}]}{1 - P_0[\delta = 0 \vee \delta = \frac{\pi}{2} | \boldsymbol{\theta}]} = \left[\frac{1}{P_0[\delta = 0 \vee \delta = \frac{\pi}{2} | \boldsymbol{\theta}]} - 1 \right]^{-1} \\ &\sim \left[1 + \frac{[1 - (p_0 + p_{\frac{\pi}{2}})]\mathcal{I}_1(\vec{\boldsymbol{\theta}})}{p_0 f(\boldsymbol{\theta}|0) + p_{\frac{\pi}{2}} f(\boldsymbol{\theta}|\frac{\pi}{2})} \right]^{-1}, \end{aligned}$$

as $\varepsilon \rightarrow 0$. With the prior odds given by

$$R_0 = \frac{p_0 + p_{\frac{\pi}{2}}}{1 - (p_0 + p_{\frac{\pi}{2}})}$$

and after algebraic manipulations, we obtain the approximation to the Bayes factor given by

$$B_{01} \sim \frac{p_0 f(\boldsymbol{\theta}|0) + p_{\frac{\pi}{2}} f(\boldsymbol{\theta}|\frac{\pi}{2})}{(p_0 + p_{\frac{\pi}{2}}) \mathcal{I}_1(\boldsymbol{\theta})}, \quad \text{as } \varepsilon \rightarrow 0.$$

2.4 Test of vM axial symmetry

We consider the Bayesian test of the null hypothesis that the sample follows a vM distribution against the alternative that it comes from an arbitrary GvM distribution. This null hypothesis implies axial symmetry in the class of vM distributions, whereas the alternative hypothesis includes both symmetric or asymmetric GvM distributions. Precisely, we have $H_0 : \kappa_2 = 0$, where μ_1, μ_2 and κ_1 are assumed known and equal to μ_1^0, μ_2^0 and κ_1^0 respectively. The GvM with $\kappa_2 = 0$ reduces to the trivially symmetric vM distribution. Formally, the GvM is defined for $\kappa_2 > 0$ only, so that the symmetry considered here is no longer within the GvM class but it is rather a vM axial symmetry. This symmetry within the GvM class should be thought as approximate, for vanishing values of κ_2 .

Symmetry with the GvM formula can also be obtained with $\kappa_1 = 0$, in which case the GvM formula reduces to an axial von Mises (vM₂) distribution given in (18), which is trivially symmetric. This case is not analysed. In what follows we focus on the case of vM axial symmetry.

Because $P[\kappa_2 = 0] = 0$, we construct the perturbed probability P_0 such that $p_0 = P_0[\kappa_2 = 0] > 0$, where $p_0 = P[\kappa_2 \in [0, \varepsilon]]$, for some $\varepsilon > 0$ small. The prior d.f. of κ_2 under the probability P is G , and under the perturbed probability P_0 it is $p_0 \Delta(\kappa'_2) + (1 - p_0)G(\kappa'_2)$, $\forall \kappa'_2 \geq 0$.

Assume $0 \notin (\kappa'_2, \kappa'_2 + d\kappa'_2)$, then

$$P_0[\kappa_2 \in (\kappa'_2, \kappa'_2 + d\kappa'_2)] = (1 - p_0)g(\kappa'_2)d\kappa'_2 = (1 - p_0)P[\kappa_2 \in (\kappa'_2, \kappa'_2 + d\kappa'_2)],$$

where g is the density of G . With algebraic manipulations similar to those of Sect. 2.2, one obtains the approximation to the Bayes factor B_{01} of posterior over prior odds given by

$$B_{01} \sim \frac{f(\boldsymbol{\theta}|0)}{\int_{\mathcal{C}_\varepsilon^c} f(\boldsymbol{\theta}|\kappa'_2)g(\kappa'_2)d\kappa'_2}, \quad \text{as } \varepsilon \rightarrow 0, \quad (16)$$

where $\mathcal{C}_\varepsilon = [0, \varepsilon]$, $\mathcal{C}_\varepsilon^c$ is its complement and where the likelihood of κ_2 is

$$f(\boldsymbol{\theta}|\kappa_2) = \left\{ 2\pi G_0 \left(\delta^0, \kappa_1^0, \kappa_2 \right) \right\}^{-n} \exp \left\{ \kappa_1^0 \sum_{i=1}^n \cos(\theta_i - \mu_1^0) + \kappa_2 \sum_{i=1}^n \cos 2(\theta_i - \mu_1^0 + \delta^0) \right\}, \quad (17)$$

with $\delta^0 = (\mu_1^0 - \mu_2^0) \bmod \pi$.

3 Numerical studies

This section provides some numerical studies for the tests introduced in Sect. 2. The major part is Sect. 3.1, which gives a simulation or Monte Carlo study of the performance of these tests. Section 3.2 provides an application to real measurements of wind directions.

3.1 Monte Carlo study

This section presents a Monte Carlo study for the tests introduced in Sect. 2: in Sect. 3.1.1 for the test of no shift between cosines, in Sect. 3.1.2 for the test axial symmetry and in Sect. 3.1.3 for the test of vM axial symmetry. The results are summarized in Sect. 3.1.4. We obtain Bayes factors for each one of these three tests for $r = 10^4$ generations of samples of size $n = 50$, that are generated from the GvM or the vM distributions. The Monte Carlo approximation to the integral (13) and to the analogue integrals of the two other tests, is computed with $s = 10^4$ generations. This simulation scheme is repeated three times and the results are compared in order to verify convergence. Confidence intervals for the Bayes factors based on the aggregation of the three simulations (with r replications each) are provided.

The axial vM distribution is used as a prior distribution for the parameter of shift between cosines δ . This distribution can be obtained by taking $\kappa_1 = 0$ in the exponent of (1) and by multiplying the density by 2, yielding

$$f(\theta \mid \mu, \kappa) = \{\pi I_0(\kappa)\}^{-1} \exp\{\kappa \cos 2(\theta - \mu)\}, \quad \forall \theta \in [0, \pi), \quad (18)$$

and for some $\mu \in [0, \pi)$ and $\kappa > 0$. We denote this distribution by $\text{vM}_2(\mu, \kappa)$.

According to the remark at the end of Sect. 2.1, we choose $\varepsilon = 0.05$ for the length of the interval of H_0 and the prior densities g as follows. For the test of no shift between cosines, we choose the $\text{vM}_2(0, \tau)$ distribution for δ , which is symmetric and unimodal with mode at $\delta = 0$. For the test of axial symmetry, we choose the mixture of $\text{vM}_2(0, \tau)$ and $\text{vM}_2(\pi/2, \tau)$ for δ . Finally, for the test of vM axial symmetry, we choose an uniform distribution for κ_2 that is highly concentrated at the boundary point 0.

3.1.1 Test of no shift between cosines of GvM

The null hypothesis considered is $H_0: \delta = 0$, with fixed $\mu_1 = \mu_1^0, \kappa_1 = \kappa_1^0, \kappa_2 = \kappa_2^0$, where $\mu_1^0 = \pi, \kappa_1^0 = 0.1, \kappa_2^0 = 5.5$. We consider three different cases, called D1, D1' and D2.

Case D1 For $i = 1, \dots, s$, we generate $\delta^{(i)}$ from the prior of δ , which is $\text{vM}_2(\nu, \tau)$ with values of the hyperparameters $\nu = 0$ and $\tau = 250$. We obtain $p_0 = 0.570$ as prior probability of the null hypothesis under the perturbed probability measure. We take the first $r = s$ of these prior values and then we obtain $\mu_2^{(i)} = (\mu_1^0 - \delta^{(i)}) \bmod \pi$ and generate the elements of the vector of n sample values $\theta^{(i)}$ independently from $\text{GvM}(\mu_1^0, \mu_2^{(i)}, \kappa_1^0, \kappa_2^0)$, for $i = 1, \dots, r$. With these simulated sample we compute

the Bayes factor $B_{01}^{(i)}$ with the approximation formula (12). We repeat this experiment three times. The fact of generating values of δ from its prior distribution, instead of taking $\delta = 0$ fixed by null hypothesis, is a way of inserting some prior uncertainty in the generated sample. If the prior is close, in some sense, to the null hypothesis, then we should obtain the Bayes factor larger than one, but smaller than the Bayes factor that would be obtained with the fixed value $\delta = 0$.

We obtained three sequences of 10^4 Bayes factors that can be summarized as follows. Figure 2a displays the three boxplots of the three simulated sequences of Bayes factors: Denote by $\bar{B}_{01}^{(j)}$ the mean of the Bayes factors of the j -th sequence, for $j = 1, 2, 3$, corresponding to left, central and right boxplot respectively. We obtained:

$$\bar{B}_{01}^{(1)} = 2.887, \bar{B}_{01}^{(2)} = 3.028 \quad \text{and} \quad \bar{B}_{01}^{(3)} = 2.955.$$

Figure 2b shows the histogram of the three generated sequences of r Bayes factors. The distribution is clearly not “bell-shaped” but it is however light-tailed: the Central limit theorem applies to the mean of the simulated Bayes factors. The asymptotic normal confidence interval for the mean value of the Bayes factors at level 0.95 and based on the three generated sequences, is given by

$$(2.937, 2.976).$$

According to Table 1 this interval indicates positive evidence for the null hypothesis: the sample has indeed increased the evidence of the null hypothesis $\delta = 0$, however to a marginal extent only. This situation can be explained by the fact that the prior density g is (highly) concentrated around 0, circularly. This can be seen in the graph of the prior density (Fig. 2c), where the histogram of 10^4 generated values of δ is shown together with the prior density. Moreover, the variability originating from the fact the sample is simulated under different values of δ leads to weaker values of the Bayes factor.

Case D1' In this other case we consider prior values of δ less concentrated around 0, by choosing $\nu = 0$ and $\tau = 50$. The resulting prior probability of H_0 is given by $p_0 = 0.276$. For $i = 1, \dots, r$, we generate the elements of the vector of n sample values $\theta^{(i)}$ independently from $\text{GvM}(\mu_1^0, \mu_2^0, \kappa_1^0, \kappa_2^0)$, with $\delta = 0$, thus with $\mu_2^0 = (\mu_1^0 - \delta) \bmod \pi = 0$. With these simulated data, we compute the Bayes factor $B_{01}^{(i)}$ with the approximation formula (12).

We obtained three sequences of $r = 10^4$ Bayes factors with means:

$$\bar{B}_{01}^{(1)} = 3.922, \bar{B}_{01}^{(2)} = 3.924 \quad \text{and} \quad \bar{B}_{01}^{(3)} = 3.921.$$

The boxplots of the three respective generated sequences are shown in Fig. 3a.

The asymptotic normal confidence interval for the mean value of the Bayes factors, at level 0.95 and based on the three generated sequences, is

$$(3.901, 3.945).$$

Fig. 2 Results of Case D1

As expected, the generated Bayes factors are larger than in case D1. The samples generated with $\delta = 0$ fixed have less uncertainty. We computed the posterior density of δ based on one generated sample. In Fig. 3b we can see the graph of that posterior density, with continuous line, together with the graph of the prior density, with dashed line. The posterior is indeed more concentrated around 0, circularly.

Case D2 We now further decrease the concentration of the prior of δ . The values of the hyperparameters are $\nu = 0$ and $\tau = 20$. We computed the prior probability of the null hypothesis under perturbation $p_0 = 0.176$. We generated the samples $\theta^{(i)}$, for $i = 1, \dots, r$, with fixed value $\mu_2^0 = 0$.

We obtained three sequences of $r = 10^4$ Bayes factors with means

$$\bar{B}_{01}^{(1)} = 5.477, \bar{B}_{01}^{(2)} = 5.539 \quad \text{and} \quad \bar{B}_{01}^{(3)} = 5.511.$$

The boxplots of the three respective generated sequences are shown in Fig. 4a.

The asymptotic normal confidence interval for the mean value of the Bayes factors, at level 0.95 and based on the three generated sequences, is

$$(5.477, 5.541).$$

(a) 3 Boxplots of the 3 sets of 10^4 simulated Bayes factors.

(b) Prior density (dashed line) and posterior density (continuous line) of δ . The posterior is based on one generated sample.

Fig. 3 Results of Case D1'

The Bayes factors are larger than they are in Cases D1 and D1'. Here they show substantial evidence for the null hypothesis. The prior distribution of δ is less favourable to the null hypothesis and so the sample brings more additional evidence for the null hypothesis. Figure 4b shows the graph of the prior density, as dashed line, together with the graph of a posterior density, as continuous line, for δ . The graph of the posterior density is based on one generated sample.

(a) 3 Boxplots of the 3 sets of 10^4 simulated Bayes factors.

(b) Prior density (dashed line) and posterior density (continuous line) of δ . The posterior is based on one generated sample.

Fig. 4 Results of case D2

3.1.2 Test of axial symmetry of GvM

In this section we consider the null hypothesis of axial symmetry, viz. $H_0: \delta = 0$ or $\delta = \pi/2$, other parameters being fixed as follows, $\mu_1 = \mu_1^0, \kappa_1 = \kappa_1^0$ and $\kappa_2 = \kappa_2^0$. We choose as before $\mu_1^0 = \pi, \kappa_1^0 = 0.1$ and $\kappa_2^0 = 5.5$. We generate δ from the prior given by the mixture of vM₂ distributions $\xi \text{vM}_2(\nu_1, \tau) + (1 - \xi) \text{vM}_2(\nu_2, \tau)$, with

$\nu_1 = 0, \nu_2 = \pi/2$ and $\xi = 0.5$. We consider three different cases, called Cases S1, S2 and S3.

Case S1 We generated δ from the prior mixture with concentration parameter $\tau = 250$. This prior distribution is close to the null distribution and Fig. 5b displays its density, together with the histogram of 10^4 generations from it. We computed the prior probabilities of the null hypothesis under the perturbed probability measure with $p_0 = p_{\pi/2} = 0.285$. We follow the principle of Case D1, where prior uncertainty is transmitted to the sample by considering generated values $\delta^{(i)}$, for $i = 1, \dots, s$, from a prior of δ close to the null hypothesis, instead of considering the fixed values of the null hypothesis, namely $\delta = 0$ or $\pi/2$. We take the first r of these prior values and we use $\mu_2^{(i)} = (\mu_1^0 - \delta^{(i)}) \bmod \pi$ for generating $\theta^{(i)}$, for $i = 1, \dots, r$. Repeating this three times, we obtained the three means of the three sequences of $r = 10^4$ Bayes factors

$$\bar{B}_{01}^{(1)} = 3.044, \bar{B}_{01}^{(2)} = 2.986 \quad \text{and} \quad \bar{B}_{01}^{(3)} = 2.950.$$

In Fig. 5a we can find the boxplots of the three respective generated sequences.

The asymptotic normal confidence interval for the mean value of the Bayes factors, at level 0.95 and based on the three generated sequences, is

$$(2.974, 3.013).$$

The conclusion is that the sample provides positive evidence of axial symmetry, even though to some smaller extent only. The same was found in Case D1.

Case S2 We generated prior values of δ from the same mixture, however with smaller concentration hyperparameter $\tau = 20$. We found $p_0 = p_{\pi/2} = 0.088$. We generated the elements of the sample vector $\theta^{(i)}$ with fixed value $\mu_2^0 = 0$, thus from $\text{GvM}(\mu_1^0, \mu_2^0, \kappa_1^0, \kappa_2^0)$, with $\mu_1^0 = \pi, \mu_2^0 = 0, \kappa_1^0 = 0.1, \kappa_2^0 = 5.5$, for $i = 1, \dots, r$. We repeated this experiment three times and obtained three sequences of Bayes factors, with respective mean values

$$\bar{B}_{01}^{(1)} = 5.322, \bar{B}_{01}^{(2)} = 5.439 \quad \text{and} \quad \bar{B}_{01}^{(3)} = 5.282.$$

The boxplots of the three sequences of Bayes factors can be found in Fig. 6a.

After aggregating the three sequences, we obtained the asymptotic normal confidence interval at level 0.95 for the mean value of the Bayes factors given by

$$(5.317, 5.378).$$

The Bayes factor is thus larger than it was in Case S1, so that the sample has brought substantial evidence of axial symmetry. Figure 6b shows the prior density of δ (with dashed line) and a posterior density of δ (with continuous line) that is based on one of

(a) 3 Boxplots of the 3 sets of 10^4 simulated Bayes factors.

(b) Prior density of δ with histogram of 10^4 generated values.

Fig. 5 Results of case S1

the previously generated samples. The posterior is highly concentrated around 0 and provides a stronger belief about symmetry than the prior.

Case S3 We retain the prior of δ of Case S2 but we generate samples $\theta^{(i)}$, for $i = 1, \dots, r$, with $\mu_1^0 = \pi$, $\mu_2^0 = \pi/2$, $\kappa_1^0 = 0.1$, and $\kappa_2^0 = 5.5$, thus from another symmetric GvM distribution. The computed values $p_0 = p_{\pi/2} = 0.088$ are the same of Case S2. We generated three sequences of $r = 10^4$ Bayes factors. The three respective boxplots of the three sequences can be found in Fig. 7a. The three respective means of these three sequences are

$$\bar{B}_{01}^{(1)} = 5.267, \bar{B}_{01}^{(2)} = 5.553 \quad \text{and} \quad \bar{B}_{01}^{(3)} = 5.395.$$

(a) 3 Boxplots of the 3 sets of 10^4 simulated Bayes factors.

(b) Prior density (dashed line) and posterior density (continuous line) of δ . The posterior is based on one generated sample.

Fig. 6 Results of case S2

By aggregating the three sequences, we obtained the asymptotic normal confidence interval at level 0.95 for the mean of the Bayes factors given by

$$(5.374, 5.436).$$

We find substantial evidence of axial symmetry. Figure 7b displays the prior density of δ (with dashed line) and a posterior density of δ (with continuous line) that is based on one of the previously generated samples. The posterior is highly concentrated around $\pi/2$ and possesses less uncertainty about symmetry than the prior.

(a) 3 Boxplots of the 3 sets of 10^4 simulated Bayes factors.

(b) Prior density (dashed line) and posterior density (continuous line) of δ . The posterior is based on one generated sample.

Fig. 7 Results of case S3

3.1.3 Test of vM axial symmetry of GvM

Now we have $H_0 : \kappa_2 = 0$, with fixed $\mu_1^0 = \pi$, $\mu_2^0 = \pi/2$ and $\kappa_1^0 = 0.1$. The prior distribution of κ_2 is uniform over $[0, 1/2]$ and the sample $\boldsymbol{\theta} = (\theta_1, \dots, \theta_n)$ is generated from the $\text{vM}(\mu_1^0, \kappa_1^0)$ distribution. The prior probability of H_0 under the perturbation is $p_0 = 0.1$. We generated three sequences of $r = 10^4$ Bayes factors: their boxplots are shown in Fig. 8. In these boxplots we removed a very small number of large values, in order to improve the readability. The three means of the three generated sequences

Fig. 8 3 boxplots of the 3 sets of 10^4 simulated Bayes factors

are

$$B_{01}^{(1)} = 3.284, \quad B_{01}^{(2)} = 3.380 \quad \text{and} \quad B_{01}^{(3)} = 3.241,$$

where the very large values that were eliminated from the boxplots were considered in the calculations of these means.

After aggregating these three sequences, we obtained the following asymptotic normal confidence interval for the mean value of the Bayes factors at level 0.95,

$$(3.268, 3.335).$$

There is a positive evidence of symmetry although rather limited. The amount of evidence is similar to the cases D1 and S1: in all these studies, the prior is much concentrated around the null hypothesis (here $\kappa_2 = 0$), so that the data have increased the evidence of the null hypothesis only to some limited extent.

3.1.4 Summary

Table 2 summarizes the simulation results that we obtained for the three tests and for the various cases.

3.2 Application to real data

The proposed Bayesian tests have been so far applied to simulated data. This section provides the application of the test of no shift between cosines of Sect. 2.2 and of axial symmetry of Sect. 2.3 to real data obtained from the study “ArticRIMS” (A Regional, Integrated Hydrological Monitoring System for the Pan Arctic Land Mass) available

Table 2 Summary of the simulation study

H_0	Case	Confidence interval for B_{01}	Evidence for H_0
No shift between cosines	D1	(2.937, 2.976)	Positive
	D1'	(3.901, 3.945)	Positive
	D2	(5.477, 5.541)	Substantial
Axial symmetry	S1	(2.974, 3.013)	Positive
	S2	(5.317, 5.378)	Substantial
	S3	(5.374, 5.436)	Substantial
vM axial symmetry	–	(3.268, 3.335)	Positive

at <http://rims.unh.edu>. The Arctic climate, its vulnerability, its relation with the terrestrial biosphere and with the recent global climate change are the subjects under investigation. For this purpose, various meteorological variables such as temperature, precipitation, humidity, radiation, vapour pressure, speed and directions of winds are measured at four different sites.

We consider wind directions measured at the site “Europe basin” and from January to December 2005. After removal of few influential measurements, the following maximum likelihood estimators are obtained: $\hat{\mu}_1 = 4.095$, $\hat{\mu}_2 = 0.869$, $\hat{\kappa}_1 = 0.304$, $\hat{\kappa}_2 = 1.910$ and thus $\hat{\delta} = (\hat{\mu}_1 - \hat{\mu}_2) \bmod \pi = 0.084$. The histogram of the sample together with the GvM density with theses values of the parameters are given in Fig. 9.

For the test of no shift between cosines, the Monte Carlo integral (13) is computed with $s = 10^6$ values of δ generated from the prior $\text{vM}_2(\nu, \tau)$, with $\nu = 0$ and $\tau = 300$. We consider $\varepsilon = 0.18$: as mentioned in Sect. 2.1, a substantial value is desirable in the practice. We obtain the Bayes factor $B_{01} = 2.550$; cf. Table 3.

For the test of symmetry, the prior of δ is the mixture of two vM of order two, i.e. $\xi \text{vM}_2(\nu_1, \tau) + (1 - \xi) \text{vM}_2(\nu_2, \tau)$, with $\nu_1 = 0$, $\nu_2 = \pi/2$, $\tau = 300$ and $\xi = 0.5$. Monte Carlo integration is done with $s = 10^6$ generations from this prior. We consider $\varepsilon = 0.18$ and obtain the Bayes factor $B_{01} = 2.252$; cf. Table 3.

The values of the two Bayes factors of Table 3 show positive evidence for the respective null hypotheses.

4 Conclusion

This article introduces three Bayesian tests relating to the symmetry of the GvM model. The first test is about the significance of the shift parameter between the cosines of frequency one and two ($H_0 : \delta = 0$). The second test is about axial symmetry ($H_0 : \delta = 0$ or $\delta = \pi/2$). The third test is about vM symmetry ($H_0 : \kappa_2 = 0$). These tests are obtained by the technique of probability perturbation. Simulation studies show the effectiveness of these three tests, in the sense that when the sample is coherent with the null hypothesis, then the Bayes factors are typically large. Applications to real data are also shown.

Fig. 9 Histogram of wind directions data together with GvM($\hat{\mu}_1, \hat{\mu}_2, \hat{\kappa}_1, \hat{\kappa}_2$) density

Table 3 Bayes factors for wind directions data

H_0	B_{01}	Evidence for H_0
No shift between cosines	2.550	Positive
Axial symmetry	2.252	Positive

Due to computational limitations, we consider null hypotheses of symmetry that concern one parameter only. The null hypotheses considered are about one or two distinct values of the parameter of interest, with all remaining parameters fixed. Composite null hypotheses that allow for unknown nuisance parameters, would require one additional dimension of Monte Carlo integration for each unknown parameter, in the computation of the marginal distribution. The computational burden would rise substantially and the Monte Carlo study, with two levels of nested generations, would become very difficult. But the essentially simple null hypotheses considered are relevant in the practice. It can happen that nuisance parameters have been accurately estimated and the question of interest is really about the parameter δ and axial symmetry. In the example of Sect. 3.2, we want to know if wind direction is axially symmetric within the GvM model. The values of the concentrations and of the axial direction are of secondary importance.

One could derive other Bayesian tests for the GvM model: a Bayesian test of bimodality is under investigation. We can also note that Navarro et al. (2017) introduced an useful multivariate GvM distribution for which similar Bayesian tests could be investigated.

The computations of this article are done with the language R, see R Development Core Team (2008), over a computing cluster with several cores. The programs are available at the software section of <http://www.stat.unibe.ch>.

Funding Open Access funding provided by Universität Bern.

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

A Proof of Proposition 2.1

The definition of axial symmetry given at the beginning of Sect. 2.3 tells that the GvM distribution is symmetric around $\alpha/2$ (or $\alpha/2 + \pi$), for some $\alpha \in [0, 2\pi)$, iff

$$f(\theta|\mu_1, \mu_2, \kappa_1, \kappa_2) = f(\alpha - \theta|\mu_1, \mu_2, \kappa_1, \kappa_2), \quad \forall \theta \in [0, 2\pi).$$

This means

$$\begin{aligned} \kappa_1 \cos(\theta - \mu_1) + \kappa_2 \cos 2(\theta - \mu_2) &= \kappa_1 \cos[(\alpha - \theta) - \mu_1] + \kappa_2 \cos 2[(\alpha - \theta) - \mu_2] \\ &= \kappa_1 \cos[\theta - (\alpha - \mu_1)] + \kappa_2 \cos 2[\theta - (\alpha - \mu_2)], \end{aligned} \quad (19)$$

$\forall \theta \in [0, 2\pi)$. By using the cosine addition formula, (19) can be re-expressed as

$$\begin{aligned} \kappa_1 \cos \theta \cos \mu_1 + \kappa_1 \sin \theta \sin \mu_1 + \kappa_2 \cos 2\theta \cos 2\mu_2 + \kappa_2 \sin 2\theta \sin 2\mu_2 \\ = \kappa_1 \cos \theta \cos(\alpha - \mu_1) + \kappa_1 \sin \theta \sin(\alpha - \mu_1) + \kappa_2 \cos 2\theta \cos 2(\alpha - \mu_2) \\ + \kappa_2 \sin 2\theta \sin 2(\alpha - \mu_2), \end{aligned}$$

$\forall \theta \in [0, 2\pi)$. This is equivalent to the equation

$$\begin{aligned} \kappa_1 [\cos \mu_1 - \cos(\alpha - \mu_1)] \cos \theta + \kappa_2 [\cos 2\mu_2 - \cos 2(\alpha - \mu_2)] \cos 2\theta \\ + \kappa_1 [\sin \mu_1 - \sin(\alpha - \mu_1)] \sin \theta + \kappa_2 [\sin 2\mu_2 - \sin 2(\alpha - \mu_2)] \sin 2\theta = 0, \end{aligned}$$

$\forall \theta \in [0, 2\pi)$. It is convenient to re-express this last equation in terms of a trigonometric polynomial of degree $N = 2$, precisely as

$$p(\theta) = \sum_{j=1}^N (a_j \cos j\theta + b_j \sin j\theta) = 0, \quad \forall \theta \in [0, 2\pi), \quad (20)$$

whose coefficients are given by

$$\begin{aligned} a_j &= \kappa_j [\cos j\mu_j - \cos j(\alpha - \mu_j)] \quad \text{and} \\ b_j &= \kappa_j [\sin j\mu_j - \sin j(\alpha - \mu_j)], \quad \text{for } j = 1, 2. \end{aligned}$$

A trigonometric polynomial of degree N has maximum $2N$ roots in $[0, 2\pi)$, unless it is the null polynomial; see e.g. p. 150 of Powell (1981). With this, (20) implies that $p(\theta)$ is the null polynomial, which means that $a_j = b_j = 0$, for $j = 1, 2$. These four equalities give the system of equations

$$\begin{cases} \cos \mu_1 = \cos(\alpha - \mu_1), \\ \sin \mu_1 = \sin(\alpha - \mu_1), \\ \cos 2\mu_2 = \cos 2(\alpha - \mu_2), \\ \sin 2\mu_2 = \sin 2(\alpha - \mu_2), \end{cases}$$

which, in terms of $\delta = (\mu_1 - \mu_2) \bmod \pi$, simplifies to

$$\begin{cases} \alpha = 2\mu_1 + 2k\pi, \\ \alpha = 2[(\mu_1 - \delta) \bmod \pi] + k_1\pi, \end{cases} \quad \text{for some } k, k_1 \in \mathbb{Z}.$$

One can eliminate the congruence symbol mod and obtain

$$\begin{cases} \alpha = 2\mu_1 + 2k\pi, \\ \alpha = 2\mu_1 - 2\delta + (k_1 + 2k_2)\pi, \end{cases} \quad \text{for some } k, k_1, k_2 \in \mathbb{Z}. \quad (21)$$

This system of simultaneous equation admits solutions iff 2δ is a multiple of π , i.e. $2\delta = 0 \bmod \pi$. Since $\delta \in [0, \pi)$, we have found the desired symmetry characterization.

References

- Abe T, Pewsey A, Shimizu K (2013) Extending circular distributions through transformation of argument. *Ann Inst Stat Math* 65(5):833–858
- Abramowitz M, Stegun I (1972) Handbook of mathematical functions, 10 edn. United States Department of Commerce
- Astfalck L, Cripps E, Gosling J, Hodkiewicz M, Milne I (2018) Expert elicitation of directional metocean parameters. *Ocean Eng* 161:268–276
- Batschelet E (1981) Circular statistics in biology. Academic Press, New York
- Berger JO (1985) Statistical decision theory and Bayesian analysis. Springer, New York
- Berger JO, Delampady M (1987) Testing precise hypotheses. *Stat Sci* 317–335
- Berger JO, Sellke T (1987) Testing a point null hypothesis: the irreconcilability of p values and evidence. *J Am Stat Assoc* 82(397):112–122
- Christmas J (2014) Bayesian spectral analysis with student-t noise. *IEEE Trans Signal Process* 62(11):2871–2878
- Gatto R (2008) Some computational aspects of the generalized von Mises distribution. *Stat Comput* 18(3):321–331
- Gatto R (2009) Information theoretic results for circular distributions. *Statistics* 43(4):409–421
- Gatto R (2021) Information theoretic results for stationary time series and the Gaussian-generalized von Mises time series. In: Arnold B, SenGupta A (eds) Selected papers for the bicentennial birth anniversary of F. Nightingale. Springer (to appear)
- Gatto R, Jammalamadaka SR (2003) Inference for wrapped symmetric α -stable circular models. *Indian J Stat* 333–355
- Gatto R, Jammalamadaka SR (2007) The generalized von Mises distribution. *Stat Methodol* 4(3):341–353
- Gatto R, Jammalamadaka SR (2014) Directional statistics: introduction. Wiley, New York, pp 1–8

- Jammalamadaka SR, SenGupta A (2001) Topics in circular statistics, vol 5. World Scientific, Singapore
- Jeffreys H (1961) Theory of probability. Oxford University Press, London
- Kass RE, Raftery AE (1995) Bayes factors. *J Am Stat Assoc* 90(430):773–795
- Kass RE, Wasserman L (1996) The selection of prior distributions by formal rules. *J Am Stat Assoc* 91(435):1343–1370
- Kato S, Jones M (2015) A tractable and interpretable four-parameter family of unimodal distributions on the circle. *Biometrika* 102(1):181–190
- Ley C, Verdebout T (2017) Modern directional statistics. CRC Press, Boca Raton
- Ley C, Verdebout T (2018) Applied directional statistics: modern methods and case studies. CRC Press, Boca Raton
- Lin Y, Dong S (2019) Wave energy assessment based on trivariate distribution of significant wave height, mean period and direction. *Appl Ocean Res* 87:47–63
- Lindley DV (1977) A problem in forensic science. *Biometrika* 64(2):207–213
- Mardia KV, Jupp PE (2000) Directional statistics, vol 494. Wiley, New York
- Navarro A, Frellsen J, Turner R (2017) The multivariate generalised von Mises distribution: inference and applications. In: Thirty-first AAAI conference on artificial intelligence
- Pewsey A (2002) Testing circular symmetry. *Can J Stat* 30(4):591–600
- Pewsey A (2004) Testing for circular reflective symmetry about a known median axis. *J Appl Stat* 31:575–585
- Pewsey A, Neuhaus M, Ruxton GD (2013) Circular statistics in R. Oxford University Press, Oxford
- Powell MJD (1981) Approximation theory and methods. Cambridge University Press, Cambridge
- R Development Core Team (2008) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Spurr BD, Koutbey MA (1991) A comparison of various methods for estimating the parameters in mixtures of von Mises distributions. *Commun Stat Simul Comput* 20(2–3):725–741
- Umbach D, Jammalamadaka SR (2009) Building asymmetry into circular distributions. *Stat Probab Lett* 79(5):659–663
- Zellner A (1984) Posterior odds ratios for regression hypotheses: general considerations and some specific results. In: Zellner A (ed) Basic issues in econometrics, pp 275–305
- Zhang L, Li Q, Guo Y, Yang Z, Zhang L (2018) An investigation of wind direction and speed in a featured wind farm using joint probability distribution methods. *Sustainability* 10(12):4338

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