


Circular autocorrelation of stationary circular Markov processes

Toshihiro Abe¹ · Hiroaki Ogata²  ·
Takayuki Shiohama³ · Hiroyuki Taniai⁴

Received: 30 April 2016 / Accepted: 22 December 2016 / Published online: 31 December 2016
© Springer Science+Business Media Dordrecht 2016

Abstract The stationary Markov process is considered and its circular autocorrelation function is investigated. More specifically, the transition density of the stationary Markov circular process is defined by two circular distributions, and we elucidate the structure of the circular autocorrelation when one of these distributions is uniform and the other is arbitrary. The asymptotic properties of the natural estimator of the circular autocorrelation function are derived. Furthermore, we consider the bivariate process of trigonometric functions and provide the explicit form of its spectral density matrix. The validity of the model was assessed by applying it to a series of wind direction data.

Keywords Circular statistics · Time series models · Toroidal data · Wind direction · Wrapped cauchy distribution

The research reported herein was supported by JSPS KAKENHI Grant Numbers 15K17593, 26380401, 15K21433, and 26870655.

✉ Hiroaki Ogata
hiroakiogata@tmu.ac.jp

Toshihiro Abe
abetosh@ss.nanzan-u.ac.jp

Takayuki Shiohama
shiohama@rs.tus.ac.jp

Hiroyuki Taniai
htaniai@aoni.waseda.jp

¹ Department of Systems and Mathematical Science, Nanzan University, 18 Yamazato-cho, Showa-ku, Nagoya 446-8673, Japan

² Department of Business Administration, Tokyo Metropolitan University, 1-1 Minami-Osawa, Hachioji-shi, Tokyo 192-0397, Japan

³ Department of Information and Computer Technology, Faculty of Engineering, Tokyo University of Science, 6-3-1 Nijjuku, Katsushika, Tokyo 125-8585, Japan

⁴ School of International Liberal Studies, Waseda University, 1-6-1 Nishiwaseda, Shinjuku, Tokyo 169-8050, Japan

1 Introduction

Directional or circular time series data arise frequently in the natural sciences such as in meteorology, biology, geography, medicine, astronomy, and many other areas. A substantial body of literature on the analysis of circular data exists, and we refer the reader to [Mardia and Jupp \(2000\)](#), [Jammalamadaka and SenGupta \(2001\)](#), and [Pewsey et al. \(2013\)](#). Typical examples are hourly or daily wind directions at a weather station ([Breckling 1989](#); [Fisher and Lee 1994](#)). Some stochastic processes have been proposed in the literature to analyze this kind of data; for example, [Wehrly and Johnson \(1980\)](#) proposed a Markov process by applying a class of bivariate circular distributions with specified marginals.

Regarding papers dedicated to circular time series analysis, we can refer to the following. In Chap. 6 of [Breckling \(1989\)](#), two stochastic processes, the von Mises process and the wrapped auto-regressive process, are fitted to time series data of wind directions. [Fisher and Lee \(1994\)](#) proposed other new models based on a projection and a link function. The models proposed in [Fisher and Lee \(1994\)](#) are known as the circular autoregressive (CAR) model and the linked autoregressive moving average (LARMA) model, which are extensions of the usual ARMA models for directional data. [Holzmann et al. \(2006\)](#) proposed hidden Markov models for circular time series. [Kato \(2010\)](#) provided a discrete time Markov process (Markov chain) on the circle by adapting the Möbius transformation. A drifting circular Markov process was considered by [Yeh et al. \(2013\)](#). [Artes et al. \(2000\)](#) provided generalized estimation equation methods for circular longitudinal data.

Because of its mathematical tractability and good fit to real-world data, the Markov property of a sequence is a basic concept in state space modeling and is widely applied in various fields. In this paper, we consider the circular Markov process of [Wehrly and Johnson \(1980\)](#) and investigate its serial correlation structure. More specifically, we provide the explicit form of the circular autocorrelation function (*cacf*) introduced in [Fisher and Lee \(1994\)](#) and [Holzmann et al. \(2006\)](#).¹ The Markov process under consideration is defined by two circular densities, f and g , and our setting for f is circular uniform, whereas g is considered to be an arbitrary circular density. The *cacf* at lag k is essentially the $2k$ th power of the mean resultant length of g . This can be considered as an extension of the result in [Holzmann et al. \(2006\)](#), in which the explicit form of *cacf* at lag k was given when f is the uniform and g is the von Mises density function. We also give the spectral density of the trigonometric functions of the Markov process.

Two kinds of estimators of *cacf* were considered, with the nonparametric and parametric estimators being a sample circular autocorrelation function and a maximum likelihood estimator, respectively. We provide the asymptotic properties of these two estimators.

The remainder of the paper is organized as follows. Section 2 presents the explicit form of the circular autocorrelation function of the stationary circular Markov process. The spectral density of the trigonometric functions of the Markov process is also given. Section 3 considers the asymptotic properties of the nonparametric and parametric estimators of *cacf*. Some numerical studies are carried out in Sect. 4. We apply the model to a time series of wind direction data in Sect. 5. Proofs of the theorems are presented in Sect. 6. Section 7 concludes.

¹ The similar term “circular serial correlation coefficient” is used in [Anderson \(1971, Section 6.5.2\)](#) in a different context, where he considered a process whose first and last observations are connected to each other.

2 Circular Markov process

Let $f_1(\theta)$ and $f_2(\xi)$ be arbitrary circular density functions on $\Pi := [0, 2\pi)$, and $F_1(\theta)$ and $F_2(\xi)$ be their distribution functions. In addition, let $g(\cdot)$ be an arbitrary density function on Π . Then, [Wehrly and Johnson \(1980\)](#) introduced both

$$\begin{aligned} f_-(\theta, \xi) &= 2\pi g[2\pi\{F_1(\theta) - F_2(\xi)\}]f_1(\theta)f_2(\xi), \\ f_+(\theta, \xi) &= 2\pi g[2\pi\{F_1(\theta) + F_2(\xi)\}]f_1(\theta)f_2(\xi) \end{aligned}$$

as being densities on Π^2 with the marginal densities $f_1(\theta)$ and $f_2(\xi)$. Bivariate distributions for circular data are referred to as being bivariate circular or toroidal data because their underlying support is the unit torus. Additional details of the analysis of toroidal data can be found in [Jones et al. \(2015\)](#) and [Kato and Pewsey \(2015\)](#).

Now, let $\{\theta_t\}_{t \geq 0}$ be a sequence of random variables on Π . Using the above fact, [Wehrly and Johnson \(1980\)](#) naturally defined the Markov process on the unit circle as

$$p(\theta_0) = f(\theta_0) \quad (\text{initial density}) \quad (1)$$

$$\begin{aligned} p(\theta_t | \theta_{t-1}, \theta_{t-2}, \dots) &= p(\theta_t | \theta_{t-1}) \\ &= 2\pi g[2\pi\{F(\theta_t) - F(\theta_{t-1})\}]f(\theta_t) \quad (\text{transition density}) \end{aligned} \quad (2)$$

where $f(\cdot)$ and $g(\cdot)$ are arbitrary densities on Π , and $F(\cdot)$ is the distribution function of $f(\cdot)$. Obviously, the transition density can also be defined via f_+ as

$$p(\theta_t | \theta_{t-1}) = 2\pi g[2\pi\{F(\theta_t) + F(\theta_{t-1})\}]f(\theta_t) \quad (3)$$

or via a combination of f_- and f_+ as

$$p(\theta_t | \theta_{t-1}) = 2\pi g[2\pi\{F(\theta_t) - qF(\theta_{t-1})\}]f(\theta_t), \quad (4)$$

where $q \in \{-1, 1\}$ is a given non-random constant. The density g is sometimes known as the binding density, as in [Jones et al. \(2015\)](#). For $q = 1$, the process generates positively autocorrelated time series, and for $q = -1$, it generates negatively autocorrelated series, which are characterized by movements from one direction to the opposite in consecutive time steps.

For $p = \pm 1, \pm 2, \dots$, the p th trigonometric moment of g is defined by

$$E_g[e^{ip\Theta}] = \rho_p e^{i\mu_p} = \alpha_p + i\beta_p = E_g[\cos(p\Theta)] + iE_g[\sin(p\Theta)]. \quad (5)$$

In particular, ρ_1 is referred to as the mean resultant length, which has range $0 \leq \rho_1 < 1$, and $\mu_1 \pmod{2\pi}$ is known as the mean direction. The proximity of the value of ρ_1 to unity implies that there was little variation.

In this paper, we consider the stationary Markov process on the unit circle with the initial density (1) and the transition density (4). For any circular stationary process, [Holzmann et al. \(2006\)](#) introduced the circular autocorrelation function (*cacf*):

$$R_C(k) := R_{\text{Corr}}(\Theta_0, \Theta_k), \quad k \geq 0,$$

where $R_{\text{Corr}}(\cdot, \cdot)$ on the right-hand side is the circular correlation coefficient in [Fisher and Lee \(1983\)](#). By using stationarity, [Holzmann et al. \(2006\)](#) wrote it as

$$R_C(k) = \frac{E[\cos \Theta_0 \cos \Theta_k]E[\sin \Theta_0 \sin \Theta_k] - E[\sin \Theta_0 \cos \Theta_k]E[\cos \Theta_0 \sin \Theta_k]}{(1 - E[\cos^2 \Theta_0])E[\cos^2 \Theta_0] - (E[\sin \Theta_0 \cos \Theta_0])^2}, \quad (6)$$

and they calculated $R_C(k)$ when f is uniform and g is a von Mises process.² Because of the symmetry of g , the calculation of theoretical circular autocorrelation becomes simple. In this work, we relax the restriction on the distributional assumption of g . We give the explicit form of the *cacf* when f is uniform and g is an arbitrary density. Let the density g be indexed by parameter vector η , where $\eta \in H \subset \mathbb{R}^q$.

Assumption 1 The function g is a continuous circular density, which is differentiable with respect to η and has a positive definite Fisher information matrix $I(\eta)$. Then there exists a mean resultant length of g such that

$$\rho_1(\eta) = \rho_1 := |E_g(e^{i\Theta})| < 1.$$

Examples of the binding density g are general circular densities such as the von Mises, wrapped Cauchy, and Jones–Pewsey distributions (Jones and Pewsey 2005). For examples of asymmetric circular distributions, we refer to Kato and Jones (2010, 2013). We obtain the following theorem:

Theorem 1 Let Assumption 1 hold, and let $\{\Theta_t\}_{t \geq 0}$ follow the stationary Markov circular process defined by (1) and (4) when f is the circular uniform density and g is an arbitrary circular density. Then, for $q \in \{1, -1\}$, the circular autocorrelation function at lag $k (\geq 0)$ is $R_C(k) = q^k \rho_1^{2k}$.

Remark 1 If we use the transition density (2), that is, $q = 1$ in (4), then we have $R_C(k) = \rho_1^{2k}$. This means the process with (2) never produces negative circular autocorrelation.

Remark 2 For the autoregressive process of order 1 (AR(1)) with autoregressive coefficient ϕ_1 , the autocorrelation of lag k becomes ϕ_1^k . Hence, the process defined by (1) and (4) can be considered as a first-order circular autoregressive process whose autoregressive coefficient is ρ_1^2 .

In the remaining part of this section, we consider the bivariate process $X_t = (\cos \Theta_t, \sin \Theta_t)'$ and the transition density (2) is used. In our setting, the marginal distribution of Θ_t is uniform, which indicates $\mu_X = (E[\cos \Theta_t], E[\sin \Theta_t])' = \mathbf{0}$. Therefore, the autocovariance matrix of X_t is

$$\Gamma_X(k) = E[(X_k - \mu_X)(X_0 - \mu_X)'] = E \begin{bmatrix} \cos \Theta_k \cos \Theta_0 & \cos \Theta_k \sin \Theta_0 \\ \sin \Theta_k \cos \Theta_0 & \sin \Theta_k \sin \Theta_0 \end{bmatrix}.$$

In comparison to (6), we can write the circular autocorrelation function under the form

$$R_C(k) = \frac{\det\{\Gamma_X(k)\}}{\det\{\Gamma_X(0)\}}.$$

The expression of a correlation coefficient for two angular variables with determinant can be found in Breckling (1989, p. 152). As discussed in Breckling (1989) and Jammalamadaka and SenGupta (2001), various circular autocorrelation coefficients are proposed. If the circular autocorrelation function $R_C(k)$ defined above satisfies the usual properties of autocorrelations, then we can define a valid circular autocovariance function of lag k , which is induced from the circular autocorrelation functions such that

$$\Gamma_\Theta(k) := \det\{\Gamma_X(k)\} = R_C(k) \det\{\Gamma_X(0)\}; \quad (7)$$

² The *cacf* function at lag $k (\geq 0)$ in Holzmann et al. (2006) should be read as $R_C(k) = (I_1(\kappa)/I_0(\kappa))^{2k}$. Here, $I_r(\kappa)$ is the modified Bessel function of the first kind of order r . See Sect. 4 for details.

see Jona-Lasinio et al. (2012). The definition of autocovariance functions differs from those in the time series context; however, the autocovariance structure based on the determinant functions is discussed by Paparoditis (1994), where he refers to it as a generalized autocovariance function that retains all the properties of usual ordinary autocovariance functions. According to Paparoditis (1994), our valid circular autocovariance function defined above is a version of the generalized autocovariance function for the circular time series.

Remark 3 A stationary process is said to be ergodic for the mean if the time series average converges to the population mean. Similarly, if the sample average provides a consistent estimate for the second moment, then the process is said to be ergodic for the second moment. The result of Theorem 1 indicates that the following absolute summable valid circular autocovariance function holds:

$$\sum_{k=-\infty}^{\infty} |\det\{\Gamma_X(k)\}| < \infty; \quad (8)$$

this ensures that the process is ergodic for all moments. See Hannan (1970, Section 4.2) and Brockwell and Davis (1991, p.379) for examples.

The following theorem provides the spectral densities for the processes $\{\Theta_t\}$ and $\{X_t\}$, where they capture the autocovariance structures for $\{\Theta_t\}$ and $\{X_t\}$:

Theorem 2 Under Assumption 1, let $\{\Theta_t\}_{t \geq 0}$ follow the stationary Markov circular process defined by (1) and (2) when f and g are the circular uniform density and an arbitrary circular density, respectively. Then, the spectral density of X_t is

$$f_X(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{-i\omega k} \Gamma_X(k) = \begin{bmatrix} f_{11}(\omega) & f_{12}(\omega) \\ f_{21}(\omega) & f_{22}(\omega) \end{bmatrix},$$

where

$$\begin{aligned} f_{11}(\omega) &= f_{22}(\omega) \\ &= \frac{1}{4\pi} \left\{ \frac{1}{2} \left(\frac{1}{1 - \rho_1 e^{-i(\omega - \mu_1)}} + \frac{1}{1 - \rho_1 e^{-i(\omega + \mu_1)}} \right. \right. \\ &\quad \left. \left. + \frac{1}{1 - \rho_1 e^{i(\omega - \mu_1)}} + \frac{1}{1 - \rho_1 e^{i(\omega + \mu_1)}} \right) - 1 \right\} \end{aligned}$$

and

$$\begin{aligned} f_{12}(\omega) &= -f_{21}(\omega) \\ &= \frac{1}{8\pi i} \left(-\frac{1}{1 - \rho_1 e^{-i(\omega - \mu_1)}} + \frac{1}{1 - \rho_1 e^{-i(\omega + \mu_1)}} \right. \\ &\quad \left. - \frac{1}{1 - \rho_1 e^{i(\omega - \mu_1)}} + \frac{1}{1 - \rho_1 e^{i(\omega + \mu_1)}} \right). \end{aligned}$$

The spectral density function for the process $\{\Theta_t\}$ can be considered as the Fourier series representation for the valid circular autocovariance functions defined by (7), which is given by

$$f_{\Theta}(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} e^{-i\omega k} \det\{\Gamma_X(k)\} = \frac{1}{8\pi} \frac{1 - \rho_1^4}{|1 - \rho_1^2 e^{-i\omega}|^2}.$$

3 Asymptotic properties of the estimators for circular autocorrelation

In this section, we consider the estimators of $R_C(k)$.

3.1 Sample circular autocorrelation function

The natural estimator is a sample version of (6). Therefore, for a fixed k and given $\Theta_1, \dots, \Theta_n$, the sample circular autocorrelation function is defined as

$$\begin{aligned} \hat{R}_C(k) &= \left\{ \left(\frac{1}{n} \sum_{t=1}^{n-k} Y_t^{(CC)}(k) \right) \left(\frac{1}{n} \sum_{t=1}^{n-k} Y_t^{(SS)}(k) \right) - \left(\frac{1}{n} \sum_{t=1}^{n-k} Y_t^{(SC)}(k) \right) \left(\frac{1}{n} \sum_{t=1}^{n-k} Y_t^{(CS)}(k) \right) \right\} \\ &\quad \times \left\{ \left(\frac{1}{n} \sum_{t=1}^n Y_t^{(CC)}(0) \right) \left(\frac{1}{n} \sum_{t=1}^n Y_t^{(SS)}(0) \right) - \left(\frac{1}{n} \sum_{t=1}^n Y_t^{(CS)}(0) \right)^2 \right\}^{-1}, \end{aligned} \quad (9)$$

where

$$\begin{aligned} &\left[Y_t^{(CC)}(k), Y_t^{(CS)}(k), Y_t^{(SC)}(k), Y_t^{(SS)}(k) \right] \\ &:= [\cos \Theta_t \cos \Theta_{t+k}, \cos \Theta_t \sin \Theta_{t+k}, \sin \Theta_t \cos \Theta_{t+k}, \sin \Theta_t \sin \Theta_{t+k}]. \end{aligned}$$

Theorem 3 Under Assumption 1, let $\{\Theta_t\}_{t \geq 0}$ follow the stationary Markov circular process defined by (1) and (4) when f and g are the circular uniform density and an arbitrary circular density, respectively. Then, the sample cacf $\hat{R}_C(k)$ has the following asymptotic properties:

1. $\hat{R}_C(k) \xrightarrow{P} R_C(k)$.
2. $\sqrt{n} \left(\hat{R}_C(k) - R_C(k) \right) \xrightarrow{d} N(0, \sigma^2)$.

The asymptotic variance σ^2 is given in the proof in Sect. 6.

3.2 Maximum likelihood estimator

When the density g is parameterized as $g(\theta; \eta)$, where $\eta \in \mathbf{H}$ is a certain nuisance parameter in an open set $\mathbf{H} \subset \mathbb{R}^q$, we can use the maximum likelihood estimation for $R_C(k)$. If f is the circular uniform density function, the joint density of $\Theta_0, \dots, \Theta_k$ from the Markov process with (1) and (4) is given by

$$p(\theta_0, \dots, \theta_k; \eta) = \frac{1}{2\pi} \prod_{i=1}^k g(\theta_i - q\theta_{i-1}; \eta).$$

Under the usual regularity conditions, the maximum likelihood estimator of η (denoted by $\hat{\eta}$) has the following asymptotic normality:

$$\sqrt{n}(\hat{\eta} - \eta) \xrightarrow{d} N(\mathbf{0}, I(\eta)^{-1})$$

where $I(\eta)$ is the Fisher information

$$I(\eta) = E \left[\frac{\partial^2}{\partial \eta \partial \eta'} \log p(\theta_0, \dots, \theta_k; \eta) \right].$$

The resultant length of g is now written in terms of η . If we denote it by $\rho_1(\eta)$, the circular autocorrelation function at lag k is written by $R_C(k) = \rho_1(\eta)^{2k}$, and we can consider the maximum likelihood estimator $\hat{R}_C^{\text{ML}}(k) = \rho_1(\hat{\eta})^{2k}$. The delta method leads to

$$\sqrt{n} \left(\hat{R}_C^{\text{ML}}(k) - R_C(k) \right) \xrightarrow{d} N(\mathbf{0}, \sigma_{\text{ML}}^2),$$

where the asymptotic variance is

$$\sigma_{\text{ML}}^2 = 4k^2 \rho_1(\eta)^{2(2k-1)} \frac{\partial \rho_1(\eta)}{\partial \eta'} I(\eta)^{-1} \frac{\partial \rho_1(\eta)}{\partial \eta}.$$

It is difficult to give closed-form expressions for the maximum likelihood estimates of the model, hence numerical methods should be used to find the solutions. We numerically solve the likelihood equations or apply numerical optimization to the log-likelihood function to get maximum likelihood estimates. For this purpose, we used the function *optim* in R with the BFGS algorithm, which is the most popular quasi-Newton method for unconstrained optimization. The function *optim* includes the argument *hessian* = *TRUE*, which returns a numerically evaluated approximate Hessian at the minimum. This can be used to estimate standard errors of the components of the MLE. We reparameterize the model such that the likelihood function satisfies the unconstrained optimization problems. The corresponding Hessian can then be obtained by the delta method.

4 Numerical examples

This section presents numerical examples of the *cacfs* of the circular Markov processes (1) and (4) together with some Monte Carlo simulations for those estimators. As for the parameter q in (4), we set $q = 1$ throughout this section.

We set the density f as being uniform. We emphasize that our theory does not require g to be symmetrical; however, we choose the binding density g as the usual von Mises and wrapped Cauchy distributions, which are given by

$$g^{(\text{VM})}(\theta; \mu, \kappa) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\}, \quad (10)$$

$$g^{(\text{WC})}(\theta; \mu, \rho) = \frac{1}{2\pi} \left(\frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)} \right), \quad (11)$$

respectively, where $\kappa > 0$ and $0 < \rho < 1$ are the concentration parameters of the von Mises and wrapped Cauchy distributions. $I_r(\kappa)$ is the modified Bessel function of the first kind of order r , defined as

$$I_r(\kappa) = \frac{1}{2\pi} \int_0^{2\pi} \cos(r\theta) e^{\kappa \cos \theta} d\theta, \quad r = 0, \pm 1, \dots$$

The mean resultant lengths of the von Mises and wrapped Cauchy distributions are given by

$$\rho_1^{(\text{VM})} = \frac{I_1(\kappa)}{I_0(\kappa)} \quad \text{and} \quad \rho_1^{(\text{WC})} = \rho,$$

respectively, and the corresponding *cacfs* become

$$R_C^{(\text{VM})}(k) = \left[\frac{I_1(\kappa)}{I_0(\kappa)} \right]^{2k} \quad \text{and} \quad R_C^{(\text{WC})}(k) = \rho^{2k}.$$

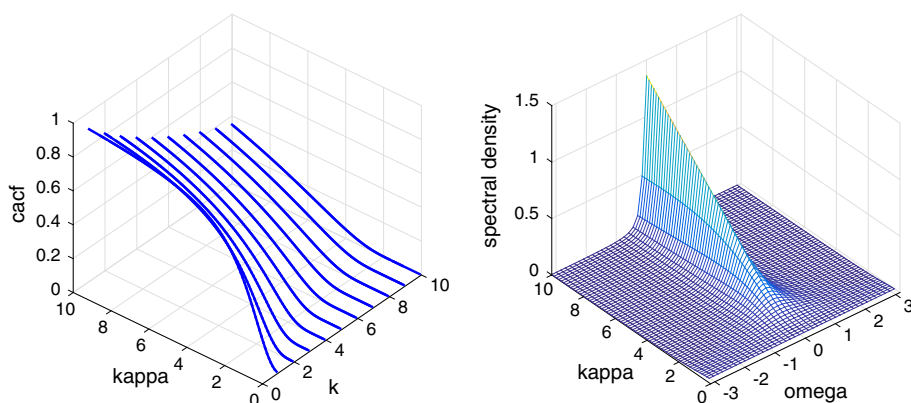


Fig. 1 The *left panel* shows the plots of the circular autocorrelation functions as a function of the lag k and κ in the domain $k \in \{1, 2, \dots, 10\}$ and $\kappa \in (0, 10)$, when g is chosen to be the von Mises density. The *right panel* shows the plots of the corresponding spectral density function

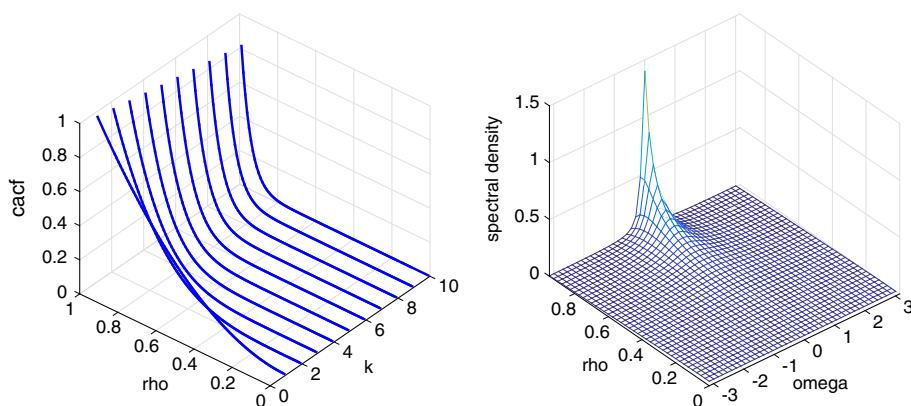


Fig. 2 The *left panel* shows the plots of the circular autocorrelation functions as a function of the lag k and ρ in the domain $k \in \{1, 2, \dots, 10\}$ and $\rho \in (0, 0.95)$, when g is chosen to be the wrapped Cauchy density. The *right panel* shows the plots of the corresponding spectral density function

Figures 1 and 2 plot the *cacf* and the spectral density function for the cases when g is chosen to be the von Mises and wrapped Cauchy densities, respectively. It is obvious that larger values in κ and ρ tend to have larger circular autocorrelations, whereas the values decrease as the lag k increases.

Next, we consider the performance of a small sample in estimating the theoretical circular autocorrelation functions (lags: $k = 1, 2, 3$). The data are generated according to the processes (1) and (2) with sample size $n = 250$. Here, we set the functions f and g as the uniform and sine-skewed von Mises densities, respectively. The density of the sine-skewed von Mises distribution is given by

$$g^{(\text{SSVM})}(\theta; \mu, \kappa, \lambda) = \frac{1}{2\pi I_0(\kappa)} \exp\{\kappa \cos(\theta - \mu)\} (1 + \lambda \sin(\theta - \mu)),$$

where $\lambda \in (-1, 1)$ is the skewness parameter (see Abe and Pewsey 2011). Here, we set the concentration parameter as $\kappa = 1$ (scattered), 2 (concentrated) and the skewness parameter as $\lambda = 0$ (symmetry), 0.5 (skewed). We set the location parameter $\mu = 0$.

Table 1 Simulation results for the means of the estimators and root mean squared errors (RMSEs) based on 1000 replications

Lag	$\kappa = 1$ (scattered)			$\kappa = 2$ (concentrated)		
	1	2	3	1	2	3
Symmetric case with $\lambda = 0$						
True <i>cacf</i>	0.199	0.040	0.008	0.487	0.237	0.115
Mean (Sample ACF)	0.185	0.034	0.006	0.444	0.207	0.095
Mean (MLE)	0.199	0.041	0.009	0.488	0.237	0.115
RMSE (Sample ACF)	0.014	0.005	0.002	0.043	0.030	0.020
RMSE (MLE)	0.001	0.001	0.001	0.001	0.002	0.002
Skewed case with $\lambda = 0.5$						
True <i>cacf</i>	0.249	0.062	0.016	0.517	0.268	0.138
Mean (Sample ACF)	0.239	0.060	0.016	0.488	0.249	0.128
Mean (MLE)	0.251	0.064	0.017	0.517	0.269	0.140
RMSE (Sample ACF)	0.010	0.003	0.000	0.030	0.019	0.010
RMSE (MLE)	0.002	0.002	0.001	0.000	0.001	0.002

We calculated both the nonparametric and parametric estimators, where the former corresponds to the sample circular autocorrelation estimator $\hat{R}_C(k)$, and the latter corresponds to the maximum likelihood estimator (MLE) $\hat{R}_C^{\text{ML}}(k) = \rho_1(\hat{\kappa}, \hat{\lambda})^{2k}$. These evaluations are based on 1000 Monte Carlo replications. The MLEs are obtained by maximizing the following log-likelihood function:

$$\ell(\kappa, \lambda) = \sum_{t=1}^n \log g^{(\text{SSVM})}(\theta_t - \theta_{t-1}; \kappa, \lambda).$$

We estimated the parameters of the von Mises distribution by setting $\lambda = 0$ in the above calculation.

The means and root mean squared errors of sample *cacfs* (upper cells) and MLEs (lower cells) are shown in Table 1. In all cases, we can confirm the unbiasedness for both estimators. Comparing the two estimators, the MLE is found to be better than the sample *cacf* in the sense of the root mean squared error.

5 Data analysis

Statistical models for wind direction prediction are important in various topics in environmental studies and energy engineering. We first describe the circular models of the sine-skewed family proposed by Abe and Pewsey (2011), and then we illustrate the data analysis using our models. The density of the sine-skewed circular distribution is defined as

$$f(\theta) = f_0(\theta - \mu)(1 + \lambda \sin(\theta - \mu)),$$

where $f_0(\cdot)$ is the symmetric circular base density and $|\lambda| < 1$ is a skewness parameter. As an example of the base density $f_0(\cdot)$, it is natural to select the Jones–Pewsey distribution (Jones and Pewsey 2005) because of its flexibility. The sine-skewed Jones–Pewsey density defined on the circle $[-\pi, \pi)$ takes the form

$$g^{(\text{SSJP})}(\theta; \mu, \kappa, \psi, \lambda) = \frac{\{\cosh(\kappa\psi) + \sinh(\kappa\psi) \cos(\theta - \mu)\}^{1/\psi}}{2\pi P_{1/\psi}(\cosh(\kappa\psi))} (1 + \lambda \sin(\theta - \mu)),$$

where $-\pi \leq \mu < \pi$ is a location parameter, $\kappa \geq 0$ is a concentration parameter, $-\infty < \psi < \infty$ controls the shape, and $|\lambda| < 1$ is a skewness parameter. Here $P_{1/\psi}$ is the associated Legendre function of the first kind of degree $1/\psi$ and order 0. Clearly, the Jones-Pewsey distribution has three well-known submodels, namely, the wrapped Cauchy ($\psi = -1$), von Mises ($\psi \rightarrow 0$), and cardioid ($\psi = 1$) distributions. For later use, we also give the sine-skewed wrapped Cauchy density

$$g^{(\text{SSWC})}(\theta; \mu, \rho, \lambda) = \frac{1}{2\pi} \left(\frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)} \right) (1 + \lambda \sin(\theta - \mu)),$$

where $0 < \rho < 1$ is given with a change of parameter $\rho = \tanh(\kappa/2)$.

The p th mean resultant length is given by

$$\rho_p = \sqrt{\alpha_{0,p}^2 + \lambda^2(\alpha_{0,p-1} - \alpha_{0,p+1})^2/4},$$

where $\alpha_{0,p}$ is the p th mean resultant length of f_0 . We see that the mean resultant length of the sine-skewed wrapped Cauchy distribution becomes $\rho_1^{(\text{SSWC})} = \sqrt{\rho^2 + \lambda^2(1 - \rho^2)/4}$. Similarly, the mean resultant length of the sine-skewed Jones-Pewsey distribution becomes

$$\rho_1^{(\text{SSJP})} = \sqrt{\alpha_{0,1}^2 + \lambda^2(1 - \alpha_{0,2})^2/4},$$

where

$$\alpha_{0,1} = \frac{|\psi|}{1 + \psi} \frac{P_{1/\psi}^1(\cosh(\kappa\psi))}{P_{1/\psi}(\cosh(\kappa\psi))} \quad \text{and} \quad \alpha_{0,2} = \frac{\psi^2}{(1 + 2\psi)(1 + \psi)} \frac{P_{1/\psi}^2(\cosh(\kappa\psi))}{P_{1/\psi}(\cosh(\kappa\psi))}.$$

The circular densities defined above are the candidates for the selection of the binding density $g(\cdot)$. We present an evaluation of our circular Markov model using a time series of wind direction data collected in Forest Grove, Oregon, U.S., on an hourly basis at an Agricultural Meteorology (AgriMet) weather station. This weather station is located at latitude 45.55305N and longitude 123.08361 W. Wind direction data are the records of the direction from which the wind originates and are recorded in terms of degrees from the north (0 degrees) and the clockwise angle increases. AgriMet stations, located mainly in agricultural areas in the Pacific Northwest region in the U.S., provide reliable meteorological information, which is applied, for example, to the improvement of agricultural productivity by being used in calculations of crop water use and requirements.

Accurate forecasting of wind energy production based on past patterns of wind speed and direction is required by wind farms. Because of the variability of the wind direction, which is greatly affected by topography, geostrophic wind, and turbulence intensity, more complicated models for wind direction should be considered; however, it is also necessary to consider simple time series models that can explain patterns of time dependency.

Samples were collected during the period from February 22 to March 15, 2015, for a sample size of 504. Figure 3 shows the time series plot of the wind direction. According to this figure, the wind direction changes every few hours and does not blow predominantly from any particular single direction. The wind direction in the observed period is also characterized by a gradual rotation in the anticlockwise direction.

Before we estimate the model parameters, we need to perform a statistical test to determine whether the observed time series follows the uniform distribution. The null hypothesis of the Rayleigh test for uniformity against the alternative hypothesis of a general unimodal

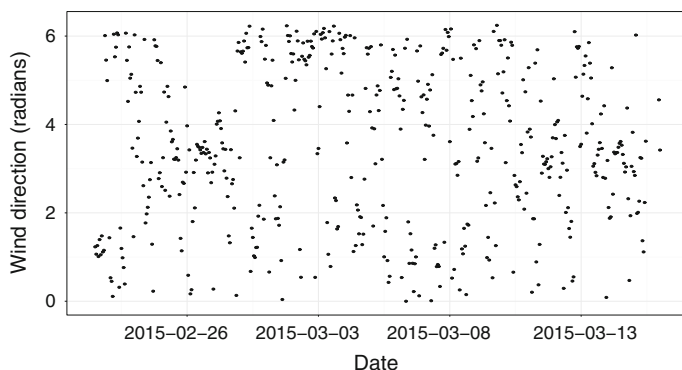


Fig. 3 Observed wind directions at AgriMet station FOGO (Forest Grove, Oregon) for the period February 22 through March 15, 2015

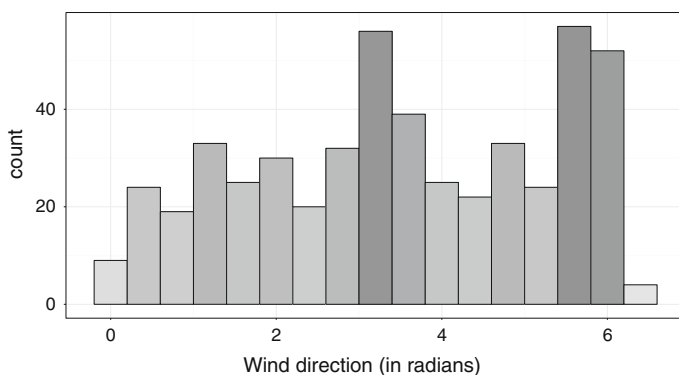


Fig. 4 Histogram of the wind directions in Forest Grove, Oregon

distribution is not rejected (test statistic 0.063, p -value 0.139). We choose a circular uniform distribution as the marginal distribution for f . Uniformity in the wind direction can be confirmed with the histogram plotted in Fig. 4, whereas some form of bimodality can be observed in which the wind is coming from the north-northwest and the south.

The time series of wind direction are moderately persistent with mild positive correlations, which can be verified via the sample circular autocorrelation plots of Fig. 5. The observed sample correlogram shows that as the lag increases, the circular autocorrelation decreases exponentially. These findings are consistent with our theoretical time series model.

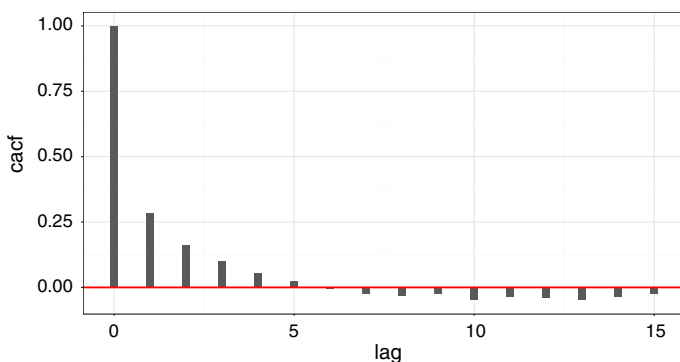
Table 2 presents the maximum likelihood estimates of the unknown parameters. According to the AIC values, the sine-skewed Jones–Pewsey distribution is the most appropriate among the models.

The estimated location parameters with values around -0.045 indicate that the time series of the wind direction has a weak negative trend in angles. However, we note that it is not different from zero at the 5% significance level.

All the concentration parameters in the wrapped Cauchy and the Jones–Pewsey distributions are estimated to be different from zero at the 1% significance level. The estimated lag 1 circular autocorrelation coefficients for the wrapped Cauchy and the sine-skewed wrapped Cauchy distributions are $\{\hat{\rho}_1^{(WC)}\}^2 = 0.309$ and $\{\hat{\rho}_1^{(SSWC)}\}^2 = 0.311$, respectively.

Table 2 MLEs, standard errors (Std. Err.), estimated mean resultant length, and AIC values for the models of the wind directions

	Wrapped Cauchy		Sine-skewed WC	
	MLE	Std. Err.	MLE	Std. Err.
μ	−0.047	0.038	−0.048	0.038
ρ	0.556	0.023	0.555	0.023
λ	–		−0.140	0.066
ρ_1^2	0.309		0.311	
AIC	1502.694		1500.262	
	Jones–Pewsey		Sine-skewed JP	
	MLEs	Std. Err.	MLEs	Std. Err.
μ	−0.045	0.024	−0.046	0.031
κ	1.225	0.065	1.227	0.043
ψ	−1.454	0.079	−1.454	0.081
λ	–		−0.141	0.066
ρ_1^2	0.276		0.298	
AIC	1500.685		1498.239	

**Fig. 5** Sample circular autocorrelation functions of wind directions in Forest Grove, Oregon

As those for the Jones–Pewsey and the sine-skewed Jones–Pewsey distributions, we have $\{\hat{\rho}_1^{(JP)}\}^2 = 0.276$ and $\{\hat{\rho}_1^{(SSJP)}\}^2 = 0.298$, respectively. These values are quite close to the value of the lag 1 sample circular autocorrelation coefficient, which is 0.286. These findings indicate that there exists a weak relationship between the hourly wind directions observed. Figure 6 plots the level curves of the joint density function of θ_t and θ_{t-1} with the sine-skewed Jones–Pewsey density as g together with the observations. The estimated skewness parameter $\lambda = -0.141$ indicates the changes in the wind direction along the azimuth tend to be skewed in an anticlockwise direction.

6 Proofs

In this section, we provide the proofs of the theorems stated in Sects. 2 and 3. Throughout the proofs of Theorems 1 and 2, the following lemma is useful for our purposes:

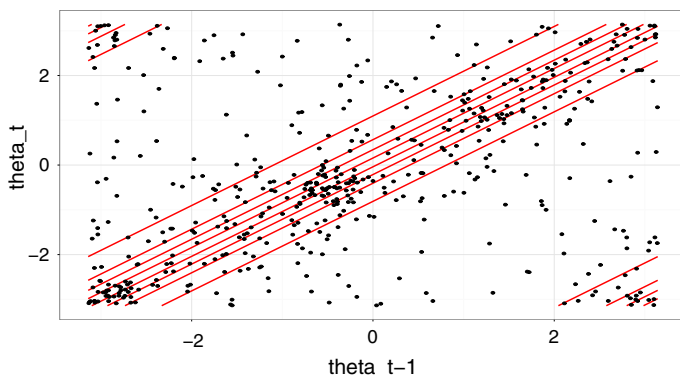


Fig. 6 Contour plots of the estimated joint density functions $p(\theta_t, \theta_{t-1})$, together with sample observations. Here $g(\cdot)$ and $f(\cdot)$ are chosen as the sine-skewed Jones–Pewsey and the circular uniform densities, respectively

Lemma 1 For any positive integer $p \geq 1$, we have

$$E \left[\begin{pmatrix} \cos(p\theta_i) \\ \sin(p\theta_i) \end{pmatrix} \middle| \Theta_{i-1} = \theta_{i-1} \right] = D_p \begin{bmatrix} \cos(p\theta_{i-1}) \\ \sin(p\theta_{i-1}) \end{bmatrix},$$

where the matrix D_p is defined by

$$D_p = \rho_p \begin{bmatrix} \cos \mu_p & -\sin \mu_p \\ \sin \mu_p & \cos \mu_p \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & q \end{bmatrix}, \quad (12)$$

with $q = \pm 1$.

Proof of Lemma 1 Direct calculation gives the desired result, such that

$$\begin{aligned} \int_{\Pi} \begin{bmatrix} \cos(p\theta_i) \\ \sin(p\theta_i) \end{bmatrix} g(\theta_i - q\theta_{i-1}) d\theta_i &= \int_{\Pi} \begin{bmatrix} \cos(p(\theta_i + q\theta_{i-1})) \\ \sin(p(\theta_i + q\theta_{i-1})) \end{bmatrix} g(\theta_i) d\theta_i \\ &= \int_{\Pi} \begin{bmatrix} \cos(p\theta_i) \cos(qp\theta_{i-1}) - \sin(p\theta_i) \sin(qp\theta_{i-1}) \\ \sin(p\theta_i) \cos(qp\theta_{i-1}) + \cos(p\theta_i) \sin(qp\theta_{i-1}) \end{bmatrix} g(\theta_i) d\theta_i \\ &= \rho_p \begin{bmatrix} \cos \mu_p & -\sin \mu_p \\ \sin \mu_p & \cos \mu_p \end{bmatrix} \begin{bmatrix} \cos(qp\theta_{i-1}) \\ \sin(qp\theta_{i-1}) \end{bmatrix} \\ &= \rho_p \begin{bmatrix} \cos \mu_p & -\sin \mu_p \\ \sin \mu_p & \cos \mu_p \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & q \end{bmatrix} \begin{bmatrix} \cos(p\theta_{i-1}) \\ \sin(p\theta_{i-1}) \end{bmatrix} \\ &= D_p \begin{bmatrix} \cos(p\theta_{i-1}) \\ \sin(p\theta_{i-1}) \end{bmatrix}, \end{aligned}$$

where we use the facts $\cos(qp\theta_{i-1}) = \cos(p\theta_{i-1})$ and $\sin(qp\theta_{i-1}) = q \sin(p\theta_{i-1})$. \square

Lemma 2 The autocovariance matrix of lag k for the bivariate process X_t is given by

$$E \begin{bmatrix} \cos \Theta_0 \cos \Theta_k & \sin \Theta_0 \cos \Theta_k \\ \cos \Theta_0 \sin \Theta_k & \sin \Theta_0 \sin \Theta_k \end{bmatrix} = \frac{1}{2} D_1^k,$$

where D_1 is given by (12) with $p = 1$.

Proof of Lemma 2 When the density function f is uniform, the joint density of $\Theta_0, \dots, \Theta_k$ given by (1) and (4) reduces to

$$p(\theta_0, \dots, \theta_k) = \frac{1}{2\pi} \prod_{i=1}^k g(\theta_i - q\theta_{i-1}).$$

Using this joint density, we obtain

$$\begin{aligned} E \begin{bmatrix} \cos \Theta_0 \cos \Theta_k & \sin \Theta_0 \cos \Theta_k \\ \cos \Theta_0 \sin \Theta_k & \sin \Theta_0 \sin \Theta_k \end{bmatrix} &= E \begin{bmatrix} \cos \Theta_k \\ \sin \Theta_k \end{bmatrix} \begin{bmatrix} \cos \Theta_0 & \sin \Theta_0 \end{bmatrix} \\ &= \frac{1}{2\pi} \int_{\Pi^{k+1}} \begin{bmatrix} \cos \theta_k \\ \sin \theta_k \end{bmatrix} \prod_{j=1}^k g(\theta_j - q\theta_{j-1}) \begin{bmatrix} \cos \theta_0 & \sin \theta_0 \end{bmatrix} \prod_{j=0}^k d\theta_j \\ &= \frac{1}{2\pi} D_1^k \int_{\Pi} \begin{bmatrix} \cos \theta_0 \\ \sin \theta_0 \end{bmatrix} \begin{bmatrix} \cos \theta_0 & \sin \theta_0 \end{bmatrix} d\theta_0 = \frac{1}{2} D_1^k. \end{aligned}$$

Since Θ_0 follows a uniform distribution, we have

$$E[\cos^2 \Theta_0] = \frac{1}{2\pi} \int_{\Pi} \cos^2 \theta_0 d\theta_0 = \frac{1}{2}, \quad E[\sin^2 \Theta_0] = \frac{1}{2\pi} \int_{\Pi} \sin^2 \theta_0 d\theta_0 = \frac{1}{2},$$

and

$$E[\sin \Theta_0 \cos \Theta_0] = \frac{1}{2\pi} \int_0^{2\pi} \sin \theta_0 \cos \theta_0 d\theta_0 = 0.$$

□

Proof of Theorem 1 From Lemma 2, we can evaluate the circular autocorrelation function as

$$\begin{aligned} R_C(k) &= \frac{E[\cos \Theta_0 \cos \Theta_k]E[\sin \Theta_0 \sin \Theta_k] - E[\sin \Theta_0 \cos \Theta_k]E[\cos \Theta_0 \sin \Theta_k]}{(1 - E[\cos^2 \Theta_0])E[\cos^2 \Theta_0] - (E[\sin \Theta_0 \cos \Theta_0])^2} \\ &= \frac{\det\{(1/2)D_1^k\}}{(1 - 1/2)(1/2) - 0^2} = (\det D_1)^k = q^k \rho_1^{2k}. \end{aligned}$$

□

Proof of Theorem 2 From Lemma 2, the autocovariance of X_t at lag $k(\geq 0)$ is given by

$$\Gamma_X(k) = \frac{1}{2} D_1^k.$$

Therefore, when $q = 1$ in (4), we have

$$D_1^k = \rho_1^k \begin{bmatrix} \cos(k\mu_1) & -\sin(k\mu_1) \\ \sin(k\mu_1) & \cos(k\mu_1) \end{bmatrix}.$$

Using $\Gamma_X(k) = \Gamma_X(-k)'$ for $k < 0$, we have

$$\begin{aligned} f_X(\omega) &= \frac{1}{2\pi} \left(\sum_{k=1}^{\infty} e^{i\omega k} \Gamma_X(k)' + \Gamma_X(0) + \sum_{k=1}^{\infty} e^{-i\omega k} \Gamma_X(k) \right) \\ &= \frac{1}{4\pi} \left(\sum_{k=0}^{\infty} e^{i\omega k} \rho_1^k \begin{bmatrix} \cos(k\mu_1) & \sin(k\mu_1) \\ -\sin(k\mu_1) & \cos(k\mu_1) \end{bmatrix} \right. \\ &\quad \left. + \sum_{k=0}^{\infty} e^{-i\omega k} \rho_1^k \begin{bmatrix} \cos(k\mu_1) & -\sin(k\mu_1) \\ \sin(k\mu_1) & \cos(k\mu_1) \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right). \end{aligned}$$

Putting

$$\cos(k\mu_1) = \frac{e^{i(k\mu_1)} + e^{-i(k\mu_1)}}{2}, \quad \sin(k\mu_1) = \frac{e^{i(k\mu_1)} - e^{-i(k\mu_1)}}{2i}$$

leads to the desired statement. The result for $f_{\Theta}(\omega)$ is obtained by similar calculations. \square

Proof of Theorem 3 First, we introduce the notations

$$\mathbf{Y}_t(k) = \left[Y_t^{(\text{CC})}(k), Y_t^{(\text{CS})}(k), Y_t^{(\text{SC})}(k), Y_t^{(\text{SS})}(k), Y_t^{(\text{CC})}(0), Y_t^{(\text{SS})}(0), Y_t^{(\text{CS})}(0) \right]',$$

$$\mathbf{v}(k) = E[\mathbf{Y}_0(k)], \quad \hat{\mathbf{v}}(k) = \frac{1}{n} \sum_{t=1}^n \mathbf{Y}_t(k),$$

and

$$\Sigma(k) = \sum_{t=1}^{n-k} \sum_{s=1}^{n-k} E \left[(\mathbf{Y}_t(k) - \mathbf{v}(k)) (\mathbf{Y}_s(k) - \mathbf{v}(k))' \right].$$

1. By the ergodic theorem, $\hat{\mathbf{v}}(k) \xrightarrow{a.s.} \mathbf{v}(k)$. Then, Slutsky's theorem leads to $\hat{R}_C(k) \xrightarrow{P} R_C(k)$.
2. Consider an arbitrary linear combination $\mathbf{c}'\hat{\mathbf{v}}(k)$, where $\mathbf{c} = (c_1, \dots, c_7)'$ is a vector constant. By Theorem 5.2 of Hall and Heyde (1980), the following asymptotic normality holds:

$$\sqrt{n} (\mathbf{c}'(\hat{\mathbf{v}}(k) - \mathbf{v}(k))) \xrightarrow{d} N(0, \mathbf{c}'\Sigma(k)\mathbf{c}).$$

Therefore, the Cramér–Wold device leads to the multivariate asymptotic normality

$$\sqrt{n} (\hat{\mathbf{v}}(k) - \mathbf{v}(k)) \xrightarrow{d} N(\mathbf{0}, \Sigma(k)).$$

Now, consider the function $h: \mathbb{R}^7 \rightarrow \mathbb{R}$ such that $h(x_1, \dots, x_7) = (x_1x_4 - x_2x_3)(x_5x_6 - x_7^2)$. Then, the sample *cacf* is expressed by $\hat{R}_C(k) = h(\hat{\mathbf{v}}(k))$. The delta method implies

$$\sqrt{n} (\hat{R}_C(k) - R_C(k)) \xrightarrow{d} N(0, \sigma^2),$$

where the asymptotic variance is

$$\sigma^2 = (\partial h / \partial \mathbf{v})' \Sigma(k) (\partial h / \partial \mathbf{v}).$$

\square

7 Summary and conclusion

We investigated the circular autocorrelation function of the circular Markov process. The Markov process was constructed with two circular densities f and g , and we derived the explicit form of the autocorrelation function at lag k when f is uniform and g is arbitrary. Specifically, it is essentially the $2k$ th power of the resultant length of g . The values of circular autocorrelation were calculated by specifying g to be von Mises and wrapped Cauchy densities, respectively. We also considered the nonparametric and parametric estimators of the circular autocorrelation function and gave their asymptotic normality. The finite sample properties were also investigated in the simulation studies. The proposed model was estimated

using time series of wind directions in Oregon. Possible future research topics include relaxing the restriction of uniformity in f and of continuity in both g and f . A thorough investigation incorporating these topics into the data analysis would be worthwhile.

Acknowledgements The authors would like to express their gratitude to the editor and two anonymous referees, whose invaluable comments improved the paper.

References

- Abe T, Pewsey A (2011) Sine-skewed circular distributions. *Stat Pap* 52:683–707
- Anderson TW (1971) *The statistical analysis of time series*. Wiley, New York
- Artes R, Paula GA, Ranvaud R (2000) Analysis of circular longitudinal data based on generalized estimating equations. *Aust N Z J Stat* 42:347–358
- Breckling J (1989) *The analysis of directional time series: application to wind speed and direction*. Springer, Berlin
- Brockwell PJ, Davis RA (1991) *Time series: theory and methods*, 2nd edn. Springer, New York
- Fisher NI, Lee AJ (1983) A correlation coefficient for circular data. *Biometrika* 70:327–332
- Fisher NI, Lee AJ (1994) Time series analysis of circular data. *J R Stat Soc B* 70:327–332
- Hall P, Heyde CC (1980) *Martingale limit theory and its application*. Academic Press, New York
- Hannan EJ (1970) *Multiple time series*. Wiley, New York
- Holzmann H, Munk A, Suster M, Zucchini W (2006) Hidden Markov models for circular and linear–circular time series. *Environ Ecol Stat* 13:325–347
- Jammalamadaka SA, SenGupta AS (2001) *Topics in circular statistics*. World Scientific Publishing Co., New York
- Jona-Lasinio G, Gelfand A, Jona-Lasinio M (2012) Spatial analysis of wave direction data using wrapped Gaussian processes. *Ann Appl Stat* 6:1478–1498
- Jones MC, Pewsey A (2005) A family of symmetric distributions on the circle. *J Am Stat Assoc* 100:1422–1428
- Jones MC, Pewsey A, Kato S (2015) On a class of circulars: copulas for circular distributions. *Ann Inst Stat Math* 67:843–862
- Kato S (2010) A Markov process for circular data. *J R Stat Soc B* 72:655–672
- 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* 105:249–262
- Kato S, Jones MC (2013) An extended family of circular distributions related to wrapped Cauchy distributions via Brownian motion. *Bernoulli* 19:154–171
- Kato S, Pewsey A (2015) A Möbius transformation-induced distribution on the torus. *Biometrika* 102:359–370
- Mardia KV, Jupp PE (2000) *Directional statistics*. Wiley, Chichester
- Papadoditis E (1994) On vector autocorrelation and generalized second-order function for time series. *J Time Ser Anal* 15:235–334
- Pewsey A, Neuhaus M, Ruxton GD (2013) *Circular statistics in R*. Oxford University Press, London
- Wehrly TE, Johnson RA (1980) Bivariate models for dependence of angular observations and a related Markov process. *Biometrika* 67:255–256
- Yeh S-Y, Harris KDM, Jupp PE (2013) A drifting Markov process on the circle, with physical applications. *Proc R Soc A* 469:20130092