ORIGINAL PAPER

Tractable circula densities from Fourier series

Shogo Kato[1] · Arthur Pewsey[2] · M. C. Jones[3]

Abstract

This article proposes an approach, based on infinite Fourier series, to constructing tractable densities for the bivariate circular analogues of copulas recently coined 'circulas'. As examples of the general approach, we consider circula densities generated by various patterns of nonzero Fourier coefficients. The shape and sparsity of such arrangements are found to play a key role in determining the properties of the resultant models. The special cases of the circula densities we consider all have simple closed-form expressions involving no computationally demanding normalizing constants and display wide-ranging distributional shapes. A highly successful model identification tool and methods for parameter estimation and goodness-of-fit testing are provided for the circula densities themselves and the bivariate circular densities obtained from them using a marginal specification construction. The modelling capabilities of such bivariate circular densities are compared with those of five existing models in a numerical experiment, and their application illustrated in an analysis of wind directions.

✉ Arthur Pewsey
  apewsey@unex.es

  Shogo Kato
  skato@ism.ac.jp

  M. C. Jones
  chris.jones@open.ac.uk

[1] Institute of Statistical Mathematics, 10-3 Midori-cho, Tachikawa 190-8562, Tokyo, Japan

[2] Department of Mathematics, Escuela Politécnica, University of Extremadura, Cáceres 10003, Spain

[3] School of Mathematics and Statistics, The Open University, Walton Hall, Milton Keynes MK7 6AA, UK

🍂 Springer

1 Introduction

A direction observed in the plane $\mathbb{R}^2$ can be represented by an angle, Θ, typically in $[0, 2\pi)$ or $[-\pi, \pi)$, measured in a specified direction from a specified origin. The natural support for such directions is the circumference of the unit circle, $\mathbb{S}^1$, data on them being referred to as *circular*. The natural support for a bivariate circular random vector with angular coordinates (Θ_1, Θ_2) is the unit torus $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$, data on such vectors sometimes being referred to as 'toroidal'.

Early constructions used to obtain models for toroidal data included maximum (Shannon) entropy characterization (or, equivalently, exponential family distributions, Mardia and Jupp 1999, p. 43), wrapping (Johnson and Wehrly 1977) and projection (Saw 1983). The latter two approaches can be applied to any bivariate distribution defined on $\mathbb{R}^2$, although the resulting toroidal densities generally cannot be expressed in closed-form or are highly convoluted.

Perhaps the best-known exponential family toroidal model is the eight-parameter bivariate von Mises distribution of Mardia (1975). Four of its parameters control the dependence between Θ_1 and Θ_2, but not in easily interpreted ways. Its submodels include the six-parameter model of Rivest (1988) and the latter's five-parameter sine, cosine and hybrid submodels, the properties of which are studied at length in Mardia and Frellsen (2012). Such five-parameter models are, to some extent, toroidal analogues of the bivariate normal distribution, with four of their parameters controlling the locations and concentrations of the marginal distributions, and the fifth the dependence between Θ_1 and Θ_2. However, their concentration and dependence parameters also control the unimodality/bimodality of their densities. Like the bivariate normal, the densities of their unimodal cases have contours that are elliptical around the mode. Their normalizing constants must be computed numerically. Their conditional distributions are von Mises, but their marginal distributions are generally not and, for some parameter values, can be bimodal.

More recently, Navarro et al. (2017) conditioned a multivariate normal distribution to obtain the twelve-parameter bivariate generalized von Mises model of order 2 (BGvM$_2$), whose conditional distributions are second-order generalized von Mises (GvM$_2$), see, for example, Gatto (2008), while Hassanzadeh and Kalaylioglu (2018) used a conditional specification approach to obtain a toroidal model with one of its marginal distributions being GvM$_2$. GvM$_2$ densities can be symmetric or asymmetric and unimodal or bimodal. A highly flexible family of toroidal models obtained by normalizing time series spectra was proposed by Taniguchi et al. (2020). In general, the interpretation of the parameters of these various models is difficult, and their normalizing constants must be computed numerically.

Ameijeiras-Alonso and Ley (2022) extended the sine-skewing approach of Umbach and Jammalamadaka (2009) and Abe and Pewsey (2011) to generate models for mildly

asymmetric toroidal data, focussing explicitly on sine-skewed extensions of the bivariate uniform, sine, cosine and wrapped Cauchy models. An appealing property of such models is that their normalizing constants are the same as those of the base models to which sine-skewing is applied.

The construction underpinning the proposal made here is a highly flexible marginal specification approach incorporating the circular analogues of bivariate copulas recently termed 'circulas' by Jones et al. (2015). Briefly put, circulas are copulas that are rescaled by $1/(2\pi)$ and satisfy periodicity constraints. Let (Θ_1, Θ_2) be a bivariate random vector of continuous circular random variables and F its distribution function on $[0, 2\pi)^2$ defined by $F(\theta_1, \theta_2) = P(\Theta_1 \in (\omega_1, \omega_1 + \theta_1], \Theta_2 \in (\omega_2, \omega_2 + \theta_2])$ $(0 \leq \theta_1, \theta_2 < 2\pi)$, where $0 \leq \omega_1, \omega_2 < 2\pi$ are arbitrary reference points. Define the marginal distribution function of Θ_j on $[0, 2\pi)$ by $F_j(\theta_j) = P(\Theta_j \in (\omega_j, \omega_j + \theta_j])$ $(0 \leq \theta_j < 2\pi)$ for $j = 1, 2$. Then, an analogous result to Sklar's theorem states that the distribution function F can be related to the marginal distribution functions, F_1 and F_2, through the *circula* $C_\circ$ as

$$F(\theta_1, \theta_2) = C_\circ(2\pi F_1(\theta_1), 2\pi F_2(\theta_2)) \quad (0 \leq \theta_1, \theta_2 < 2\pi).$$

$C_\circ$ is itself a bivariate circular distribution function, but with the specific reference points $\omega_1 = \omega_2 = 0$ and marginal distributions for $\Psi_1 = 2\pi F_1(\Theta_1)$ and $\Psi_2 = 2\pi F_2(\Theta_2)$ that are circular uniform rather than the (linear) uniform ones of their copula counterparts. Taking partial derivatives, the density of (Θ_1, Θ_2) is given by

$$f(\theta_1, \theta_2) = 4\pi^2 f_1(\theta_1) f_2(\theta_2) c_\circ(2\pi F_1(\theta_1), 2\pi F_2(\theta_2)), \tag{1}$$

where f_1 and f_2 are the marginal densities of Θ_1 and Θ_2, and $c_\circ$ is the circula density corresponding to $C_\circ$. Clearly, $c_\circ$ controls the dependence between Θ_1 and Θ_2, the circula density for independence being $c_\circ(\psi_1, \psi_2) = 1/(4\pi^2)$ for any $\psi_1, \psi_2 \in [0, 2\pi)$, where $\psi_1 = 2\pi F_1(\theta_1)$ and $\psi_2 = 2\pi F_2(\theta_2)$. The domain of $c_\circ$ is extended to $\mathbb{R}^2$ through the assumption $c_\circ(\psi_1 \pm 2\pi k, \psi_2 \pm 2\pi l) = c_\circ(\psi_1, \psi_2)$ $(k, l = 0, 1, \ldots)$. In addition, $c_\circ$ is usually assumed to be continuous at $(2\pi k, \psi_2)$ and $(\psi_1, 2\pi l)$ for any $\psi_1, \psi_2 \in [0, 2\pi)$ as well as all other points on the torus.

Jones et al. (2015) provided an in-depth treatment of the Wehrly and Johnson (1980) class of circulas, with density

$$c_\circ(\psi_1, \psi_2) = \frac{1}{2\pi} g(\psi_1 \pm \psi_2), \tag{2}$$

where g is a circular density. See Jones et al. (2015) for details of the evolution of this class, Shieh and Johnson (2005) and Kato and Pewsey (2015) for special cases, and Pewsey and Kato (2016) for work on goodness-of-fit testing. The wider copula-related literature is summarized in Jones et al. (2015), and generalizations of copulas to other compact Riemannian manifolds have been considered in Jupp (2015); see also Jupp and Kume (2020).

Here we propose a general, Fourier series-based, approach to constructing tractable circula densities for use within (1) that includes (2) as a special case and focus on $c_\circ$

generated using five basic patterns of Fourier coefficients. The construction is attractive because: (i) various new circula models can be generated having simple closed-form expressions for their densities and flexible forms of dependence structure controlled by relatively few parameters; (ii) three well-known circular dependence measures are simple functions of just two of the Fourier coefficients; (iii) the conditional mean directions and mean resultant lengths depend only on a limited number of Fourier coefficients; (iv) methods for simulation, model identification, parameter estimation and testing goodness-of-fit and independence are available for both the $c_\circ$ and the bivariate circular distributions generated from using them within (1).

Although any forms of f_1 and f_2 can be employed within (1) to generate bivariate circular distributions, throughout the paper we illustrate the construction's application using marginal circular densities from the Kato and Jones (2015) family. The latter is the most flexible four-parameter family of unimodal circular densities presently available, has closed-form expressions for both its density and distribution functions, and all four of its parameters have clear interpretations. As we shall see, most of the conditional distributions of the specific circula densities considered in Sect. 3 are also members of the Kato–Jones family. We note that the $c_\circ$ that we explore here might also be used in a semiparametric approach to modelling in which kernel density estimates and empirical distribution functions are used in place of f_1, f_2, F_1 and F_2 in (1). Such an approach would inherit the tractable and interpretable parametric modelling of dependence.

We stress that the approach based on *infinite* Fourier series introduced here differs in important ways from previously proposed constructions incorporating *truncated* (or *partial*) Fourier series. Fernández-Durán (2007) considered bivariate circular models based on (1), the Wehrly and Johnson (1980) circula density (2) and non-negative trigonometric sum densities (i.e. truncated Fourier series constrained to be non-negative) for f_1, f_2 and g. More generally, Pertsemlidis et al. (2005) and Fernández-Durán and Gregorio-Domínguez (2014) proposed toroidal densities obtained from truncated bivariate Fourier series, the latter constrained to be non-negative. When fitted to toroidal data, such models generally include large numbers of parameters that are difficult to interpret (Mardia et al. 2007) and, due to harmonic arte-facts, manifest modes, sometimes multiple, unsupported by the data (see, for example, Fig. 4 of Fernández-Durán and Gregorio-Domínguez 2014). In contrast, our tractable $c_\circ$, as well as the toroidal models generated by using them within (1) in combination with the highly flexible unimodal circular marginal densities of Kato and Jones (2015), have relatively few parameters, all of which have clear interpretations, and closed-form expressions for their densities which involve no computationally demanding normalizing constants. As an important consequence of the latter property, numerically implemented maximum likelihood estimation is swift. While truncated bivariate Fourier series with Fourier coefficients satisfying the constraints in (6) and (4) to fol-low might be used as circulas, they will generally not be as tractable as the ones based on infinite Fourier series that we propose here.

The rest of the article is organized as follows. In Sect. 2, we provide the details of our proposed construction together with general results for three circular dependence measures and the conditional mean directions and mean resultant lengths of circulas generated using it. In Sect. 3, we consider $c_\circ$ generated using five basic patterns of

Fourier coefficients within the proposed construction and provide details of their basic properties. In Sect. 4, we explain how model identification, parameter estimation and goodness-of-fit testing can be performed, for both the considered $c_\circ$ and the models for toroidal data derived from using them within (1). The results from a numerical experiment, designed to compare the large-sample modelling capabilities of such toroidal models with those of five existing bivariate circular models, are presented in Sect. 5. In Sect. 6, a new Wehrly and Johnson (1980) circula density and our proposed inferential methods are applied in the analysis of wind directions. Lastly, we offer some concluding remarks in Sect. 7. All equations, figures and tables with numbers preceded by the letter S are contained in an accompanying online supplementary materials document.

2 Circula densities from Fourier series

2.1 Circula densities

It is well known (Mardia and Jupp 1999, Sect. 3.3.2) that any continuous circular density, f, can be expressed in the form of a Fourier series as

$$f(\theta) = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \phi(m) e^{-im\theta} \quad (-\pi \le \theta < \pi),$$

where $i = \sqrt{-1}$, for appropriately chosen Fourier coefficients $\phi(m)$ ($m \in \mathbb{Z}$). Now consider the family of continuous distributions on the torus whose density can be expressed analogously as

$$f(\theta_1, \theta_2) = \frac{1}{4\pi^2} \sum_{m,n=-\infty}^{\infty} \phi(m, n) e^{-i(m\theta_1+n\theta_2)} \quad (-\pi \le \theta_1, \theta_2 < \pi), \quad (3)$$

where the Fourier coefficients $\phi(m, n) \in \mathbb{C}$ ($m, n \in \mathbb{Z}$) are appropriately defined so that $f(\theta_1, \theta_2) \ge 0$ and $\int_{-\pi}^{\pi} \int_{-\pi}^{\pi} f(\theta_1, \theta_2) d\theta_1 d\theta_2 = 1$.

Proposition 1 *The following hold for density* (3).

(i) *If a random vector* (Θ_1, Θ_2) *has density* (3) *then*

$$E\left\{e^{i(m\Theta_1+n\Theta_2)}\right\} = \phi(m, n) \quad (m, n \in \mathbb{Z}).$$

(ii) *A density in family* (3) *is a circula density if and only if*

$$\phi(m, 0) = \begin{cases} 1, & m = 0, \\ 0, & m \ne 0, \end{cases} \quad \phi(0, n) = \begin{cases} 1, & n = 0, \\ 0, & n \ne 0, \end{cases} \quad (m, n \in \mathbb{Z}). \quad (4)$$

Proof (i) It is clear from the following well-known equation that (i) holds:

$$\int_{-\pi}^{\pi} e^{it\theta} \, d\theta = \begin{cases} 2\pi, & t = 0, \\ 0, & t \in \mathbb{Z}\backslash\{0\}. \end{cases} \tag{5}$$

(ii) If $\phi(m, n)$ satisfies (4), the marginal density $f_1(\theta_1)$ can be expressed as

$$f_1(\theta_1) = \int_{-\pi}^{\pi} f(\theta_1, \theta_2) d\theta_2 = \frac{1}{2\pi} \sum_{m=-\infty}^{\infty} \phi(m, 0) e^{-im\theta_1} = \frac{1}{2\pi}.$$

The second equality follows from (5). Similarly, $f_2(\theta_2) = 1/(2\pi)$, and thus (3) is a circula density. Next, assume that a density in family (3) is a circula density. Since this assumption implies that each marginal density of (3) is the circular uniform density, the equations in (4) follow from (i). □

It follows from Proposition 1(i) that, for any $\phi(m, n)$ in (3),

$$\overline{\phi(m, n)} = \phi(-m, -n), \quad |\phi(m, n)| \leq 1, \quad \phi(0, 0) = 1, \tag{6}$$

where $\bar{z}$ denotes the complex conjugate of z. The circula densities of Proposition 1(ii) can be expressed as

$$c_{\circ}(\psi_1, \psi_2) = \frac{1}{4\pi^2} \left[1 + 2 \, \mathrm{Re} \left\{ \sum_{m=1}^{\infty} \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \phi(m, n) \, e^{-i(m\psi_1 + n\psi_2)} \right\} \right]$$

$$= \frac{1}{4\pi^2} \left[1 + 2 \, \mathrm{Re} \left\{ \sum_{n=1}^{\infty} \sum_{\substack{m=-\infty \\ m \neq 0}}^{\infty} \phi(m, n) \, e^{-i(m\psi_1 + n\psi_2)} \right\} \right]. \tag{7}$$

If $\phi(m, n) = 0$ for all $m, n \neq 0$, the density is that of the bivariate circular uniform distribution. Circula densities of the form (7) are the main focus of the paper.

2.2 Circular dependence measures

Here we provide general results for three signed circular dependence measures when applied to circula densities of the form (7). Let (Θ_1, Θ_2) be a bivariate circular random vector. Then, the dependence measures of Fisher and Lee (1983), Jammalamadaka and Sarma (1988) and, for a circula, that of Rivest (1982) are defined as

$$\rho_{\mathrm{FL}} = \frac{\det\{E(X_1 X_2^T)\}}{[\det\{E(X_1 X_1^T)\}\det\{E(X_2 X_2^T)\}]^{1/2}},$$

$$\rho_{\mathrm{JS}} = \frac{E\{\sin(\Theta_1 - \mu_1) \sin(\Theta_2 - \mu_2)\}}{[E\{\sin^2(\Theta_1 - \mu_1)\} E\{\sin^2(\Theta_2 - \mu_2)\}]^{1/2}},$$

$$\rho_{\mathrm{R}} = 2\lambda_2,$$

respectively, where, for $j = 1, 2$, $X_j = (\cos \Theta_j, \sin \Theta_j)^T$, μ_j is the mean direction of Θ_j, and λ_2 denotes the smallest singular value of $E(X_1 X_2^T)$.

For circula densities (7), it can easily be shown that the dependence measures between the circular uniform random variables Ψ_1 and Ψ_2 reduce to

$$\rho_R = \rho_{JS} = |\phi(1, -1)| - |\phi(1, 1)|, \quad \rho_{FL} = |\phi(1, -1)|^2 - |\phi(1, 1)|^2.$$

Clearly, when only one of $\phi(1, -1)$ and $\phi(1, 1)$ is nonzero, then ρ_R, ρ_{JS} and ρ_{FL} are simple functions of either $\phi(1, -1)$ or $\phi(1, 1)$.

2.3 Conditional mean directions and resultant lengths

The first trigonometric moment of $\Psi_1 | \Psi_2 = \psi_2$ for a circula density of the form (7) can be expressed as

$$E(e^{i\Psi_1} | \psi_2) = \int_{-\pi}^{\pi} e^{i\psi_1} c_{1|2}(\psi_1 | \psi_2) \mathrm{d}\psi_1 = \sum_{n=-\infty}^{\infty} \phi(1, n) e^{-in\psi_2}, \qquad (8)$$

where $c_{1|2}(\psi_1 | \psi_2)$ denotes the conditional density of $\Psi_1 | \Psi_2 = \psi_2$. Expression (8) implies that, in order to calculate the mean direction and mean resultant length of $\Psi_1 | \Psi_2 = \psi_2$, it suffices to focus on the nonzero Fourier coefficients in the $m = 1$ column of a plot of the $\{\phi(m, n)\}_{m,n \in \mathbb{Z}}$. Figure 1 provides examples of such plots. It then follows that:

(a) If all the coefficients in the $m = 1$ column are zero, the mean resultant length of the conditional distribution of $\Psi_1 | \Psi_2 = \psi_2$, $R(\Psi_1 | \psi_2) \equiv |E(e^{i\Psi_1} | \psi_2)|$, is zero;

(b) If there is only one nonzero coefficient in the $m = 1$ column, say at $(m, n) = (1, d)$, the mean direction of the conditional distribution of $\Psi_1 | \Psi_2 = \psi_2$, $M(\Psi_1 | \psi_2) \equiv \arg(E(e^{i\Psi_1} | \psi_2))$, is $-d\psi_2$ and $R(\Psi_1 | \psi_2) = |\phi(1, d)|$;

(c) If there are two or more nonzero coefficients in the $m = 1$ column, $M(\Psi_1 | \psi_2)$ is nonlinear and $R(\Psi_1 | \psi_2)$ is in general heteroscedastic.

Analogous results hold for the conditional distribution of $\Psi_2 | \Psi_1 = \psi_1$.

3 Circula models generated by five patterns of Fourier coefficients

3.1 Preliminaries

In this section we consider examples of (7) generated by five simple patterns of nonzero Fourier coefficients and provide details of the basic properties of certain specific models. Given the constraints on the Fourier coefficients in (6) and (4) , unless explicitly stated otherwise, we consider circula densities defined through nonzero Fourier coefficients on the $\mathbb{Z}^+ \times (\mathbb{Z}^+ \cup \mathbb{Z}^-)$ lattice. Figure 1 illustrates the five patterns of nonzero Fourier coefficients. We emphasize the parametric specification of these coefficients. In particular, the specific models generated by Patterns 2–5 are derived using geometric series of nonzero Fourier coefficients which have the added benefit of generating
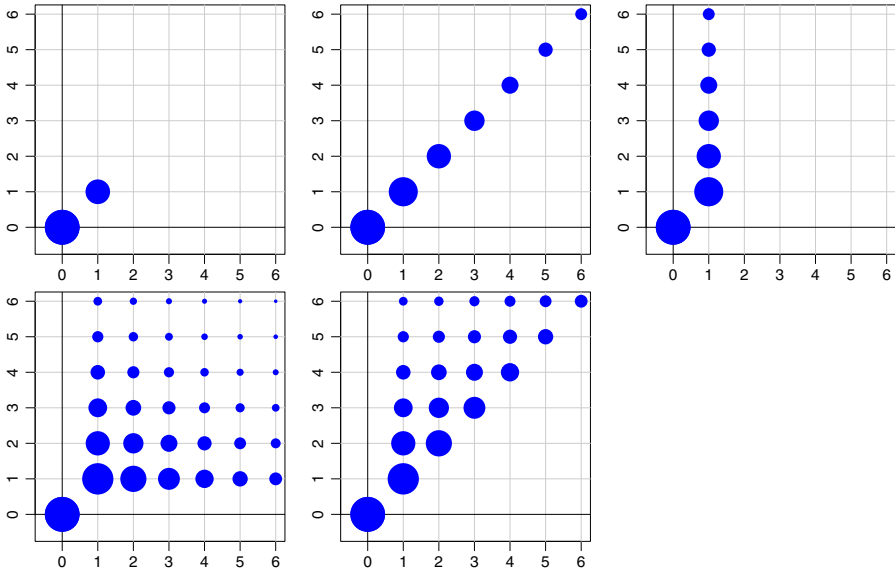
Fig. 1 The nonzero Fourier coefficients of Patterns 1–5 with $q = -1$ and $m, n \in \{0, 1, \ldots, 6\}$. The values of m appear on the horizontal axes, and those for n on the vertical axes. The area of a dot at (m, n) is proportional to the value of $\phi(m, n)$. The panels for the five patterns are ordered from top left to bottom right, reading from left to right

circulas with closed-form expressions for their densities. In order to obtain circula densities exhibiting pointwise symmetry about the origin, all of the nonzero Fourier coefficients are assumed to be real.

Since all of the patterns we consider have at most one of $\phi(1, -1)$ and $\phi(1, 1)$ which is nonzero, we only quote values of ρ_R, those for ρ_{JS} and ρ_{FL} following from the relations identified in Sect. 2.2. The conditional mean directions and resultant lengths of all of the models we consider can be easily calculated using the results presented in Sect. 2.3; see Appendix A of the supplementary materials document. It follows that the conditionals of the models considered afford uniform, linear-homoscedastic, or nonlinear-heteroscedastic circular–circular regression. These properties and expressions for modes and antimodes of the circula densities we consider can be found in Table 1.

Appendix B of the supplementary materials document explains how data from any $c_\circ$ or bivariate circular density of the form (1) can be simulated.

3.2 Pattern 1: single (nonzero) point

Let

$$\phi(m, n) = \begin{cases} \gamma, & (m, n) = (1, -q), \\ 0, & \text{otherwise}, \end{cases}$$

Table 1 Some properties of six Fourier-based circula models. Formulae for regression quantities can be found in Appendix A

Model	Mode(s)	Antimode(s)	ρR	Form of $\Psi_1\|\Psi_2$ Regression	Form of $\Psi_2\|\Psi_1$ Regression
Equation (9)	Along $\psi_1 = q\psi_2$	Along $\psi_1 = q\psi_2 + \pi$	$q\gamma$	Linear-homoscedastic	Linear-homoscedastic
Equation (10)	Along $\psi_1 = q\psi_2 + \omega_g^m$ (where ω_g^m is mode of unimodal g)	Along $\psi_1 = q\psi_2 + \omega_g^a$ (where ω_g^a is antimode of unimodal g)	R_g (where R_g is mean resultant length of g)	Linear-homoscedastic	Linear-homoscedastic
Equation (11)	Along $\psi_1 = q\psi_2$	Along $\psi_1 = q\psi_2 + \pi$	$q\gamma$	Linear-homoscedastic	Linear-homoscedastic
Equation (12)	$(0, 0)$	$(-\pi, 0)$	$q\gamma$	Nonlinear-heteroscedastic	Linear-homoscedastic
Equation (13)	$(0, 0)$	See text	$q\gamma$	Nonlinear-heteroscedastic	Linear-homoscedastic
Equation (14)	$(0, 0)$	See text	$q\gamma$	Nonlinear-heteroscedastic	Linear-homoscedastic

Fig. 2 Planar contour plot of circula density (9) with parameter vector $(q, \gamma) = (1, 0.3)$. Here and in all other planar contour plots, the horizontal axis represents ψ_1 and the vertical axis ψ_2. The cross identifies $(\psi_1, \psi_2) = (0, 0)$
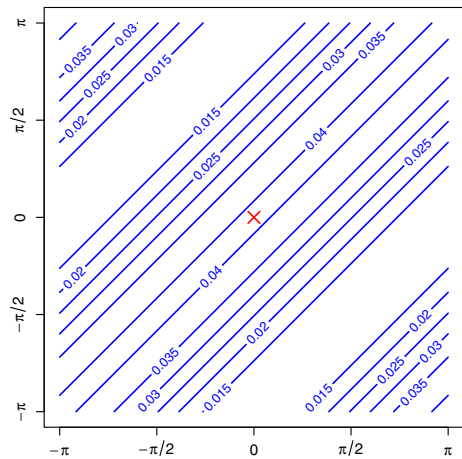
where $0 \leq \gamma \leq 1/2$ and, here and henceforth, $q \in \{-1, 1\}$. Thus, this pattern includes just a single nonzero real-valued Fourier coefficient. In this case, the circula density (7) reduces to

$$c_\circ(\psi_1, \psi_2) = \frac{1}{4\pi^2} \{1 + 2\gamma \cos(\psi_1 - q\psi_2)\}. \tag{9}$$

Figure 2 portrays a planar contour plot of density (9) when $q = 1$ and $\gamma = 0.3$. For this model, and all but one of the other models in this section, $\rho_R = q\gamma$ (see Table 1). The strength of dependence between Ψ_1 and Ψ_2 is thus controlled by γ, and its sign by q. The conditional distributions of $\Psi_1 | \Psi_2 = \psi_2$ and $\Psi_2 | \Psi_1 = \psi_1$ are cardioid distributions on the circle.

3.3 Pattern 2: diagonal line

Consider circula densities generated using nonzero Fourier coefficients on a diagonal of the $\mathbb{Z}^+ \times (\mathbb{Z}^+ \cup \mathbb{Z}^-)$ lattice. Specifically, let

$$\phi(m, n) = \begin{cases} \varphi(m), & n = -qm, \\ 0, & n \neq -qm, \end{cases}$$

where the $\varphi(m)$ are the Fourier coefficients of any circular distribution, for which $\varphi(0) = 1$. Then, the circula density is of the form

$$c_\circ(\psi_1, \psi_2) = \frac{1}{2\pi} g(\psi_1 - q\psi_2), \tag{10}$$

where $g(\psi) = (2\pi)^{-1}[1 + 2\text{Re}\{\sum_{m=1}^{\infty} \varphi(m)e^{-im\psi}\}]$ is a density on the unit circle. This is the Wehrly and Johnson (1980) class of circula densities discussed in Jones et al. (2015). For all but the case when g is circular uniform, (10) has linear contours parallel to the $q\pi/4$ diagonal. Density (10) reduces to (9) when $\varphi(1) = \gamma$ and $\varphi(m) = 0$ otherwise, and $0 \leq \gamma \leq 1/2$.

When $\varphi(m) = \gamma \rho^{m-1}$ $(m = 1, 2, \ldots, 0 \leq \gamma < 1, \ 2\gamma - 1 \leq \rho < 1)$, density (10) can be expressed in closed form as

$$c_\circ(\psi_1, \psi_2) = \frac{1}{4\pi^2} \left\{ 1 + 2\gamma \frac{\cos(\psi_1 - q\psi_2) - \rho}{1 + \rho^2 - 2\rho \cos(\psi_1 - q\psi_2)} \right\}. \qquad (11)$$

For this model, all three parameters, $\gamma = \phi(1, -q) = \varphi(1)$, ρ and q, affect the dependence between ψ_1 and ψ_2 although, as elsewhere, $\rho_R = q\gamma$. Figure S8 presents planar contour plots of density (11) for $q = -1$, $\gamma = 0.7$ and three values of ρ. The parameter ρ regulates the concentration of the wrapped Cauchy-like distribution. Both conditional distributions are special cases of the unimodal circular distributions of Kato and Jones (2015) which are two component mixtures with circular uniform and wrapped Cauchy components.

3.4 Pattern 3: vertical line

Consider the pattern formed by nonzero Fourier coefficients on the vertical line of the $\mathbb{Z}^+ \times (\mathbb{Z}^+ \cup \mathbb{Z}^-)$ lattice given by

$$\phi(m, n) = \begin{cases} \varphi(n), & m = 1, \ qn \leq -1, \\ 0, & \text{otherwise}, \end{cases}$$

for some nonzero Fourier coefficients φ.

For example, if $\varphi(n) = \gamma \rho^{|n|-1}$ $(0 \leq \gamma \leq 1/2, \ 0 \leq \rho \leq 1 - 2\gamma)$, then

$$c_\circ(\psi_1, \psi_2) = \frac{1}{4\pi^2} \left\{ 1 + 2\gamma \frac{\cos(\psi_1 - q\psi_2) - \rho \cos \psi_1}{1 + \rho^2 - 2\rho \cos \psi_2} \right\}. \qquad (12)$$

Panels (a) and (b) of Fig. S3 present planar contour plots of density (12) with $q = 1$ and two (γ, ρ) combinations. The dependence between Ψ_1 and Ψ_2 is clearly regulated by all three parameters: q, γ and ρ. In particular, γ regulates the strength of the dependence between Ψ_1 and Ψ_2, and ρ the degree of deformation of the density's shape around the main diagonal. The conditional distribution of $\Psi_1 | \Psi_2 = \psi_2$ is a cardioid distribution and that of $\Psi_2 | \Psi_1 = \psi_1$ is a special case of the Kato–Jones family.

Generalizations of Patterns 1–3, for which the nonzero Fourier coefficients are positioned more generically, are considered in Appendix C of the supplementary materials document.

3.5 Pattern 4: square

Now consider a square pattern of nonzero Fourier coefficients. As an attractive three-parameter example of this, let

$$\phi(m, n) = \begin{cases} \gamma \rho_1^{m-1} \rho_2^{|n|-1}, & m, -qn \in \mathbb{Z}^+, \\ 0, & \text{otherwise}, \end{cases}$$

where $0 < \rho_1, \rho_2 < 1$. Density (7) then becomes

$$c_\circ(\psi_1, \psi_2) = \frac{1}{4\pi^2}\left\{1 + 2\gamma\frac{\cos(\psi_1 - q\psi_2) - \rho_2\cos\psi_1 - \rho_1\cos\psi_2 + \rho_1\rho_2}{(1 + \rho_1^2 - 2\rho_1\cos\psi_1)(1 + \rho_2^2 - 2\rho_2\cos\psi_2)}\right\},$$
(13)

where the range of γ (≥ 0) is $\gamma \leq \rho_1\rho_2/[\{(1-\rho_1^2)(1-\rho_2^2)\}^{-1}-1]$ if $|\rho_1-\rho_2| \leq \rho_1\rho_2$, $\gamma \leq (1-\rho_1)(1+\rho_2)/2$ if $|\rho_1-\rho_2| > \rho_1\rho_2$ and $\rho_1 > \rho_2$, and $\gamma \leq (1+\rho_1)(1-\rho_2)/2$ if $|\rho_1-\rho_2| > \rho_1\rho_2$ and $\rho_1 < \rho_2$. Figure S9 presents a contour plot of the maximum value of γ as a function of ρ_1 and ρ_2. The maximum of γ tends to 0.5 as $(\rho_1, \rho_2) \to (0, 0)$.

The mode of density (13) occurs at $(\psi_1, \psi_2) = (0, 0)$. If $|\rho_1 - \rho_2| > \rho_1\rho_2$, then there is only one antimode of (13) which occurs at $(\psi_1, \psi_2) = (0, -\pi)$ for $\rho_1 > \rho_2$ and $(\psi_1, \psi_2) = (-\pi, 0)$ for $\rho_1 < \rho_2$. When $|\rho_1 - \rho_2| \leq \rho_1\rho_2$, there are two antimodes which occur at $(\psi_1, \psi_2) = \pm(\alpha_1, -q\alpha_2)$, where $\alpha_j = \arg\{\rho_j + (1 - \rho_j^2)/(\rho_j + e^{ix_j})\} = \text{atan2}\{-(1 - \rho_j^2)\sin x_j, 2\rho_j + (1 + \rho_j^2)\cos x_j\}$, $x_j = \arccos\{(\rho_j^2 - \rho_j^2\rho_k^2 - \rho_k^2)/(2\rho_j\rho_k^2)\}$ $(j, k = 1, 2, \ j \neq k)$ and $\text{atan2}(y, x)$ returns the angle measured anticlockwise from the x-axis to the vector connecting the origin with (x, y). The conditional densities of $\Psi_1|\Psi_2 = \psi_2$ and $\Psi_2|\Psi_1 = \psi_1$ are special cases of the Kato–Jones family.

Planar contour plots of circula density (13) designed to illustrate the roles of ρ_1 and ρ_2 are displayed in Fig. 3. When $\rho_1 = \rho_2$, the density is symmetric about the main diagonal and increasingly concentrated in the neighbourhood of the origin as $\rho_1 = \rho_2$ increases. For a fixed value of ρ_1, as ρ_2 increases the main axis of the central elliptical contour tilts increasingly away from the main diagonal towards $\psi_2 = 0$ and the dispersion increases in the neighbourhood of $(-\pi, -\pi) = (\pi, \pi)$. Due to the symmetry of (13), for a fixed value of ρ_2 the main axis tilts increasingly towards $\psi_1 = 0$ as ρ_1 increases.
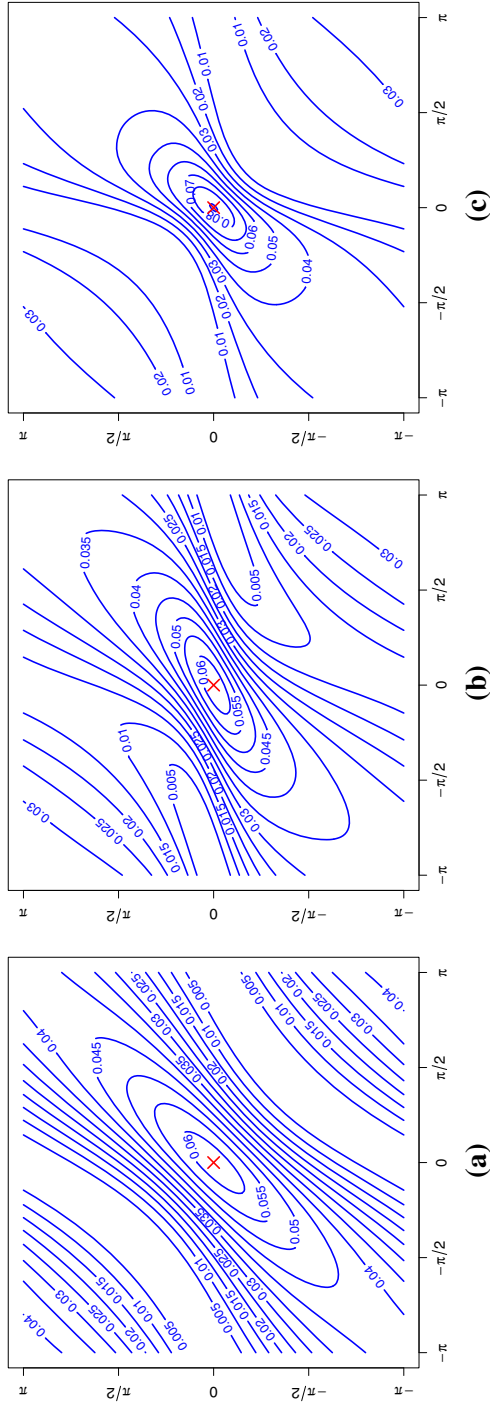
3.6 Pattern 5: triangle

As a final basic pattern, consider an arrangement of nonzero Fourier coefficients forming a triangular pattern. In particular, the pattern of nonzero Fourier coefficients underpinning circula density (11) can be extended to

$$\phi(m, n) = \begin{cases} \gamma\rho^{m-1}\lambda^{-(m+qn)}, & 1 \leq m \leq -qn, \\ 0, & \text{otherwise,} \end{cases}$$

where $0 < \rho < 1$, $0 \leq \lambda < 1$. The constraints on γ (≥ 0) are $\gamma \leq \rho(1 - \rho^2)(1 - \lambda^2)/[1 - \rho^2(1 - \lambda^2)]$ if $\lambda \geq (1-\rho)/\rho$ and $\gamma \leq (1-\lambda)(1+\rho)/2$ otherwise. A contour plot of the maximum value of γ as a function of ρ and λ is presented in Fig. S10. The maximum of γ tends to 1 as $(\rho, \lambda) \to (1, 0)$. For this model, the circula density is

$$c_\circ(\psi_1, \psi_2) = \frac{1}{4\pi^2}\left[1 + 2\gamma\frac{\cos(\psi_1 - q\psi_2) - \lambda\cos\psi_1 + \rho\lambda\cos\psi_2 - \rho}{\{1 + \rho^2 - 2\rho\cos(\psi_1 - q\psi_2)\}(1 + \lambda^2 - 2\lambda\cos\psi_2)}\right].$$
(14)

Fig. 3 Planar contour plots of circula density (13) with $q = 1$ and (ρ_1, ρ_2, γ) equal to: **a** (0.2, 0.2, 0.470); **b** (0.2, 0.5, 0.300); **c** (0.5, 0.5, 0.321). For each panel, the value of γ corresponds to the upper bound of its parameter space

The mode of the density is located at $(\psi_1, \psi_2) = (0, 0)$, and the antimode(s) at $(\psi_1, \psi_2) = (-\pi, 0)$ if $\lambda < (1 - \rho)/\rho$ or at $(\psi_1, \psi_2) = \pm(\alpha_1, -q\alpha_2)$ if $\lambda \geq (1 - \rho)/\rho$. Here $\alpha_1 = -\alpha_2 + \text{atan2}\{(1 - \rho^2) \sin x_1, \rho(2 + \rho \cos x_1)\}$, $\alpha_2 = \text{atan2}\{-(1 - \lambda^2) \sin x_2, 2\lambda + (1 + \lambda^2) \cos x_2\}$, $x_1 = \arccos\{(-1 - \rho^2 + \rho^2\lambda^2)/(2\rho)\}$, and $x_2 = \arccos\{(1 - \rho^2 - \rho^2\lambda^2)/(2\rho^2\lambda)\}$. The conditional distribution of $\Psi_1|\Psi_2 = \psi_2$ is a special case of the Kato–Jones family, while that of $\Psi_2|\Psi_1 = \psi_1$ does not seem to be a well-known distribution in general.

Figure 4 presents planar contour plots of circula density (14) for $q = 1$, $\rho = 0.6$ and three combinations of λ and γ. When $\lambda = 0$ the contours are straight lines because (14) reduces to (11). As λ increases, the axis of the central elliptical contour tilts increasingly away from the main diagonal towards $\psi_2 = 0$ and the other contours tend to be increasingly asymmetric about the main diagonal. Although the conditional distribution of $\Psi_2|\Psi_1 = \psi_1$ can have more than one mode, the joint distribution of (Ψ_1, Ψ_2) with this unique dependence structure, as well as the conditional distribution of $\Psi_1|\Psi_2 = \psi_2$, is always unimodal.

Appendix D of the supplementary materials document briefly discusses the shapes of bivariate circular densities obtained using (1) with circula densities generated by Patterns 2 and 5.

4 Model fitting

4.1 Circula densities

Let $\{(\psi_{1k}, \psi_{2k}); k = 1, \ldots, n\}$ denote an i.i.d. sample of random vectors from a circula density c_o where, from here onwards, n denotes sample size. If the form of c_o is unknown and n is moderate to large, an inspection of a planar scatterplot of the data will usually be sufficient to identify q and provide insight as to the form of c_o. However, the absolute values of the empirical Fourier coefficients, $\tilde{\phi}(r, -qs) = \frac{1}{n} \sum_{k=1}^{n} e^{i(r\psi_{1k} - qs\psi_{2k})}$, for $r, s = 1, 2, \ldots, 6$, say, represented graphically in a level plot, will generally prove more revealing, showing patterns, like those in Fig. 1, indicative of the structure of the Fourier coefficients of c_o. Computation of the empirical Fourier coefficients is extremely fast. In practice, a range of potential c_o's might be explored and the best fitting model established using information criteria, model reduction based on likelihood ratio testing and formal goodness-of-fit testing.

All of the densities of the five specific circula models in Sect. 3 have, by design, relatively simple closed-form expressions involving no computationally demanding normalizing constants, so computation of their log-likelihood functions is straightforward. Maximum likelihood (ML) estimation is then conducted using standard constrained optimization techniques. Method of moments (MM) estimates, calculated sequentially using the relations in Table 2, $q = \text{sgn}(|\phi(1, -1)| - |\phi(1, 1)|)$, and $\tilde{\phi}(r, -qs)$ substituted for $\phi(r, -qs)$, can be used as starting values.

The distribution function of a circula can be represented as $C_o(\psi_1, \psi_2) = C_{1|2}(\psi_1|\psi_2) C_2(\psi_2)$, where $C_2(\psi_2) = \psi_2/(2\pi)$ is the marginal, uniform, distribution function of Ψ_2 and $C_{1|2}$ is the conditional distribution function of $\Psi_1|\Psi_2 = \psi_2$. It follows that Ψ_2 and $2\pi C_{1|2}(\Psi_1|\Psi_2 = \psi_2)$ are independent and follow circular uniform
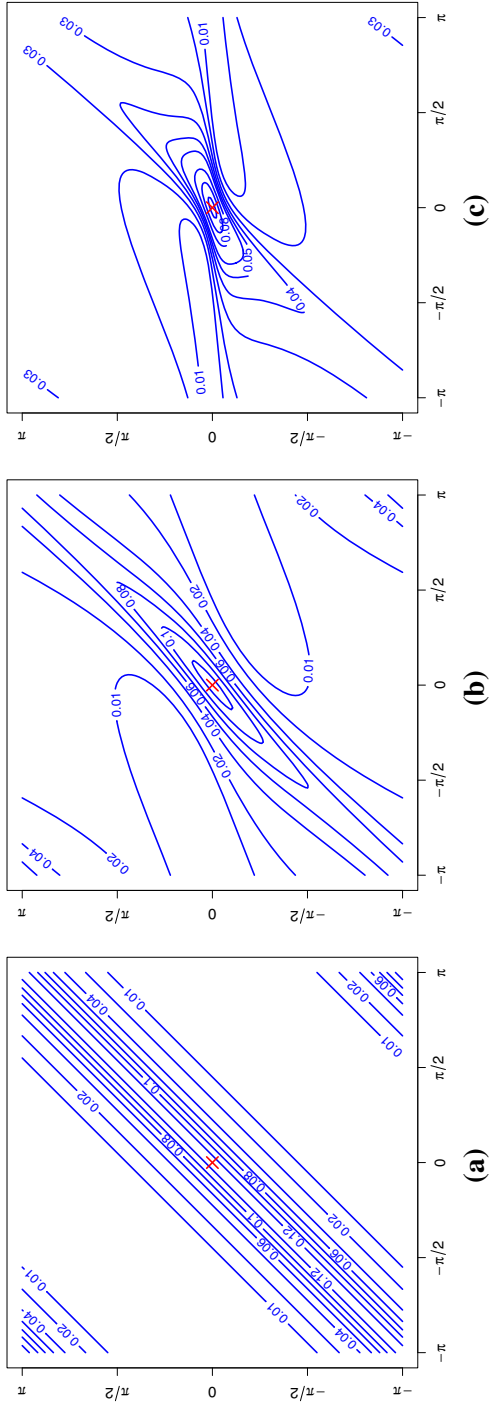
Fig. 4 Planar contour plots of circula density (14) for $q = 1$, $\rho = 0.6$ and (λ, γ) equal to: **a** $(0, 0.8)$; **b** $(0.4, 0.48)$; **c** $(0.8, 0.16)$. For each panel, the value of γ corresponds to the upper bound of its parameter space

Table 2 Relations between parameters and Fourier coefficients for five specific models

Model	Relations
Equation (9)	$\gamma = |\phi(1, -q)|$
Equations (11), (12)	$\gamma = |\phi(1, -q)|, \rho = |\phi(1, -2q)|/\gamma$
Equation (13)	$\gamma = |\phi(1, -q)|, \rho_1 = |\phi(2, -q)|/\gamma, \rho_2 = |\phi(1, -2q)|/\gamma$
Equation (14)	$\gamma = |\phi(1, -q)|, \rho = |\phi(2, -2q)|/\gamma, \lambda = |\phi(1, -2q)|/\gamma$

distributions. Similar results hold for Ψ_1 and $2\pi C_{2|1}(\Psi_2|\Psi_1 = \psi_1)$. When the parameters of the conditional distributions are estimated, $\{(\psi_{2k}, 2\pi \hat{C}_{1|2}(\psi_{1k}|\psi_{2k})); k = 1, \ldots, n\}$ and $\{(\psi_{1k}, 2\pi \hat{C}_{2|1}(\psi_{2k}|\psi_{1k})); k = 1, \ldots, n\}$ will be samples of pseudo-uniform vectors on the torus. Systematic patterns in scatterplots of such samples are indicative of lack-of-fit. To test goodness-of-fit, we apply the obvious adaptation of the parametric bootstrap approach described in Sect. 3.2 of Pewsey and Kato (2016) to the two pseudo-uniform samples.

4.2 Bivariate circular densities

The shapes of the bivariate circular densities obtained using density (1) depend on the reference points from which the marginal densities are integrated in the definitions of the marginal distribution functions F_1 and F_2. Traditionally, the reference point used has been the origin, 0. However, for this choice, changes in the location parameters of the marginal distributions result in shape changes, not just location shifts, in the densities obtained using Eq. (1). To avoid such shape changes, we define $F_j(\theta)$ ($j = 1, 2$) as $F_j(\theta) = \int_{\omega_j^A}^{\theta} f_j(\varphi)\mathrm{d}\varphi$ ($\omega_j^A \leq \theta < \omega_j^A + 2\pi$), where ω_j^A denotes the antimode of the circular density f_j.

We advocate the following sequential approach to modelling i.i.d. samples of random vectors $\{(\theta_{1k}, \theta_{2k}); k = 1, \ldots, n\}$ exploiting the three-component structure of density (1). If the distributional forms of the marginal densities, f_1 and f_2, are not specified beforehand, histograms and/or kernel density estimates can provide insight into forms for them. Then, their parameters are initially estimated separately using ML. Denoting the marginal distribution functions corresponding to those parameter estimates by $\hat{F}_1^s$ and $\hat{F}_2^s$, next the 'pseudo-sample' $\{(\hat{\Psi}_{1k}^s, \hat{\Psi}_{2k}^s); k = 1, \ldots, n\} = \{(2\pi \hat{F}_1^s(\theta_{1k}), 2\pi \hat{F}_2^s(\theta_{2k})); k = 1, \ldots, n\}$ is computed. The procedures described in Sect. 4.1 are then applied to this pseudo-sample to obtain initial estimates of the form of the underlying circula density and its parameters. Finally, the estimates from the previous two stages are used as starting values in the maximization of the full log-likelihood function derived from (1). We denote the marginal distribution functions corresponding to the parameter estimates obtained in this final estimation stage by $\hat{F}_1$ and $\hat{F}_2$.

When the parameters of (1) are estimated as above, the samples $\{(2\pi \hat{F}_2(\theta_{2k}), 2\pi \hat{C}_{1|2}(2\pi \hat{F}_1(\theta_{1k})|2\pi \hat{F}_2(\theta_{2k}))); k = 1, \ldots, n\}$ and $\{(2\pi \hat{F}_1(\theta_{1k}), 2\pi \hat{C}_{2|1}(2\pi \hat{F}_2(\theta_{2k})| 2\pi \hat{F}_1(\theta_{1k}))); k = 1, \ldots, n\}$ will be pseudo-uniformly distributed on the torus, and the

goodness-of-fit of the fitted bivariate circular density can be tested using the parametric bootstrap approach referred to in Sect. 4.1.

Appendix E of the supplementary materials document explains how data from a circula density or a bivariate circular density of the form (1) can be tested for independence using the permutation approach proposed in Sect. 3.3 of Kato and Pewsey (2015).

5 Numerical experiment

As part of our investigations into the performance of the modelling approach based on (1) and our proposed circula densities, we performed an experiment designed to compare its large-sample modelling capabilities with those of five existing bivariate circular models. We simulated single samples of size $n = 2000$ from each of the six models in Table 3, ranging from the 6-parameter pointwise symmetric model of Rivest (1988) to models capable of describing very varied distributional shapes. The parameter values in the same table were selected to produce representative unimodal cases of each model.

For each simulated sample we fitted all six models by ML, without pursuing model reduction, using R code developed by us and the `CircNNTSR` package of Fernández-Durán and Gregorio-Domínguez (2016) to fit the bivariate non-negative trigonometric sum (BNNTS) models. Table 4 contains the results obtained, the AIC values for the BNNTS models being those for the BIC-adjudged best fits, BIC being the model selection criterion advocated by Fernández-Durán and Gregorio-Domínguez (2014).

Using the column-wise sum of the ranks of the off-diagonal row-wise BIC values in Table 4 as a simple measure of overall performance, the following ordering, from best to worst, is obtained, with the sums of the ranks appearing between square brackets: $BGvM_2$ [7], BNNTS [15], FBvM [15], KJ^2Circ [16], R [18], SsS [19]. The analogous ordering for the AIC values is: $BGvM_2$ [7], BNNTS [11], FBvM [16], KJ^2Circ [17], R [19], SsS [20]. Both orderings identify the $BGvM_2$ model as having the best overall performance. The BIC-based measure identifies the KJ^2(13) model as having an overall performance similar to those of the BNNTS and FBvM models, with a marginally better overall performance than the R and SsS models. According to the AIC measure, BNNTS models perform second best, with the KJ^2(13) model performing similarly to the FBvM model and, again, marginally better than the R and SsS models.

We note that, for the samples from all five alternative models, the best-fitting case of (1) with Kato–Jones marginal densities and one of our proposed circula densities (KJ^2Circ) was the one incorporating circula model (13) (KJ^2(13)). This fact indicates that the circula structure underpinning the five alternative models can be approximately matched using just one of the circula models considered. Figures S11 and S12 present planar scatterplots of the six simulated data sets with contour plots of the best-fitting alternative and KJ^2Circ densities superimposed. They provide insight into the ability of the KJ^2Circ models to mimic the alternative models, and vice versa.

As the number of parameters of the underlying model increases, the performance of the KJ^2Circ and BNNTS models becomes increasingly competitive, although the latter generally have far more parameters than their KJ^2Circ counterparts. Also, while

Table 3 Densities and parameter values of the six models used in simulations: Rivest (R); sine-skewed Sine (SsS); full bivariate von Mises (FBvM); bivariate generalized von Mises of order 2 (BGvM$_2$); (1) with Kato–Jones $f_j(\theta_j)$ and (14) as circula density (KJ^2(14)); bivariate non-negative trigonometric sum of order (1,2) (BNNTS(1,2)). Throughout, $c_j = \cos(\theta_j - \mu_j)$, $s_j = \sin(\theta_j - \mu_j)$ and $j = 1, 2$. The models are ordered, from top to bottom, by increasing number of parameters

Model	Bivariate density, or marginal densities, and parameter values
R	$f(\theta_1, \theta_2) \propto \exp(\kappa_1 c_1 + \kappa_2 c_2 + \alpha_{11} c_1 c_2 + \alpha_{22} s_1 s_2)$
	$\mu_1 = 0, \mu_2 = \pi, \kappa_1 = 2, \kappa_2 = 1, \alpha_{11} = -3, \alpha_{22} = 2$
SsS	$f(\theta_1, \theta_2) \propto (1 + \lambda_1 s_1 + \lambda_2 s_2) \exp(\kappa_1 c_1 + \kappa_2 c_2 + \alpha_{22} s_1 s_2)$
	$\mu_1 = 0, \mu_2 = \pi/8, \kappa_1 = 1, \kappa_2 = 1.5, \alpha_{22} = 1, \lambda_1 = -0.5, \lambda_2 = 0.5$
FBvM	$f(\theta_1, \theta_2) \propto \exp(\kappa_1 c_1 + \kappa_2 c_2 + \alpha_{11} c_1 c_2 + \alpha_{21} c_2 s_1 + \alpha_{12} c_1 s_2 + \alpha_{22} s_1 s_2)$
	$\mu_1 = 0, \mu_2 = -\pi/16, \kappa_1 = 1.8, \kappa_2 = 1,$
	$\alpha_{11} = 0.5, \alpha_{12} = 0.6, \alpha_{21} = -0.5, \alpha_{22} = -1$
BGvM$_2$	$f(\theta_1, \theta_2) \propto \exp\{\kappa_{11} c_1 + \kappa_{12} c_2 + \kappa_{21} \cos(2(\theta_1 - \nu_1)) + \kappa_{22} \cos(2(\theta_2 - \nu_2))$
	$\quad + \eta_1 \cos(\theta_1 - \theta_2 - \omega_1) + \eta_2 \cos(\theta_1 + \theta_2 - \omega_2)\}$
	$\mu_1 = 0, \mu_2 = -0.3, \nu_1 = -0.9, \nu_2 = 0.9, \omega_1 = -1.3, \omega_2 = 0,$
	$\kappa_{11} = 1.9, \kappa_{12} = 1, \kappa_{21} = 0.3, \kappa_{22} = -0.7, \eta_1 = -0.7, \eta_2 = 0.8$
KJ^2(14)	$f_j(\theta_j) = \frac{1}{2\pi}\left[1 + 2\gamma_j^2 \frac{\gamma_j c_j - \bar{\alpha}_{2,j}}{\gamma_j^2 + \bar{\alpha}_{2,j}^2 + \bar{\beta}_{2,j}^2 - 2\gamma_j\{\bar{\alpha}_{2,j} c_j + \bar{\beta}_{2,j} s_j\}}\right]$
	$\mu_1 = 0, \mu_2 = -\pi/4, \gamma_1 = 0.58, \gamma_2 = 0.4, \bar{\beta}_{2,1} = 0.16, \bar{\beta}_{2,2} = -0.11,$
	$\bar{\alpha}_{2,1} = 0.16, \bar{\alpha}_{2,2} = 0.11, \rho = 0.6, \lambda = 0.4, \gamma = 0.48$
BNNTS(1,2)	$f(\theta_1, \theta_2) = \lvert \sum_{j=0}^{1} \sum_{k=0}^{2} \tau_{jk} \exp\{i(j\theta_1 + k\theta_2)\}\rvert^2$
	$\tau_{00} = 0.09659105, \tau_{01} = 0.06458246 - 0.00823441i,$
	$\tau_{02} = 0.01863562 + 0.01807478i, \tau_{10} = 0.06865395 + 0.01224350i,$
	$\tau_{11} = 0.06970889 + 0.00037061i, \tau_{12} = 0.02990480 + 0.02169680i$

the interpretation of the parameters of KJ^{2}Circ models is straightforward, the interpretation of BNNTS parameters is generally impossible. The percentage increase in BIC for the best-fitting KJ^2(13) model above that for the best-fitting alternative model is largest for the sample from Rivest's 6-parameter pointwise symmetric model (4.6%) and lowest for the data from the 10-(free)parameter BNNTS(1,2) model (0.6%).

6 Illustrative application

In contrast with the large sample size employed in Sect. 5, our illustrative analysis considers a much smaller sample of $n = 80$ pairs of wind directions, $\{(\theta_{1k}, \theta_{2k}), k = 1, \ldots, n\}$, recorded daily at a Milwaukee meteorological station at 4am and 6am, respectively, during the last three months of 2020. (This analysis was inspired by a related data analysis in Wehrly and Johnson (1980), which we have updated and expanded.) The data were extracted from the vast Milwaukee Met Data Archive series for 2020 available at https://www.glerl.noaa.gov/metdata/mil/archive/mil2020.01t.txt, and files containing them are included in the zip file linked to the supplementary

Table 4 AIC (top) and BIC (bottom) values of the specified fitted models for samples of size $n = 2000$ simulated using the model and parameter combinations in Table 3. The lowest off-diagonal values are highlighted in bold. The circula density of the best-fitting case of (1) with Kato–Jones marginal densities and one of our proposed circula densities (KJ^2Circ) is identified as (cd). The number of (free) parameters (np) for each fitted BNNTS model is that for the BIC-adjudged best fit

| Simulated model	Fitted model							
	R	SsS	FBvM	BGvM_2	KJ^2Circ	(cd)	BNNTS	np
R	10871.5	11757.1	11052.1	**10877.6**	11340.7	(13)	10938.6	38
	10905.1	11796.4	11096.9	**10944.8**	11402.3		11151.4	
SsS	11923.2	11861.0	11874.9	**11873.5**	12105.7	(13)	11949.1	22
	11956.8	11900.2	**11919.7**	11940.7	12167.3		12072.3	
FBvM	10645.7	10548.1	10501.3	**10509.2**	10672.3	(13)	10589.8	22
	10679.3	10587.3	10546.1	**10576.4**	10733.9		10713.0	
BGvM_2	10516.4	10376.8	10410.1	10099.4	10245.3	(13)	**10180.6**	38
	10550.0	10416.0	10454.9	10166.6	**10306.9**		10393.4	
KJ^2(14)	11862.4	12025.1	11680.1	11372.9	10829.5	(14)	**11118.7**	58
	11896.0	12064.3	11724.9	**11440.1**	10891.1		11443.5	
BNNTS(1,2)	12372.8	12380.6	12376.6	12235.3	**12178.2**	(13)	12105.2	10
	12406.4	12419.8	12421.4	12302.6	**12239.8**		12161.2	

materials document. Note that the archive does not include pairs of observations for the other 12 days of the trimester.

Given the two-hour separation between the measurements in each pair, one would expect them to be correlated. However, as the pairs span a trimester, it is not necessarily obvious what form of relation might exist between the pairs, nor what shapes the marginal distributions might exhibit. For the univariate time series, the only sample autocorrelation coefficient identified as being significantly different from 0 using the randomization version of the approach of Fisher (1993, Sect. 7.2.2) was the lag 1 coefficient for the 6am series, with an estimated p-value of 0.03 based on 9999 randomizations. However, the value of that autocorrelation coefficient is low (0.11) and in our analysis of the pairs of observations we treat them as forming an i.i.d. sample of toroidal data. Figure 5 includes a planar scatterplot of the data converted to radians in $[-\pi, \pi)$, and Fig. 6a linear histograms of the θ_1 and θ_2 values. In the latter, the superimposed circular densities are those of Kato–Jones ML fits.

The pseudo-sample $\{(2\pi \hat{F}_1^s(\theta_{1k}), 2\pi \hat{F}_2^s(\theta_{2k})); k = 1, \ldots, n\}$ is portrayed in Fig. 6b. From an inspection of it, the dependence between the pseudo-circular uniform variates is positive, i.e. $q = 1$. Figure 6c is a level plot of the absolute values of the empirical Fourier coefficients for that pseudo-sample and $q = 1$. The largest absolute values can be seen to form a diagonal pattern, indicative of circula density (11) as a potential model for the underlying $c_\circ$. Figure 6d presents a planar scatterplot of $\{(2\pi \hat{F}_1(\theta_{1k}), 2\pi \hat{F}_2(\theta_{2k})); k = 1, \ldots, n\}$ superimposed upon a contour plot of the circula density from the fitted KJ^2(11) model.

Fig. 5 Planar scatterplot of $\{(\theta_{1k}, \theta_{2k}), k = 1, \ldots, n\}$ superimposed on a contour plot of the ML-fitted density for the full model (1) with marginal Kato–Jones densities and the circula density (11) with $q = 1$
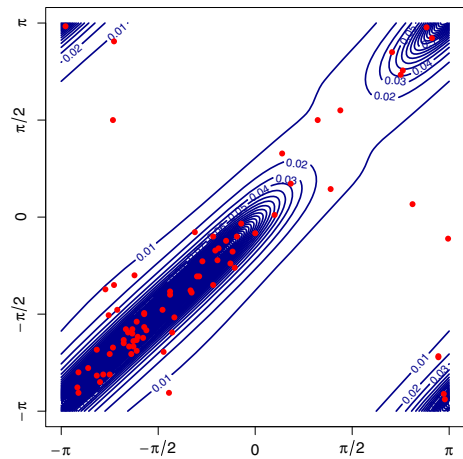

Table 5 Components of model (1) fitted to the wind direction data, parameters of the two marginal Kato–Jones densities and the circula density (11), MM and ML estimates for the parameters of the Kato–Jones densities fitted separately to the θ_1 and θ_2 values ($\mathrm{MM_s}$ and $\mathrm{ML_s}$), MM and ML likelihood estimates for the parameters of the circula density (11) with $q = 1$ fitted to the values of $\{(2\pi \hat{F}_1^s(\theta_{1k}), 2\pi \hat{F}_2^s(\theta_{2k})); k = 1, \ldots, n\}$ ($\mathrm{MM_{c_o}}$ and $\mathrm{ML_{c_o}}$), and ML estimates for the full model ($\mathrm{ML_f}$)

Component	$f_1(\theta_1)$				$f_2(\theta_2)$				c_o	
Parameter	μ_1	γ_1	$\bar{\beta}_{2,1}$	$\bar{\alpha}_{2,1}$	μ_2	γ_2	$\bar{\beta}_{2,2}$	$\bar{\alpha}_{2,2}$	γ	ρ
$\mathrm{MM_s}$	-1.89	0.49	-0.07	0.12	-1.70	0.50	-0.05	0.12		
$\mathrm{ML_s}$	-1.89	0.50	-0.06	0.14	-1.72	0.50	-0.06	0.16		
$\mathrm{MM_{c_o}}$									0.75	0.07
$\mathrm{ML_{c_o}}$									0.73	0.72
$\mathrm{ML_f}$	-1.77	0.39	0.00	0.03	-1.73	0.43	0.00	0.05	0.78	0.76

The point estimates from the different stages of the fitting process are given in Table 5. Table 6 contains the AIC and BIC values for the $\mathrm{KJ}^2(11)$ model and the five alternative bivariate circular models employed in Sect. 5. Both criteria select the $\mathrm{KJ}^2(11)$ model, a member of the Wehrly and Johnson (1980) class with Kato and Jones (2015) circular marginal distributions and circula density (11), as being best. Contour plots of all six fitted densities are presented in Fig. S13. The densities of the models with the highest BIC-values, SsS and BNNTS, are bimodal.

The contour plot of the ML-fitted $\mathrm{KJ}^2(11)$ density included in Fig. 5 appears to describe the distribution of the observations well, and the scatterplots of the pseudo-toroidal uniform values in panels (e) and (f) provide little evidence of systematic departures from toroidal uniformity. Also, the goodness-of-fit testing approach of Pewsey and Kato (2016), using the Bingham-type test statistic for toroidal uniformity of Wellner (1979) and $B = 999$ parametric bootstrap samples, returned a test statistic value of 2.63 and an estimated p-value of 0.33. Thus, there is no significant evidence against the $\mathrm{KJ}^2(11)$ model being a good model for the data.
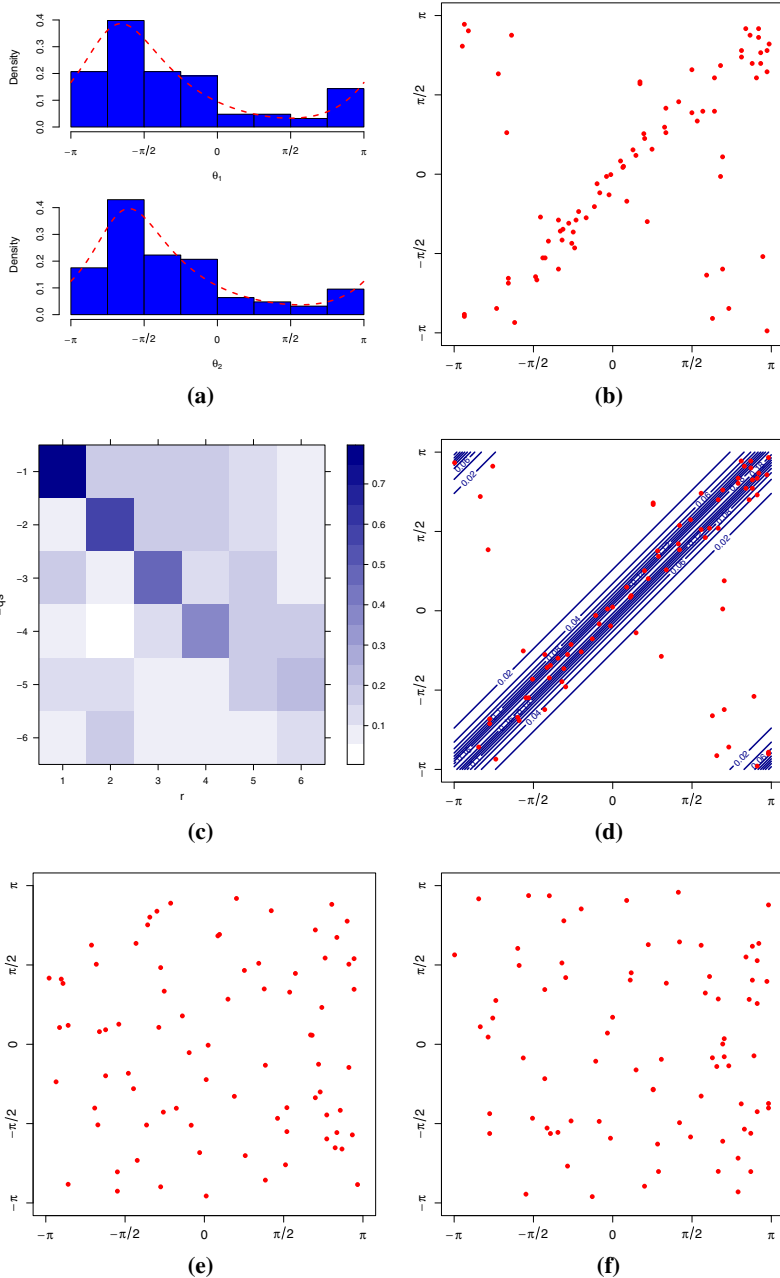
Fig. 6 a Linear histograms of the wind directions with ML-fitted Kato–Jones circular densities (dashed curves) superimposed. Planar scatterplots of: **b** $\{(2\pi \hat{F}_1^s(\theta_{1k}), 2\pi \hat{F}_2^s(\theta_{2k})); k = 1, \ldots, n\}$, **d** $\{(2\pi \hat{F}_1(\theta_{1k}), 2\pi \hat{F}_2(\theta_{2k})); k = 1, \ldots, n\}$ superimposed on a contour plot of the fitted circula density (11) from the full model fit, **e** $\{(2\pi \hat{F}_2(\theta_{2k}), 2\pi \hat{C}_{1|2}(2\pi \hat{F}_1(\theta_{1k})|2\pi \hat{F}_2(\theta_{2k}))); k = 1, \ldots, n\}$, **f** $\{(2\pi \hat{F}_1(\theta_{1k}), 2\pi \hat{C}_{2|1}(2\pi \hat{F}_2(\theta_{2k})|2\pi \hat{F}_1(\theta_{1k}))); k = 1, \ldots, n\}$. **c** Level plot of the absolute values of the empirical Fourier coefficients for the pseudo-sample in **b** and $q = 1$

Table 6 AIC (top) and BIC (bottom) values of the specified fitted models for the $n = 80$ pairs of wind directions. The number of (free) parameters (np) for the BNNTS model is that for the BIC-adjudged best fit. The lowest values are highlighted in bold

Fitted model						
R	SsS	FBvM	BGvM$_2$	KJ^2(11)	BNNTS	np
402.1	454.2	402.5	407.4	**380.6**	441.2	16
416.4	470.9	421.5	435.9	**404.4**	479.3	

The KJ^2(11) model identified in our structured sequential analysis postulates the underlying marginal circular distributions of the wind directions at 4am and 6am to be symmetric cases of the unimodal Kato–Jones family with very similar parameter values. Given the 2-hour separation between the two wind direction measurements, the similarity between the marginal distributions is perhaps to be expected. The fitted model also postulates that the positive dependence between the pairs of wind directions can be modelled using a circula with density (11) and $q = 1$. From the results in Jones et al. (2015), this implies that $\Psi_2 = 2\pi F_2(\Theta_2)$ is the result of a random rotation, Ω, from $\Psi_1 = 2\pi F_1(\Theta_1)$, Ω being independent of Ψ_1 and having circular density $g(\omega) = 2\pi \times (11)$ with $\psi_1 - \psi_2 = \omega$.

7 Conclusion

As the specific models introduced in Sect. 3 illustrate, the proposed general Fourier series construction provides a means of generating tractable parametric circula densities with simple closed-form expressions and wide-ranging distributional shapes. Moreover, when combined with Kato and Jones (2015) circular densities through (1) they provide highly flexible models for toroidal data for which the marginal distributions are unimodal. Multimodal toroidal data can be modelled using mixtures of such models.

The flexibility of the circula densities can be further enhanced by allowing the Fourier series coefficients to be complex. The main effect of such an extension is to skew the circula distributions in various ways.

Stationary Markov models for circular time series can be defined from our circula densities using an analogous approach to that of Wehrly and Johnson (1980).

In principle, our bivariate circula construction can be extended to produce d-dimensional circula densities using the multivariate analogue of (3) and patterns of nonzero Fourier coefficients distributed in d dimensions. Another possibility is, as mentioned in Jones et al. (2015), to model multivariate circular data using the circular analogues of pair copulas.

Our proposed methodology makes use of level plots of the absolute values of empirical Fourier coefficients computed from pseudo-samples as a highly successful model identification tool. We stress again that such plots are generally easier to interpret than scatterplots of the pseudo-samples themselves because they usually provide insight into the structure of the Fourier coefficients of the underlying circula density.

The advantage of the modelling approach based on (1) is that it facilitates the separate modelling of the circular marginals and a circula density in a structured sequential way. Also, as we have illustrated, it accommodates formal goodness-of-fit testing, an issue neglected in the literature related to the application of existing models for toroidal data.

Electronic supplementary material

The online supplementary materials document contains the appendices, and the equations, tables and figures with numbers preceded by the letter S. It also includes a link to a zip file containing all the data sets and our R code developed to analyse them.

Declarations

Conflict of interest   The authors declare that they have no conflict of interest.

Data and code availability   The data sets and the R code used to analyse them can be accessed via the link in Appendix G of the supplementary materials document.

References

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

Ameijeiras-Alonso J, Ley C (2022) Sine-skewed toroidal distributions and their application in protein bioinformatics. Biostatistics to appear

Fernández-Durán JJ (2007) Models for circular-linear and circular-circular data constructed from circular distributions based on nonnegative trigonometric sums. Biometrics 63(2):579–585

Fernández-Durán JJ, Gregorio-Domínguez MM (2014) 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

Fisher NI (1993) Statistical analysis of circular data. Cambridge University Press, Cambridge

Fisher NI, Lee AJ (1983) A correlation coefficient for circular data. Biometrika 70(2):327–332

Gatto R (2008) Some computational aspects of the generalized von Mises distribution. Stat Comput 18(3):321–331

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

Jammalamadaka SR, Sarma Y (1988) A correlation coefficient for angular variables. In: Matusita K (ed) Statistical theory and data analysis II. North-Holland, Amsterdam, pp 349–364

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

Jones MC, Pewsey A, Kato S (2015) On a class of circulas: copulas for circular distributions. Ann Inst Stat Math 67(5):843–862

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

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, Pewsey A (2015) A Möbius transformation-induced distribution on the torus. Biometrika 102(2):359–370

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

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. Springer, Berlin, pp 159–178

Mardia KV, Jupp PE (1999) Directional statistics. Wiley, Chichester

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

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

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):35

Pewsey A, Kato S (2016) Parametric bootstrap goodness-of-fit testing for Wehrly–Johnson bivariate circular distributions. Stat Comput 26(6):1307–1317

Rivest LP (1982) Some statistical methods for bivariate circular data. J R Stat Soc Ser B Methodol 44(1):81–90

Rivest LP (1988) A distribution for dependent unit vectors. Commun Stat Theory Methods 17(2):461–483

Saw JG (1983) Dependent unit vectors. Biometrika 70(3):665–671

Shieh GS, Johnson RA (2005) Inference based on a bivariate distribution with von Mises marginals. Ann Inst Stat Math 57(4):789–802

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

Umbach D, Jammalamadaka SR (2009) Building asymmetry into circular distributions. Stat Probab Lett 79(5):659–663

Wehrly TE, Johnson RA (1980) Bivariate models for dependence of angular observations and a related Markov process. Biometrika 67(1):255–256

Wellner JA (1979) Permutation tests for directional data. Ann Stat 7(5):929–943