TEST (2021) 30:1–58
https://doi.org/10.1007/s11749-021-00759-x

INVITED PAPER

Recent advances in directional statistics

Arthur Pewsey[1] · **Eduardo García-Portugués**[2]

Received: 25 April 2020 / Accepted: 30 January 2021 / Published online: 19 March 2021

Abstract
Mainstream statistical methodology is generally applicable to data observed in Euclidean space. There are, however, numerous contexts of considerable scientific interest in which the natural supports for the data under consideration are Riemannian manifolds like the unit circle, torus, sphere, and their extensions. Typically, such data can be represented using one or more directions, and directional statistics is the branch of statistics that deals with their analysis. In this paper, we provide a review of the many recent developments in the field since the publication of Mardia and Jupp (Wiley 1999), still the most comprehensive text on directional statistics. Many of those developments have been stimulated by interesting applications in fields as diverse as astronomy, medicine, genetics, neurology, space situational awareness, acoustics, image analysis, text mining, environmetrics, and machine learning. We begin by considering developments for the exploratory analysis of directional data before progressing to distributional models, general approaches to inference, hypothesis testing, regression, nonparametric curve estimation, methods for dimension reduction, classification and clustering, and the modelling of time series, spatial and spatio-temporal data. An

We are most grateful to three anonymous referees and, in alphabetical order, to Davide Buttarazzi, Marco Di Marzio, Miguel Fernández, Stephan Huckemann, Peter Jupp, Shogo Kato, Christophe Ley, Kanti Mardia, and Louis-Paul Rivest, for their enthusiastic feedback on our original submission and helpful suggestions as to how it might be further improved. This work was supported by Grants PGC2018-097284-B-100, IJCI-2017-32005 and MTM2016-76969-P from the Spanish Ministry of Economy and Competitiveness, and GR18016 from the Junta de Extremadura. All four grants were co-funded with FEDER funds from the European Union.

This invited paper is discussed in the comments available at: https://doi.org/10.1007/s11749-021-00760-4, https://doi.org/10.1007/s11749-021-00761-3, https://doi.org/10.1007/s11749-021-00763-1, https://doi.org/10.1007/s11749-021-00764-0.

✉ Arthur Pewsey
apewsey@unex.es

Eduardo García-Portugués
edgarcia@est-econ.uc3m.es

1 Department of Mathematics, University of Extremadura, Cáceres, Spain

2 Department of Statistics, Carlos III University of Madrid, Leganés, Spain

overview of currently available software for analysing directional data is also provided, and potential future developments are discussed.

Keywords Classification · Clustering · Dimension reduction · Distributional models · Exploratory data analysis · Hypothesis tests · Nonparametric methods · Regression · Serial dependence · Software · Spatial statistics

Mathematics Subject Classification 62H11

1 Introduction

Directional statistics is that branch of statistical methodology specifically designed for use with observations that are directions. A direction observed in the plane $\mathbb{R}^2$, like wind direction, can be represented by an angle, θ, typically in $[0, 2\pi)$ or $[-\pi, \pi)$, measured in a specified direction from a specified origin, or by the unit vector $\boldsymbol{x} = (\cos\theta, \sin\theta)'$ for which $||\boldsymbol{x}|| = \sqrt{\boldsymbol{x}'\boldsymbol{x}} = 1$. The natural support for such directions is the circumference of the unit circle, $\mathbb{S}^1$, data on it being referred to as *circular*. The term "circular data" is also used to distinguish them from data with the real line $\mathbb{R}$ (or some subset of it) as their support, which henceforth we will refer to as *linear* data. Certain calculations can be performed more efficiently using the complex representation $z = \mathrm{e}^{\mathrm{i}\theta}$, where $\mathrm{i} = \sqrt{-1}$, for which $|z| = 1$ and $\mathrm{Arg}(z) = \theta \in [-\pi, \pi)$. Closely related to circular data are *axial data*, which arise when axes, for which the angles θ and $\theta+\pi$ are indistinguishable, are observed. Observations made on directions in $\mathbb{R}^3$, like the positions of stars on the celestial sphere, can be represented by pairs of angles or 3×1 unit column vectors, have natural support the unit sphere, $\mathbb{S}^2$, and are referred to as being *spherical*. Circular and spherical data are the most commonly occurring forms of directional data. Since their supports are compact manifolds, it is (generally) inappropriate, and can prove thoroughly misleading, to apply standard statistical methods, designed for observations with more familiar supports like $\mathbb{R}^d$, $d \geq 1$, to them.

Other data types that fall within the remit of directional statistics include *toroidal* and *cylindrical* data: toroidal data, with support the unit torus, $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$, arising when observations on a pair of circular variables are made, and cylindrical data, with support the cylinder $\mathbb{S}^1 \times \mathbb{R}$ or some subset of it, when observations are made on a pair consisting of one circular and one linear variable. For example, toroidal data are obtained when wind direction is recorded at two different meteorological stations, and cylindrical data if, instead, wind direction and velocity are jointly measured at the same station.

In applications, data on these various manifolds, or their generalisations, such as the unit d-dimensional sphere, $\mathbb{S}^d$, and the d-torus, $\mathbb{T}^d = (\mathbb{S}^1)^d$, $d \geq 1$, might be observed and analysed as regression, time series, spatial or spatio-temporal data. Henceforth, we will use the term "spherical data" to refer to data on any $\mathbb{S}^d$ with $d \geq 1$, not just $\mathbb{S}^2$, unless specifically mentioned otherwise.

Directional statistics can also be applied to data that are not originally directions but which can be represented on, or transformed to, one of the manifolds referred to previously. For instance, times on the 24 h clock can be analysed as circular data after transferring them to the unit circle (see, e.g., Gill and Hangartner 2010). More generally, methods for spherical data can be applied to data originally observed in Euclidean space, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n \in \mathbb{R}^{d+1}$, after their Euclidean normalisation to $\boldsymbol{x}_1/||\boldsymbol{x}_1||, \ldots, \boldsymbol{x}_n/||\boldsymbol{x}_n|| \in \mathbb{S}^d$, a form of transformation often encountered (see, e.g., Banerjee et al. 2005).

Rotation groups, Stiefel and Grassmann manifolds, the elements of which are orthonormal frames and subspaces of $\mathbb{R}^d$, respectively, and other sample spaces such as hyperboloids, complex projective spaces, and general manifolds, are also important to the field but here, because of length restrictions, we refer only tangentially to certain developments related to them. Specifically, we do not consider models for rotations in $\mathbb{R}^3$ despite the fact that a 3×3 rotation matrix can be represented as a 4×1 unit vector called a quaternion, and modelling such rotations is equivalent to modelling axial data on $\mathbb{S}^3$. We would direct the reader interested in these topics to Mardia and Jupp (1999, Chapter 13), Mardia and Patrangenaru (2005), Chirikjian and Kyatkin (2001), Chikuse (2003), Arnold and Jupp (2018), and Rivest and Oualkacha (2018). An important related field is shape analysis (Kendall et al. 1999; Dryden and Mardia 2016), where a *preshape* corresponding to a configuration of k landmarks in $\mathbb{R}^d$ can be regarded as a point on $\mathbb{S}^{d(k-1)-1}$.

The last article-length review of directional statistics was Jupp and Mardia (1989). It contained an extensive bibliography which included virtually all publications on directional statistics between 1975 and 1988. In their review, the authors sought to unify the theory of directional statistics from a mathematical perspective. In attempting to doing so, they referred to five key underpinning ideas: (i) exponential families; (ii) transformation structure; (iii) tangent-normal decomposition; (iv) transformation (of a directional problem) to a multivariate one; and (v) the central limit theorem (CLT), and three basic approaches to directional statistics, termed the embedding, wrapping, and intrinsic approaches. All of these underlying principles have been fundamental to the ongoing development of the field, apart perhaps from the first. Whilst exponential models have certain appealing mathematical and inferential properties, insistence on them has largely been abandoned in recent years, primarily because of an increasing awareness of the need to model distributional features beyond location and concentration, such as the varying levels of skewness and peakedness frequently exhibited by real data. Moreover, directional data are often multimodal and finite mixture distributions, which do not belong to the exponential family, are natural choices with which to model them.

Books covering numerous facets of directional statistics published prior to the review of Jupp and Mardia (1989) include Mardia (1972), Batschelet (1981), Watson (1983), Fisher et al. (1987), and Fisher (1993). Those published after that review include Mardia and Jupp (1999), Jammalamadaka and SenGupta (2001), Pewsey et al. (2013), Ley and Verdebout (2017a), and Ley and Verdebout (2018), the latter being an excellent overview of interesting and important modern applications of directional statistics. We take as our definition of "recent developments" those that have appeared in the literature since the publication of Mardia and Jupp (1999), still the most comprehensive book-length treatment of the field. Whilst many, but certainly not all, of the themes we discuss are addressed in the books of Ley and Verdebout, our aim has been to provide a concise review of the most important developments since the publication of Mardia and Jupp (1999) which is as exhaustive as possible, subject to length constraints. Given the latter, we have concentrated on describing key ideas and directing the interested reader to relevant original sources where more detailed information can be found. With the increasing pace of advances in the field, it is perhaps inevitable that we will have overlooked some developments. We hope that the number of such omissions is minimal and apologise in advance for any that might have arisen.

Important areas of application that have stimulated much of the recent research activity in the field include bioinformatics (Boomsma et al. 2008; Mardia et al. 2018), astronomy (Cabella and Marinucci 2009; Marinucci and Peccati 2011), medicine (Vuollo et al. 2016; Pardo et al. 2017), genetics (Eisen et al. 1998; Dortet-Bernadet and Wicker 2008), neurology (Gu et al. 2004; Kaufman et al. 2005), space situational awareness (Horwood and Poore 2014; Kent et al. 2016), acoustics (McMillan et al. 2013; Traa and Smaragdis 2013), image analysis (Jung et al. 2011; Esteves et al. 2020), text mining (Dhillon and Modha 2001; Banerjee et al. 2005), machine learning (Hamsici and Martinez 2007; Sra 2018), and the modelling of wildfires (García-Portugués et al. 2014; Ameijeiras-Alonso et al. 2018) and sea conditions (Jona-Lasinio et al. 2012, 2018; Lagona 2018).

The remainder of the paper is structured as follows. In Sect. 2 we review advances in exploratory data analysis before proceeding to distributional models in Sect. 3, general approaches to inference in Sect. 4, and to hypothesis testing in Sect. 5. Section 6 discusses developments for correlation and regression. Section 7 focuses on advances in nonparametric curve estimation, Sect. 8 on methods for dimension reduction, and Sect. 9 on classification and clustering. Developments in modelling serial dependence, and spatial and spatio-temporal data, are reviewed in Sects. 10 and 11, respectively. Advances in data depth, the design and analysis of experiments, order-restricted analysis, outlier detection, and compositional data analysis are considered more briefly in Sect. 12. An overview of the software currently available for analysing directional data is provided in Sect. 13. The paper ends with the brief Sect. 14 in which conclusions are drawn and potential future developments discussed.

2 Exploratory data analysis

As for other types of data, the exploratory analysis of directional data usually begins with an inspection of some graphical summary of the data. Various adaptations of the popular rose diagram have been developed recently. Munro and Blenkinsop (2012) introduced a moving rose diagram and applied it to circular datasets from the Earth sciences. Rodgers et al. (2014) proposed the wrap-around time series plot for displaying time series exhibiting periodic patterns. Morphet and Symanzik (2010) proposed the circular dataimage, a graphical tool that uses a colour wheel to encode directions over a map. Rose diagrams, circular histograms, and other circular plots were adapted in Xu and Wang (2020) so as to obtain area-proportional displays.

Circular boxplots have been investigated only relatively recently (but see Anderson 1993). For $\theta_1, \ldots, \theta_n \in [0, 2\pi)$, Abuzaid et al. (2012) advocated one centred on the circular median, $M = \arg\min_{\phi\in[0,2\pi)} \sum_{i=1}^{n} d_c(\phi, \theta_i)$, where

$$d_c(\phi, \theta) = \pi - |\pi - |\theta - \phi|| \tag{1}$$

is the shortest arc length distance between the two angles $\phi, \theta \in [0, 2\pi)$ when represented as points on the circumference of the unit circle. In an attempt to mimic more closely Tukey's original construction, Buttarazzi et al. (2018) proposed a depth-based boxplot in which the observations are ranked from the antimedian to the median. For both proposals, the fences are calibrated assuming an underlying von Mises distribution (see Sect. 3.1).

The SiZer, an abbreviation for "significant zero crossings of derivatives" (Chaudhuri and Marron 1999), is a handy tool used to identify statistically significant features at different scales, such as modes and antimodes, in univariate linear data. A circular adaptation of the SiZer, the CircSiZer, based on the kernel density estimator (18) and bootstrap confidence intervals to assess the significance of smoothed derivatives, was proposed in Oliveira et al. (2014). It can also be employed to explore significant features in linear-circular regression. In both contexts, smoothing is based on a von Mises kernel with concentration parameter κ (see (18) and (20)). This kernel was shown not to be "causal" by Huckemann et al. (2016), in the sense that its convolution with a circular density function is not guaranteed to maintain or reduce the number of modes as the level of smoothing, $1/\kappa$, increases. They proved that, amongst all the circular kernels satisfying certain mild assumptions, the wrapped normal (see Sect. 3.1) is the only one that yields circular causality. Employing such a kernel, they proposed the Wrapped SiZer (WiZer), with asymptotic confidence intervals used to assess the significance of smoothed derivatives. Extension of the SiZer methodology to spherical data led to the SphereSiZer of Vuollo and Holmstrom (2018), itself inspired by the adaptation of the SiZer to bivariate linear data by Godtliebsen et al. (2002). The SphereSiZer uses a von Mises–Fisher kernel density estimator for data on $\mathbb{S}^2$ (see (17)), and bootstrap confidence intervals to assess the significance of smoothed gradients. It produces a movie, indexed by the smoothing scale, that displays statistically significant density gradients as a vector field and highlights spherical regions with high density.

3 Distributional models

3.1 Circular models

The probability density function (pdf) of an absolutely continuous circular random variable (rv) Θ, f_Θ, is such that $f_\Theta(\theta) \geq 0$ and $f_\Theta(\theta+2\pi) = f_\Theta(\theta)$ for almost all $\theta \in \mathbb{R}$. Also, $\int_\theta^{\theta+2\pi} f_\Theta(\omega)\,\mathrm{d}\omega = 1$. Thus, f_Θ is nonnegative, 2π-periodic, and integrates to 1 over any interval of length 2π. As a consequence of this latter property, it is usual to define a circular pdf through its values on $[0, 2\pi)$ or $[-\pi, \pi)$. For instance, the circular uniform distribution, the most fundamental model for circular data corresponding to there being no preferred direction, has pdf $f_\Theta(\theta) = 1/(2\pi)$, $\theta \in [0, 2\pi)$. The circular cumulative distribution function (cdf) is defined as the non-periodic function $F_\Theta(\theta) = \int_{\theta_0}^{\theta} f(\omega)\,\mathrm{d}\omega$, with θ_0 typically being 0 or $-\pi$.

Six general approaches have often been used to generate models for circular data (Mardia and Jupp 1999, Section 3.5): wrapping, projection, perturbation, conditioning, diffusion, and characterisations such as maximum likelihood or maximum entropy. The latter leads to distributions whose entropy is maximal under certain constraints, usually on their moments. The classical von Mises (vM) model with pdf

$$f_\Theta(\theta;\mu,\kappa) = \frac{1}{2\pi\mathcal{I}_0(\kappa)}\exp\{\kappa\cos(\theta-\mu)\}, \tag{2}$$

where $\mu \in [0, 2\pi)$ is the mean direction, $\kappa > 0$ its concentration parameter, and $\mathcal{I}_\nu$ denotes the modified Bessel function of the first kind and order ν, can be derived using no less than five of these constructions (Mardia and Jupp 1999, Section 3.5.4). Due to their relevance in the sequel, below we give brief descriptions of wrapping, projection, and perturbation.

If X is a linear rv, then $\Theta = X \pmod{2\pi} \in [0, 2\pi)$ is its wrapped circular counterpart. Alternatively, using a complex representation, $\Theta = \mathrm{Arg}\{\exp(\mathrm{i}X)\} \in [-\pi, \pi)$. If ϕ_X is the characteristic function (cf) of X, then the cf of Θ is the set $\{\phi_k : k = 0, \pm 1, \ldots\}$ where $\phi_k = \mathrm{E}(\mathrm{e}^{\mathrm{i}k\Theta}) = \phi_X(k)$, the ϕ_k being referred to as the Fourier coefficients or trigonometric moments (Pewsey et al. 2013, Section 4.2.2) of Θ. Thus, Θ inherits the cf of X. If f_X is the pdf of X, then the pdf of Θ is $f_\Theta(\theta) = \sum_{k=-\infty}^{\infty} f_X(\theta + 2\pi k)$, the infinite sum generally not reducing to a closed-form expression. An important exception is the pdf of the wrapped Cauchy (WC) distribution,

$$f_\Theta(\theta;\mu,\rho) = \frac{1}{2\pi}\frac{1-\rho^2}{1+\rho^2-2\rho\cos(\theta-\mu)},$$

where $\mu = \mathrm{Arg}\{\mathrm{E}(\mathrm{e}^{\mathrm{i}\Theta})\} \in [-\pi, \pi)$ is the mean direction and $\rho = |\mathrm{E}(\mathrm{e}^{\mathrm{i}\Theta})| \in [0, 1]$ the mean resultant length. More generally, the trigonometric moments of the WC model are given by

$$\phi_k = (\rho \mathrm{e}^{\mathrm{i}\mu})^k, \quad k = 1, 2, \ldots. \tag{3}$$

Perhaps the best-known wrapped model is the wrapped normal distribution, obtained when $X \sim \mathrm{N}(\mu, \sigma^2)$. It can be used to closely approximate the vM distribution, and vice versa (Pewsey and Jones 2005). Appealing wrapped circular models investigated recently include the wrapped: skew-normal (Pewsey 2000, 2006), exponential and Laplace (Jammalamadaka and Kozubowski 2004), t (Pewsey et al. 2007), stable (Pewsey 2008), and generalised normal-Laplace (Reed and Pewsey 2009).

Projection involves projecting univariate or bivariate linear random variables onto $\mathbb{S}^1$. For example, stereographic projection of the linear random variable X produces the circular rv $\Theta = 2\tan^{-1}(X)$ (Abe et al. 2010). Radial projection of a bivariate linear random vector $\boldsymbol{X} = (X_1, X_2)'$ onto $\mathbb{S}^1$ results in the circular rv $\Theta = \mathrm{Arg}(X_1 + \mathrm{i}X_2)$ or, equivalently, the random point $\boldsymbol{X}/||\boldsymbol{X}||$ on $\mathbb{S}^1$. Perhaps the most popular distribution of this latter type is the projected normal, also known as offset normal or angular Gaussian, the pdf of which can be symmetric or asymmetric, unimodal, or bimodal in shape (Mardia and Jupp 1999, Section 3.5.6). Projection is a natural construction when modelling measurements relative to the position of an observer.

Perturbation involves multiplying a pdf by a suitable function so as to modulate its shape in some desired way. The cardioid distribution, with pdf

$$f_\Theta(\theta; \mu, \rho) = \frac{1}{2\pi}\{1 + 2\rho\cos(\theta - \mu)\},$$

where $|\rho| < 1/2$, is an example of perturbation of the circular uniform model. Umbach and Jammalamadaka (2009) adapted the perturbation approach of Azzalini (1985) to the circular context, a special case of which is the sine-skewed family of distributions studied by Abe and Pewsey (2011). If g_Θ denotes a base symmetric unimodal circular pdf with mean direction μ, then the pdf of its sine-skewed extension is

$$f_\Theta(\theta; \mu, \lambda) = g_\Theta(\theta - \mu)\{1 + \lambda\sin(\theta - \mu)\},$$

where $\lambda \in [-1, 1]$ is a skewing parameter. The symmetric base pdf is unperturbed when $\lambda = 0$; otherwise, it is skewed in the anticlockwise direction ($\lambda > 0$) or the clockwise direction ($\lambda < 0$). Sine-skewed densities have the same normalising constants as their base symmetric densities, but can model only moderate departures from symmetry and are not necessarily unimodal.

An overarching family of symmetric unimodal circular distributions containing, amongst others, the circular uniform, cardioid, vM, and WC distributions, was proposed by Jones and Pewsey (2005). Its pdf is

$$f_\Theta(\theta; \mu, \rho, \psi) \propto \{1 + \tanh(\kappa\psi)\cos(\theta - \mu)\}^{1/\psi},$$

where $\mu \in [0, 2\pi)$ is the mean direction, $\kappa \geq 0$ is a concentration parameter, and $\psi \in \mathbb{R}$ is a shape index.

Recently, Kato and Jones (2015) proposed a highly flexible extension of the WC model obtained by broadening the trigonometric moments in (3) to $\gamma(\rho e^{i\lambda})^{-1}\{\rho e^{i(\mu+\lambda)}\}^k$. The resulting family is unimodal and has pdf

$$f_\Theta(\theta; \mu, \rho, \gamma, \lambda) = \frac{1}{2\pi}\left\{1 + 2\gamma\frac{\cos(\theta - \mu) - \rho\cos\lambda}{1 + \rho^2 - 2\rho\cos(\theta - \mu - \lambda)}\right\}, \tag{4}$$

where $\mu \in [0, 2\pi)$, $\rho \in [0, 1)$, $\gamma \in [0, (1 + \rho)/2]$, and $\lambda \in [-\pi, \pi)$ satisfies $\rho\gamma\cos\lambda \geq (\rho^2 + 2\gamma - 1)/2$. Its cdf also has a closed form. Its reparametrisation in terms of standard trigonometric moments (Pewsey 2004a) has parameters with clear interpretations and is the one generally used to perform inference.

Constructions based on Möbius transformation, Brownian motion, and transformation of argument have also been used recently to generate more flexible families of circular models. A Möbius transformation preserving the unit circle maps a point on the unit circle, Θ, to another, Θ^*, via

$$e^{i\Theta^*} = e^{i\phi}\frac{e^{i\Theta} + re^{i\omega}}{re^{i(\Theta-\omega)} + 1},$$

where $\phi, \omega \in [-\pi, \pi)$ and $r \in [0, 1)$, or equivalently via

$$\Theta^* = \phi + \omega + 2\tan^{-1}[w_r \tan\{(\Theta - \omega)/2\}], \tag{5}$$

where $w_r = (1 - r)/(1 + r)$. Applying this Möbius transformation to a circular uniform rv results in a WC rv (McCullagh 1996). Kato and Jones (2010) and Wang and Shimizu (2012) studied families obtained by applying the same transformation to vM and cardioid random variables, respectively. Jacimovic and Crnkić (2017) related the former family to the dynamics of coupled oscillators. Kato and Jones (2013) varied the Brownian motion specification leading to the WC distribution so as to generate a four-parameter extension of it. The families of Kato and Jones (2010), Wang and Shimizu (2012), and Kato and Jones (2013) have pdfs that can be symmetric or asymmetric and unimodal or bimodal in shape.

Transformation of argument involves replacing the argument of an existing circular pdf, f_Θ, by some function of θ. Jones and Pewsey (2012) used this approach to derive inverse Batschelet distributions. The resulting four-parameter distributions are unimodal and highly flexible in shape. Unlike the smooth unimodal models of Kato and Jones (2015), inverse Batschelet distributions can adopt Laplace-like shapes. The most flexible unimodal circular models currently available are those of Jones and Pewsey (2012) and Kato and Jones (2015).

Of the modelling approaches available for multimodal circular data, finite mixtures have proven the most popular. Mixture models with vM components have recently received renewed attention (Mooney et al. 2003; Fu et al. 2008; see, also, Sect. 3.4). The pdf of an m component vM mixture is

$$f_\Theta(\theta; \boldsymbol{p}, \boldsymbol{\mu}, \boldsymbol{\kappa}) = \sum_{j=1}^{m} p_j f_\Theta(\theta; \mu_j, \kappa_j), \tag{6}$$

where $\boldsymbol{p} = (p_1, \ldots, p_m)'$ is a vector of mixing probabilities satisfying $\sum_{j=1}^{m} p_j = 1$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)'$, $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_m)'$, and $f_\Theta(\theta; \mu_j, \kappa_j)$ is as in (2). When the interpretation of the parameters of the component densities is straightforward, so is the interpretation of the parameters of a mixture.

More generally, Holzmann et al. (2004) established conditions for the identifiability of mixtures of location-scale extensions of wrapped circular models including the wrapped symmetric α-stable, wrapped normal, and WC distributions. Mixtures with circular triangular (McVinish and Mengersen 2008), skew-rotationally symmetric (Miyata et al. 2020), and power Batschelet (Mulder et al. 2020b) components have also been considered.

Regarding alternative approaches to modelling multimodal circular data, generalised von Mises models (Gatto 2008, 2009), with the density of the generalised vM distribution of order m being

$$f_\Theta(\theta; \boldsymbol{\mu}, \boldsymbol{\kappa}) = \exp\left\{\kappa_0 + \sum_{j=1}^{m} \kappa_j \cos(j(\theta - \mu_j))\right\}, \tag{7}$$

where $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_m)'$, $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_m)'$, $\mu_j \in [0, 2\pi/j)$, and $\kappa_j \geq 0$, have a long history dating back to Maksimov (1967). The normalising constant, e^{κ_0}, must generally be computed numerically. Fernández-Durán (2004) revisited work by Fejér (1916) when defining a family of circular distributions based on nonnegative trigonometric (i.e. truncated Fourier) sums. Whilst they do not require the calculation of normalising constants, fitted densities of this type tend to have many parameters and display minor harmonic modes that need not be supported by the data. Recently, Taniguchi et al. (2020) proposed flexible models for circular data obtained by normalising the spectra of stochastic processes, the residue theorem being used to calculate their normalising constants. The interpretation of the parameters of all three of these types of model is, however, generally difficult.

Recently, novel applications have stimulated renewed interest in models for discrete circular data (Mastrantonio et al. 2019; Mardia and Sriram 2020).

3.2 Models for toroidal data

Let (Θ_1, Θ_2) denote the angular coordinates of a random vector distributed on the torus $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$. Some of the approaches used to generate models for toroidal data are extensions of those introduced in Sect. 3.1. These include maximum entropy characterisation, projection, and wrapping (Johnson and Wehrly 1977; Baba 1981; Mardia et al. 2008). Models for univariate and bivariate axial data were proposed by Arnold and SenGupta (2006).

The bivariate von Mises model of Mardia (1975) is a maximum entropy (equivalently, an exponential family) distribution with pdf

$$\begin{aligned} f_{\Theta_1,\Theta_2}(\theta_1, \theta_2) \propto \; & \exp\{\kappa_1 \cos(\theta_1 - \mu_1) + \kappa_2 \cos(\theta_2 - \mu_2) \\ & + (\cos(\theta_1 - \mu_1), \sin(\theta_1 - \mu_1))\boldsymbol{A}(\cos(\theta_2 - \mu_2), \sin(\theta_2 - \mu_2))'\}, \end{aligned} \tag{8}$$

where $\mu_1, \mu_2 \in [-\pi, \pi)$, $\kappa_1, \kappa_2 \geq 0$, and $\boldsymbol{A}$ is a 2×2 matrix. The most compact form for its normalising constant involves a doubly infinite sum (Mardia 2010). The model has eight parameters, three more than the minimum of five required to control the locations and concentrations of the two marginal variables and the dependence between them. Moreover, their interpretation is difficult (Mardia et al. 2007). In the search for five-parameter analogous of the bivariate normal distribution, Singh et al. (2002), Mardia et al. (2007), and Kent et al. (2008) proposed the sine, cosine, and hybrid submodels of (8), respectively. The properties of these three submodels were compared in Kent et al. (2008) and Mardia and Frellsen (2012). Their normalising constants are available as infinite sums. Their conditional distributions are vM, but their marginal pdfs are generally not and, for some parameter values, can be bimodal.

Extensions of (8) and its cosine submodel, for use with data on $\mathbb{T}^d$, $d \geq 2$, were proposed by Mardia et al. (2008) and Mardia and Patrangenaru (2005), respectively. No simple closed analytic form is generally available for the normalising constant of the sine multivariate von Mises model of Mardia et al. (2008), but its conditional distributions are vM and thus its parameters can be estimated by maximising the pseudo-likelihood. Conditions on its parameters to ensure unimodality were established in Mardia and Voss (2014) and pseudo-likelihood regularised approaches were given in Rodriguez-Lujan et al. (2015, 2017). A multivariate extension of the second-order generalised vM (GvM_2) distribution (with $m = 2$ in (7)), obtained by conditioning a multivariate Gaussian distribution on $\mathbb{R}^{2d}$ to $\mathbb{T}^d$, was proposed by Navarro et al. (2017). Its one-dimensional conditional distributions are GvM_2, and the sine multivariate vM model is a special case of it. Hassanzadeh and Kalaylioglu (2018) recently proposed a model for data on $\mathbb{T}^2$ obtained using a conditional specification construction involving GvM_2 pdfs.

The range of available models can be expanded beyond toroidal analogues of the bivariate normal distribution using the projection approach of Saw (1983) to construct models with more flexible specified marginal distributions. A simpler marginal specification approach can be traced back to Wehrly and Johnson (1980). They proposed toroidal pdfs of the form

$$f_{\Theta_1,\Theta_2}(\theta_1, \theta_2) = 2\pi f_{\Theta_1}(\theta_1) f_{\Theta_2}(\theta_2) f_{\Omega}(2\pi[F_{\Theta_2}(\theta_2) - q F_{\Theta_1}(\theta_1)]), \tag{9}$$

where f_{Θ_j} and F_{Θ_j} are the marginal pdf and cdf of Θ_j, $j = 1, 2$, f_Ω is a circular *binding* pdf, and $q = \pm 1$ determines whether the dependence is positive or negative. Various models obtained using (9) with different choices for f_{Θ_1}, f_{Θ_2}, and f_Ω are referred to in Jones et al. (2015).

Kato and Pewsey (2015) considered a case of (9) having a closed-form pdf that is unimodal and pointwise symmetric, and marginal and conditional distributions that are all WC. This bivariate WC model can also be obtained by applying a Möbius transformation to a tractable toroidal model with circular uniform marginal distributions derived by Kato (2009) using a Brownian motion construction.

Shieh and Johnson (2005) were the first to note the relationship between (9) and copulas (Sklar 1959), toroidal pdfs being generated through

$$f_{\Theta_1,\Theta_2}(\theta_1, \theta_2) = f_{\Theta_1}(\theta_1) f_{\Theta_2}(\theta_2) c(F_{\Theta_1}(\theta_1), F_{\Theta_2}(\theta_2)), \tag{10}$$

where c is a *copula* pdf. García-Portugués et al. (2013a) imposed periodic restrictions on c to construct alternatives to (9). Jones et al. (2015) revisited (9) and considered, instead of (10),

$$f_{\Theta_1,\Theta_2}(\theta_1, \theta_2) = 4\pi^2 f_{\Theta_1}(\theta_1) f_{\Theta_2}(\theta_2) c_\circ(2\pi F_{\Theta_1}(\theta_1), 2\pi F_{\Theta_2}(\theta_2)),$$

where now $c_\circ$ is what they coined a *circula* density, with arguments that are circular uniform. For any circula density, $c_\circ(\theta_1, \theta_2) = c_\circ(\theta_1 \pm 2\pi k, \theta_2 \pm 2\pi l), k, l \in \mathbb{Z}^+$. Jones et al. (2015) showed that (9) corresponds to $c_\circ(\phi_1, \phi_2) = \frac{1}{2\pi} f_\Omega(\phi_2 - q\phi_1)$, the pdf of (Φ_1, Φ_2), where Φ_1 and $\Phi_2 = \Phi_1 - q\Omega \pmod{2\pi}$ are circular uniform random variables and Ω follows the circular pdf f_Ω independently of Φ_1. This circula density is a particular case of the (infinite) Fourier series approach to obtaining circula densities proposed recently by Kato et al. (2018). They considered six cases of their general construction, all having simple closed-form expressions for their densities. Jupp (2015) extended the idea of copulas to compact Riemannian manifolds.

Recently, Ameijeiras-Alonso and Ley (2020) used a sine-skewing approach (see Sect. 3.1) to generate models for asymmetric data on $\mathbb{T}^d$, $d \geq 2$. Alternative approaches to modelling such data make use of pdfs obtained from truncated Fourier series (Pertsemlidis et al. 2005; Fernández-Durán and Gregorio-Domínguez 2014b) or normalised spectra of stochastic processes (Taniguchi et al. 2020). These models have properties analogous to those mentioned in Sect. 3.1 for their circular counterparts.

3.3 Models for cylindrical data

Let (Θ, X) denote the coordinates of a random vector distributed on the cylinder $\mathbb{S}^1 \times \mathbb{R}$. Approaches used to generate models for cylindrical data have included wrapping (Johnson and Wehrly 1977), conditioning, marginal specification, and maximum entropy characterisation.

Mardia and Sutton (1978) conditioned a trivariate normal distribution to obtain a six-parameter cylindrical model for which the marginal distribution of Θ is vM and the conditional distribution of $X|\Theta = \theta$ is normal. More recently, Kato and Shimizu (2008) proposed an eight-parameter extension of it having generalised vM distributions for Θ and $\Theta|X = x$.

An analogous marginal specification approach to that used to derive pdf (9) can be employed to obtain a cylindrical model with pdf

$$f_{\Theta,X}(\theta, x) = 2\pi f_\Theta(\theta) f_X(x) f_\Omega(2\pi[F_\Theta(\theta) - qF_X(x)]), \tag{11}$$

where f_Θ and f_X are the marginal pdfs of Θ and X, respectively, and F_Θ and F_X their cdfs. Johnson and Wehrly (1978) considered cases of (11) with X normally distributed, and circular uniform or vM distributions for Θ. Recently, other cases have been applied to model cylindrical data from disciplines such as wind energy analysis (Carta et al. 2008; Zhang et al. 2018a), ocean engineering (Soukissian 2014), and image analysis (Roy et al. 2017).

Johnson and Wehrly (1978) also proposed three maximum entropy cylindrical models, with conditional distributions that are vM and normal or exponential. The dependence structures of all three models are, however, severely constrained. Their model having vM and exponential marginal distributions when Θ and X are independent was recently extended by Abe and Ley (2017), Imoto et al. (2019), and Abe and Shimatani (2018) so as to admit skew and more flexible models for Θ and X, respectively.

Recently, Mastrantonio (2018) proposed the joint projected normal and skew-normal distribution, the first model for multivariate cylindrical data. It is highly flexible and closed under marginalisation.

3.4 Models for spherical data

Suppose $\boldsymbol{X}$ is a unit random vector on $\mathbb{S}^d$. Perhaps the best-known model for spherical data is the von Mises–Fisher (vMF) distribution, with pdf

$$f_{\boldsymbol{X}}(\boldsymbol{x};\boldsymbol{\mu},\kappa)=\frac{\kappa^{(d-1)/2}}{(2\pi)^{(d+1)/2}\mathcal{I}_{(d-1)/2}(\kappa)}\exp\{\kappa\boldsymbol{x}'\boldsymbol{\mu}\},\tag{12}$$

where $\boldsymbol{\mu}\in\mathbb{S}^d$ is the mean direction vector and $\kappa\geq 0$ is a concentration parameter. Other classical models are the Kent, Fisher–Watson, Bingham–Mardia, Bingham, and Watson distributions (Mardia and Jupp 1999, Chapter 9), the last two being models for axial data. They are all submodels of the Fisher–Bingham exponential family of distributions with pdf

$$f_{\boldsymbol{X}}(\boldsymbol{x};\boldsymbol{\mu},\kappa,\boldsymbol{A})\propto\exp\{\kappa\boldsymbol{x}'\boldsymbol{\mu}+\boldsymbol{x}'\boldsymbol{A}\boldsymbol{x}\},\tag{13}$$

where $\boldsymbol{A}$ is a symmetric $(d+1)\times(d+1)$ matrix, and $\boldsymbol{\mu}$ and κ play the same roles as in (12). Evaluation of the distribution's normalising constant is challenging: Kume and Sei (2018) showed how it can be calculated using the holonomic gradient method, and Yuan (2021) provided an algorithm for computing it, when $d=2$, based on an infinite series expansion involving hypergeometric functions. Kent et al. (2018) developed an efficient acceptance-rejection method of simulating variates from Fisher–Bingham distributions on spheres and related manifolds. Kent et al. (2016) introduced a five-parameter special case of the Fisher–Bingham model for use with data patterns that are unimodal and concentrated near a great circle. More recently, Kim et al. (2019) proposed two kinds of small-sphere distributions, one of which is a member of the Fisher–Bingham family. Previously, Oualkacha and Rivest (2009) had developed an alternative to the Bingham distribution for modelling symmetric axial data, with a simple closed-form normalising constant.

Rotationally symmetric (RS) spherical pdfs depend on $\boldsymbol{x}$ only through $\boldsymbol{x}'\boldsymbol{\mu}$ and, as a consequence, have contours that are circular when $\boldsymbol{x}\in\mathbb{S}^2$. Historically, the vMF has been the most important such model. In recent years, numerous other RS families have been proposed in the literature: Section 2.3.2 of Ley and Verdebout (2017a) summarises many of them. The spherical logistic distribution of Moghimbeygi and Golalizadeh (2020) provides a multimodal and RS extension of the vMF, with a closed-form normalising constant when $d=2$. Another recent addition is the highly tractable spherical Cauchy distribution of Kato and McCullagh (2020), which extends the WC to $\mathbb{S}^d$ and has a very simple normalising constant.

The pdf of the Kent distribution, i.e. (13) constrained to have $\boldsymbol{A}\boldsymbol{\mu}=\boldsymbol{0}$, has elliptical contours and hence can be used to model certain departures from isotropy. Recently, Paine et al. (2018) proposed the elliptically symmetric angular Gaussian (ESAG) as an alternative. As simulation from it and the computation of its pdf are far quicker than for the Kent model, this model is a particularly appealing alternative when the use of computer intensive methods is being contemplated. Other Kent-like alternatives with the advantages of the ESAG model are the scaled vMF family of Scealy and Wood

(2019), which has an additional parameter controlling tail-weight, and the tangent models of García-Portugués et al. (2020b).

The sine-skewed circular distributions of Sect. 3.1 are special cases of the skew-rotationally symmetric (SRS) distributions proposed as models for asymmetric spherical data by Ley and Verdebout (2017b). In turn, SRS models are spherical analogues of the skew-symmetric linear models of Wang et al. (2004). The perturbation of spherical distributions was studied in greater generality by Jupp et al. (2016).

Asymmetric or bimodal spherical data can be modelled using the general projected normal family of distributions referred to in Mardia and Jupp (1999, Section 9.3.3) and advocated more recently from a Bayesian perspective in Hernandez-Stumpfhauser et al. (2017). Núñez-Antonio and Geneyro (2020) proposed a projected gamma distribution to model data on the positive orthant of $\mathbb{S}^d$.

In Sect. 9 we summarise recent developments in the use of mixture distributions with spherical component pdfs as a means of modelling multimodal spherical data. The flexible directional log-spline pdfs of Ferreira et al. (2008), based on thin-plate splines on $\mathbb{S}^d$ (Taijeron et al. 1994), provide an alternative means of modelling multimodality and skewness. They are given by

$$f_X(\boldsymbol{x}; \boldsymbol{c}, \mathcal{K}, m) = \exp\left\{c_0 + \sum_{j=1}^{m} c_j R_d(\boldsymbol{x}; \boldsymbol{k}_j)\right\}, \tag{14}$$

where $m \geq 1$, $(c_0, c_1, \ldots, c_m)' \in \mathbb{R}^{m+1}$, $\mathcal{K} = \{\boldsymbol{k}_1, \ldots, \boldsymbol{k}_m\}$ is a set of knot vectors in $\mathbb{S}^d$, and $R_d(\cdot; \boldsymbol{k}_j)$ are real-valued spline basis functions on $\mathbb{S}^d$ that, when evaluated at $\boldsymbol{x} \in \mathbb{S}^d$, are functions of $\boldsymbol{x}'\boldsymbol{k}_j$, for $j = 1, \ldots, m$. Ferreira et al. (2008) proposed a Bayesian inferential approach for (14). Fernández-Durán and Gregorio-Domínguez (2014a) constructed pdfs on $\mathbb{S}^2$ through nonnegative trigonometric sum expansions in terms of spherical angles.

4 General approaches to inference

Historically, inference for the models in Sect. 3 has generally been frequentist: sometimes using the method of (trigonometric) moments but more generally being likelihood-based. The maximum likelihood (ML) estimators of full exponential family models are moment estimators (van der Vaart 2000, Chapter 4) and, as a consequence, closed-form expressions exist for the ML estimators of, e.g., the vM distribution and the cylindrical model of Mardia and Sutton (1978). Exact ML inference for the highly challenging Fisher–Bingham model and its submodels identified in Sect. 3.4 was developed recently in Kume and Sei (2018). More generally, maximisation of the log-likelihood has to be performed numerically. When available, method of moments estimates can be used as starting values for that process. For some models, statistical inference based on the full log-likelihood is intractable and pseudo-likelihood methods have been employed (Kent et al. 2008). Score matching estimators, inspired by the Hyvärinen (2005) scoring rule, circumvent the need to calculate normalising constants for directional distributions (Mardia et al. 2016; Mardia 2018; Takasu et al. 2018).

Large-sample results for ML-based inference generally assume standard regularity conditions to derive the asymptotic normality of ML estimators (van der Vaart 2000, Chapter 5). Pewsey (2004a) employed the delta method to obtain the asymptotic distribution of the fundamental measures of central location, concentration, skewness, and kurtosis used in the analysis of circular data. For some models, large-sample ML-based inference for parameter values on the boundary of the parameter space will be of interest, and the results of Self and Liang (1987) can be employed (see, e.g., Shieh and Johnson 2005). For small-sized samples, bootstrap confidence interval constructions have become increasingly popular (Pewsey et al. 2013, Chapter 5). Computer-intensive resampling methods in hypothesis testing are mentioned in Sect. 5.

Recently, Le Cam's local asymptotic normality approach to inference has been adapted to problems in directional statistics: see Sect. 5, and Section 5 of Ley and Verdebout (2017a). The first such adaptation appeared in Ley et al. (2013), where optimal rank-based estimators of the location parameter of rotationally symmetric spherical distributions were proposed. More recently, Paindaveine and Verdebout (2020) considered inference under high concentration for the spherical location of a semi-parametric class of rotationally symmetric distributions.

Bayesian inferential techniques have become increasingly popular in recent years, often being implemented using Markov chain Monte Carlo (MCMC) methods. A general approach to MCMC simulation on embedded Riemannian manifolds was introduced by Byrne and Girolami (2013) and illustrated for the Fisher–Bingham distribution. Bayesian approaches to inference have been developed for the: vM (Damien and Walker 1999) and mixtures thereof (Mulder et al. 2020a), WC (Ghosh et al. 2019), vMF (Núñez-Antonio and Gutiérrez-Peña 2005b; Hornik and Grün 2013), bivariate vM (Mardia 2010), and projected normal (Núñez-Antonio and Gutiérrez-Peña 2005a) distributions. Bhattacharya and SenGupta (2009b) considered Bayesian inference for circular distributions with unknown normalising constants. Fallaize and Kypraios (2016) gave a Monte Carlo exact Bayesian method of inference for the Bingham distribution. Bayesian approaches based on projected and wrapped models have become popular for a range of applications: see Sects. 3.4 and 11. Scoring rules provide an alternative to the traditional Bayesian formulation and have been applied for the vMF distribution (Giummolè et al. 2019).

Robust estimators have been proposed for the parameters of the vM and wrapped normal distributions (Agostinelli 2007), the vMF distribution (Kato and Eguchi 2016), and a range of other circular distributions in a series of papers referred to by Laha et al. (2019).

Asymptotic results for extrinsic and intrinsic means on manifolds, including $\mathbb{S}^d$, were obtained in Bhattacharya and Patrangenaru (2003, 2005); see Bhattacharya and Patrangenaru (2014) for a review on the topic. Hotz (2013) gave a detailed comparison between extrinsic and intrinsic means on $\mathbb{S}^1$. Nonparametric inference on intrinsic means on circles and spheres, however, can be fundamentally different from its Euclidean analogues due to the effect of *smeariness* (asymptotic rates are slower than $n^{-1/2}$) present on $\mathbb{S}^1$ (Hotz and Huckemann 2015) and $\mathbb{S}^d$ (Eltzner and Huckemann 2019). The related effect of *finite sample smeariness* has been studied on the circle by Hundrieser et al. (2020) and may affect all of the circular distributions mentioned previously. As a consequence, quantile-based tests may be inappropriate, whilst suitable bootstrap tests remain valid.

The inferential impact of the reference systems used for circular distributions was explored recently in Mastrantonio et al. (2019).

5 Hypothesis testing

Here we consider hypothesis tests for uniformity, symmetry, location, concentration, goodness-of-fit, and other testing scenarios. Calibration of the tests has generally been based on asymptotic theory and, for small- to moderate-sized samples, the use of resampling methods.

5.1 Uniformity

Uniformity (or isotropy), corresponding to there being no preferred direction, is the most important dividing hypothesis in directional statistics. García-Portugués and Verdebout (2018) provide a review of tests for it.

Sobolev tests (Beran 1968, 1969; Giné 1975) form, by far, the most extensive class of tests for uniformity on $\mathbb{S}^d$. Given a sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ on $\mathbb{S}^d$, Sobolev statistics take the form

$$S_n(\{v_k^2\}) = \frac{1}{n} \sum_{i,j=1}^{n} \sum_{k=1}^{\infty} v_k^2 h_k(\boldsymbol{X}_i, \boldsymbol{X}_j), \tag{15}$$

where

$$h_k(\boldsymbol{u}, \boldsymbol{v}) = \begin{cases} 2\cos(k\cos^{-1}(\boldsymbol{u}'\boldsymbol{v})), & d = 1, \\ \left(1 + \frac{2k}{d}\right) C_k^{(d-1)/2}(\boldsymbol{u}'\boldsymbol{v}), & d > 1, \end{cases}$$

$C_k^{(d-1)/2}$ is the k-th Gegenbauer polynomial of index $(d-1)/2$, and the v_k^2 should decay fast enough to ensure convergence in (15). Different choices for $\{v_k^2\}$ give different local optimality properties, consistencies, and powers against specific kinds of alternatives. For example, the choices $v_k = \delta_{kj}$, $j = 1, 2$, give, respectively, the test statistics of Rayleigh (1919) and Bingham (1974). Both were modified by Jupp (2001) to improve their convergence under the null hypothesis. The alternatives for which the Rayleigh and Bingham tests are inconsistent were identified by Ehler and Galanis (2011) as the minimisers of certain potentials over $\mathbb{S}^d$. The Rayleigh test plays a key role in the CUSUM-based test for circular uniformity developed by Lombard and Maxwell (2012). Pycke (2007, 2010) proposed uniformity tests on $\mathbb{S}^2$ and $\mathbb{S}^1$ based on the geometric mean of pairwise chordal distances, whilst Bakshaev (2010) gave an analogous approach based on the arithmetic mean. The Bayesian optimality of Sobolev tests on $\mathbb{S}^1$ was studied by Sun and Lockhart (2019).

"Data-driven" Sobolev tests are obtained by using an information criterion to truncate the infinite series in (15). This approach was used to obtain tests of uniformity on $\mathbb{S}^1$ by Bogdan et al. (2002), and on compact Riemannian manifolds by Jupp (2008, 2009). Such truncation simplifies the computation of (15) and its asymptotic distribution, the latter effectively being chi-squared. A variation on this approach was pursued recently by Jammalamadaka et al. (2020), who proposed increased levels of truncation of (15) on $\mathbb{S}^1$ and $\mathbb{S}^2$, so as to obtain a normal limit rather than the usual weighted sum of chi-squared random variables appearing in the asymptotic null distribution of (15).

Su and Wu (2011) considered spherical harmonics and exponential models as alternatives to uniformity, and derived score tests strongly related to Sobolev tests. Also related to Sobolev tests, García-Portugués et al. (2020a) proposed a class of tests based on the projected empirical cdf that yields extensions for data on $\mathbb{S}^d$ of the Watson (1961) and Rothman (1972) tests for circular uniformity, and a novel Anderson–Darling-like test for uniformity on $\mathbb{S}^d$.

Recent non-Sobolev tests for circular uniformity include the four-point Cramér–von Mises test of Feltz and Goldin (2001), the likelihood-ratio test against a mixture with symmetric wrapped stable and circular uniform components of SenGupta and Pal (2001), the spacings-based Gini mean difference test of Tung and Jammalamadaka (2013), and the Bayesian tests of Mulder and Klugkist (2021) against the vM distribution and the kernel density estimator (18). Tests for uniformity on $\mathbb{S}^d$ include that of Faÿ et al. (2013), based on needlets (see Sect. 7.1.2), and those of Lacour and Pham Ngoc (2014) and Kim et al. (2016) for "noisy" data on $\mathbb{S}^2$, i.e. where the density of the observations is a convolution of an error pdf with a true underlying pdf. Cuesta-Albertos et al. (2009) proposed a projection-based test, and Ebner et al. (2018) one based on a sum of weighted nearest-neighbour distances. Cutting et al. (2020) investigated tests for uniformity on $\mathbb{S}^d$ against axial alternatives.

Tests for uniformity when $d \to \infty$ as $n \to \infty$ are scarcer. Cai and Jiang (2012) and Cai et al. (2013) proposed tests based on $\max_{i<j} |\boldsymbol{X}_i' \boldsymbol{X}_j|$. Paindaveine and Verdebout (2016) and Cutting et al. (2017a) studied a standardised Rayleigh statistic under uniformity and vMF alternatives, respectively.

Simulation studies comparing the performance of various tests for uniformity on $\mathbb{S}^1$ have been carried out by Landler et al. (2018), on $\mathbb{S}^d$, $d \geq 1$, by García-Portugués et al. (2020a), and on $\mathbb{S}^d$, $d \geq 2$, by Figueiredo and Gomes (2003) and Figueiredo (2007). Humphreys and Ruxton (2017) and Landler et al. (2019) performed simulation experiments to compare the performance of tests for circular uniformity when the data are grouped.

5.2 Symmetry

There are at least four forms of symmetry that might be of interest in the analysis of circular data: reflective symmetry about an unknown direction, reflective symmetry about a known median axis, reflective symmetry about a specified median axis, and ℓ-fold symmetry. Pewsey (2002, 2004b) described these various forms of symmetry and proposed simple, trigonometric moment-based, omnibus tests for the first two scenarios. More recently, Ameijeiras-Alonso et al. (2020), Meintanis and Verdebout (2019), and Ley and Verdebout (2014b) proposed tests for the same two setups that are optimal against the k-sine-skewed models of Umbach and Jammalamadaka (2009).

As we saw in Sect. 3.4, many of the models for spherical data are rotationally symmetric. Recently, García-Portugués et al. (2020b) developed semi-parametric tests for rotational symmetry when the axis of symmetry is known or unknown. Previously, Ley and Verdebout (2017b) had developed two tests for rotational symmetry about a known centre within the class of skew-rotationally symmetric distributions.

For data on $\mathbb{S}^2$, Jammalamadaka and Terdik (2019) proposed tests for various types of symmetry, as well as uniformity, based on spherical harmonics.

5.3 Location and concentration

Various tests for the parameters of vM and vMF models have been developed in recent years. Larsen et al. (2002) proposed improved likelihood-ratio tests for the: (i) mean direction of a vM distribution; (ii) equality of mean directions of two vM distributions; (iii) concentration of a vMF distribution; and (iv) equality of concentrations of two vMF distributions. Watamori and Jupp (2005) introduced improved likelihood-ratio and score tests for homogeneity of concentration in vMF distributions. Their score tests were derived and studied from an alternative perspective in the review of Ley and Verdebout (2014a). Laha and Mahesh (2015) investigated the robustness of tests for the locations of the vM and vMF models. Gatto (2017) proposed a simultaneous test for the mean direction and concentration of a vMF distribution. For data from vMF distributions with a common unknown concentration, Rumcheva and Presnell (2017) proposed an improved version of the multisample likelihood-ratio test for the equality of mean directions.

Widening the scope to rotationally symmetric spherical distributions, Tsai (2009) introduced asymptotically efficient rank tests for the equality of the modal direction vectors of two unimodal rotationally symmetric spherical distributions. Ley et al. (2015) investigated the high dimensional robustness of Watson's test for the mean direction. Paindaveine and Verdebout (2015) proposed optimal rank-based tests for the mean direction, and Ley et al. (2017) used the invariance principle to construct rank-based semi-parametric tests for the homogeneity of mean directions. Paindaveine and Verdebout (2017) investigated the problem of testing for a specified mean direction when the underlying distribution tends to uniformity. Cutting et al. (2017b) proposed tests for concentration in low and high dimensions, and Verdebout (2015, 2017) tests for homogeneity of concentration.

A simultaneous saddlepoint test for the mean direction and dispersion of the wrapped symmetric stable model was proposed in Gatto (2000), and a nonparametric extension of it, for an assumed underlying unimodal circular distribution, in Gatto (2006). Amaral et al. (2007) proposed nonparametric bootstrap and permutation tests for the equality of the mean directions of directional distributions.

For axial data, Figueiredo (2017) proposed and explored the performance of bootstrap and permutation counterparts of a high-concentration test for the homogeneity of principal axes of Watson distributions.

5.4 Goodness-of-fit

Given an independent and identically distributed circular sample, $\Theta_1, \ldots, \Theta_n$, with underlying cdf F, the goodness-of-fit testing problem of $H_0 : F = F_0$ versus $H_1 : F \neq F_0$, where F_0 is a fully specified cdf, can be addressed using tests for circular uniformity. Appealing to the probability integral transform, testing H_0 is equivalent to testing the sample $2\pi F_0(\Theta_1), \ldots, 2\pi F_0(\Theta_n)$ on $[0, 2\pi)$ for circular uniformity. When the parameters of a model are estimated, the sampling distributions of the test statistics are affected. However, those sampling distributions can be approximated using the parametric bootstrap (Pewsey et al. 2013, Chapter 6).

For data on other supports, the situation is more complicated as there is no canonical transformation to uniformity. Examples of bootstrap goodness-of-fit tests for models fitted to toroidal data appear in Jones et al. (2015), Pewsey and Kato (2016), and Kato et al. (2018): the approach used in the latter essentially being based on the multivariate probability integral transform. Almost-canonical transformations for other Riemannian manifolds have been proposed recently in Jupp and Kume (2020).

Using spherical harmonic expansions, Boulerice and Ducharme (1997) developed goodness-of-fit tests for the vMF and Watson distributions. Jupp (2005) considered weighted Sobolev goodness-of-fit tests for distributions on compact Riemannian manifolds. Deschepper et al. (2008) proposed a lack-of-fit test for linear-circular regression models based on the arcs generated by the circular observations. Wouters et al. (2009) proposed data-driven goodness-of-fit tests for the vM model based on orthonormal polynomials.

Smoothing-based approaches to testing the goodness of fit of parametric models to directional data have also been developed. Boente et al. (2014) and García-Portugués et al. (2015) used kernel density estimators (see Sect. 7.1) to test the goodness of fit of spherical, and spherical-linear/spherical models, respectively. In the regression context, García-Portugués et al. (2016) proposed a goodness-of-fit test for linear-spherical models based on an extension of (19).

5.5 Other testing scenarios

For bivariate circular data, the asymptotic sampling properties of likelihood-ratio tests of independence were considered by Shieh and Johnson (2005), and their permutation analogues by Kato and Pewsey (2015) and Kato et al. (2018). Nonparametric, kernel-based, tests for the independence of spherical-linear/spherical variables were proposed in García-Portugués et al. (2014, 2015).

Several tests for change-point detection in circular series have been introduced. Nonparametric tests include the permutation test of Byrne et al. (2009) and the CUSUM test of Lombard et al. (2017). In vM-distributed series, likelihood (Ghosh et al. 1999; Hawkins and Lombard 2015) and CUSUM (Hawkins and Lombard 2017) procedures have been advocated. Bayesian approaches have also been considered (Ghosh et al. 1999; SenGupta and Laha 2008).

Other tests for circular data include the kernel density estimate-based tests of Fisher and Marron (2001) and Ameijeiras-Alonso et al. (2019a) for assessing the number of modes, and the test of Ducharme et al. (2012) for detecting vortices in two-dimensional vector fields. Recently, bootstrap-based tests using smoothing have been introduced in Zhang et al. (2019), for comparing two samples, and in Alonso-Pena et al. (2020), for testing in circular regression.

6 Correlation and regression

6.1 Correlation

Mardia and Jupp (1999, Section 11.2) provide details of some of the correlation coefficients for toroidal, cylindrical, and spherical data that have been proposed in the literature. Others are considered in SenGupta (2001, Chapter 8). Recently, Zhan et al. (2019) reviewed the correlation coefficients available for toroidal data and proposed two new ones. In the context of Bayesian network modelling, Leguey et al. (2019b) and Leguey et al. (2019a) introduced mutual information measures of the dependence between circular and linear variables, and between two circular variables, respectively.

6.2 Regression

Here we consider parametric regression models. Throughout, we use the generic notation $\mathcal{Y}$-$\mathcal{X}$ to denote that $\mathcal{Y}$ is the response variable and $\mathcal{X}$ is the explanatory variable. Recent developments in nonparametric regression are described in Sect. 7.2.

6.2.1 Circular-circular regression

Circular-circular regression is used to model the relationship between a circular response variable, Ψ, and a circular explanatory variable, Θ. Polsen and Taylor (2015) reviewed parametric circular-circular regression models and considered inference and diagnostic analysis for them. They focused on the general inverse tangent link-based regression model

$$\Psi = \mathrm{Arg}\{g_1(\Theta; \boldsymbol{\eta}) + \mathrm{i}g_2(\Theta; \boldsymbol{\eta})\} + \varepsilon \pmod{2\pi}, \tag{16}$$

where the first term represents the conditional mean direction of Ψ given Θ, g_1 and g_2 are non-uniquely identifiable functions, $\boldsymbol{\eta}$ is a vector of parameters, and ε is a circular error variable. The decentred model of Rivest (1997) and the Möbius transformation-based models of Downs and Mardia (2002), Kato et al. (2008), and Kato and Jones (2010) are all special cases of (16). The first two have vM errors, whilst those of the models of Kato et al. (2008) and Kato and Jones (2010) are WC and a four-parameter highly flexible unimodal extension of the WC different from (4), respectively. Polsen and Taylor (2015) also related the bivariate regression model of Sarma and Jammalamadaka (1993) (see also Jammalamadaka and SenGupta 2001, Section 8.6), having finite-order trigonometric polynomials for g_1 and g_2, to the model in (16).

McMillan et al. (2013) proposed a hierarchical Bayesian approach for repeated measures circular data that are bimodal, based on a two-component circular-circular regression model with parameters that change according to a function expressed in a finite circular B-spline basis (see Sect. 7.2.1).

6.2.2 Circular-linear regression

Circular-linear regression is used to model the relationship between a circular response variable, Θ, and a vector containing one or more covariates denoted here by $\boldsymbol{X}$. Mardia and Jupp (1999, Section 11.3.2) refer to the use of the link-based models of Fisher and Lee (1992) in this context. It would appear that the most popular such link function has been $2\tan^{-1}$. For this choice, the conditional mean is

$$\mathrm{E}(\Theta|\boldsymbol{X}=\boldsymbol{x}) = 2\tan^{-1}(\boldsymbol{\beta}'\boldsymbol{x}),$$

where $\boldsymbol{\beta}$ is a vector of regression coefficients. This link function maps the origin of $\mathbb{R}$ to the angle 0, and the two extremes of $\mathbb{R}$ to the angle furthest from 0, namely $-\pi \equiv \pi$. Presnell et al. (1998) identified important practical difficulties with estimating the parameters of such models using ML methods. As a means of circumventing those inherent inferential problems, George and Ghosh (2006) proposed a semi-parametric Bayesian approach. Artes (2008) developed analysis of covariance tests for link-based models.

Instead of using link functions, Presnell et al. (1998) proposed an alternative modelling approach based on projecting (see Sect. 3.1) the unobserved responses from a multivariate linear model onto $\mathbb{S}^1$. This approach has become increasingly popular, particularly in Bayesian applications (see, e.g., Núñez-Antonio et al. 2011; Wang and Gelfand 2013; Hernandez-Stumpfhauser et al. 2016). The interpretation of predictor effects in projected normal regression models has been considered recently by Cremers et al. (2018).

A different tack was taken by Lund (2002), who evaded the problem of devising a meaningful regression function through the use of a tree-based approach to predicting a circular response from a combination of circular and linear predictors.

Various approaches to modelling longitudinal data have been developed recently. Artes et al. (2000) considered the use of estimating equations when the angular response is assumed to follow a circular distribution parametrised by its mean direction and mean resultant length. D'Elia (2001) proposed a variance components model with fixed and random effects. Lagona (2016) introduced a regression model for correlated circular data which assumes that the angular measurements arise from the sine multivariate vM distribution of Mardia et al. (2008). All three of these proposals make use of the link function approach of Fisher and Lee (1992). The model of Lagona (2016) was extended by Mulder and Klugkist (2017), who employed weakly informative priors within a Bayesian framework to elude the problems with ML estimation for link-based models. Other researchers have adapted projected normal models. Núñez-Antonio and Gutiérrez-Peña (2014) investigated one in which the components are specified as mixed linear models. Maruotti (2016) considered a mixed linear model with correlated random coefficients controlling dependence that can be represented as a finite mixture of projected normal distributions. Maruotti et al. (2016) proposed a time-dependent extension of the projected normal regression model with a hidden Markov heterogeneity structure.

Many of the above proposals have been used to model animal orientation data. Other models for such data include that of Rivest et al. (2016), which features a consensus model for the angular response, based on circular and linear covariates, combined with vM errors. Recently, Rivest and Kato (2019) proposed a random effects circular regression model for clustered circular data where both the cluster effects and the regression errors have vM distributions. Their model is based on the multivariate angular pdf with vM-distributed cluster-level random effects of Holmquist and Gustafsson (2017). Other approaches to modelling animal orientation data are considered in Sect. 11.

6.2.3 Linear-circular regression

Linear-circular regression is applied to model the relationship between a linear response variable and one or more circular covariates. The standard approach is to regress the linear variable on sums of trigonometric polynomials of the circular variables, using least squares to estimate the parameters (Johnson and Wehrly 1978). Bhattacharya and SenGupta (2009a) and SenGupta and Bhattacharya (2015) have considered Bayesian approaches to linear-circular modelling.

Recently, Cremers et al. (2020) proposed several regression models for a cylindrical response variable with linear and circular components.

6.2.4 Spherical response

Spherical regression was first considered by Chang (1986). Downs (2003) made use of Möbius transformation, stereographic projection and link functions to develop $\mathbb{S}^2$-$\mathbb{S}^2$ regression models with the conditional distribution between response and predictor being vMF. Hinkle et al. (2014) proposed polynomial models for manifold-linear regression. For $\mathbb{S}^d$-$\mathbb{S}^d$ regression, Rosenthal et al. (2014) employed projective linear transformations to model the conditional mean direction of the response, combined with a vMF error structure. Cornea et al. (2017) proposed a more general semi-parametric intrinsic manifold-manifold regression model that incorporates parametric link functions and a nonparametric error structure. Very recently, Paine et al. (2020) introduced a very general regression model for an $\mathbb{S}^2$-valued response with covariates that can be spherical, linear or categorical, and two kinds of anisotropic error distributions. In its most general formulation, a preliminary orthogonal transformation of the response is assumed to follow an anisotropic distribution with covariate-dependent parameters. For $\mathbb{S}^d$-$\mathbb{R}^q$ regression, Scealy and Wood (2019) proposed a flexible heteroscedastic model with scaled vMF errors.

Related regression problems for a $\mathbb{S}^2$-valued response include the fitting of small circles to spherical data (Rivest 1999) and the analysis of rotational deformations through fitting small circles on the sphere nonparametrically (Schulz et al. 2015) and parametrically (Kim et al. 2019).

7 Nonparametric curve estimation

Here we review advances in nonparametric curve estimation. See Sect. 5 for nonparametric tests, and later sections for other nonparametric methods.

7.1 Density estimation

7.1.1 Smoothing-based

Kernel density estimation (KDE) on $\mathbb{S}^d$ dates back to Beran (1979), Hall et al. (1987), and Bai et al. (1988). In the latter's formulation, the kernel estimator for a sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ from the target pdf f is given by

$$\hat{f}(\boldsymbol{x}; h) = \frac{c_L(h)}{n} \sum_{j=1}^{n} L\left(\frac{1 - \boldsymbol{x}'\boldsymbol{X}_j}{h^2}\right), \quad \boldsymbol{x} \in \mathbb{S}^d, \tag{17}$$

where $h > 0$ denotes the bandwidth, $L : [0, \infty) \to [0, \infty)$ a kernel, and $c_L(h)$ a normalising constant. For the vMF kernel $L(r) = e^{-r}$ and $d = 1$, (17) reduces to (6) with common concentration $\kappa = 1/h^2$, namely

$$\hat{f}(\theta; \kappa) = \frac{1}{2\pi \mathcal{I}_0(\kappa) n} \sum_{j=1}^{n} \exp\{\kappa \cos(\theta - \Theta_j)\}, \quad \theta \in [-\pi, \pi). \tag{18}$$

Several extensions and modifications of (17) and (18) have been proposed. For $d \geq 2$, Klemelä (2000) used $L(\kappa \cos^{-1}(\boldsymbol{x}'\boldsymbol{X}_j))$ in (17) to analyse estimators of f and its derivatives. Extending (18) to $[-\pi, \pi)^d$, Di Marzio et al. (2011) introduced a class of Fourier-based sine-order circular kernels containing many well-known circular pdfs. García-Portugués et al. (2013b) considered the extension of (17) to $\mathbb{S}^d \times \mathbb{R}$. Amiri et al. (2017) transformed (17) into a sequentially updating estimator. Tsuruta and Sagae (2017a) showed that using a WC kernel instead of a vM kernel in (18) worsens the optimal asymptotic mean integrated squared error (AMISE) rate from $n^{-4/5}$ to $n^{-2/3}$, despite both kernels being second sine-order. This motivated Tsuruta and Sagae (2017b) to propose a class of p-th order kernels with an optimal AMISE rate of $n^{-2p/(2p+1)}$.

Bandwidth selection is crucial to KDE and hence was also addressed in most of the aforementioned contributions. Plug-in selectors as alternatives to cross-validation (CV) have received most attention. Taylor (2008) proposed the first plug-in selector for (18) by deriving the AMISE under the assumption that f is vM. The plug-in rule of Oliveira et al. (2012) employed the AMISE of Di Marzio et al. (2011), but used a two-component vM mixture in its curvature term. García-Portugués (2013) gave plug-in selectors for (17) using the AMISE and MISE for mixtures of vMF pdfs. Recently, Tsuruta and Sagae (2020) studied the convergence rates of direct plug-in and CV selectors for KDE on $\mathbb{S}^1$. Pham Ngoc (2019) proposed a bandwidth selector for (17) with a convergence rate of $n^{-2p/(2p+d)}$ for p-th order kernels.

Asymptotic results obtained for (17) include: CLTs for the integrated squared error of KDEs on $\mathbb{S}^d$ (Zhao and Wu 2001), and $\mathbb{S}^d \times \mathbb{R}$ and $\mathbb{S}^{d_1} \times \mathbb{S}^{d_2}$ (García-Portugués et al. 2015); lower bounds for asymptotic minimax risks (Klemelä 2003); laws for the iterated logarithm (Wang and Zhao 2001, 2003); and large and moderate deviations (Gao and Li 2010; Li 2014).

Convolutions on $\mathbb{S}^d$ are intimately related with KDE and are key to fast computation. They have been studied in Eğecioğlu and Srinivasan (2000), Dokmanic and Petrinovic (2010), and Le Bihan et al. (2016).

As alternatives to (17) and (18), Wang and Ma (2000) introduced a nearest-neighbour estimator of f and Park (2012, 2013) considered KDE via the tangent space of $\mathbb{S}^d$. Di Marzio et al. (2017) matched trigonometric moments of f with their smoothed sample versions to derive pdf estimators. As in high-order KDE, such estimators lower the bias and retain the variance order of (18), although negative values are possible. Using a different approach, Di Marzio et al. (2016b, 2018a) investigated local likelihood (Loader 1996) for pdfs on $[-\pi, \pi)^d$ by using local approximation of $\log f$. KDE based on the heat kernel on $\mathbb{S}^d$ (see Hartman and Watson 1974), the $d = 1$ case of which being the wrapped normal kernel, was applied in Zhang et al. (2019).

Extensions of (18) enable the construction of smoothed estimators for circular cdfs (Di Marzio et al. 2012b) and conditional pdfs (Di Marzio et al. 2016a). More generally, KDE has also been developed for compact Riemannian manifolds (Pelletier 2005; Henry and Rodriguez 2009), with inherent reduced tractability.

7.1.2 Series-based

An alternative approach to estimating a circular pdf is to use sample trigonometric moments as estimates of coefficients in its Fourier series expansion. Such estimates generally exhibit harmonic peaks and troughs and can be negative, although the latter defect can be circumvented by imposing constraints (Fernández-Durán 2004). Instead, periodic Bernstein polynomials might be considered. However, as Carnicero et al. (2018) have shown, imposing periodicity on such polynomials increases the error rate from $n^{-4/5}$ to $n^{-2/3}$. An interesting connection between Fourier-based estimation and (18) arises through the use of the WC kernel (Chaubey 2018).

Spherical harmonics (see, e.g., Dai and Xu 2013) extend Fourier orthogonal bases to $\mathbb{S}^d$ with increasing complexity as d grows. Hence, pdf estimation through spherical harmonics inherits both the advantages and disadvantages of Fourier series estimation on the circle. A compelling alternative are *needlets* (Narcowich et al. 2006; see also Marinucci et al. 2008), a class of spherical wavelets. Needlets build on spherical harmonics to form a *tight frame* on $L^2(\mathbb{S}^d)$ that is not a basis, as redundancy is allowed, but has superior localisation properties. Needlet coefficients can be estimated from sample spherical harmonic coefficients. Baldi et al. (2009a) approached adaptive pdf estimation on $\mathbb{S}^d$ by thresholding needlet coefficients, and Kueh (2012) studied the latter estimator under varying local pdf regularity. Like Fourier-based estimates, needlet-based pdf estimates can take negative values.

Circular deconvolution, i.e. the estimation of a pdf on $\mathbb{S}^1$ from noisy observations (see Sect. 5.1), has been tackled with increasing generality in Efromovich (1997), Comte and Taupin (2003), and Johannes and Schwarz (2013). Spherical deconvolution has been studied through spherical harmonics (Healy et al. 1998; Kim and Koo 2002; Kim et al. 2004) and needlets (Kerkyacharian et al. 2011).

7.1.3 Bayesian-based

Density estimation using Dirichlet process mixtures (DPMs) is a popular nonparametric Bayesian approach and has been employed with directional variables too. Lennox et al. (2009) provided a DPM model having sine bivariate vM distributions to model pairs of dihedral angles on $\mathbb{T}^2$. Straub et al. (2015) proposed a DPM model of Gaussian distributions in distinct tangent spaces to $\mathbb{S}^d$. DPM models with projected normal distributions have been advocated by Núñez-Antonio et al. (2018) on $\mathbb{S}^1$ and by Abraham et al. (2019) on $\mathbb{T}^d$. Density estimation through DPM on manifolds was addressed in Bhattacharya and Bhattacharya (2012, Chapter 13).

7.2 Regression estimation

7.2.1 Linear response

The Nadaraya–Watson estimator for linear-spherical regression is

$$\hat{m}(\boldsymbol{x}; h) = \frac{c_L(h)}{n\hat{f}(\boldsymbol{x}; h)} \sum_{j=1}^{n} Y_j L\left(\frac{1 - \boldsymbol{x}'\boldsymbol{X}_j}{h^2}\right), \quad \boldsymbol{x} \in \mathbb{S}^d, \tag{19}$$

which, for the vMF kernel and $\mathbb{S}^1$, reduces to

$$\hat{m}(\theta; \kappa) = \frac{\sum_{j=1}^{n} Y_j \exp\{\kappa \cos(\theta - \Theta_j)\}}{\sum_{j=1}^{n} \exp\{\kappa \cos(\theta - \Theta_j)\}}, \quad \theta \in [-\pi, \pi). \tag{20}$$

As an extension of (20), Di Marzio et al. (2009) introduced local polynomial regression for predictors on $\mathbb{T}^d$ through a sine term-based Taylor expansion. Their approach was extended further by Qin et al. (2011) to accommodate circular and multivariate predictors using product kernels, a broadly applicable approach to combine different predictors. Tsuruta and Sagae (2018) showed the different optimal error rates for the (second sine-order) WC and vM kernels.

On $\mathbb{S}^d$, (19) was considered by Wang et al. (2000) and Wang (2002) when deriving laws for iterated logarithm and exponential error bounds, respectively. Di Marzio et al. (2014) extended (20) to local polynomial regression using a Taylor expansion within the tangent-normal decomposition. García-Portugués et al. (2016) used a different Taylor expansion yielding a local linear estimator that, for $d = 1$, coincides with the Di Marzio et al. (2009) proposal. Di Marzio et al. (2019b) built on their construction in Di Marzio et al. (2014) to perform local polynomial logistic regression with a spherical predictor.

Monnier (2011) proposed needlet-based regression for a uniformly distributed predictor on $\mathbb{S}^d$ and Gaussian noise, whilst Lin (2019) weakened those assumptions and introduced regularisation on the needlet coefficients.

Thin-plate splines on $\mathbb{S}^d$ (Taijeron et al. 1994) offer an alternative smoothing approach to kernel methods. Such splines have been considered for improving brain conformal mapping to $\mathbb{S}^2$ (Zou et al. 2007). Kaufman et al. (2005) introduced circular Bayesian adaptive regression splines for modelling the firing rates of neurons activated by movements of a monkey's wrist. Quadratic B-splines on the circle were constructed in McMillan et al. (2013).

Related to regression for a $\mathbb{S}^1$ predictor, Hall et al. (2000), Hall and Yin (2003), and Genton and Hall (2007) studied the estimation of periodic functions over an (unwrapped) time domain. Klemelä (1999) considered the estimation of a function on $\mathbb{S}^d$ observed in Gaussian continuous time white noise.

7.2.2 Circular or spherical response

Boente and Fraiman (1991) considered estimators for $\mathbb{S}^d$-$\mathbb{R}^q$ regression based on locally weighted spherical means, with nearest-neighbour or Nadaraya–Watson weights. Their construction was generalised to local polynomial $\mathbb{S}^1$-$\mathbb{S}^1$ and $\mathbb{S}^1$-$\mathbb{R}$ (Di Marzio et al. 2013), $\mathbb{S}^1$-$\mathbb{R}^q$ (Meilán-Vila et al. 2020), and $\mathbb{S}^d$-$\mathbb{S}^q$ (Di Marzio et al. 2014) regression through local circular and spherical means. A novel approach to $\mathbb{S}^d$-$\mathbb{S}^d$ regression, based on local polynomial expansions of the *rotation* function, was advocated by Di Marzio et al. (2019c).

Quantile $\mathbb{S}^1$-$\mathbb{S}^1$ and $\mathbb{S}^1$-$\mathbb{R}$ regression was developed by Di Marzio et al. (2016c) through inversion of the conditional circular distribution and smoothing a circular check function.

From a Bayesian perspective, Scott (2011) estimated the regression function on $\mathbb{S}^2$ by imposing shrinkage priors on its needlet coefficients. Navarro et al. (2017) proposed multivariate generalised vM circular processes as a replacement for Gaussian processes in circular regression.

More generally, Cheng and Wu (2013) addressed linear-manifold regression through local linear regression on the tangent space, and Lin et al. (2017) gave an extrinsic Nadaraya–Watson estimator for manifold-linear regression.

8 Dimension reduction methods

8.1 Principal component analysis

8.1.1 General manifolds

Principal component analysis (PCA) for data on a Riemannian manifold $\mathcal{M}$ of dimension d, such as $\mathbb{S}^d$ or $\mathbb{T}^d$, has received considerable attention lately. Approaches to manifold PCA can be classified using two broad dichotomies: (i) *extrinsic* (based on tangent space) versus *intrinsic* (geodesic-based); (ii) *forward* (sequential computation of the j-th principal component, $j = 1, \ldots, d$) versus *backward* (computation of a sequence of nested subspaces of decreasing dimension within $\mathcal{M}$ based on constraints; Damon and Marron (2014)). Huckemann et al. (2010) gave a detailed review of the topic, and Marron and Alonso (2014) and Pennec (2018) more recent overviews.

Fletcher et al. (2004) introduced principal geodesic analysis (PGA) as an analogue of PCA in symmetric spaces such as $\mathbb{S}^d$ and $\mathbb{T}^d$. It is centred upon the intrinsic sample mean on $\mathcal{M}$, $\hat{\boldsymbol{\mu}}$, and defines the first principal geodesic as the one passing through $\hat{\boldsymbol{\mu}}$ that minimises the sum of squared intrinsic residuals. Other principal geodesics are obtained sequentially by imposing orthogonality at $\hat{\boldsymbol{\mu}}$. PGA involves a complex optimisation process, only solved later by Sommer et al. (2014). This complexity led Fletcher et al. (2004) to propose tangent PCA (tPCA) as an approximation. tPCA performs PCA with the log-mapped data onto the tangent plane at $\hat{\boldsymbol{\mu}}$ and then obtains the principal geodesics on $\mathcal{M}$ spanned by the tangent principal directions. When $\mathcal{M} = \mathbb{S}^2$, the principal components of PGA and tPCA are great circles that pass through $\hat{\boldsymbol{\mu}}$.

Two limitations of PGA are exemplified on $\mathbb{S}^2$: (a) great circles are forced to cross at $\hat{\boldsymbol{\mu}}$; (b) great circles are unable to describe certain forms of variation (see Sect. 8.1.2). Huckemann and Ziezold (2006) tackled (a) by introducing geodesic PCA (GPCA) for Riemannian manifolds, a forward-type method with a backward shift that locates a data centre $\tilde{\boldsymbol{\mu}}$ *after* finding the best fitting geodesic. The other components cross orthogonally at $\tilde{\boldsymbol{\mu}}$, a restriction circumvented by horizontal component analysis (Sommer 2013). Curry et al. (2019) recently proposed principal symmetric space approximation (PSSA), which considers totally geodesic subspaces (great subspheres, on $\mathbb{S}^d$) and is computationally tractable on certain manifolds.

A non-geodesic approach to PCA on $\mathcal{M}$ is barycentric subspace analysis (Pennec 2018). It considers k-dimensional *affine spans* (great subspheres if $\mathcal{M} = \mathbb{S}^d$) spanned by $k+1$ $\mathcal{M}$-affinely independent points, whose successive addition/removal yields a forward/backward-type sequence of nested subspaces.

Zhang and Fletcher (2013) proposed probabilistic PGA, in which the normal distribution used in probabilistic PCA (Tipping and Bishop 1999) is replaced by what the authors refer to as the *Riemannian normal distribution*, with pdf $f(\boldsymbol{x}; \boldsymbol{\mu}, \sigma^2) \propto \exp\{-d_g(\boldsymbol{x}, \boldsymbol{\mu})^2/(2\sigma^2)\}$, where $\boldsymbol{x}, \boldsymbol{\mu} \in \mathcal{M}$ and d_g is the intrinsic distance on $\mathcal{M}$ ($d_g(\boldsymbol{x}, \boldsymbol{\mu}) = \cos^{-1}(\boldsymbol{x}'\boldsymbol{\mu})$ if $\mathcal{M} = \mathbb{S}^d$). Sommer (2019) advocated an alternative to PGA based on an anisotropic normal distribution over $\mathcal{M}$, generated from the marginal distributions of a diffusion process on $\mathcal{M}$ with a constant infinitesimal covariance.

The previous approaches assume a parametric form for the first principal curve on $\mathcal{M}$. Instead, the *principal flow* of Panaretos et al. (2014) is defined as the curve of maximal data variation on $\mathcal{M}$ that, starting at $\hat{\boldsymbol{\mu}}$, is tangential to the vector field formed by the first eigenvector of the local tangent covariance matrix. Higher-order principal flows, which are always curves, are defined analogously.

Dai and Müller (2018) adapted tPCA for functional data on $\mathcal{M}$ (e.g. flight trajectories on $\mathbb{S}^2$) by replacing PCA by functional PCA on the tangent plane.

Nonparametric inference on backward nested principal component subspaces, generalising the result of Anderson (1963) on asymptotic inference for classical PCA, has been provided by Huckemann and Eltzner (2018).

8.1.2 Methods for spherical data

In relation to limitation (b) of Sect. 8.1.1, and for the specific case of $\mathbb{S}^2$, Jung et al. (2011) advocated principal arc analysis (PAA), a non-geodesic approach designed to improve the flexibility of GPCA. PAA employs small circles on $\mathbb{S}^2$ as the primary modes of data variation, an idea generalised to $\mathbb{S}^d$ by Jung et al. (2012) as principal nested spheres (PNS). By iteratively performing a series of tangent-normal decompositions on $\mathbb{S}^d$, PNS is a backward-type approach that produces a sequence of subspheres isomorphic to $\mathbb{S}^j$, $j = d-1, \ldots, 1$, that none of the methods in Sect. 8.1.1 are able to match in terms of flexibility.

Despite the generality of the approaches in Sect. 8.1.1, the success of PNS highlights the advantages of focusing on specific manifolds, such as $\mathbb{S}^d$ or $\mathbb{T}^d$, and exploiting their peculiarities so as to obtain more informative methods. Other examples of the benefits of specificity include PAA on direct product manifolds such as $(\mathbb{R}^3 \times \mathbb{R}^+ \times \mathbb{S}^2 \times \mathbb{S}^2)^m$, introduced by Jung et al. (2011), and composite PNS for skeletal models, proposed by Pizer et al. (2013).

8.1.3 Methods for toroidal data

Torus-specific PCA proposals have been stimulated by the need to analyse dihedral angles in bioinformatics, and the inapplicability of most of the methods in Sect. 8.1.1 due to the pathological behaviour of geodesics on $\mathbb{T}^d$.

The majority of toroidal PCA methods resort to some sort of transformation prior to applying classical PCA. Mu et al. (2005) proposed dihedral PCA (dPCA) by mapping angles from $[-\pi, \pi)^d$ to $\mathbb{T}^d$ and then performing PCA. Complex dPCA (Altis et al. 2007) performs PCA on the complex representation of angles. Riccardi et al. (2009) proposed angular PCA (aPCA), based on applying PCA to toroidal data centred on their circular means. Kent and Mardia (2009) gave a trigonometric moment characterisation of the covariance matrix in a wrapped normal model on $\mathbb{T}^d$, facilitating PCA on it. Nodehi et al. (2015) applied PGA on $\mathbb{T}^d$ in what they called dPGA. The latter two approaches yield principal directions that almost surely wrap around infinitely. With regard to this issue, Kent and Mardia (2015) discussed desirable properties for principal component curves on $[-\pi, \pi)^2$. Sittel et al. (2017) introduced a variation on aPCA, called dPCA+, that shifts each variable so that $-\pi \equiv \pi$ is located at the lowest pdf region for minimising the distortion when PCA is applied. Sargsyan et al. (2012, 2015) used a non-injective mapping from $\mathbb{T}^d$ to $\mathbb{S}^d$ that equates toroidal angles in $[-\pi, \pi)^d$ to hyperspherical coordinate angles, even though the latter are defined in $[0, \pi]^{d-1} \times [-\pi, \pi)$, then applied GPCA.

A better-grounded approach to torus PCA is T-PCA (Eltzner et al. 2018), which maps $\mathbb{T}^d$ to $\mathbb{S}^d$ with a deformation, which, for $d = 2$, corresponds to cutting $\mathbb{T}^2$ at a data-driven point to form a cylinder, contracting the circles at its ends to single points, and reconnecting at those points. Principal nested deformed spheres (Eltzner et al. 2015) is an extension of the T-PCA approach to data on a polysphere $\mathbb{S}^{d_1} \times \cdots \times \mathbb{S}^{d_m}$. PSSA can also be applied to toroidal data, with geodesics on $\mathbb{T}^d$ identified using model selection.

8.2 Other dimension reduction methods

Nonlinear dimension reduction methods for Euclidean data have also been adapted to the directional context. Lunga and Ersoy (2013) modified t-stochastic neighbour embedding (t-SNE) of van der Maaten and Hinton (2008) to obtain a dimension reduction method, from $\mathbb{S}^d$ to $\mathbb{S}^q$, $q \ll d$, using neighbourhoods formed by the extension of the WC distribution to $\mathbb{S}^d$ in Kato (2009). Wang and Wang (2016) proposed a modification of t-SNE with vMF neighbourhoods.

Wilson et al. (2014) revisited multidimensional scaling on $\mathbb{S}^d$, proposing a new approach that obviated the minimisation of stress functions based on spherical distances inherent in former approaches. They used their approach to map textures of 3D objects onto spheres (Elad et al. 2005), and model normalised time-warping functions (Veeraraghavan et al. 2009). Lu et al. (2019) adapted t-SNE for dimension reduction from $\mathbb{R}^d$ to $\mathbb{S}^q$, $q \ll d$, highlighting the benefits of the clusterings obtained on the spherical geometry. Note that these transformations of multivariate data into spherical data, termed *spherical embeddings*, call for the use of directional statistics with data which were not originally directions.

9 Classification and clustering

9.1 Classification

SenGupta and Roy (2005) introduced a discrimination rule based on the chordal distance between a new circular observation and those from two known circular populations. More recently, Di Marzio et al. (2018b) considered nonparametric circular classification based on KDE and local logistic regression. Pandolfo et al. (2018a) studied the depth-versus-depth classifier for circular data. Leguey et al. (2019a) proposed Bayesian classification algorithms for WC-distributed circular predictors.

SenGupta and Ugwuowo (2011) proposed generalised likelihood-ratio tests for classifying toroidal and cylindrical data into two populations when one of the misclassification probabilities is assumed to be known. Fernandes and Cardoso (2016) developed a logistic classifier for use with circular and linear predictors.

Classification rules for data on $\mathbb{S}^d$ from Watson and vMF populations were developed by Figueiredo and Gomes (2006) and Figueiredo (2009), respectively. Bhattacharya and Dunson (2012) proposed a Dirichlet process mixture model classifier comprising vMF kernels. López-Cruz et al. (2015) considered the naive Bayes classifier with vMF-related conditional distributions of directional predictors. Techniques for the classification of image textures, based on multiresolution directional filters, were proposed by Kim and So (2018). Di Marzio et al. (2019a) considered KDE-based nonparametric classification. The cosine depth *distribution* classifier was introduced in Demni et al. (2019). A comparison of different classification rules on $\mathbb{S}^2$ was performed by Tsagris and Alenazi (2019).

9.2 Clustering

The development of clustering methods for directional data has been a major research theme lately, particularly amongst the machine learning community.

The two most popular approaches to clustering data on $\mathbb{S}^d$ are spherical k-means and the use of (finite) mixture models with vMF components, the vMF extension of (6). Both have often been applied after projecting data in $\mathbb{R}^{d+1}$ to $\mathbb{S}^d$. The spherical k-means approach (Dhillon and Modha 2001) maximises the cosine similarity measure $\sum_{j=1}^{n} \boldsymbol{X}'_j \boldsymbol{c}_{(j)}$ between the sample $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$ and k centroids $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_k \in \mathbb{S}^d$, where $\boldsymbol{c}_{(j)}$ is the centroid of the cluster containing $\boldsymbol{X}_j$.

Dhillon and Sra (2003) and Banerjee et al. (2003, 2005) gave expectation-maximisation (EM) algorithms for fitting vMF mixtures. Such mixtures accommodate spherical k-means as a particular limiting case (Banerjee et al. 2005). Other approaches to fitting vMF mixtures include those of: Yang and Pan (1997), based on embedding fuzzy partitions in the mixtures; Taghia et al. (2014), a Bayesian approach employing variational inference; Gopal and Yang (2014), who proposed Bayesian graphical modelling approaches based on variational inference and collapsed Gibbs sampling; Qiu et al. (2015), which used a new information criterion to determine the number of clusters; Kasarapu and Allison (2015), based on the Bayesian minimum message length criterion to determine the optimal number of components; Mash'al and Hosseini (2015), a k-means++ method for identifying favourable mixture starting values; and Salah and Nadif (2019), a co-clustering approach based on diagonal block mixtures of vMF distributions. Mixtures with vMF components have been employed to model data from photometry (Hara et al. 2008), text mining (Banerjee et al. 2009), speech recognition (Tang et al. 2009), radiation therapy (Bangert et al. 2010), pattern recognition (Calderara et al. 2011), multichannel array processing (Costa et al. 2014), and collaborative filtering (Salah and Nadif 2017). vMF mixtures have been used to cluster pdf objects based on their spherical-valued wavelet coefficients (Montanari and Calò 2013).

Clustering approaches based on mixtures with other types of directional distributions have also been advocated. To increase cluster shape modelling flexibility on $\mathbb{S}^d$, Peel et al. (2001) used mixtures of Kent distributions, whereas Dortet-Bernadet and Wicker (2008) proposed ones with inverse stereographic projections of multivariate normal distributions. Mixture models with wrapped normal components were investigated by Agiomyrgiannakis and Stylianou (2009), and used to cluster X-ray position data in Abraham et al. (2013). Bayesian approaches to fitting projected normal mixtures have been proposed by Wang and Gelfand (2014) and Rodríguez et al. (2020), and for general projected normal mixtures by Hernandez-Stumpfhauser et al. (2017). Franke et al. (2016) developed an EM algorithm to fit the latter type of mixtures to data on $\mathbb{S}^2$. For data on $\mathbb{T}^d$, Mardia et al. (2012) proposed mixtures of concentrated sine multivariate vM components, with approximated normalising constants, to cluster dihedral angles of an amino acid. For cylindrical data, Lagona and Picone (2011, 2012) developed latent-class mixture models to cluster incomplete environmental data.

Mixture model-based approaches for clustering axial data have also been proposed. Bijral et al. (2007), Souden et al. (2013), and Sra and Karp (2013) developed EM-based algorithms to fit mixtures of Watson distributions. Hasnat et al. (2014) provided an alternative approach to fitting such models based on Bregman divergence. A clustering approach based on mixtures of Bingham distributions was developed by Yamaji and Sato (2011).

Variations of spherical k-means include the spherical fuzzy and possibilistic c-means proposed by Kesemen et al. (2016) and Benjamin et al. (2019), respectively, and the adaptations by Maitra and Ramler (2010) for computational efficiency. On $\mathbb{S}^1$, Baragona (2003) further investigated an alternative partitioning based on the statistic of Lund (1999). A nonparametric alternative to k-means is kernel mean shift clustering, introduced on $\mathbb{S}^d$ by Oba et al. (2005). It was then extended by Cetingul and Vidal (2009) to $\mathbb{S}^d$ and other specific manifolds, using intrinsic and extrinsic perspectives, and later reintroduced on $\mathbb{S}^d$ with minor variations (Chang-Chien et al. 2010; Yang et al. 2014). A modification of kernel mean shift that uses time-varying bandwidths was adapted to spherical data by Hung et al. (2015).

10 Modelling serial dependence

10.1 Discrete-time processes

Let $\{\Theta_t\}_{t=1,2,\ldots}$ denote a discrete-time circular process, and $\{\theta_t\}_{t=1,2,\ldots,n}$ a corresponding circular time series. Mardia and Jupp (1999) provide a summary of the projected normal, wrapped, linked autoregressive moving average (ARMA), and circular autoregressive (CAR) models for circular time series considered in Fisher and Lee (1994) and Fisher (1993, Section 7.2).

If $\{(X_t, Y_t)\}_{t=1,2,\ldots}$ is a stationary bivariate normal process, then $\{\Theta_t\}_{t=1,2,\ldots}$, where $\Theta_t = \operatorname{Arg}(X_t + \mathrm{i}Y_t)$, is a projected normal process. If $\{X_t\}_{t=1,2,\ldots}$ is a process on $\mathbb{R}$, then $\{\Theta_t\}_{t=1,2,\ldots}$, where $\Theta_t = X_t \pmod{2\pi}$, is the corresponding wrapped process. The wrapped AR processes of Breckling (1989) provide an example. A linked process is defined through $\Theta_t = \mu + g(X_t)$, where $\mu \in [-\pi, \pi)$ and g is a link function defined in Fisher (1993, Section 7.2.4) as a mapping from $\mathbb{R}$ to $(-\pi, \pi)$. A linked ARMA(p, q) process is obtained if $X_t = g^{-1}(\Theta_t)$ is an ARMA(p, q) process. A CAR process is one for which $\Theta_t | \Theta_{t-1} = \theta_{t-1}, \ldots, \Theta_{t-p} = \theta_{t-p}$ is vM-distributed with mean direction

$$\mu_t = \mu + g[\alpha_1 g^{-1}(\theta_{t-1} - \mu) + \cdots + \alpha_p g^{-1}(\theta_{t-p} - \mu)]$$

and concentration parameter κ, where $\mu \in [-\pi, \pi)$ and $\alpha_1, \ldots, \alpha_p \in \mathbb{R}$. Artes and Toloi (2009) proposed an extension of the CAR model with covariates. Processes with distributions other than the vM can be defined analogously to CAR. Erdem and Shi (2011) consider four ARMA-based approaches to the short-term forecasting of wind speed and direction.

Markov models can be constructed using a transition pdf, $f_{\Theta_t|\Theta_{t-1}=\theta_{t-1}}$, derived from a bivariate circular pdf. Hughes (2007) used this approach, dating back to Wehrly and Johnson (1980), to obtain stationary Markov processes from the sine and cosine bivariate vM models referred to in Sect. 3.3, and the Möbius transformation-based regression model of Downs and Mardia (2002). Similarly, Kato (2010) employed the regression model of Kato et al. (2008) to derive a stationary Markov process. Yeh et al. (2013) proposed a circular Markov process based on a transition pdf that belongs to the class of generalised vM distributions referred to in Sect. 3.1. Le Bihan et al. (2016) studied Markov processes with rotationally symmetric transition pdfs on $\mathbb{S}^d$, specifically analysing the vMF case.

In a paper that has stimulated much research into the modelling of time series, spatial, and spatio-temporal data (see Sect. 11), Holzmann et al. (2006) introduced hidden Markov models (HMMs) (Zucchini et al. 2016) for circular as well as cylindrical time series. Such models offer considerable flexibility in their serial dependence properties and use mixtures of varying distributions to model different underlying regimes. More specifically, Holzmann et al. (2006) considered circular HMMs with state-dependent vM, wrapped normal or WC distributions, and marginal distributions that are mixtures of each. They also proposed a cylindrical HMM with state-dependent vM distributions for the circular component. Bulla et al. (2012) extended this approach to develop a multivariate hidden Markov model for bivariate circular and bivariate linear data with sine bivariate vM and bivariate skew-normal pdfs. An HMM for toroidal time series using sine bivariate vM pdfs and allowing for missing observations was proposed in Lagona and Picone (2013). Hokimoto and Shimizu (2014) developed a model incorporating a non-homogeneous HMM with cylindrical covariates. Ailliot et al. (2015) proposed Markov-switching autoregressive models based on a non-homogeneous hidden Markov chain for circular time series with vM innovations. HMMs for use with cylindrical time series have been proposed by Lagona et al. (2015a), Mastrantonio et al. (2015), and Mastrantonio and Calise (2016). Mastrantonio et al. (2015) considered a projected normal-based extension of the model of Bulla et al. (2012) that allows for conditional correlation between the circular and linear variables. The Dirichlet process mixture model of Mastrantonio and Calise (2016) is designed for use with discrete cylindrical variables.

HMMs with toroidal components have also been employed in protein structure modelling. Boomsma et al. (2008) proposed one with cosine bivariate vM distributions to model the pairs of dihedral angles describing protein backbones. Lennox et al. (2010) considered a Dirichlet process mixture of HMMs with sine bivariate vM distributions for the dihedral angles. Golden et al. (2017) developed an HMM to model the evolution of pairs of proteins, with bivariate wrapped normal diffusions (García-Portugués et al. 2019) used to describe dihedral angle evolution.

Recently, Mazumder and Bhattacharya (2017) proposed a state-space model for circular time series, with a circular latent process, based on wrapped Gaussian processes (Mazumder and Bhattacharya 2016). Beran and Ghosh (2020) introduced a class of linked processes for circular time series, allowing for long-range dependence, obtained by transforming Gaussian processes. Hokimoto and Shimizu (2008) extended the multiple regression model of Johnson and Wehrly (1978) to develop a time series model for data on $\mathbb{T}^d \times \mathbb{R}^q$.

Nonparametric kernel-based trend estimation in circular time series was tackled in Di Marzio et al. (2012a). Beran (2016) proposed a class of nonparametric normalised symmetric linear estimators for the trend of $\mathbb{S}^d$-valued time series.

The modelling of longitudinal data on smooth Riemannian manifolds, such as bird-migration trajectories and hurricane paths on $\mathbb{S}^2$, has been addressed by Su et al. (2014) and Zhang et al. (2018b).

10.2 Continuous-time processes

Continuous-time processes involving directional data have received considerably less attention than discrete-time processes. Hill and Häder (1997) introduced a random walk model whose reorientation process follows a vM diffusion (Kent 1975). Several variations of such random walk models have been developed for biological purposes: see Codling et al. (2008) for a review. García-Portugués et al. (2019) proposed various Langevin diffusions on the torus that can be viewed as analogues of Ornstein–Uhlenbeck processes and studied likelihood-based estimation approaches for them. Sommer (2019) considered anisotropic diffusion processes on Riemannian manifolds, and Jensen et al. (2019) simulated diffusion bridges on $\mathbb{T}^d$. Ball et al. (2008) introduced Brownian motion and Ornstein–Uhlenbeck processes on the shape space of $\mathbb{R}^2$.

Kurz et al. (2019) provided an overview of various recursive filtering algorithms involving a variety of circular, toroidal, and spherical distributions. Analogues of the Kalman filter based on the vMF and Bingham distributions on $\mathbb{S}^d$ were proposed by Chiuso and Picci (1998) and Kurz et al. (2014), respectively. Filtering using the wrapped normal distribution on $\mathbb{S}^1$ was considered by Traa and Smaragdis (2013). Pitt and Shephard (1999) proposed auxiliary particle filter methods and applied them to a ship tracking problem modelled using a WC process. Recently, the use of various types of filtering in the tracking of space debris has received considerable attention (Kent et al. 2020; Bhattacharjee 2020, and references therein).

11 Spatial and spatio-temporal modelling

The modelling of spatial and space-time directional data is one of the branches of directional statistics that has experienced particularly important advances in recent years. Many contributions involve fitting hierarchical Bayesian spatial models to meteorological data using MCMC methods. As an approach to modelling hurricane winds, Modlin et al. (2012) proposed a Bayesian hierarchical model for vector fields featuring a wrapped normal conditional autoregressive model. Jona-Lasinio et al. (2012) formulated a similar model incorporating, instead, a wrapped Gaussian spatial process to model wave directions at different sea locations. Using a different perspective, Wang (2013), Wang and Gelfand (2014), and Wang et al. (2015) considered models based on projected normal processes for modelling wave direction and height at different sites.

As Lagona et al. (2015b) has pointed out, such Bayesian hierarchical models require specific assumptions on the prior distributions of the parameters of interest and ad hoc MCMC for fitting. Instead, Lagona et al. (2015b) developed an HMM to model the temporal evolution of the sea surface in terms of time-varying circular-linear patterns that arise through latent environmental conditions. Fitting is performed using a pseudo-likelihood approach.

Extending the wrapped normal-based Bayesian approach of Jona-Lasinio et al. (2012), Mastrantonio et al. (2016a) introduced a wrapped skew-normal process, for use with spatio-temporal circular data, which is capable of modelling asymmetric marginal distributions. Mastrantonio et al. (2016b) extended and compared the processes of Jona-Lasinio et al. (2012) and Wang and Gelfand (2014) to the spatio-temporal setting by introducing space-time dependence and space- and time-varying covariate information.

In Lagona and Picone (2016) and Ranalli et al. (2018), hidden Markov random field models were proposed for the analysis of cylindrical spatial series, enabling segmentation of latent environmental conditions. Jona-Lasinio et al. (2018) and Lagona (2018) provided overviews of many of the developments discussed above. Ameijeiras-Alonso et al. (2019b) extended the approach of Ranalli et al. (2018) to develop a hidden Markov random field for the spatial segmentation of wildfires, using a mixture of Kato and Jones (2015) pdfs with parameters varying according to a latent nonhomogeneous Potts model.

Next, we consider models for animal orientation data based on random walks and HMMs that provide alternatives to those in Sect. 6.2.2. Morales et al. (2004) proposed a Bayesian approach to fitting multiple random walks to animal movement data with paths composed of random step lengths and turning angles. Each step and turn is assigned to a random walk characteristic of a hidden behavioural state. A similar approach was proposed by McClintock et al. (2012), with movement paths considered to be movement strategies between which animals switch in response to environmental factors. The authors combined a variety of methodologies to develop a suite of models based on biased and correlated random walks that allow for different forms of movement. Nicosia et al. (2017) proposed a hidden-state random walk model in which a circular-linear process models the direction and distance between consecutive positions of an animal, and the hidden states describe the animal's behaviour.

Random fields on $\mathbb{S}^2$ are discussed in depth in Marinucci and Peccati (2011). Amongst many other advances, their monograph analyses recent high-frequency limit results and tests for Gaussianity and isotropy of scalar-valued random fields, and considers applications in the analysis of the cosmic microwave background. Recent research into isotropic Gaussian random fields on $\mathbb{S}^2$ has developed CLTs for functionals of needlet coefficients (Baldi et al. 2009b), limit results for the first Minkowski functional (Leonenko and Ruiz-Medina 2018), isotropy tests based on spherical harmonics (Sahoo et al. 2019), and tests for the detection of local maxima on isotropic fields (Cheng et al. 2020). The construction of valid covariance functions on $\mathbb{S}^d$, for use in geostatistics, has been summarised in the excellent overview of Gneiting (2013). New covariance functions on $\mathbb{S}^d$ include the spatio-temporal covariance functions of Porcu et al. (2016) and the matrix-valued covariance functions of Guella et al. (2018); see also the review by Porcu et al. (2020). A review of advances in the construction of covariance functions and process models on $\mathbb{S}^2$ was given in Jeong et al. (2017).

Irwin et al. (2002) gave a review of spatio-temporal nonlinear filtering and illustrated the use of cylindrical filtering in the analysis of battlespace data.

12 Other topics

12.1 Statistical depth

Agostinelli and Romanazzi (2013) studied, mainly for $\mathbb{S}^1$ and $\mathbb{S}^2$, the angular simplicial depth of Liu and Singh (1992) and the angular Tukey depth of Small (1987). Within the class of rotationally symmetric distributions on $\mathbb{S}^d$, Ley et al. (2014) defined a depth based on the quantiles of the sample projections onto the mean direction. Pandolfo et al. (2018b) introduced computationally tractable distance-based depths on $\mathbb{S}^d$, illustrating their use in location estimation and classification. A nonparametric approach to constructing tolerance regions for spherical data was proposed by Mushkudiani (2002).

12.2 Design and analysis of experiments

Otieno and Anderson-Cook (2012) provided an overview of the design of experiments involving directional variables, and methods available for analysing the data obtained from them. Recently, optimal designs for linear-spherical regression, based on Fourier series and spherical harmonics, have been established for $\mathbb{S}^1$ by Dette and Melas (2003), for $\mathbb{S}^2$ by Dette et al. (2005) and Dette and Wiens (2009), and for $\mathbb{S}^d$, with $d > 2$, by Dette et al. (2019).

12.3 Order-restricted analysis

The random-periods model (RPM) of Liu et al. (2004) is a nonlinear regression model used to estimate the phase angles of periodically expressed genes. Rueda et al. (2009) developed circular isotonic regression estimation to infer the relative order of phase angles from the unconstrained estimates of the RPM. Fernández et al. (2012) proposed a test for a specified ordering of phase angles assuming the unconstrained estimators of the RPM to be vM-distributed. Barragán et al. (2015) developed methods for estimating and testing for a common ordering of phase angles across multiple experiments. A review of such developments was provided by Rueda et al. (2015). Subsequently, Rueda et al. (2016) proposed a piecewise circular regression model for the relationship between the phase angles of cell-cycle genes in two species with differing periods, and Barragán et al. (2017) considered the problem of aggregating different circular orders for the peak expressions of genes coming from heterogeneous datasets. Recently, Larriba et al. (2020) proposed a circular signal plus error model for identifying components of systems displaying rhythmic temporal patterns.

Independently of these developments, Klugkist et al. (2012), Baayen et al. (2012), and Baayen and Klugkist (2014) proposed ANOVA tests under order restrictions on the mean directions of vM distributions.

12.4 Outlier detection

New tests for detecting outliers in circular data, based on the circular distance (1), sums of such distances, and gaps, were introduced and compared with existing procedures in a series of papers referred to by Mahmood et al. (2017). Sau and Rodriguez (2018) developed a minimum distance approach to estimating the parameters of spherical models that provides an outlier detection tool. Outlier detection tests for cylindrical, simple circular regression, and circular time series data were proposed in Sadikon et al. (2019), Abuzaid et al. (2013), and Abuzaid et al. (2014), respectively.

Eigenvalue, likelihood-ratio, and geodesic distance-based tests for detecting outliers in axial data from an assumed underlying Watson distribution were developed in Figueiredo and Gomes (2005), Figueiredo (2007), and Barros et al. (2017).

12.5 Compositional data analysis

Compositional data analysis is used when the data under consideration are vectors of nonnegative proportions summing to one. The most popular approach to analysing such data is that of Aitchison (1986). However, an alternative approach, based on the square-root transformation from a unit simplex to $\mathbb{S}^d$, was discussed in Stephens (1982). Recently, that approach has been further developed by Scealy and Welsh (2011, 2014a, b, 2017). The relationship between compositional and directional data was further exploited by Cuesta-Albertos et al. (2009) in the context of testing for uniformity.

13 Software

Historically, a major impediment to the application of directional statistics was a lack of software implementing the methodology particular to it. In recent years, the advent of the R statistical computing environment (R Core Team 2020) and its ecosystem of contributed packages has partially addressed that paucity. An overview of many such packages was given by Pewsey (2018). Relevant packages written in other languages include `CircStat` (Berens 2009) and `PyCircStat` (Sinz et al. 2018) for data on $\mathbb{S}^1$, `libDirectional` (Kurz et al. 2019) for data on $\mathbb{T}^d$ and $\mathbb{S}^d$, `Mocapy++` for constructing probabilistic models of biomolecular structure (Paluszewski and Hamelryck 2010), and the promising `geomstats` (Miolane et al. 2020) for manifold-valued data.

13.1 General-purpose packages

There are two main R packages designed for use with directional data: `circular` (Agostinelli and Lund 2017) and `Directional` (Tsagris et al. 2020). Both include functions for the analysis of data on $\mathbb{S}^1$, $\mathbb{T}^2$, and $\mathbb{S}^1 \times \mathbb{R}$. `Directional` also has routines for data on $\mathbb{S}^d$.

For data on $\mathbb{S}^1$, the circular package has functions for: descriptive statistics; KDE; pdf evaluation, simulation, and estimation for a range of classical and more recently proposed circular distributions; tests for uniformity, homogeneity, goodness-of-fit, and change points; and one-way ANOVA. It also has functions for $\mathbb{S}^1$-$\mathbb{S}^1$ and $\mathbb{S}^1$-$\mathbb{R}$ regression and includes a variety of datasets. Many of circular's capabilities were illustrated in Pewsey et al. (2013). The latter's companion workspace, CircStatsInR, includes over 150 routines for techniques not implemented in circular.

Amongst its more specific capabilities, Directional implements techniques on $\mathbb{S}^d$ for: descriptive statistics; spherical data visualisation; constructing convenient rotation matrices and transformations; KDE; pdf computation, simulation, and ML estimation for various spherical distributions; and $\mathbb{S}^d$-$\mathbb{S}^d$ correlation and regression, ANOVA, classification, and clustering.

13.2 More specific R packages

Here we provide an overview of more specific R packages and their functionality, following the order used to present themes in the previous sections.

Various graphical representations for data on $\mathbb{S}^1$ are supported in bpDir (Buttarazzi 2020), season (Barnett and Baker 2020), and bReeze (Graul and Poppinga 2018). Visualisation of data on $\mathbb{S}^2$ is facilitated by the outstanding rgl (Adler et al. 2020) and plot3D (Soetaert 2019) packages.

Efficient modelling with vMF mixtures on $\mathbb{S}^d$ is implemented in movMF (Hornik and Grün 2014). Several mixture models can be fitted using Bayesian methods to data on $\mathbb{S}^1$ and $\mathbb{T}^2$ with BAMBI (Chakraborty and Wong 2019). Nonnegative trigonometric sums can be fitted to data on $\mathbb{T}^d$ and $\mathbb{S}^2$ using CircNNTSR (Fernández-Durán and Gregorio-Domínguez 2016).

Tests for uniformity and rotational symmetry on $\mathbb{S}^d$ are available in sphunif (García-Portugués and Verdebout 2020) and rotasym (García-Portugués et al. 2020c), respectively.

Bayesian projected normal regression models for data on $\mathbb{S}^1$ are implemented in bpnreg (Cremers 2020). Also for data on $\mathbb{S}^1$, nonparametric kernel methods for density and regression estimation are available in NPCirc (Oliveira et al. 2014). KDE and bandwidth selection on $\mathbb{S}^d$ are supported in DirStats (García-Portugués 2020a). Nonparametric $\mathbb{S}^d$-$\mathbb{S}^d$ regression is implemented in nprotreg (Taylor et al. 2018). Smoothing splines on $\mathbb{S}^2$ are supported in mgcv (Wood 2017).

Principal nested spheres and spherical k-means clustering can be performed with shapes (Dryden 2019) and skmeans (Hornik et al. 2012), respectively.

Markov switching autoregressive models with vM innovations are implemented in NHMSAR (Monbet 2020). Animal orientation data can be analysed using CircMLE (Fitak and Johnsen 2017), FLightR (Rakhimberdiev et al. 2017), move (Kranstauber et al. 2020), and moveHMM (Michelot et al. 2016). Tools for toroidal diffusions are provided in sdetorus (García-Portugués 2020b).

Bayesian methods for fitting spatial and spatio-temporal models to circular data are implemented in CircSpaceTime (Jona Lasinio et al. 2020). Spherical random fields can be analysed using `RandomFields` (Schlather et al. 2015). Routines for the management and analysis of cosmic microwave background data on $\mathbb{S}^2$ are available in `rcosmo` (Fryer et al. 2020).

Methods for analysing data on $\mathbb{S}^1$ under order restrictions are supported by `isocir` (Barragán et al. 2013). Outlier detection methods for $\mathbb{S}^1$-$\mathbb{S}^1$ regression are available in `CircOutlier` (Ghazanfarihesari and Sarmad 2016). Depths on $\mathbb{S}^1$ and $\mathbb{S}^2$ can be computed using `depth` (Genest et al. 2019).

Intrinsic means and fundamental geodesic tools for $\mathbb{S}^d$ and other manifolds are available in `RiemBase` (You 2020).

14 Conclusions and future developments

We hope that the previous sections provide both seasoned and neophyte researchers with a concise, comprehensive, and useful overview of the widespread developments in directional statistics that have taken place over the last two decades. As often happens in research, most of those developments evolved in an uncoordinated way through the efforts of individuals and research groups working independently of one another. Given this background, predicting how the field might develop over the next 20 years is essentially impossible. That said, the further development of models with greater flexibility, techniques for high-dimensional and complex directional data involving combinations of different data types, as well as Bayesian, nonparametric, and resampling methods, would appear highly probable in the short term as such developments would be consistent with current trends. More generally, progress in all the areas covered in the previous sections is certainly possible and will no doubt evolve through responses to interesting new applications and the exigencies of the Riemannian supports of directional data, often incorporating appropriate adaptations of methodologies from other fields of statistics. The development of software to implement new techniques will continue to be crucial to the wider and proper application of directional statistics.

Supplementary Materials

The BibTeX file `DirectionalStats.bib` includes entries for over 1700 references related to directional statistics. Future updated versions of it will be available at https://github.com/egarpor/DirectionalStatsBib. We hope researchers in the field will find this resource useful.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11749-021-00759-x.

References

Abe T, Ley C (2017) A tractable, parsimonious and flexible model for cylindrical data, with applications. Econom Stat 4:91–104

Abe T, Pewsey A (2011) Sine-skewed circular distributions. Stat Pap 52(3):683–707

Abe T, Shimatani IK (2018) Cylindrical distributions and their applications to biological data. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 163–185

Abe T, Shimizu K, Pewsey A (2010) Symmetric unimodal models for directional data motivated by inverse stereographic projection. J Jpn Stat Soc 40(1):45–61

Abraham C, Molinari N, Servien R (2013) Unsupervised clustering of multivariate circular data. Stat Med 32(8):1376–1382

Abraham C, Servien R, Molinari N (2019) A clustering Bayesian approach for multivariate non-ordered circular data. Stat Model 19(6):595–616

Abuzaid AH, Mohamed IB, Hussin AG (2012) Boxplot for circular variables. Comput Stat 27(3):381–392

Abuzaid AH, Hussin AG, Mohamed IB (2013) Detection of outliers in simple circular regression models using the mean circular error statistic. J Stat Comput Simul 83(2):269–277

Abuzaid AH, Mohamed IB, Hussin AG (2014) Procedures for outlier detection in circular time series models. Environ Ecol Stat 21(4):793–809

Adler D, Murdoch D, et al. (2020) rgl: 3D visualization using OpenGL. R package version 0.100.54 https://CRAN.R-project.org/package=rgl

Agiomyrgiannakis Y, Stylianou Y (2009) Wrapped Gaussian mixture models for modeling and high-rate quantization of phase data of speech. IEEE Trans Audio Speech Lang Process 17(4):775–786

Agostinelli C (2007) Robust estimation for circular data. Comput Stat Data Anal 51(12):5867–5875

Agostinelli C, Lund U (2017) circular: circular statistics. R package version 0.4-93 https://CRAN.R-project.org/package=circular

Agostinelli C, Romanazzi M (2013) Nonparametric analysis of directional data based on data depth. Environ Ecol Stat 20(2):253–270

Ailliot P, Bessac J, Monbet V, Pène F (2015) Non-homogeneous hidden Markov-switching models for wind time series. J Stat Plan Inference 160:75–88

Aitchison J (1986) The statistical analysis of compositional data. Monographs on statistics and applied probability, vol 25. Chapman and Hall, London

Alonso-Pena M, Ameijeiras-Alonso J, Crujeiras RM (2020) Nonparametric tests for circular regression. J Stat Comput Simul 91(3):1–24

Altis A, Nguyen PH, Hegger R, Stock G (2007) Dihedral angle principal component analysis of molecular dynamics simulations. J Chem Phys 126(24):244111

Amaral GJA, Dryden IL, Wood ATA (2007) Pivotal bootstrap methods for k-sample problems in directional statistics and shape analysis. J Am Stat Assoc 102(478):695–707

Ameijeiras-Alonso J, Ley C (2020) Sine-skewed toroidal distributions and their application in protein bioinformatics. Biostatistics to appear

Ameijeiras-Alonso J, Crujeiras RM, Rodríguez Casal A (2018) Directional statistics for wildfires. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 187–210

Ameijeiras-Alonso J, Benali A, Crujeiras RM, Rodríguez-Casal A, Pereira JM (2019a) Fire seasonality identification with multimodality tests. Ann Appl Stat 13(4):2120–2139

Ameijeiras-Alonso J, Lagona F, Ranalli M, Crujeiras RM (2019b) A circular nonhomogeneous hidden Markov field for the spatial segmentation of wildfire occurrences. Environmetrics 30(2):e2501

Ameijeiras-Alonso J, Ley C, Pewsey A, Verdebout T (2020) On optimal tests for circular reflective symmetry about an unknown central direction. Stat Pap to appear

Amiri A, Thiam B, Verdebout T (2017) On the estimation of the density of a directional data stream. Scand J Stat 44(1):249–267

Anderson CM (1993) Graphical methods for circular and cylindrical data. Technical report, Univeristy of Waterloo

Anderson TW (1963) Asymptotic theory for principal component analysis. Ann Math Stat 34(1):122–148

Arnold BC, SenGupta A (2006) Recent advances in the analyses of directional data in ecological and environmental sciences. Environ Ecol Stat 13(3):253–256

Arnold R, Jupp P (2018) Orientations of symmetrical objects. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 25–44

Artes R (2008) Hypothesis tests for covariance analysis models for circular data. Commun Stat Theory Methods 37(8–10):1632–1640

Artes R, Toloi CMC (2009) An autoregressive model for time series of circular data. Commun Stat Theory Methods 39(1):186–194
Artes R, Paula GA, Ranvaud R (2000) Analysis of circular longitudinal data based on generalized estimating equations. Aust N Z J Stat 42(3):347–358
Azzalini A (1985) A class of distributions which includes the normal ones. Scand J Stat 12(2):171–178
Baayen C, Klugkist I (2014) Evaluating order-constrained hypotheses for circular data from a between-within subjects design. Psychol Methods 19(3):398–408
Baayen C, Klugkist I, Mechsner F (2012) A test of order-constrained hypotheses for circular data with applications to human movement science. J Mot Behav 44(5):351–363
Baba Y (1981) Statistics of angular data: wrapped normal distribution model. Proc Inst Stat Math 28(1):41–54
Bai ZD, Rao CR, Zhao LC (1988) Kernel estimators of density function of directional data. J Multivar Anal 27(1):24–39
Bakshaev A (2010) N-distance tests of uniformity on the hypersphere. Nonlinear Anal Model Control 15(1):15–8
Baldi P, Kerkyacharian G, Marinucci D, Picard D (2009a) Adaptive density estimation for directional data using needlets. Ann Stat 37(6A):3362–3395
Baldi P, Kerkyacharian G, Marinucci D, Picard D (2009b) Asymptotics for spherical needlets. Ann Stat 37(3):1150–1171
Ball FG, Dryden IL, Golalizadeh M (2008) Brownian motion and Ornstein–Uhlenbeck processes in planar shape space. Methodol Comput Appl Probab 10(1):1–22
Banerjee A, Dhillon I, Ghosh J, Sra S (2003) Generative model-based clustering of directional data. In: KDD '03. Association for Computing Machinery, New York, pp 19–28
Banerjee A, Dhillon IS, Ghosh J, Sra S (2005) Clustering on the unit hypersphere using von Mises–Fisher distributions. J Mach Learn Res 6(Sep):1345–1382
Banerjee A, Dhillon IS, Ghosh J, Sra S (2009) Text clustering with mixture of von Mises–Fisher distributions. In: Srivastava AN, Sahami M (eds) Text mining. Chapman and Hall/CRC data mining and knowledge discovery series. CRC Press, New York, pp 151–184
Bangert M, Hennig P, Oelfke U (2010) Using an infinite von Mises-Fisher mixture model to cluster treatment beam directions in external radiation therapy. In: ICMLA '10. IEEE Computer Society, Washington, DC, pp 746–751
Baragona R (2003) Further results on Lund's statistic for identifying cluster in a circular data set with application to time series. Commun Stat Simul Comput 32(3):943–952
Barnett A, Baker P (2020) season: analysing seasonal data R functions. R package version 0.3.12 https://CRAN.R-project.org/package=season
Barragán S, Fernández MA, Rueda C, Peddada S (2013) isocir: an R package for constrained inference using isotonic regression for circular data, with an application to cell biology. J Stat Softw 54(4):1–17
Barragán S, Rueda C, Fernández MA, Peddada SD (2015) Determination of temporal order among the components of an oscillatory system. PLOS ONE 10:e0124842
Barragán S, Rueda C, Fernández MA (2017) Circular order aggregation and its application to cell-cycle genes expressions. IEEE/ACM Trans Comput Biol Bioinform 14(4):819–829
Barros CM, Amaral GJA, Nascimento ADC, Cysneiros AHMA (2017) Detecting influential observations in Watson data. Commun Stat Simul Comput 46(14):6882–6898
Batschelet E (1981) Circular statistics in biology. Mathematics in biology series. Academic Press, London
Benjamin JBM, Hussain I, Yang MS (2019) Possibilistic c-means clustering on directional data. In: 2019 12th international congress on image and signal processing, biomedical engineering and informatics (CISP-BMEI). IEEE, New York, pp 1–6
Beran J, Ghosh S (2020) Estimating the mean direction of strongly dependent circular time series. J Time Ser Anal 41:210–228
Beran R (2016) Nonparametric estimation of trend in directional data. Stoch Process Their Appl 126(12):3808–3827
Beran RJ (1968) Testing for uniformity on a compact homogeneous space. J Appl Probab 5(1):177–195
Beran RJ (1969) Asymptotic theory of a class of tests for uniformity of a circular distribution. Ann Math Stat 40(4):1196–1206
Beran RJ (1979) Exponential models for directional data. Ann Stat 7(6):1162–1178
Berens P (2009) CircStat: a MATLAB toolbox for circular statistics. J Stat Softw 31(10):1–21
Bhattacharjee S (2020) Tracking space debris using directional statistics. Ph.D. Thesis, University of Leeds

Bhattacharya A, Bhattacharya R (2012) Nonparametric inference on manifolds. Institute of mathematical statistics monographs, vol 2. Cambridge University Press, Cambridge
Bhattacharya A, Dunson D (2012) Nonparametric Bayes classification and hypothesis testing on manifolds. J Multivar Anal 111:1–19
Bhattacharya R, Patrangenaru V (2003) Large sample theory of intrinsic and extrinsic sample means on manifolds. Ann Stat 31(1):1–29
Bhattacharya R, Patrangenaru V (2005) Large sample theory of intrinsic and extrinsic sample means on manifolds-II. Ann Stat 33(3):1225–1259
Bhattacharya R, Patrangenaru V (2014) Statistics on manifolds and landmarks based image analysis: a nonparametric theory with applications. J Stat Plan Inference 145:1–22
Bhattacharya S, SenGupta A (2009a) Bayesian analysis of semiparametric linear-circular models. J Agric Biol Environ Stat 14(1):33–65
Bhattacharya S, SenGupta A (2009b) Bayesian inference for circular distributions with unknown normalising constants. J Stat Plan Inference 139(12):4179–4192
Bijral A, Breitenbach M, Grudic GZ (2007) Mixture of Watson distributions: a generative model for hyperspherical embeddings. In: Meila M, Shen X (eds) Proceedings of the eleventh international conference on artificial intelligence and statistics. Proceedings of Machine Learning Research (PMLR), San Juan, Puerto Rico, pp 35–42
Bingham C (1974) An antipodally symmetric distribution on the sphere. Ann Stat 2(6):1201–1225
Boente G, Fraiman R (1991) Nonparametric regression for directional data. Trab Mat 176:1–13
Boente G, Rodriguez D, González-Manteiga W (2014) Goodness-of-fit test for directional data. Scand J Stat 41(1):259–275
Bogdan M, Bogdan K, Futschik A (2002) A data driven smooth test for circular uniformity. Ann Inst Stat Math 54(1):29–44
Boomsma W, Mardia KV, Taylor CC, Ferkinghoff-Borg J, Krogh A, Hamelryck T (2008) A generative, probabilistic model of local protein structure. Proc Natl Acad Sci USA 105(26):8932–8937
Boulerice B, Ducharme GR (1997) Smooth tests of goodness-of-fit for directional and axial data. J Multivar Anal 60(1):154–174
Breckling J (1989) The analysis of directional time series: applications to wind speed and direction. Lecture notes in statistics, vol 61. Springer, London
Bulla J, Lagona F, Maruotti A, Picone M (2012) A multivariate hidden Markov model for the identification of sea regimes from incomplete skewed and circular time series. J Agric Biol Environ Stat 17(4):544–567
Buttarazzi D (2020) bpDir: boxplots for directional data. R package version 0.1.1 https://CRAN.R-project.org/package=bpDir
Buttarazzi D, Pandolfo G, Porzio GC (2018) A boxplot for circular data. Biometrics 74(4):1492–1501
Byrne RW, Noser R, Bates LA, Jupp PE (2009) How did they get here from there? detecting changes of direction in terrestrial ranging. Anim Behav 77(3):619–631
Byrne S, Girolami M (2013) Geodesic Monte Carlo on embedded manifolds. Scand J Stat 40(4):825–845
Cabella P, Marinucci D (2009) Statistical challenges in the analysis of cosmic microwave background radiation. Ann Appl Stat 3(1):61–95
Cai T, Jiang T (2012) Phase transition in limiting distributions of coherence of high-dimensional random matrices. J Multivar Anal 107:24–39
Cai T, Fan J, Jiang T (2013) Distributions of angles in random packing on spheres. J Mach Learn Res 14(21):1837–1864
Calderara S, Prati A, Cucchiara R (2011) Mixtures of von Mises distributions for people trajectory shape analysis. IEEE Trans Circuits Syst Video Technol 21(4):457–471
Carnicero JA, Wiper MP, Ausín MC (2018) Density estimation of circular data with Bernstein polynomials. Hacet J Math Stat 47(2):273–286
Carta JA, Ramirez P, Bueno C (2008) A joint probability density function of wind speed and direction for wind energy analysis. Energy Convers Manag 49(6):1309–1320
Cetingul HE, Vidal R (2009) Intrinsic mean shift for clustering on Stiefel and Grassmann manifolds. In: 2009 IEEE conference on computer vision and pattern recognition. IEEE, New York, pp 1896–1902
Chakraborty S, Wong SWK (2019) BAMBI: bivariate angular mixture models. R package version 2.3.0 https://CRAN.R-project.org/package=BAMBI
Chang T (1986) Spherical regression. Ann Stat 14(3):907–924
Chang-Chien SJ, Yang MS, Hung WL (2010) Mean shift-based clustering for directional data. In: Third international workshop on advanced computational intelligence. IEEE, New York, pp 367–372

Chaubey YP (2018) Smooth kernel estimation of a circular density function: a connection to orthogonal polynomials on the unit circle. J Probab Stat 2018:1–4
Chaudhuri P, Marron JS (1999) SiZer for exploration of structures in curves. J Am Stat Assoc 94(447):807–823
Cheng D, Cammarota V, Fantaye Y, Marinucci D, Schwartzman A (2020) Multiple testing of local maxima for detection of peaks on the (celestial) sphere. Bernoulli 26(1):31–60
Cheng MY, Wu HT (2013) Local linear regression on manifolds and its geometric interpretation. J Am Stat Assoc 108(504):1421–1434
Chikuse Y (2003) Statistics on special manifolds. Lecture notes in statistics, vol 174. Springer, Heidelberg
Chirikjian GS, Kyatkin A (2001) Engineering applications of noncommutative harmonic analysis. CRC Press, Boca Raton
Chiuso A, Picci G (1998) Visual tracking of points as estimation on the unit sphere. In: Kriegman DJ, Hager GD, Morse AS (eds) The confluence of vision and control. Lecture notes in control and information sciences, vol 237. Springer, London, pp 90–105
Codling EA, Plank MJ, Benhamou S (2008) Random walk models in biology. J R Soc Interface 5(25):813–834
Comte F, Taupin ML (2003) Adaptive density deconvolution for circular data. Technical report MAP5 2003–10. Université Paris Descartes
Cornea E, Zhu H, Kim P, Ibrahim JG (2017) Regression models on Riemannian symmetric spaces. J R Stat Soc Ser B Stat Methodol 79(2):463–482
Costa M, Koivunen V, Poor HV (2014) Estimating directional statistics using wavefield modeling and mixtures of von-Mises distributions. IEEE Signal Process Lett 21(12):1496–1500
Cremers J (2020) bpnreg: Bayesian projected normal regression models for circular data. R package version 1.0.3 https://CRAN.R-project.org/package=bpnreg
Cremers J, Mulder KT, Klugkist I (2018) Circular interpretation of regression coefficients. Br J Math Stat Psychol 71(1):75–95
Cremers J, Pennings HJM, Ley C (2020) Regression models for cylindrical data in psychology. Multivar Behav Res 55(6):910–925
Cuesta-Albertos JA, Cuevas A, Fraiman R (2009) On projection-based tests for directional and compositional data. Stat Comput 19(4):367–380
Curry C, Marsland S, McLachlan RI (2019) Principal symmetric space analysis. J Comput Dyn 6(2):251–276
Cutting C, Paindaveine D, Verdebout T (2017a) Testing uniformity on high-dimensional spheres against monotone rotationally symmetric alternatives. Ann Stat 45(3):1024–1058
Cutting C, Paindaveine D, Verdebout T (2017b) Tests of concentration for low-dimensional and high-dimensional directional data. In: Ahmed SE (ed) Big and complex data analysis. Contributions to statistics. Springer, New York
Cutting C, Paindaveine D, Verdebout T (2020) On the power of axial tests of uniformity on spheres. Electron J Stat 14(1):2123–2154
Dai F, Xu Y (2013) Approximation theory and harmonic analysis on spheres and balls. Springer monographs in mathematics. Springer, New York
Dai X, Müller HG (2018) Principal component analysis for functional data on Riemannian manifolds and spheres. Ann Stat 46(6B):3334–3361
Damien P, Walker S (1999) A full Bayesian analysis of circular data using the von Mises distribution. Can J Stat 27(2):291–298
Damon J, Marron JS (2014) Backwards principal component analysis and principal nested relations. J Math Imaging Vis 50(1):107–114
D'Elia A (2001) A statistical model for orientation mechanism. Stat Methods Appl 10(1–3):157–174
Demni H, Messaoud A, Porzio GC (2019) The cosine depth distribution classifier for directional data. In: Bauer N, Ickstadt K, Lübke K, Szepannek G, Trautmann H, Vichi M (eds) Applications in statistical computing. Studies in classification, data analysis, and knowledge organization. Springer, Cham, pp 49–60
Deschepper E, Thas O, Ottoy JP (2008) Tests and diagnostic plots for detecting lack-of-fit for circular-linear regression models. Biometrics 64(3):912–920
Dette H, Melas VB (2003) Optimal designs for estimating individual coefficients in Fourier regression models. Ann Stat 31(5):1669–1692

Dette H, Wiens DP (2009) Robust designs for 3D shape analysis with spherical harmonic descriptors. Stat Sin 19(1):83–102
Dette H, Melas VB, Pepelyshev A (2005) Optimal designs for three-dimensional shape analysis with spherical harmonic descriptors. Ann Stat 33(6):2758–2788
Dette H, Konstantinou M, Schorning K, Gösmann J (2019) Optimal designs for regression with spherical data. Electron J Stat 13(1):361–390
Dhillon I, Sra S (2003) Modeling data using directional distributions. Technical report TR-03-06, Department of Computer Sciences, University of Texas at Austin
Dhillon IS, Modha DS (2001) Concept decompositions for large sparse text data using clustering. Mach Learn 42(1):143–175
Di Marzio M, Panzera A, Taylor CC (2009) Local polynomial regression for circular predictors. Stat Probab Lett 79(19):2066–2075
Di Marzio M, Panzera A, Taylor CC (2011) Kernel density estimation on the torus. J Stat Plan Inference 141(6):2156–2173
Di Marzio M, Panzera A, Taylor CC (2012a) Non-parametric smoothing and prediction for nonlinear circular time series. J Time Ser Anal 33(4):620–630
Di Marzio M, Panzera A, Taylor CC (2012b) Smooth estimation of circular cumulative distribution functions and quantiles. J Nonparametr Stat 24(4):935–949
Di Marzio M, Panzera A, Taylor CC (2013) Non-parametric regression for circular responses. Scand J Stat 40(2):238–255
Di Marzio M, Panzera A, Taylor CC (2014) Nonparametric regression for spherical data. J Am Stat Assoc 109(506):748–763
Di Marzio M, Fensore S, Panzera A, Taylor CC (2016a) A note on nonparametric estimation of circular conditional densities. J Stat Comput Simul 86(13):2573–2582
Di Marzio M, Fensore S, Panzera A, Taylor CC (2016b) Practical performance of local likelihood for circular density estimation. J Stat Comput Simul 86(13):2560–2572
Di Marzio M, Panzera A, Taylor CC (2016c) Nonparametric circular quantile regression. J Stat Plan Inference 170:1–14
Di Marzio M, Fensore S, Panzera A, Taylor CC (2017) Nonparametric estimating equations for circular probability density functions and their derivatives. Electron J Stat 11(2):4323–4346
Di Marzio M, Fensore S, Panzera A, Taylor CC (2018a) Circular local likelihood. TEST 27(4):921–945
Di Marzio M, Fensore S, Panzera A, Taylor CC (2018b) Nonparametric classification for circular data. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 241–257
Di Marzio M, Fensore S, Panzera A, Taylor CC (2019a) Kernel density classification for spherical data. Stat Probab Lett 144:23–29
Di Marzio M, Fensore S, Panzera A, Taylor CC (2019b) Local binary regression with spherical predictors. Stat Probab Lett 144:30–36
Di Marzio M, Panzera A, Taylor CC (2019c) Nonparametric rotations for sphere-sphere regression. J Am Stat Assoc 114(525):466–476
Dokmanic I, Petrinovic D (2010) Convolution on the n-sphere with application to pdf modeling. IEEE Trans Signal Process 58(3):1157–1170
Dortet-Bernadet JL, Wicker N (2008) Model-based clustering on the unit sphere with an illustration using gene expression profiles. Biostatistics 9(1):66–80
Downs TD (2003) Spherical regression. Biometrika 90(3):655–668
Downs TD, Mardia KV (2002) Circular regression. Biometrika 89(3):683–697
Dryden IL (2019) shapes: statistical shape analysis. R package version 1.2.5 https://CRAN.R-project.org/package=shapes
Dryden IL, Mardia KV (2016) Statistical shape analysis with applications in R. Wiley series in probability and statistics, 2nd edn. Wiley, Chichester
Ducharme GR, Vincent C, Aliaume C (2012) A statistical test to detect vortices in the current fields of bodies of water. Environ Ecol Stat 19(3):345–367
Ebner B, Henze N, Yukich JE (2018) Multivariate goodness-of-fit on flat and curved spaces via nearest neighbor distances. J Multivar Anal 165:231–242
Efromovich S (1997) Density estimation for the case of supersmooth measurement error. J Am Stat Assoc 92(438):526–535
Ehler M, Galanis J (2011) Frame theory in directional statistics. Stat Probab Lett 81(8):1046–1051

Eisen MB, Spellman PT, Brown PO, Botstein D (1998) Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci USA 95(25):14863–14868
Elad A, Keller Y, Kimmel R (2005) Texture mapping via spherical multi-dimensional scaling. In: Kimmel R, Sochen NA, Weickert J (eds) Scale space and PDE methods in computer vision. Lecture notes in computer science, vol 3459. Springer, Berlin, pp 443–455
Eltzner B, Huckemann SF (2019) A smeary central limit theorem for manifolds with application to high-dimensional spheres. Ann Stat 47(6):3360–3381
Eltzner B, Jung S, Huckemann S (2015) Dimension reduction on polyspheres with application to skeletal representations. In: Nielsen F, Barbaresco F (eds) Geometric science of information. Lecture notes in computer science, vol 9389. Springer, Cham, pp 22–29
Eltzner B, Huckemann S, Mardia KV (2018) Torus principal component analysis with applications to RNA structure. Ann Appl Stat 12(2):1332–1359
Erdem E, Shi J (2011) Comparison of bivariate distribution construction approaches for analysing wind speed and direction data. Wind Energy 14(1):27–41
Esteves C, Allen-Blanchette C, Makadia A, Daniilidis K (2020) Learning SO(3) equivariant representations with spherical CNNs. Int J Comput Vis 128:588–600
Eğecioğlu O, Srinivasan A (2000) Efficient nonparametric density estimation on the sphere with applications in fluid mechanics. SIAM J Sci Comput 22(1):152–176
Fallaize CJ, Kypraios T (2016) Exact Bayesian inference for the Bingham distribution. Stat Comput 26(1–2):349–360
Faÿ G, Delabrouille J, Kerkyacharian G, Picard D (2013) Testing the isotropy of high energy cosmic rays using spherical needlets. Ann Appl Stat 7(2):1040–1073
Fejér L (1916) Über trigonometrische Polynome. J Reine Angew Math 146:53–82
Feltz CJ, Goldin GA (2001) Partition-based goodness-of-fit tests on the line and the circle. Aust N Z J Stat 43(2):207–220
Fernandes K, Cardoso JS (2016) Discriminative directional classifiers. Neurocomputing 207:141–149
Fernández MA, Rueda C, Peddada SD (2012) Identification of a core set of signature cell cycle genes whose relative order of time to peak expression is conserved across species. Nucleic Acids Res 40(7):2823–2832
Fernández-Durán JJ (2004) Circular distributions based on nonnegative trigonometric sums. Biometrics 60(2):499–503
Fernández-Durán JJ, Gregorio-Domínguez MM (2014a) Distributions for spherical data based on nonnegative trigonometric sums. Stat Pap 55(4):983–1000
Fernández-Durán JJ, Gregorio-Domínguez MM (2014b) Modeling angles in proteins and circular genomes using multivariate angular distributions based on multiple nonnegative trigonometric sums. Stat Appl Genet Mol Biol 13(1):1–18
Fernández-Durán JJ, Gregorio-Domínguez MM (2016) CircNNTSR: an R package for the statistical analysis of circular, multivariate circular, and spherical data using nonnegative trigonometric sums. J Stat Softw 70(6):1–19
Ferreira JTAS, Juárez MA, Steel MFJ (2008) Directional log-spline distributions. Bayesian Anal 3(2):297–316
Figueiredo A (2007) Comparison of tests of uniformity defined on the hypersphere. Stat Probab Lett 77(3):329–334
Figueiredo A (2009) Discriminant analysis for the von Mises-Fisher distribution. Commun Stat Simul Comput 38(9):1991–2003
Figueiredo A (2017) Bootstrap and permutation tests in ANOVA for directional data. Comput Stat 32(4):1213–1240
Figueiredo A, Gomes P (2003) Power of tests of uniformity defined on the hypersphere. Commun Stat Simul Comput 32(1):87–94
Figueiredo A, Gomes P (2005) Discordancy test for the bipolar Watson distribution defined on the hypersphere. Commun Stat Simul Comput 34(1):145–153
Figueiredo A, Gomes P (2006) Discriminant analysis based on the Watson distribution defined on the hypersphere. Statistics 40(5):435–445
Fisher NI (1993) Statistical analysis of circular data. Cambridge University Press, Cambridge
Fisher NI, Lee AJ (1992) Regression models for an angular response. Biometrics 48(3):665–677
Fisher NI, Lee AJ (1994) Time series analysis of circular data. J R Stat Soc Ser B Methodol 56(2):327–339
Fisher NI, Marron JS (2001) Mode testing via the excess mass estimate. Biometrika 88(2):499–517

Fisher NI, Lewis T, Embleton BJ (1987) Statistical analysis of spherical data. Cambridge University Press, Cambridge

Fitak RR, Johnsen S (2017) Bringing the analysis of animal orientation data full circle: model-based approaches with maximum likelihood. J Exp Biol 220(21):3878–3882

Fletcher PT, Lu C, Pizer SM, Joshi S (2004) Principal geodesic analysis for the study of nonlinear statistics of shape. IEEE Trans Med Imaging 23(8):995–1005

Franke J, Redenbach C, Zhang N (2016) On a mixture model for directional data on the sphere. Scand J Stat 43(1):139–155

Fryer D, Olenko A, Li M, Wang Y (2020) rcosmo: cosmic microwave background data analysis. R package version 1.1.2 https://CRAN.R-project.org/package=rcosmo

Fu Y, Chen J, Li P (2008) Modified likelihood ratio test for homogeneity in a mixture of von Mises distributions. J Stat Plan Inference 138(3):667–681

Gao FQ, Li LN (2010) Large deviations and moderate deviations for kernel density estimators of directional data. Acta Math Sin Engl Ser 26(5):937–950

García-Portugués E (2013) Exact risk improvement of bandwidth selectors for kernel density estimation with directional data. Electron J Stat 7:1655–1685

García-Portugués E (2020a) DirStats: nonparametric methods for directional data. R package version 0.1.6 https://CRAN.R-project.org/package=DirStats

García-Portugués E (2020b) sdetorus: statistical tools for toroidal diffusions. R package version 0.1.7 https://CRAN.R-project.org/package=sdetorus

García-Portugués E, Verdebout T (2018) A review of uniformity tests on the hypersphere. arXiv:1804.00286

García-Portugués E, Verdebout T (2020) sphunif: uniformity tests on the circle, sphere, and hypersphere. https://github.com/egarpor/sphunif

García-Portugués E, Crujeiras RM, González-Manteiga W (2013a) Exploring wind direction and SO_2 concentration by circular-linear density estimation. Stoch Environ Res Risk Assess 27(5):1055–1067

García-Portugués E, Crujeiras RM, González-Manteiga W (2013b) Kernel density estimation for directional-linear data. J Multivar Anal 121:152–175

García-Portugués E, Barros AMG, Crujeiras RM, González-Manteiga W, Pereira J (2014) A test for directional-linear independence, with applications to wildfire orientation and size. Stoch Environ Res Risk Assess 28(5):1261–1275

García-Portugués E, Crujeiras RM, González-Manteiga W (2015) Central limit theorems for directional and linear random variables with applications. Stat Sin 25(3):1207–1229

García-Portugués E, Van Keilegom I, Crujeiras RM, González-Manteiga W (2016) Testing parametric models in linear-directional regression. Scand J Stat 43(4):1178–1191

García-Portugués E, Sørensen M, Mardia KV, Hamelryck T (2019) Langevin diffusions on the torus: estimation and applications. Stat Comput 29(1):1–22

García-Portugués E, Navarro-Esteban P, Cuesta-Albertos JA (2020a) On a projection-based class of uniformity tests on the hypersphere. arXiv:2008.09897

García-Portugués E, Paindaveine D, Verdebout T (2020b) On optimal tests for rotational symmetry against new classes of hyperspherical distributions. J Am Stat Assoc 115(532):1873–1887

García-Portugués E, Paindaveine D, Verdebout T (2020c) rotasym: tests for rotational symmetry on the hypersphere. R package version 1.0.9 https://CRAN.R-project.org/package=rotasym

Gatto R (2000) Multivariate saddlepoint test for the wrapped normal model. J Stat Comput Simul 65(1–4):271–285

Gatto R (2006) A bootstrap test for circular data. Commun Stat Theory Methods 35(1–3):281–292

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 (2017) Multivariate saddlepoint tests on the mean direction of the von Mises–Fisher distribution. Metrika 80(6–8):733–747

Genest M, Masse JC, Plante JF (2019) depth: nonparametric depth functions for multivariate analysis. R package version 2.1-1.1 https://CRAN.R-project.org/package=depth

Genton MG, Hall P (2007) Statistical inference for evolving periodic functions. J R Stat Soc Ser B Stat Methodol 69(4):643–657

George BJ, Ghosh K (2006) A semiparametric Bayesian model for circular-linear regression. Commun Stat Simul Comput 35(4):911–923

Ghazanfarihesari A, Sarmad M (2016) CircOutlier: detection of outliers in circular-circular regression. R package version 3.2.3 https://CRAN.R-project.org/package=CircOutlier

Ghosh K, Jammalamadaka SR, Vasudaven M (1999) Change-point problems for the von Mises distribution. J Appl Stat 26(4):423–434

Ghosh M, Zhong X, SenGupta A, Zhang R (2019) Non-subjective priors for wrapped Cauchy distributions. Stat Probab Lett 153:90–97

Gill J, Hangartner D (2010) Circular data in political science and how to handle it. Polit Anal 18(3):316–336

Giné E (1975) Invariant tests for uniformity on compact Riemannian manifolds based on Sobolev norms. Ann Stat 3(6):1243–1266

Giummolè F, Mameli V, Ruli E, Ventura L (2019) Objective Bayesian inference with proper scoring rules. TEST 28(3):728–755

Gneiting T (2013) Strictly and non-strictly positive definite functions on spheres. Bernoulli 19(4):1327–1349

Godtliebsen F, Marron JS, Chaudhuri P (2002) Significance in scale space for bivariate density estimation. J Comput Graph Stat 11(1):1–21

Golden M, García-Portugués E, Sørensen M, Mardia KV, Hamelryck T, Hein J (2017) A generative angular model of protein structure evolution. Mol Biol Evol 34(8):2085–2100

Gopal S, Yang Y (2014) Von Mises-Fisher clustering models. In: Xing EP, Jebara T (eds) Proceedings of the 31st international conference machine learning, vol 32. Proceedings of Machine Learning Research (PMLR), Bejing, pp 154–162

Graul C, Poppinga C (2018) bReeze: functions for wind resource assessment. R package version 0.4-3 https://CRAN.R-project.org/package=bReeze

Gu X, Wang Y, Chan TF, Thompson PM, Yau ST (2004) Genus zero surface conformal mapping and its application to brain surface mapping. IEEE Trans Med Imaging 23(8):949–958

Guella JC, Menegatto VA, Porcu E (2018) Strictly positive definite multivariate covariance functions on spheres. J Multivar Anal 166:150–159

Hall P, Yin J (2003) Nonparametric methods for deconvolving multiperiodic functions. J R Stat Soc Ser B Stat Methodol 65(4):869–886

Hall P, Watson GS, Cabrera J (1987) Kernel density estimation with spherical data. Biometrika 74(4):751–762

Hall P, Reimann J, Rice J (2000) Nonparametric estimation of a periodic function. Biometrika 87(3):545–557

Hamsici OC, Martinez AM (2007) Spherical-homoscedastic distributions: the equivalency of spherical and normal distributions in classification. J Mach Learn Res 8(Jul):1583–1623

Hara K, Nishino K, Ikeuchi K (2008) Mixture of spherical distributions for single-view relighting. IEEE Trans Pattern Anal Mach Intell 30(1):25–35

Hartman P, Watson GS (1974) "Normal" distribution functions on spheres and the modified Bessel functions. Ann Probab 2(4):593–607

Hasnat MA, Alata O, Trémeau A (2014) Unsupervised clustering of depth images using Watson mixture model. In: 2014 22nd international conference on pattern recognition. IEEE, New York, pp 214–219

Hassanzadeh F, Kalaylioglu Z (2018) A new multimodal and asymmetric bivariate circular distribution. Environ Ecol Stat 25(3):363–385

Hawkins DM, Lombard F (2015) Segmentation of circular data. J Appl Stat 42(1):88–97

Hawkins DM, Lombard F (2017) Cusum control for data following the von Mises distribution. J Appl Stat 44(8):1319–1332

Healy DMJ, Hendriks H, Kim PT (1998) Spherical deconvolution. J Multivar Anal 67(1):1–22

Henry G, Rodriguez D (2009) Kernel density estimation on Riemannian manifolds: asymptotic results. J Math Imaging Vis 34(3):235–239

Hernandez-Stumpfhauser D, Breidt FJ, Opsomer JD (2016) Hierarchical Bayesian small area estimation for circular data. Can J Stat 44(4):416–430

Hernandez-Stumpfhauser D, Breidt FJ, van der Woerd MJ (2017) The general projected normal distribution of arbitrary dimension: modeling and Bayesian inference. Bayesian Anal 12(1):113–133

Hill NA, Häder DP (1997) A biased random walk model for the trajectories of swimming micro-organisms. J Theor Biol 186(4):503–526

Hinkle J, Fletcher PT, Joshi S (2014) Intrinsic polynomials for regression on Riemannian manifolds. J Math Imaging Vis 50(1):32–52

Hokimoto T, Shimizu K (2008) An angular-linear time series model for waveheight prediction. Ann Inst Stat Math 60(4):781–800
Hokimoto T, Shimizu K (2014) A non-homogeneous hidden Markov model for predicting the distribution of sea surface elevation. J Appl Stat 41(2):294–319
Holmquist B, Gustafsson P (2017) A two-level directional model for dependence in circular data. Can J Stat 45(4):461–478
Holzmann H, Munk A, Stratmann B (2004) Identifiability of finite mixtures-with applications to circular distributions. Sankhyā 66(3):440–449
Holzmann H, Munk A, Suster M, Zucchini W (2006) Hidden Markov models for circular and linear-circular time series. Environ Ecol Stat 13(3):325–347
Hornik K, Grün B (2013) On conjugate families and Jeffreys priors for von Mises-Fisher distributions. J Stat Plan Inference 143(5):992–999
Hornik K, Grün B (2014) movMF: an R package for fitting mixtures of von Mises-Fisher distributions. J Stat Softw 58(10):1–31
Hornik K, Feinerer I, Kober M, Buchta C (2012) Spherical *k*-means clustering. J Stat Softw 50(10):1–22
Horwood JT, Poore AB (2014) Gauss von Mises distribution for improved uncertainty realism in space situational awareness. SIAM/ASA J Uncertain Quantif 2(1):276–304
Hotz T (2013) Extrinsic vs intrinsic means on the circle. In: Nielsen F, Barbaresco F (eds) Geometric science of information. Lecture notes in computer science, vol 8085. Springer, Berlin, pp 433–440
Hotz T, Huckemann S (2015) Intrinsic means on the circle: uniqueness, locus and asymptotics. Ann Inst Stat Math 67(1):177–193
Huckemann S, Ziezold H (2006) Principal component analysis for Riemannian manifolds, with an application to triangular shape spaces. Adv Appl Probab 38(2):299–319
Huckemann S, Hotz T, Munk A (2010) Intrinsic shape analysis: geodesic PCA for Riemannian manifolds modulo isometric Lie group actions. Stat Sin 20(1):1–58
Huckemann S, Kim KR, Munk A, Rehfeldt F, Sommerfeld M, Weickert J, Wollnik C (2016) The circular SiZer, inferred persistence of shape parameters and application to early stem cell differentiation. Bernoulli 22(4):2113–2142
Huckemann SF, Eltzner B (2018) Backward nested descriptors asymptotics with inference on stem cell differentiation. Ann Stat 46(5):1994–2019
Hughes G (2007) Multivariate and time series models for circular data with applications to protein conformational angles. Ph.D. Thesis, University of Leeds
Humphreys RK, Ruxton GD (2017) Consequences of grouped data for testing for departure from circular uniformity. Behav Ecol Sociobiol 71(11):167
Hundrieser S, Eltzner B, Huckemann SF (2020) Finite sample smeariness of Fréchet means and application to climate. arXiv:2005.02321
Hung WL, Chang-Chien SJ, Yang MS (2015) An intuitive clustering algorithm for spherical data with application to extrasolar planets. J Appl Stat 42(10):2220–2232
Hyvärinen A (2005) Estimation of non-normalized statistical models by score matching. J Mach Learn Res 6(Apr):695–709
Imoto T, Shimizu K, Abe T (2019) A cylindrical distribution with heavy-tailed linear part. Jpn J Stat Data Sci 2(1):129–154
Irwin ME, Cressie N, Johannesson G (2002) Spatial-temporal nonlinear filtering based on hierarchical statistical models. TEST 11(2):249–302
Jacimovic V, Crnkić A (2017) Collective motions of globally coupled oscillators and some probability distributions on circle. Phys Lett A 381(24):1989–1994
Jammalamadaka SR, Kozubowski TJ (2004) New families of wrapped distributions for modeling skew circular data. Commun Stat Theory Methods 33(9):2059–2074
Jammalamadaka SR, SenGupta A (2001) Topics in circular statistics. Series on multivariate analysis, vol 5. World Scientific, Singapore
Jammalamadaka SR, Terdik GH (2019) Harmonic analysis and distribution-free inference for spherical distributions. J Multivar Anal 171:436–451
Jammalamadaka SR, Meintanis S, Verdebout T (2020) On new Sobolev tests of uniformity on the circle with extension to the sphere. Bernoulli 26(3):2226–2252
Jensen MH, Mallasto A, Sommer S (2019) Simulation of conditioned diffusions on the flat torus. In: Nielsen F, Barbaresco F (eds) Geometric science of information. Lecture notes in computer science, vol 11712. Springer, Cham, pp 685–694

Jeong J, Jun M, Genton MG (2017) Spherical process models for global spatial statistics. Stat Sci 32(4):501–513
Johannes J, Schwarz M (2013) Adaptive circular deconvolution by model selection under unknown error distribution. Bernoulli 19(5A):1576–1611
Johnson RA, Wehrly TE (1977) Measures and models for angular correlation and angular-linear correlation. J R Stat Soc Ser B Methodol 39(2):222–229
Johnson RA, Wehrly TE (1978) Some angular-linear distributions and related regression models. J Am Stat Assoc 73(363):602–606
Jona-Lasinio G, Gelfand A, Jona-Lasinio M (2012) Spatial analysis of wave direction data using wrapped Gaussian processes. Ann Appl Stat 6(4):1478–1498
Jona-Lasinio G, Gelfand AE, Mastrantonio G (2018) Spatial and spatio-temporal circular processes with application to wave directions. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 129–162
Jona Lasinio G, Santoro M, Mastrantonio G (2020) CircSpaceTime: an R package for spatial and spatio-temporal modelling of circular data. J Stat Comput Simul 90(7):1315–1345
Jones MC, Pewsey A (2005) A family of symmetric distributions on the circle. J Am Stat Assoc 100(472):1422–1428
Jones MC, Pewsey A (2012) Inverse Batschelet distributions for circular data. Biometrics 68(1):183–193
Jones MC, Pewsey A, Kato S (2015) On a class of circulas: copulas for circular distributions. Ann Inst Stat Math 67(5):843–862
Jung S, Foskey M, Marron JS (2011) Principal arc analysis on direct product manifolds. Ann Appl Stat 5(1):578–603
Jung S, Dryden IL, Marron JS (2012) Analysis of principal nested spheres. Biometrika 99(3):551–568
Jupp PE (2001) Modifications of the Rayleigh and Bingham tests for uniformity of directions. J Multivar Anal 77(1):1–20
Jupp PE (2005) Sobolev tests of goodness of fit of distributions on compact Riemannian manifolds. Ann Stat 33(6):2957–2966
Jupp PE (2008) Data-driven Sobolev tests of uniformity on compact Riemannian manifolds. Ann Stat 36(3):1246–1260
Jupp PE (2009) Data-driven tests of uniformity on product manifolds. J Stat Plan Inference 139(11):3820–3829
Jupp PE (2015) Copulae on products of compact Riemannian manifolds. J Multivar Anal 140:92–98
Jupp PE, Kume A (2020) Measures of goodness of fit obtained by almost-canonical transformations on Riemannian manifolds. J Multivar Anal 176:104579
Jupp PE, Mardia KV (1989) A unified view of the theory of directional statistics. Int Stat Rev 57(3):261–294
Jupp PE, Regoli G, Azzalini A (2016) A general setting for symmetric distributions and their relationship to general distributions. J Multivar Anal 148:107–119
Kasarapu P, Allison L (2015) Minimum message length estimation of mixtures of multivariate Gaussian and von Mises-Fisher distributions. Mach Learn 100(2–3):333–378
Kato S (2009) A distribution for a pair of unit vectors generated by Brownian motion. Bernoulli 15(3):898–921
Kato S (2010) A Markov process for circular data. J R Stat Soc Ser B Stat Methodol 72(5):655–672
Kato S, Eguchi S (2016) Robust estimation of location and concentration parameters for the von Mises–Fisher distribution. Stat Pap 57(1):205–234
Kato S, Jones MC (2010) A family of distributions on the circle with links to, and applications arising from, Möbius transformation. J Am Stat Assoc 105(489):249–262
Kato S, Jones MC (2013) An extended family of circular distributions related to wrapped Cauchy distributions via Brownian motion. Bernoulli 19(1):154–171
Kato S, Jones MC (2015) A tractable and interpretable four-parameter family of unimodal distributions on the circle. Biometrika 102(1):181–190
Kato S, McCullagh P (2020) Some properties of a Cauchy family on the sphere derived from the Möbius transformations. Bernoulli 266(4):3224–3248
Kato S, Pewsey A (2015) A Möbius transformation-induced distribution on the torus. Biometrika 102(2):359–370
Kato S, Shimizu K (2008) Dependent models for observations which include angular ones. J Stat Plan Inference 138(11):3538–3549
Kato S, Shimizu K, Shieh GS (2008) A circular-circular regression model. Stat Sin 18(2):633–645

Kato S, Pewsey A, Jones MC (2018) Circulas from Fourier series. Technical report 7, School of Mathematics and Statistics, Open University
Kaufman CG, Ventura V, Kass RE (2005) Spline-based non-parametric regression for periodic functions and its application to directional tuning of neurons. Stat Med 24(14):2255–2265
Kendall DG, Barden D, Carne TK, Le H (1999) Shape and shape theory. Wiley series in probability and statistics. Wiley, Chichester
Kent JT (1975) Discussion of "Statistics of directional data". J R Stat Soc Ser B Methodol 37(3):377–378
Kent JT, Mardia KV (2009) Principal component analysis for the wrapped normal torus model. In: Gusnanto A, Mardia KV, Fallaize CJ (eds) LASR 2009: statistical tools for challenges in bioinformatics. Department of Statistics, University of Leeds, Leeds, pp 39–41
Kent JT, Mardia KV (2015) The winding number for circular data. In: Mardia KV, Gusnanto A, Nooney C, Voss J (eds) LASR 2015: geometry-driven statistics and its cutting edge applications: celebrating four decades of Leeds statistics workshops. Department of Statistics, University of Leeds, Leeds, pp 47–50
Kent JT, Mardia KV, Taylor CC (2008) Modelling strategies for bivariate circular data. In: Barber S, Baxter PD, Gusnanto A, Mardia KV (eds) LASR 2008: the art and science of statistical bioinformatics. Department of Statistics, University of Leeds, Leeds, pp 70–73
Kent JT, Hussein I, I, Jah MK, (2016) Directional distributions in tracking of space debris. In: 2016 19th international conference on information fusion (FUSION). IEEE, New York, pp 2081–2086
Kent JT, Ganeiber AM, Mardia KV (2018) A new unified approach for the simulation of a wide class of directional distributions. J Comput Graph Stat 27(2):291–301
Kent JT, Bhattacharjee S, Faber WR, Hussein II (2020) A unified approach to the orbital tracking problem. In: 2020 IEEE international conference on multisensor fusion and integration for intelligent systems (MFI). IEEE, New York, pp 82–87
Kerkyacharian G, Pham Ngoc TM, Picard D (2011) Localized spherical deconvolution. Ann Stat 39(2):1042–1068
Kesemen O, Tezel Ö, Özkul E (2016) Fuzzy *c*-means clustering algorithm for directional data (FCM4DD). Expert Syst Appl 58:76–82
Kim B, Huckemann S, Schulz J, Jung S (2019) Small-sphere distributions for directional data with application to medical imaging. Scand J Stat 46(4):1047–1071
Kim NC, So HJ (2018) Directional statistical Gabor features for texture classification. Pattern Recognit Lett 112:18–26
Kim PT, Koo JY (2002) Optimal spherical deconvolution. J Multivar Anal 80(1):21–42
Kim PT, Koo JY, Park HJ (2004) Sharp minimaxity and spherical deconvolution for super-smooth error distributions. J Multivar Anal 90(2):384–392
Kim PT, Koo JY, Pham Ngoc TM (2016) Supersmooth testing on the sphere over analytic classes. J Nonparametr Stat 28(1):84–115
Klemelä J (1999) Asymptotic minimax risk for the white noise model on the sphere. Scand J Stat 26(3):465–473
Klemelä J (2000) Estimation of densities and derivatives of densities with directional data. J Multivar Anal 73(1):18–40
Klemelä J (2003) Lower bounds for the asymptotic minimax risk with spherical data. J Stat Plan Inference 113(1):113–136
Klugkist I, Bullens J, Postma A (2012) Evaluating order-constrained hypotheses for circular data using permutation tests. Br J Math Stat Psychol 65(2):222–236
Kranstauber B, Smolla M, Scharf AK (2020) move: visualizing and analyzing animal track data. R package version 4.0.4 https://CRAN.R-project.org/package=move
Kueh A (2012) Locally adaptive density estimation on the unit sphere using needlets. Constr Approx 36(3):433–458
Kume A, Sei T (2018) On the exact maximum likelihood inference of Fisher–Bingham distributions using an adjusted holonomic gradient method. Stat Comput 28(4):835–847
Kurz G, Gilitschenski I, Julier SJ, Hanebeck UD (2014) Recursive Bingham filter for directional estimation involving 180 degree symmetry. J Adv Inf Fusion 9(2):90–105
Kurz G, Gilitschenski I, Pfaff F, Drude L, Hanebeck UD, Haeb-Umbach R, Siegwart RY (2019) Directional statistics and filtering using libDirectional. J Stat Softw 89(4):1–31
Lacour C, Pham Ngoc TM (2014) Goodness-of-fit test for noisy directional data. Bernoulli 20(4):2131–2168

Lagona F (2016) Regression analysis of correlated circular data based on the multivariate von Mises distribution. Environ Ecol Stat 23(1):89–113
Lagona F (2018) Correlated cylindrical data. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 45–59
Lagona F, Picone M (2011) A latent-class model for clustering incomplete linear and circular data in marine studies. J Data Sci 9(4):585–605
Lagona F, Picone M (2012) Model-based clustering of multivariate skew data with circular components and missing values. J Appl Stat 39(5):927–945
Lagona F, Picone M (2013) Maximum likelihood estimation of bivariate circular hidden Markov models from incomplete data. J Stat Comput Simul 83(7):1223–1237
Lagona F, Picone M (2016) Model-based segmentation of spatial cylindrical data. J Stat Comput Simul 86(13):2598–2610
Lagona F, Picone M, Maruotti A (2015a) A hidden Markov model for the analysis of cylindrical time series. Environmetrics 26(8):534–544
Lagona F, Picone M, Maruotti A, Cosoli S (2015b) A hidden Markov approach to the analysis of space-time environmental data with linear and circular components. Stoch Environ Res Risk Assess 29(2):397–409
Laha AK, Mahesh KC (2015) Robustness of tests for directional mean. Statistics 49(3):522–536
Laha AK, Raja ACP, Mahesh KC (2019) SB-robust estimation of mean direction for some new circular distributions. Stat Pap 60(3):527–552
Landler L, Ruxton GD, Malkemper EP (2018) Circular data in biology: advice for effectively implementing statistical procedures. Behav Ecol Sociobiol 72(8):128
Landler L, Ruxton GD, Malkemper EP (2019) Circular statistics meets practical limitations: a simulation-based Rao's spacing test for non-continuous data. Mov Ecol 7(1):15
Larriba Y, Rueda C, Fernández MA, Peddada SD (2020) Order restricted inference in chronobiology. Stat Med 39(3):265–278
Larsen PV, Blaesild P, Sørensen MK (2002) Improved likelihood ratio tests on the von Mises–Fisher distribution. Biometrika 89(4):947–951
Le Bihan N, Chatelain F, Manton JH (2016) Isotropic multiple scattering processes on hyperspheres. IEEE Trans Inf Theory 62(10):5740–5752
Leguey I, Bielza C, Larrañaga P (2019a) Circular Bayesian classifiers using wrapped Cauchy distributions. Data Knowl Eng 122:101–115
Leguey I, Larrañaga P, Bielza C, Kato S (2019b) A circular-linear dependence measure under Johnson–Wehrly distributions and its application in Bayesian networks. Inf Sci 486:240–253
Lennox KP, Dahl DB, Vannucci M, Tsai JW (2009) Density estimation for protein conformation angles using a bivariate von Mises distribution and Bayesian nonparametrics. J Am Stat Assoc 104(486):586–596
Lennox KP, Dahl DB, Vannucci M, Day R, Tsai JW (2010) A Dirichlet process mixture of hidden Markov models for protein structure prediction. Ann Appl Stat 4(2):916–942
Leonenko NN, Ruiz-Medina MD (2018) Increasing domain asymptotics for the first Minkowski functional of spherical random fields. Theory Probab Math Stat 97:127–149
Ley C, Verdebout T (2014a) Local powers of one- and multi-sample tests for the concentration of Fisher-von Mises-Langevin distributions. Int Stat Rev 82(3):440–456
Ley C, Verdebout T (2014b) Simple optimal tests for circular reflective symmetry about a specified median direction. Stat Sin 24(3):1319–1339
Ley C, Verdebout T (2017a) Modern directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton
Ley C, Verdebout T (2017b) Skew-rotationally-symmetric distributions and related efficient inferential procedures. J Multivar Anal 159:67–81
Ley C, Verdebout T (eds) (2018) Applied directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton
Ley C, Swan Y, Thiam B, Verdebout T (2013) Optimal R-estimation of a spherical location. Stat Sin 23(1):305–332
Ley C, Sabbah C, Verdebout T (2014) A new concept of quantiles for directional data and the angular Mahalanobis depth. Electron J Stat 8(1):795–816
Ley C, Paindaveine D, Verdebout T (2015) High-dimensional tests for spherical location and spiked covariance. J Multivar Anal 139:79–91
Ley C, Swan Y, Verdebout T (2017) Efficient ANOVA for directional data. Ann Inst Stat Math 69(1):39–62

Li L (2014) Moderate deviations results for a symmetry testing statistic based on the kernel density estimator for directional data. Commun Stat Theory Methods 43(14):3007–3018

Lin L, St Thomas B, Zhu H, Dunson DB (2017) Extrinsic local regression on manifold-valued data. J Am Stat Assoc 112(519):1261–1273

Lin SB (2019) Nonparametric regression using needlet kernels for spherical data. J Complexity 50:66–83

Liu D, Umbach DM, Peddada SD, Li L, Crockett PW, Weinberg CR (2004) A random-periods model for expression of cell-cycle genes. Proc Natl Acad Sci USA 101(19):7240–7245

Liu RY, Singh K (1992) Ordering directional data: concepts of data depth on circles and spheres. Ann Stat 20(3):1468–1484

Loader CR (1996) Local likelihood density estimation. Ann Stat 24(4):1602–1618

Lombard F, Maxwell RK (2012) A cusum procedure to detect deviations from uniformity in angular data. J Appl Stat 39(9):1871–1880

Lombard F, Hawkins DM, Potgieter CJ (2017) Sequential rank CUSUM charts for angular data. Comput Stat Data Anal 105:268–279

López-Cruz PL, Bielza C, Larrañaga P (2015) Directional naive Bayes classifiers. Pattern Anal Appl 18(2):225–246

Lu Y, Corander J, Yang Z (2019) Doubly stochastic neighbor embedding on spheres. Pattern Recognit Lett 128:100–106

Lund U (1999) Cluster analysis for directional data. Commun Stat Simul Comput 28(4):1001–1009

Lund U (2002) Tree-based regression for a circular response. Commun Stat Theory Methods 31(9):1549–1560

Lunga D, Ersoy O (2013) Spherical stochastic neighbor embedding of hyperspectral data. IEEE Trans Geosci Remote Sens 51(2):857–871

Mahmood EA, Rana S, Midi H, Hussin AG (2017) Detection of outliers in univariate circular data using robust circular distance. J Mod Appl Stat Methods 16(2):22

Maitra R, Ramler IP (2010) A k-mean-directions algorithm for fast clustering of data on the sphere. J Comput Graph Stat 19(2):377–396

Maksimov VM (1967) Necessary and sufficient statistics for the family of shifts of probability distributions on continuous bicompact groups (in Russian). Theor Verojatna 12(2):307–321

Mardia K (2018) A new estimation methodology for standard directional distributions. In: 2018 21st international conference on information fusion (FUSION). IEEE, New York, pp 724–729

Mardia KV (1972) Statistics of directional data. Probability and mathematical statistics. Academic Press, London

Mardia KV (1975) Statistics of directional data. J R Stat Soc Ser B Methodol 37(3):349–393

Mardia KV (2010) Bayesian analysis for bivariate von Mises distributions. J Appl Stat 37(3):515–528

Mardia KV, Frellsen J (2012) Statistics of bivariate von Mises distributions. In: Hamelryck T, Mardia K, Ferkinghoff-Borg J (eds) Bayesian methods in structural bioinformatics. Statistics for biology and health. Springer, Berlin, pp 159–178

Mardia KV, Jupp PE (1999) Directional statistics. Wiley series in probability and statistics. Wiley, Chichester

Mardia KV, Patrangenaru V (2005) Directions and projective shapes. Ann Stat 33(4):1666–1699

Mardia KV, Sriram K (2020) Families of discrete circular distributions with some novel applications. arXiv:2009.05437

Mardia KV, Sutton TW (1978) A model for cylindrical variables with applications. J R Stat Soc Ser B Methodol 40(2):229–233

Mardia KV, Voss J (2014) Some fundamental properties of a multivariate von Mises distribution. Commun Stat Theory Methods 43(6):1132–1144

Mardia KV, Taylor CC, Subramaniam GK (2007) Protein bioinformatics and mixtures of bivariate von Mises distributions for angular data. Biometrics 63(2):505–512

Mardia KV, Hughes G, Taylor CC, Singh H (2008) A multivariate von Mises distribution with applications to bioinformatics. Can J Stat 36(1):99–109

Mardia KV, Kent JT, Zhang Z, Taylor CC, Hamelryck T (2012) Mixtures of concentrated multivariate sine distributions with applications to bioinformatics. J Appl Stat 39(11):2475–2492

Mardia KV, Kent JT, Laha AK (2016) Score matching estimators for directional distributions. arXiv:1604.08470

Mardia KV, Foldager JI, Frellsen J (2018) Directional statistics in protein bioinformatics. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman and Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 1–23

Marinucci D, Peccati G (2011) Random fields on the sphere. London mathematical society lecture note series. Cambridge University Press, Cambridge

Marinucci D, Pietrobon D, Balbi A, Baldi P, Cabella P, Kerkyacharian G, Natoli P, Picard D, Vittorio N (2008) Spherical needlets for cosmic microwave background data analysis. Mon Not R Astron Soc 383(2):539–545

Marron JS, Alonso AM (2014) Overview of object oriented data analysis. Biomet J 56(5):732–753

Maruotti A (2016) Analyzing longitudinal circular data by projected normal models: a semi-parametric approach based on finite mixture models. Environ Ecol Stat 23(2):257–277

Maruotti A, Punzo A, Mastrantonio G, Lagona F (2016) A time-dependent extension of the projected normal regression model for longitudinal circular data based on a hidden Markov heterogeneity structure. Stoch Environ Res Risk Assess 30(6):1725–1740

Mash'al M, Hosseini R (2015) K-means++ for mixtures of von Mises-Fisher distributions. In: 2015 7th conference on information and knowledge technology (IKT). IEEE, New York, pp 1–6

Mastrantonio G (2018) The joint projected normal and skew-normal: a distribution for poly-cylindrical data. J Multivar Anal 165:14–26

Mastrantonio G, Calise G (2016) Hidden Markov model for discrete circular-linear wind data time series. J Stat Comput Simul 86(13):2611–2624

Mastrantonio G, Maruotti A, Jona Lasinio G (2015) Bayesian hidden Markov modelling using circular-linear general projected normal distribution. Environmetrics 26(2):145–158

Mastrantonio G, Gelfand AE, Jona Lasinio G (2016a) The wrapped skew Gaussian process for analyzing spatio-temporal data. Stoch Environ Res Risk Assess 30(8):2231–2242

Mastrantonio G, Jona Lasinio G, Gelfand AE (2016b) Spatio-temporal circular models with non-separable covariance structure. TEST 25(2):331–350

Mastrantonio G, Jona Lasinio G, Maruotti A, Calise G (2019) Invariance properties and statistical inference for circular data. Stat Sin 29(1):67–80

Mazumder S, Bhattacharya S (2016) Bayesian nonparametric dynamic state space modeling with circular latent states. J Stat Theory Pract 10(1):154–178

Mazumder S, Bhattacharya S (2017) Nonparametric dynamic state space modeling of observed circular time series with circular latent states: a Bayesian perspective. J Stat Theory Pract 11(4):693–718

McClintock BT, King R, Thomas L, Matthiopoulos J, McConnell BJ, Morales JM (2012) A general discrete-time modeling framework for animal movement using multistate random walks. Ecol Monogr 82(3):335–349

McCullagh P (1996) Möbius transformation and Cauchy parameter estimation. Ann Stat 24(2):787–808

McMillan GP, Hanson TE, Saunders G, Gallun FJ (2013) A two-component circular regression model for repeated measures auditory localization data. J R Stat Soc Ser C Appl Stat 62(4):515–534

McVinish R, Mengersen K (2008) Semiparametric Bayesian circular statistics. Comput Stat Data Anal 52(10):4722–4730

Meilán-Vila A, Francisco-Fernández M, Crujeiras RM, Panzera A (2020) Nonparametric multiple regression estimation for circular responses. TEST to appear

Meintanis S, Verdebout T (2019) Le Cam maximin tests for symmetry of circular data based on the characteristic function. Stat Sin 29(3):1301–1320

Michelot T, Langrock R, Patterson T, McClintock B (2016) moveHMM: an R package for the statistical modelling of animal movement data using hidden Markov models. Methods Ecol Evol 7(11):1308–1315

Miolane N, Le Brigant A, Mathe J, Hou B, Guigui N, Thanwerdas Y, Heyder S, Peltre O, Koep N, Zaatiti H, Hajri H, Cabanes Y, Gerald T, Chauchat P, Shewmake C, Kainz B, Donnat C, Holmes S, Pennec X (2020) geomstats: a Python package for Riemannian geometry in machine learning. arXiv: 2004.04667

Miyata Y, Shiohama T, Abe T (2020) Estimation of finite mixture models of skew-symmetric circular distributions. Metrika 83:895–922

Modlin D, Fuentes M, Reich B (2012) Circular conditional autoregressive modeling of vector fields. Environmetrics 23(1):46–53

Moghimbeygi M, Golalizadeh M (2020) Spherical logistic distribution. Commun Math Stat 8(2):151–166

Monbet V (2020) NHMSAR: non-homogeneous Markov switching autoregressive models. R package version 1.17 https://CRAN.R-project.org/package=NHMSAR

Monnier JB (2011) Non-parametric regression on the hypersphere with uniform design. TEST 20(2):412–446

Montanari A, Calò DG (2013) Model-based clustering of probability density functions. Adv Data Anal Classif 7(3):301–319
Mooney JA, Helms PJ, Jolliffe IT (2003) Fitting mixtures of von Mises distributions: a case study involving sudden infant death syndrome. Comput Stat Data Anal 41(3–4):505–513
Morales JM, Haydon DT, Frair J, Holsinger KE, Fryxell JM (2004) Extracting more out of relocation data: building movement models as mixtures of random walks. Ecology 85(9):2436–2445
Morphet WJ, Symanzik J (2010) The circular dataimage, a graph for high-resolution circular-spatial data. Int J Digit Earth 3(1):47–71
Mu Y, Nguyen PH, Stock G (2005) Energy landscape of a small peptide revealed by dihedral angle principal component analysis. Proteins 58(1):45–52
Mulder K, Klugkist I (2017) Bayesian estimation and hypothesis tests for a circular generalized linear model. J Math Psychol 80:4–14
Mulder K, Jongsma P, Klugkist I (2020a) Bayesian inference for mixtures of von Mises distributions using reversible jump MCMC sampler. J Stat Comput Simul 90(9):1539–1556
Mulder K, Klugkist I, van Renswoude D, Visser I (2020b) Mixtures of peaked power Batschelet distributions for circular data with application to saccade directions. J Math Psychol 95:102309
Mulder KT, Klugkist I (2021) Bayesian tests for circular uniformity. J Stat Plan Inference 211:315–325
Munro MA, Blenkinsop TG (2012) MARD–A moving average rose diagram application for the geosciences. Comput Geosci 49:112–120
Mushkudiani NA (2002) Small nonparametric tolerance regions for directional data. J Stat Plan Inference 100(1):67–80
Narcowich FJ, Petrushev P, Ward JD (2006) Localized tight frames on spheres. SIAM J Math Anal 38(2):574–594
Navarro AKW, Frellsen J, Turner RE (2017) The multivariate generalised von Mises distribution: inference and applications. In: Proceedings of the thirty-first AAAI conference on artificial intelligence (AAAI-17). Association for the Advancement of Artificial Intelligence, San Francisco, pp 2394–2400
Nicosia A, Duchesne T, Rivest LP, Fortin D (2017) A general hidden state random walk model for animal movement. Comput Stat Data Anal 105:76–95
Nodehi A, Golalizadeh M, Heydari A (2015) Dihedral angles principal geodesic analysis using nonlinear statistics. J Appl Stat 42(9):1962–1972
Núñez-Antonio G, Geneyro E (2020) A multivariate projected gamma model for directional data. Commun Stat: Case Stud Data Anal Appl to appear
Núñez-Antonio G, Gutiérrez-Peña E (2005a) A Bayesian analysis of directional data using the projected normal distribution. J Appl Stat 32(10):995–1001
Núñez-Antonio G, Gutiérrez-Peña E (2005b) A Bayesian analysis of directional data using the von Mises-Fisher distribution. Commun Stat Simul Comput 34(4):989–999
Núñez-Antonio G, Gutiérrez-Peña E (2014) A Bayesian model for longitudinal circular data based on the projected normal distribution. Comput Stat Data Anal 71:506–519
Núñez-Antonio G, Gutiérrez-Peña E, Escarela G (2011) A Bayesian regression model for circular data based on the projected normal distribution. Stat Model 11(3):185–201
Núñez-Antonio G, Mendoza M, Contreras-Cristán A, Gutiérrez-Peña E, Mendoza E (2018) Bayesian nonparametric inference for the overlap of daily animal activity patterns. Environ Ecol Stat 25(4):471–494
Oba S, Kato K, Ishii S (2005) Multi-scale clustering for gene expression profiling data. In: Fifth IEEE symposium on bioinformatics and bioengineering (BIBE ’05). IEEE, New York, pp 210–217
Oliveira M, Crujeiras RM, Rodríguez-Casal A (2012) A plug-in rule for bandwidth selection in circular density estimation. Comput Stat Data Anal 56(12):3898–3908
Oliveira M, Crujeiras RM, Rodríguez-Casal A (2014) CircSiZer: an exploratory tool for circular data. Environ Ecol Stat 21(1):143–159
Otieno SB, Anderson-Cook CM (2012) Design and analysis of experiments for directional data. In: Hinkelmann K (ed) Design and analysis of experiments. Wiley series in probability and statistics. Wiley, Hoboken, pp 501–532
Oualkacha K, Rivest LP (2009) A new statistical model for random unit vectors. J Multivar Anal 100(1):70–80
Paindaveine D, Verdebout T (2015) Optimal rank-based tests for the location parameter of a rotationally symmetric distribution on the hypersphere. In: Hallin M, Mason D, Pfeifer D, Steinebach J (eds) Mathematical statistics and limit theorems. Springer, Cham, pp 249–269
Paindaveine D, Verdebout T (2016) On high-dimensional sign tests. Bernoulli 22(3):1745–1769

Paindaveine D, Verdebout T (2017) Inference on the mode of weak directional signals: a Le Cam perspective on hypothesis testing near singularities. Ann Stat 45(2):800–832
Paindaveine D, Verdebout T (2020) Inference for spherical location under high concentration. Ann Stat 48(5):2982–2998
Paine PJ, Preston SP, Tsagris M, Wood ATA (2018) An elliptically symmetric angular Gaussian distribution. Stat Comput 28(3):689–697
Paine PJ, Preston SP, Tsagris M, Wood ATA (2020) Spherical regression models with general covariates and anisotropic errors. Stat Comput 30(1):153–165
Paluszewski M, Hamelryck T (2010) Mocapy++ - a toolkit for inference and learning in dynamic Bayesian networks. BMC Bioinform 11(126):1–6
Panaretos VM, Pham T, Yao Z (2014) Principal flows. J Am Stat Assoc 109(505):424–436
Pandolfo G, D'Ambrosio A, Porzio GC (2018a) A note on depth-based classification of circular data. Electron J Appl Stat Anal 11(2):447–462
Pandolfo G, Paindaveine D, Porzio GC (2018b) Distance-based depths for directional data. Can J Stat 46(4):593–609
Pardo A, Real E, Krishnaswamy V, López-Higuera JM, Pogue BW, Conde OM (2017) Directional kernel density estimation for classification of breast tissue spectra. IEEE Trans Med Imaging 36(1):64–73
Park HS (2012) Asymptotic behavior of the kernel density estimator from a geometric viewpoint. Commun Stat Simul Comput 41(19):3479–3496
Park HS (2013) Comparison of relative efficiency of kernel density estimator with the exponential map. J Korean Stat Soc 42(2):267–275
Peel D, Whiten WJ, McLachlan GJ (2001) Fitting mixtures of Kent distributions to aid in joint set identification. J Am Stat Assoc 96(453):56–63
Pelletier B (2005) Kernel density estimation on Riemannian manifolds. Stat Probab Lett 73(3):297–304
Pennec X (2018) Barycentric subspace analysis on manifolds. Ann Stat 46(6A):2711–2746
Pertsemlidis A, Zelinka J, Fondon JW, Henderson RK, Otwinowski Z (2005) Bayesian statistical studies of the Ramachandran distribution. Stat Appl Genet Mol Biol 4(1)
Pewsey A (2000) The wrapped skew-normal distribution on the circle. Commun Stat Theory Methods 29(11):2459–2472
Pewsey A (2002) Testing circular symmetry. Can J Stat 30(4):591–600
Pewsey A (2004a) The large-sample joint distribution of key circular statistics. Metrika 60(1):25–32
Pewsey A (2004b) Testing for circular reflective symmetry about a known median axis. J Appl Stat 31(5):575–585
Pewsey A (2006) Modelling asymmetrically distributed circular data using the wrapped skew-normal distribution. Environ Ecol Stat 13(3):257–269
Pewsey A (2008) The wrapped stable family of distributions as a flexible model for circular data. Comput Stat Data Anal 52(3):1516–1523
Pewsey A (2018) Applied directional statistics with R: an overview. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman & Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 277–290
Pewsey A, Jones MC (2005) Discrimination between the von Mises and wrapped normal distributions: just how big does the sample size have to be? Statistics 39(2):81–89
Pewsey A, Kato S (2016) Parametric bootstrap goodness-of-fit testing for Wehrly–Johnson bivariate circular distributions. Stat Comput 26(6):1307–1317
Pewsey A, Lewis T, Jones MC (2007) The wrapped t family of circular distributions. Aust N Z J Stat 49(1):79–91
Pewsey A, Neuhäuser M, Ruxton GD (2013) Circular statistics in R. Oxford University Press, Oxford
Pham Ngoc TM (2019) Adaptive optimal kernel density estimation for directional data. J Multivar Anal 173:248–267
Pitt MK, Shephard N (1999) Filtering via simulation: auxiliary particle filters. J Am Stat Assoc 94(446):590–599
Pizer SM, Jung S, Goswami D, Vicory J, Zhao X, Chaudhuri R, Damon JN, Huckemann S, Marron JS (2013) Nested sphere statistics of skeletal models. In: Breuß M, Bruckstein A, Maragos P (eds) Innovations for shape analysis, mathematics and visualization. Springer, Berlin, pp 93–115
Polsen O, Taylor CC (2015) Parametric circular-circular regression and diagnostic analysis. In: Dryden IL, Kent JT (eds) Geometry driven statistics. Wiley series in probability and statistics. Wiley, Chichester, pp 115–128

Porcu E, Bevilacqua M, Genton MG (2016) Spatio-temporal covariance and cross-covariance functions of the great circle distance on a sphere. J Am Stat Assoc 111(514):888–898
Porcu E, Furrer R, Nychka D (2020) 30 years of space-time covariance functions. WIREs Comp Stat to appear:e1512
Presnell B, Morrison SP, Littell RC (1998) Projected multivariate linear models for directional data. J Am Stat Assoc 93(443):1068–1077
Pycke JR (2007) A decomposition for invariant tests of uniformity on the sphere. Proc Am Math Soc 135(9):2983–2993
Pycke JR (2010) Some tests for uniformity of circular distributions powerful against multimodal alternatives. Can J Stat 38(1):80–96
Qin X, Zhang JS, Yan XD (2011) A nonparametric circular-linear multivariate regression model with a rule-of-thumb bandwidth selector. Comput Math Appl 62(8):3048–3055
Qiu X, Wu S, Wu H (2015) A new information criterion based on Langevin mixture distribution for clustering circular data with application to time course genomic data. Stat Sin 25(4):1459–1476
R Core Team (2020) R: A language and environment for statistical computing. R foundation for statistical computing. https://www.R-project.org/
Rakhimberdiev E, Saveliev A, Piersma T, Karagicheva J (2017) FLightR: an R package for reconstructing animal paths from solar geolocation loggers. Methods Ecol Evol 8(11):1482–1487
Ranalli M, Lagona F, Picone M, Zambianchi E (2018) Segmentation of sea current fields by cylindrical hidden Markov models: a composite likelihood approach. J R Stat Soc Ser C Appl Stat 67(3):575–598
Rayleigh Lord (1919) On the problem of random vibrations, and of random flights in one, two, or three dimensions. Lond Edinb Dublin Philos Mag J Sci 37(220):321–347
Reed WJ, Pewsey A (2009) Two nested families of skew-symmetric circular distributions. TEST 18(3):516–528
Riccardi L, Nguyen PH, Stock G (2009) Free-energy landscape of RNA hairpins constructed via dihedral angle principal component analysis. J Phys Chem B 113(52):16660–16668
Rivest LP (1997) A decentred predictor for circular-circular regression. Biometrika 84(3):717–726
Rivest LP (1999) Some linear model techniques for analyzing small-circle spherical data. Can J Stat 27(3):623–638
Rivest LP, Kato S (2019) A random-effects model for clustered circular data. Can J Stat 47(4):712–728
Rivest LP, Oualkacha K (2018) On modeling of SE(3) objects. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman & Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 111–127
Rivest LP, Duchesne T, Nicosia A, Fortin D (2016) A general angular regression model for the analysis of data on animal movement in ecology. J R Stat Soc Ser C Appl Stat 65(3):445–463
Rodgers JL, Beasley WH, Schuelke M (2014) Graphical data analysis on the circle: wrap-around time series plots for (interrupted) time series designs. Multivar Behav Res 49(6):571–580
Rodríguez CE, Núñez-Antonio G, Escarela G (2020) A Bayesian mixture model for clustering circular data. Comput Stat Data Anal 143:106842
Rodriguez-Lujan L, Bielza C, Larrañaga P (2015) Regularized multivariate von Mises distribution. In: Puerta JM, Gámez JA, Dorronsoro B, Barrenechea E, Troncoso A, Baruque B, Galar M (eds) Advances in artificial intelligence. Lecture notes in computer science, vol 9422. Springer, Cham, pp 25–35
Rodriguez-Lujan L, Bielza C, Larrañaga P (2017) Frobenius norm regularization for the multivariate von Mises distribution. Int J Intell Syst 32(2):153–176
Rosenthal M, Wu W, Klassen E, Srivastava A (2014) Spherical regression models using projective linear transformations. J Am Stat Assoc 109(508):1615–1624
Rothman ED (1972) Tests for uniformity of a circular distribution. Sankhyā, Ser A 34(1):23–32
Roy A, Pal A, Garain U (2017) JCLMM: A finite mixture model for clustering of circular-linear data and its application to psoriatic plaque segmentation. Pattern Recognit 66:160–173
Rueda C, Fernández MA, Peddada SD (2009) Estimation of parameters subject to order restrictions on a circle with application to estimation of phase angles of cell cycle genes. J Am Stat Assoc 104(485):338–347
Rueda C, Fernández MA, Barragán S, Peddada SD (2015) Some advances in constrained inference for ordered circular parameters in oscillatory systems. In: Dryden IL, Kent JT (eds) Geometry driven statistics. Wiley series in probability and statistics. Wiley, Chichester, pp 97–114
Rueda C, Fernández MA, Barragán S, Mardia KV, Peddada SD (2016) Circular piecewise regression with applications to cell-cycle data. Biometrics 72(4):1266–1274

Rumcheva P, Presnell B (2017) An improved test of equality of mean directions for the Langevin-von Mises-Fisher distribution. Aust N Z J Stat 59(1):119–135
Sadikon NH, Ibrahim AIN, Mohamed I, Shimizu K (2019) A new test of discordancy in cylindrical data. Commun Stat Simul Comput 48(8):2512–2522
Sahoo I, Guinness J, Reich BJ (2019) A test for isotropy on a sphere using spherical harmonic functions. Stat Sin 29(3):1253–1276
Salah A, Nadif M (2017) Social regularized von Mises–Fisher mixture model for item recommendation. Data Min Knowl Discov 31(5, SI):1218–1241
Salah A, Nadif M (2019) Directional co-clustering. Adv Data Anal Classif 13(3):591–620
Sargsyan K, Wright J, Lim C (2012) GeoPCA: a new tool for multivariate analysis of dihedral angles based on principal component geodesics. Nucleic Acids Res 40(3):e25–e25
Sargsyan K, Hua YH, Lim C (2015) Clustangles: an open library for clustering angular data. J Chem Inf Modeling 55(8):1517–1520
Sarma YR, Jammalamadaka SR (1993) Circular regression. In: Matsusita K, Puri ML, Hayakawa T (eds) Statistical science and data analysis. VSP, Utrecht, pp 109–128
Sau MF, Rodriguez D (2018) Minimum distance method for directional data and outlier detection. Adv Data Anal Classif 12(3):587–603
Saw JG (1983) Dependent unit vectors. Biometrika 70(3):665–671
Scealy JL, Welsh AH (2011) Regression for compositional data by using distributions defined on the hypersphere. J R Stat Soc Ser B Stat Methodol 73(3):351–375
Scealy JL, Welsh AH (2014a) Colours and cocktails: compositional data analysis: 2013 Lancaster lecture. Aust N Z J Stat 56(2):145–169
Scealy JL, Welsh AH (2014b) Fitting Kent models to compositional data with small concentration. Stat Comput 24(2):165–179
Scealy JL, Welsh AH (2017) A directional mixed effects model for compositional expenditure data. J Am Stat Assoc 112(517):24–36
Scealy JL, Wood ATA (2019) Scaled von Mises–Fisher distributions and regression models for paleomagnetic directional data. J Am Stat Assoc 114(528):1547–1560
Schlather M, Malinowski A, Menck PJ, Oesting M, Strokorb K (2015) Analysis, simulation and prediction of multivariate random fields with package RandomFields. J Stat Softw 63(8):1–25
Schulz J, Jung S, Huckemann S, Pierrynowski M, Marron JS, Pizer SM (2015) Analysis of rotational deformations from directional data. J Comput Graph Stat 24(2):539–560
Scott JG (2011) Bayesian estimation of intensity surfaces on the sphere via needlet shrinkage and selection. Bayesian Anal 6(2):307–327
Self SG, Liang KY (1987) Asymptotic properties of maximum likelihood estimators and likelihood ratio tests under nonstandard conditions. J Am Stat Assoc 82(398):605–610
SenGupta A, Bhattacharya S (2015) Finite mixture-based Bayesian analysis of linear-circular models. Environ Ecol Stat 22(4):667–679
SenGupta A, Laha AK (2008) A likelihood integrated method for exploratory graphical analysis of change point problem with directional data. Commun Stat Theory Methods 37(11–12):1783–1791
SenGupta A, Pal C (2001) On optimal tests for isotropy against the symmetric wrapped stable-circular uniform mixture family. J Appl Stat 28(1):129–143
SenGupta A, Roy S (2005) A simple classification rule for directional data. In: Balakrishnan N, Nagaraja HN, Kannan N (eds) Advances in ranking and selection, multiple comparisons, and reliability. Statistics for industry and technology. Birkhäuser, Boston, pp 81–90
SenGupta A, Ugwuowo FI (2011) A classification method for directional data with application to the human skull. Commun Stat Theory Methods 40(3):457–466
Shieh GS, Johnson RA (2005) Inference based on a bivariate distribution with von Mises marginals. Ann Inst Stat Math 57(4):789–802
Singh H, Hnizdo V, Demchuk E (2002) Probabilistic model for two dependent circular variables. Biometrika 89(3):719–723
Sinz F, Berens B, Kuemmerer M, Wallis T (2018) PyCircStat: circular statistics with Python. https://github.com/circstat/pycircstat
Sittel F, Filk T, Stock G (2017) Principal component analysis on a torus: theory and application to protein dynamics. J Chem Phys 147(24):244101
Sklar M (1959) Fonctions de répartition à n dimensions et leurs marges. Publ Inst Stat Univ Paris 8:229–231
Small CG (1987) Measures of centrality for multivariate and directional distributions. Can J Stat 15(1):31–39

Soetaert K (2019) plot3D: plotting multi-dimensional data. R package version 1.3 https://CRAN.R-project.org/package=plot3D

Sommer S (2013) Horizontal dimensionality reduction and iterated frame bundle development. In: Nielsen F, Barbaresco F (eds) Geometric science of information. Lecture notes in computer science, vol 8085. Springer, Berlin, pp 76–83

Sommer S (2019) An infinitesimal probabilistic model for principal component analysis of manifold valued data. Sankhyā, Ser A 81(1):37–62

Sommer S, Lauze F, Nielsen M (2014) Optimization over geodesics for exact principal geodesic analysis. Adv Comput Math 40(2):283–313

Souden M, Kinoshita K, Nakatani T (2013) An integration of source location cues for speech clustering in distributed microphone arrays. In: 2013 IEEE International conference on acoustics, speech and signal processing. IEEE, New York, pp 111–115

Soukissian TH (2014) Probabilistic modeling of directional and linear characteristics of wind and sea states. Ocean Eng 91:91–110

Sra S (2018) Directional statistics in machine learning: a brief review. In: Ley C, Verdebout T (eds) Applied directional statistics. Chapman & Hall/CRC interdisciplinary statistics series. CRC Press, Boca Raton, pp 259–276

Sra S, Karp D (2013) The multivariate Watson distribution: maximum-likelihood estimation and other aspects. J Multivar Anal 114:256–269

Stephens MA (1982) Use of the von Mises distribution to analyse continuous proportions. Biometrika 69(1):197–203

Straub J, Chang J, Freifeld O, Fisher JWI (2015) A Dirichlet process mixture model for spherical data. In: Lebanon G, Vishwanathan SVN (eds) Proceedings of the eighteenth international conference on artificial intelligence and statistics. Proceedings of machine learning research, vol 38. PMLR, San Diego, pp 930–938

Su J, Kurtek S, Klassen E, Srivastava A (2014) Statistical analysis of trajectories on Riemannian manifolds: bird migration, hurricane tracking and video surveillance. Ann Appl Stat 8(1):530–552

Su Y, Wu XK (2011) Smooth test for uniformity on the surface of a unit sphere. In: 2011 International conference on machine learning and cybernetics, IEEE, New York, pp 867–872

Sun SZ, Lockhart RA (2019) Bayesian optimality for Beran's class of tests of uniformity around the circle. J Stat Plan Inference 198:79–90

Taghia J, Ma Z, Leijon A (2014) Bayesian estimation of the von-Mises Fisher mixture model with variational inference. IEEE Trans Pattern Anal Mach Intell 36(9):1701–1715

Taijeron HJ, Gibson AG, Chandler C (1994) Spline interpolation and smoothing on hyperspheres. SIAM J Sci Comput 15(5):1111–1125

Takasu Y, Yano K, Komaki F (2018) Scoring rules for statistical models on spheres. Stat Probab Lett 138:111–115

Tang H, Chu SM, Huang TS (2009) Generative model-based speaker clustering via mixture of von Mises-Fisher distributions. In: 2009 IEEE international conference on acoustics, speech and signal processing, IEEE, New York, pp 4101–4104

Taniguchi M, Kato S, Ogata H, Pewsey A (2020) Models for circular data from time series spectra. J Time Ser Anal 41(6):809–829

Taylor CC (2008) Automatic bandwidth selection for circular density estimation. Comput Stat Data Anal 52(7):3493–3500

Taylor CC, Lafratta G, Fensore S (2018) nprotreg: nonparametric rotations for sphere-sphere regression. R package version 1.0.1 https://CRAN.R-project.org/package=nprotreg

Tipping ME, Bishop CM (1999) Probabilistic principal component analysis. J R Stat Soc Ser B Stat Methodol 61(3):611–622

Traa J, Smaragdis P (2013) A wrapped Kalman filter for azimuthal speaker tracking. IEEE Signal Process Lett 20(12):1257–1260

Tsagris M, Alenazi A (2019) Comparison of discriminant analysis methods on the sphere. Commun Stat: Case Stud Data Anal Appl 5(4):467–491

Tsagris M, Athineou G, Sajib A, Amson E, Waldstein MJ (2020) Directional: directional statistics. R package version 4.4 https://CRAN.R-project.org/package=Directional

Tsai MT (2009) Asymptotically efficient two-sample rank tests for modal directions on spheres. J Multivar Anal 100:445–458

Tsuruta Y, Sagae M (2017a) Asymptotic property of wrapped Cauchy kernel density estimation on the circle. Bull Inform Cybernet 49:1–10
Tsuruta Y, Sagae M (2017b) Higher order kernel density estimation on the circle. Stat Probab Lett 131:46–50
Tsuruta Y, Sagae M (2018) Properties for circular nonparametric regressions by von Miese and wrapped Cauchy kernels. Bull Inform Cybernet 50:1–13
Tsuruta Y, Sagae M (2020) Theoretical properties of bandwidth selectors for kernel density estimation on the circle. Ann Inst Stat Math 72(2):511–530
Tung DD, Jammalamadaka SR (2013) On the Gini mean difference test for circular data. Commun Stat Theory Methods 42(11):1998–2008
Umbach D, Jammalamadaka SR (2009) Building asymmetry into circular distributions. Stat Probab Lett 79(5):659–663
van der Maaten LJP, Hinton GE (2008) Visualizing high-dimensional data using t-SNE. J Mach Learn Res 9(Nov):2579–2605
van der Vaart AW (2000) Asymptotic statistics. Cambridge series in statistical and probabilistic mathematics. Cambridge University Press, Cambridge
Veeraraghavan A, Srivastava A, Roy-Chowdhury AK, Chellappa R (2009) Rate-invariant recognition of humans and their activities. IEEE Trans Image Process 18(6):1326–1339
Verdebout T (2015) On some validity-robust tests for the homogeneity of concentrations on spheres. J Nonparametr Stat 27(3):372–383
Verdebout T (2017) On the efficiency of some rank-based test for the homogeneity of concentrations. J Stat Plan Inference 191:101–109
Vuollo V, Holmstrom L (2018) A scale space approach for exploring structure in spherical data. Comput Stat Data Anal 125:57–69
Vuollo V, Holmström L, Aarnivala H, Harila V, Heikkinen T, Pirttiniemi P, Valkama AM (2016) Analyzing infant head flatness and asymmetry using kernel density estimation of directional surface data from a craniofacial 3D model. Stat Med 35(26):4891–4904
Wang F (2013) Space and space-time modeling of directional data. Ph.D. Thesis, Duke University
Wang F, Gelfand AE (2013) Directional data analysis under the general projected normal distribution. Stat Methodol 10(1):113–127
Wang F, Gelfand AE (2014) Modeling space and space-time directional data using projected Gaussian processes. J Am Stat Assoc 109(508):1565–1580
Wang F, Gelfand AE, Jona-Lasinio G (2015) Joint spatio-temporal analysis of a linear and a directional variable: space-time modeling of wave heights and wave directions in the Adriatic Sea. Stat Sin 25(1):25–39
Wang J, Boyer J, Genton MG (2004) A skew-symmetric representation of multivariate distributions. Stat Sin 14(4):1259–1270
Wang M, Shimizu K (2012) On applying Möbius transformation to cardioid random variables. Stat Methodol 9(6):604–614
Wang M, Wang D (2016) VMF-SNE: embedding for spherical data. In: 2016 IEEE international conference on acoustics, speech and signal processing (ICASSP). IEEE, New York, pp 2344–2348
Wang X, Zhao L (2001) Laws of the iterated logarithm for kernel estimator of density function of spherical data. J Syst Sci Math Sci 21(3):264–273
Wang X, Zhao L, Wu Y (2000) Distribution free laws of the iterated logarithm for kernel estimator of regression function based on directional data. Chinese Ann Math Ser B 21(4):489–498
Wang XM (2002) Exponential bounds of mean error for the kernel regression estimates with directional data. Chinese Ann Math Ser A 23(1):55–62
Wang XM, Ma L (2000) Nearest neighbor estimator for density function of directional data. J Biomath 15(3):332–338
Wang XM, Zhao LC (2003) A law of logarithm for kernel density estimators with directional data. Acta Math Sinica Chin Ser 46(5):865–874
Watamori Y, Jupp PE (2005) Improved likelihood ratio and score tests on concentration parameters of von Mises–Fisher distributions. Stat Probab Lett 72(2):93–102
Watson GS (1961) Goodness-of-fit tests on a circle. Biometrika 48(1/2):109–114
Watson GS (1983) Statistics on spheres. University of Arkansas lecture notes in the mathematical sciences. Wiley, New York
Wehrly TE, Johnson RA (1980) Bivariate models for dependence of angular observations and a related Markov process. Biometrika 67(1):255–256

Wilson RC, Hancock ER, Pekalska E, Duin RPW (2014) Spherical and hyperbolic embeddings of data. IEEE Trans Pattern Anal Mach Intell 36(11):2255–2269

Wood SN (2017) Generalized additive models, 2nd edn. Chapman & Hall/CRC texts in statistical science series. CRC Press, Boca Raton

Wouters H, Thas O, Ottoy JP (2009) Data-driven smooth tests and a diagnostic tool for lack-of-fit for circular data. Aust N Z J Stat 51(4):461–480

Xu D, Wang Y (2020) Area-proportional visualization for circular data. J Comput Graph Stat 29(2):351–357

Yamaji A, Sato K (2011) Clustering of fracture orientations using a mixed Bingham distribution and its application to paleostress analysis from dike or vein orientations. J Struct Geol 33(7):1148–1157

Yang MS, Pan JA (1997) On fuzzy clustering of directional data. Fuzzy Sets Syst 91(3):319–326

Yang MS, Chang-Chien SJ, Kuo HC (2014) On mean shift clustering for directional data on a hypersphere. In: Rutkowski L, Korytkowski M, Scherer R, Tadeusiewicz R, Zadeh LA, Zurada JM (eds) Artificial intelligence and soft computing. Lecture notes in compututer scence, vol 8468. Springer, Cham, pp 809–818

Yeh SY, Harris KDM, Jupp PE (2013) A drifting Markov process on the circle, with physical applications. Proc R Soc A 469(2156):20130092

You K (2020) RiemBase: functions and C++ header files for computation on manifolds. R package version 0.2.4 https://CRAN.R-project.org/package=RiemBase

Yuan T (2021) The 8-parameter Fisher–Bingham distribution on the sphere. Comput Stat 36:409–420

Zhan X, Ma T, Liu S, Shimizu K (2019) On circular correlation for data on the torus. Stat Pap 60(6):1827–1847

Zhang L, Li Q, Guo Y, Yang Z, Zhang L (2018a) An investigation of wind direction and speed in a featured wind farm using joint probability distribution methods. Sustainability 10(12):4338

Zhang M, Fletcher T (2013) Probabilistic principal geodesic analysis. In: Burges CJC, Bottou L, Welling M, Ghahramani Z, Weinberger KQ (eds) Advances in neural information processing systems 26. Curran Associates, Red Hook, pp 1178–1186

Zhang Z, Klassen E, Srivastava A (2018b) Phase-amplitude separation and modeling of spherical trajectories. J Comput Graph Stat 27(1):85–97

Zhang Z, Klassen E, Srivastava A (2019) Robust comparison of kernel densities on spherical domains. Sankhyā, Ser A 81(1):144–171

Zhao L, Wu C (2001) Central limit theorem for integrated squared error of kernel estimators of spherical density. Sci China, Ser A Math 44(4):474–483

Zou G, Hua J, Muzik O (2007) Non-rigid surface registration using spherical thin-plate splines. In: Ayache N, Ourselin S, Maeder A (eds) Medical image computing and computer-assisted intervention - MICCAI 2007. Springer, Berlin, pp 367–374

Zucchini W, MacDonald IL, Langrock R (2016) Hidden Markov models for time series. Monographs on statistics and applied probability, vol 150, 2nd edn. CRC Press, Boca Raton