Particle Filtering for Non-linear State-Space Models for Wind Speeds and Directions

Naoya Hieda[a], Takayuki Shiohama[b,*]

[a] *Graduate School of Engineering, Tokyo University of Science*
6-3-1 Niijuku, Katsushika, Tokyo, 125-8585 JAPAN
[b] *Department of Information and Computer Technology*
Tokyo University of Science
6-3-1 Niijuku, Katsushika, Tokyo, 125-8585 JAPAN

Abstract

The state space form is a useful framework for estimating unobserved state variables from some given observations. The applications can be found in diverse areas of natural science and engineering such as ecology, epidemiology, meteorology and economics and finance. The wind speeds and directions have complex time series probability structures involving highly non-Gaussian and nonlinear transition. In this study, we consider a simulation-based inference using the sequential Monte Carlo methods for computing the posterior distributions for the state variables given all available observations. We propose an alternative approach that allows us to extend the methods of importance sampling distributions incorporating with the class of circular Markov transition densities. The resulting methods are compared with various resampling schemes with real data applications.

Keywords: circular data, EM-algorithm, state-space model, particle filter

1. Introduction

Circular (or directional) data refer to data recorded as points for which directions are measured, typically in the fields of biology, geography, medicine, and astronomy. For such data, which are usually expressed in terms of compass angles or pairs of sine and cosine variables, the beginning and end of the scale in the domain coincide. Owing to this periodicity, analyzing circular data is challenging because traditional statistics are not meaningful, and may even be misleading when the particular definition of the domain is ignored. Recent developments in circular data analysis using the statistical computing software, R, are summarized in Pewsey et al. (2013). Although most circular data are in the form of time series, little research has been carried out in the field of circular time series analysis compared with the number of circular time series modeling approaches.

In general, three main approaches are used to model circular time series. The first method is used to obtain circular-valued random variables by wrapping; one example is the wrapped autoregressive process of Breckling (2012). The second approach is based on a link function that maps a line onto a circular domain, called a linked autoregressive moving average process. This model was proposed by Fisher & Lee (1994). The last approach specifies the density of the conditional distribution, including the Markov process of Wehrly & Johnson (1980), Möbius transformation of Kato (2010), and hidden Markov models of Holzmann et al. (2006). Abe et al. (2017) studied the circular Markov process of Wehrly & Johnson (1980) and obtained theoretical circular autocorrelation structures under simple model assumptions. According to their results, circular autocorrelations are determined by the mean

*Corresponding author
Email addresses: `4417621@ed.tus.ac.jp` (Naoya Hieda), `shiohama@rs.tus.ac.jp` (Takayuki Shiohama)

resultant length of the underlying circular density of the process. Abe et al. (2018) considered the circular Markov processes whose concentration parameter could be time-varying.

Many data in directional time series applications display nonlinear features such as heteroskedasticity and a nonlinear relationship between wind direction and speed. These features become more and more relevant as the length of the observed time series increases and as the series itself is subject to changes in the dynamic structure. In this paper, we address the circular process of Wehrly & Johnson (1980), which allows time-varying concentration parameters. The proposed model can incorporate the time-varying autocorrelations of the observed circular time series. For this purpose, we introduce a simple nonparametric regression model to the model parameter with time-varying observed exogenous variables, which cause a reasonable fit of the observed time series. The proposed models are then used to illustrate how wind direction and speed are related to the time-varying parameters. For further detail on the time series analysis of wind direction, see, for example, Breckling (2012), Ailliot et al. (2006), and Fuentes et al. (2005).

In an applications in meteorology, bivariate data consists of wind speeds and directions are often modeled by using projected normal distributions (Mardia & Jupp (2009)). However, time series modeling for such a bivariate dataset is not sufficiently studied in the literature. The dataset of wind speeds and directions are called cylindrical data because circular wind direction data are observed along with linear wind speeds ones. Lagona et al. (2015) proposed a hidden Markov model for analyzing cylindrical time series, and the proposed model can adequately explain circular–linear correlation, and temporal autocorrelation of the observed data. The state space modeling using circular random variable is considered in Mazumder & Bhattacharya (2017) and Kurz et al. (2016). In this study, we extend existing circular state-space models to cope with cylindrical time series.

Sequential Monte Carlo (SMC) methods are the set of simulation-based methods which provide a convenient and attractive approach to computing posterior distributions. Over the last few years, there has been a proliferation of scientific papers on SMC methods and their applications. Several closely related algorithms, under the names of bootstrap filters, condensation, particle filters, Monte Carlo filters, interacting particle approximations and survival of the fittest, have appeared in several research fields.

In general, the parameter estimation in Gaussian and linear state-space model can be done by usual maximum likelihood estimation (MLE). However, the EM algorithm is used for estimating model parameters in the state space model, and it turns out be more robust than the direct solution of the MLE. The recent development of EM algorithms in SMC methods are explained in Kantas et al. (2015).

The remainder of this paper is organized as follows. Section 2 introduces our non-linear and non-Gaussian state- space models, then explains some of the model properties. Section 3 describes the estimation procedures unknown model parameters by sequential Monte Carlo methods. Section 4 investigates maximum likelihood based inference by a Monte Carlo simulation. Section 5 presents the data analysis, which illustrates how wind speed and direction are related through the single latent process. Finally, Section 6 concludes the study.

2. Models and Assumptions

The proposed model in this paper belongs to the Markovian, nonlinear, non-Gaussian state-space models, which can be estimated via particle filters. Let $\{\Theta_t\}_{t\in\mathbb{Z}}$, $\Theta \in \mathcal{S}^1$, be a sequence of random variables on a unit circle $\mathcal{S}^1$, and $\{\psi(\alpha_t)\}_{t\in\mathbb{Z}}$ be the sequence of time varying concentration parameters of the density function of $\{\Theta_t\}$ that are determined by suitable transformation of stationary autoregressive process $\{\alpha_t\}$. We consider the stationary Markov process on the circle defined in Wehrly & Johnson (1980) with initial probability condition $p_{\boldsymbol{\eta}}(\theta_0, \alpha_0) = f_{\boldsymbol{\eta}}(\theta_0, \psi(\alpha_0))$ and

$$\begin{aligned}
&p_{\boldsymbol{\eta}}(\theta_t|\theta_0, \dots \theta_{t-1}, \alpha_0, \dots, \alpha_{t-1}) \\
=&f_{\boldsymbol{\eta}}(\theta_t|\theta_{t-1}, \psi(\alpha_{t-1})) = 2\pi g_{\boldsymbol{\eta}}[2\pi\{F_{\boldsymbol{\eta}}(\theta_t; \psi(\alpha_t)) - F_{\boldsymbol{\eta}}(\theta_{t-1}; \psi(\alpha_{t-1}))\}],
\end{aligned}$$

where $f_{\boldsymbol{\eta}}(\cdot)$ and $g_{\boldsymbol{\eta}}(\cdot)$ are arbitrary densities on the circle with unknown parameter vectors $\boldsymbol{\eta}$. The function $\psi : \mathbb{R} \to H_\rho$ is the map from $\mathbb{R}$ to suitable concentration parameter space H_ρ of the circular density function g, where the density function g is often called as binding density, for example, Jones et al. (2015). The circular probability distribution function $F(\cdot)$ is defined by

$$F_{\boldsymbol{\eta}}(\theta; \psi(\alpha_t)) = \int_0^\theta f_{\boldsymbol{\eta}}(\xi; \psi(\alpha_t)) d\xi, \quad \theta \in [-\pi, \pi).$$

The unobserved states variable $\{\alpha_t\}_{t\in\mathbb{Z}}$, $\alpha_t \in \mathbb{R}$ is modeled with an autoregressive process with unit variance as follows:

$$\alpha_{t+1} = \phi\alpha_t + \varepsilon_{t+1}, \quad \varepsilon_t \sim \text{i.i.d.} N(0, 1/(1-\phi^2)),$$

where $\phi \in (-1, 1)$ is the autoregressive parameter. The latent process $\{\alpha_t\}$ determines both the time series of wind speeds and the concentration parameter in circular Markov transition density. Let the process $\{V_t\}_{t\in\mathbb{Z}}$ be the time series of wind speed. We assume that the process of wind speed follows a version of stochastic volatility models as:

$$V_t = \beta \exp(\alpha_t^2/2) w_t \quad w_t \sim \text{Gamma}(s_0, s_0^{-1})$$

where $\beta > 0$ and $\varepsilon_{v,t}$ is the i.i.d. gamma random variables with shape and scale parameters s_0 and s_0^{-1}, respectively, which has unit mean and variance s_0^{-1}. For details on the stochastic volatility models, we refer Taylor (1982) for simplest model, and Kim et al. (1998) for Markov chain Monte Carlo (MCMC) methods.

Hereafter we consider the wrapped Cauchy distribution for the choice of the circular Markov transition densities f and g. Then, the time-varying concentration parameter ρ_t are given by the following transformation of the latent process $\{\alpha_t\}$ as

$$\rho_t = \psi(\alpha_t) = \{\tanh(\mu_\rho + \sigma_\rho\alpha_t) + 1\}/2,$$

where $\mu_\rho \in \mathbb{R}$ and $\sigma_\rho \in \mathbb{R}_+$ are the parameters of the link function which control the relationship between wind speeds and concentration in wind directions via sigmoid function. The model parameter vector $\boldsymbol{\eta}$ considered in this study is denoted by $\boldsymbol{\eta} = (\mu_f, \rho_f, \mu_g, \phi, \mu_\rho, \sigma_\rho, \beta, s_0)^T$, where μ_f and μ_g are the location parameters in the marginal and binding distributions of wind directions, and $\rho_f \in (0, 1)$ is the concentration parameter of the marginal distribution of the wind direction. The corresponding parameter space becomes $\boldsymbol{H} = [0, 2\pi]^2 \times (0, 1) \times (-1, 1) \times \mathbb{R} \times \mathbb{R}_+^3 \subset \mathbb{R}^8$.

Denote the state variable as $\boldsymbol{x}_t = (\Theta_t, \alpha_t)^\top$ and observation variable as $\boldsymbol{y}_t = (\Theta_t, V_t)^\top$. We also denote by $\boldsymbol{x}_{0:t} \equiv \{\boldsymbol{x}_0, \ldots, \boldsymbol{x}_t\}$ and $\boldsymbol{y}_{1:t} \equiv \{\boldsymbol{y}_1, \ldots, \boldsymbol{y}_t\}$, respectively, the latent variable and the observations up to time t. We use the notation p_η for the density function which depends on $\boldsymbol{\eta}$.

The above defined models belong to the class of the nonlinear and non-Gaussian time series models as

$$\boldsymbol{x}_{t+1} = \psi_\eta(\boldsymbol{x}_t, \boldsymbol{v}_{t+1}), \quad \boldsymbol{y}_t = \phi_\eta(\boldsymbol{x}_t, \boldsymbol{w}_t),$$

where $\{\boldsymbol{v}_t\}_{t\geq 1}$ and $\{\boldsymbol{w}_t\}_{t\geq 0}$ are independent sequences of independent random variables and (ψ_η, ϕ_η) are a pair of nonlinear functions. These models are known as general state-space models as in, for example, Doucet et al. (2001), Diaconis (2003), and Douc et al. (2014).

Our aim is to estimate recursively in time the posterior distribution $p_\eta(\boldsymbol{x}_{0:t}|\boldsymbol{y}_{1:t})$, the filtering distribution $p_\eta(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$ with their expectations as well as estimating unknown parameter vector $\boldsymbol{\eta}$.

3. Particle Filtering and Smoothing

In estimating non-linear and non-Gaussian state space model, the particle filters are implemented to approximate the sequence of the conditional distribution of the latent processes $p_\eta(\boldsymbol{x}_{1:t}|\boldsymbol{y}_{1:n})$. This distribution is approximated by a large cloud of N random samples called particles in SMC. These particles are propagated over time using an importance sampling and resampling mechanisms. The Monte Carlo approximation for marginal and joint posterior distribution can be conducted in a recursive way so called prediction and update stage.

The prediction stage forecasts the state variable at time t using observations up to time $t-1$. Using prior density $p_\eta(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1})$, the predictive density $p_\eta(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})$ can be obtained by

$$p_\eta(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1}) = \int p_\eta(\boldsymbol{x}_t, \boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1})d\boldsymbol{x}_{t-1} = \int p_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})p_\eta(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1})d\boldsymbol{x}_{t-1},$$

where the transition density $p_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$ is given by equation ().

With the new observation $\boldsymbol{y}_t$ becomes available, the marginal posterior density is obtained by using Bayes theorem as follows

$$\begin{aligned} p_\eta(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) =& \frac{p_\eta(\boldsymbol{y}_t|\boldsymbol{x}_t)p_\eta(\boldsymbol{x}_t|\boldsymbol{y}_{1:t-1})}{p_\eta(\boldsymbol{y}_t|\boldsymbol{y}_{1:t-1})} \\ &\propto p_\eta(\boldsymbol{y}_t|\boldsymbol{x}_t)\int p_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})p_\eta(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1})d\boldsymbol{x}_{t-1} \end{aligned}$$

A Sequential Monte Carlo algorithm is a numerical approximation of the sequence of posterior densities $\{p_\eta(\boldsymbol{x}_t|\boldsymbol{y}_{1:t})\}$. Instead of sampling from the transition density $p_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})$, we can approximate it by sampling from a known importance density $q_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{y}_{1:t})$, then the posterior density can be expressed by using importance sampling distribution as

$$p_\eta(\boldsymbol{x}_t|\boldsymbol{y}_{1:t}) \propto p_\eta(\boldsymbol{y}_t|\boldsymbol{x}_t)\int \frac{p_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}{q_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{y}_{1:t})} q_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{y}_{1:t})p_\eta(\boldsymbol{x}_{t-1}|\boldsymbol{y}_{1:t-1})d\boldsymbol{x}_{t-1}$$

Define the weights at time $t-1$ by $\tilde{\omega}_{\eta,t-1}$, then the importance weights $\tilde{\omega}_{\eta,t}$ called particle weight, becomes

$$\tilde{\omega}_{\eta,t} = \tilde{\omega}_{\eta,t-1}\frac{p_\eta(\boldsymbol{y}_t|\boldsymbol{x}_t)p_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1})}{q_\eta(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}, \boldsymbol{y}_{1:t})}.$$

3.1. Particle Smoothing

3.2. Parameter Estimation

4. Data Analysis

To implement out proposed model as well as estimating model parameter procedures, we use the wind speeds and directions data. The data are taken from AgriMet station at Forest Grove in Oregon U. S. form 1st February to 28th February, 2015. Wind direction are the records of the direction from which the wind originates, and are in terms of degrees from north (0 degrees), and angle increases in a clockwise direction. Time Series plots for the wind directions and speeds are shown in Figure 4.1.

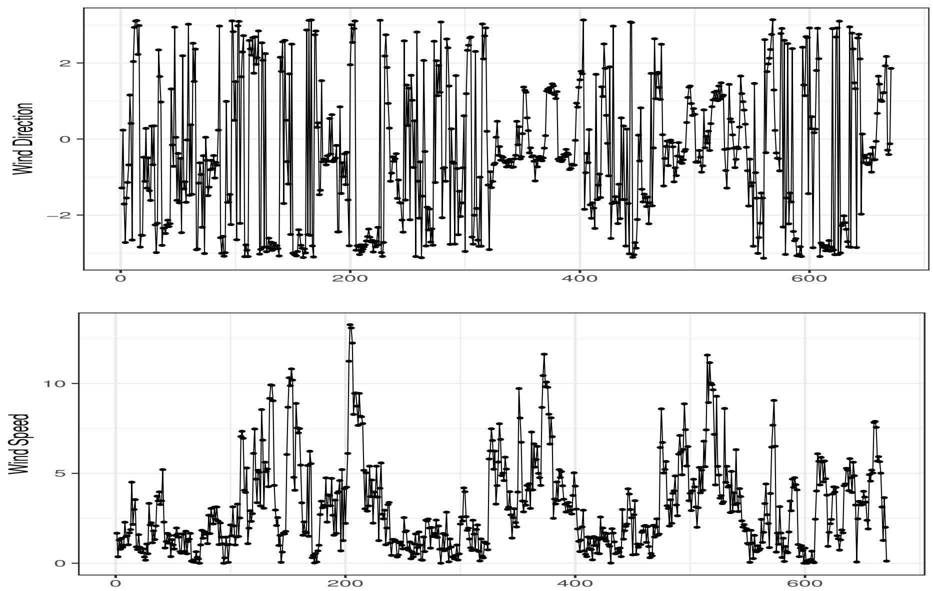

Figure 4.1: Time series plots for the wind directions (top) and speeds (bottom) in Forest Grove, Oregon, U.S. from Feb 1 to 28, 2015.

The estimated model parameters by using EM algorithm are summarized in Table=4.1.

Table 4.1: Parameter estimates by EM algorithms.

μ_f	ρ_f	μ_g	ϕ	μ_ρ	σ_ρ	β	s
-1.6533	0.2494	0.0032	0.8743	0.2209	0.5942	2.5539	2.5358

Figure 4.2 plots the observed wind directions and speeds data together with particle filtering and smoothing series.

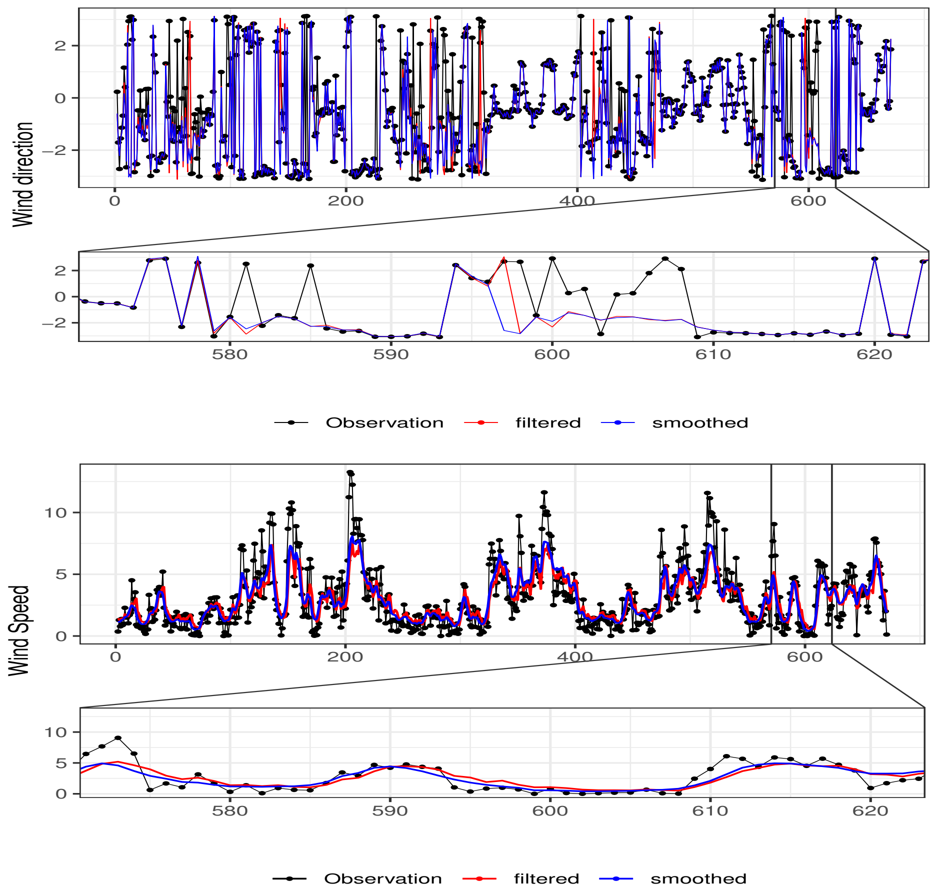

Figure 4.2: Time series plots for the wind directions (top) and speeds (bottom) together with its filtered and smoothed series.

The estimated latent autoregressive process together with time-varying concentration parameters are plotted in Figure 4.3.

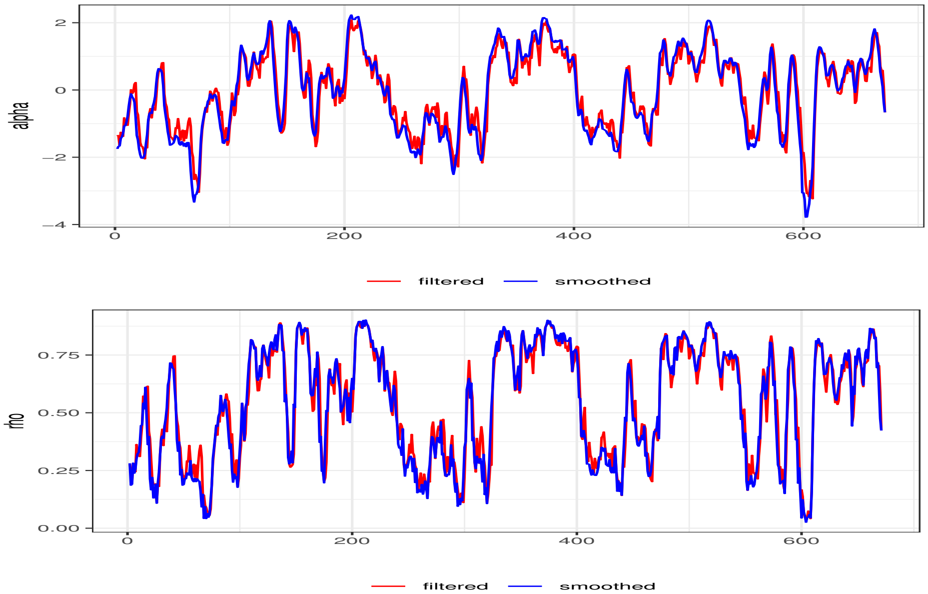

Figure 4.3: Time series plots for the latent process $\{\alpha_t\}$(top) and the time-varying concentration parameters $\{\rho_t\}$ (bottom).

References

Abe, T., Ogata, H., Shiohama, T., & Taniai, H. (2017). Circular autocorrelation of stationary circular markov processes. *Statistical Inference for Stochastic Processes*, *20*, 275–290.

Abe, T., Ogata, H., Shiohama, T., & Taniai, H. (2018). Modeling circular markov processes with time-varying concentration. *Manuscript Submitted for Publication*, .

Ailliot, P., Monbet, V., & Prevosto, M. (2006). An autoregressive model with time-varying coefficients for wind fields. *Environmetrics*, *17*, 107–117.

Breckling, J. (2012). *The analysis of directional time series: applications to wind speed and direction* volume 61. Springer Science & Business Media.

Diaconis, P. (2003). Sequential monte carlo methods in practice.

Douc, R., Moulines, E., & Stoffer, D. (2014). *Nonlinear time series: Theory, methods and applications with R examples*. Chapman and Hall/CRC.

Doucet, A., De Freitas, N., & Gordon, N. (2001). An introduction to sequential monte carlo methods. In *Sequential Monte Carlo methods in practice* (pp. 3–14). Springer.

Fisher, N., & Lee, A. (1994). Time series analysis of circular data. *Journal of the Royal Statistical Society. Series B (Methodological)*, (pp. 327–339).

Fuentes, M., Chen, L., Davis, J. M., & Lackmann, G. M. (2005). Modeling and predicting complex space–time structures and patterns of coastal wind fields. *Environmetrics: The official journal of the International Environmetrics Society*, *16*, 449–464.

Holzmann, H., Munk, A., Suster, M., & Zucchini, W. (2006). Hidden markov models for circular and linear-circular time series. *Environmental and Ecological Statistics*, *13*, 325–347.

Jones, M., Pewsey, A., & Kato, S. (2015). On a class o circulas: copulas or circular distributions. *Annals of the Institute of Statistical Mathematics*, *67*, 843–862.

Kantas, N., Doucet, A., Singh, S. S., Maciejowski, J., Chopin, N. et al. (2015). On particle methods for parameter estimation in state-space models. *Statistical science*, *30*, 328–351.

Kato, S. (2010). A markov process for circular data. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *72*, 655–672.

Kim, S., Shephard, N., & Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with arch models. *The review of economic studies*, *65*, 361–393.

Kurz, G., Gilitschenski, I., & Hanebeck, U. D. (2016). Recursive bayesian filtering in circular state spaces. *IEEE Aerospace and Electronic Systems Magazine*, *31*, 70–87.

Lagona, F., Picone, M., & Maruotti, A. (2015). A hidden markov model for the analysis of cylindrical time series. *Environmetrics*, *26*, 534–544.

Mardia, K. V., & Jupp, P. E. (2009). *Directional statistics* volume 494. John Wiley & Sons.

Mazumder, S., & Bhattacharya, S. (2017). Nonparametric dynamic state space modeling of observed circular time series with circular latent states: A bayesian perspective. *Journal of Statistical Theory and Practice*, *11*, 693–718.

Pewsey, A., Neuhäuser, M., & Ruxton, G. D. (2013). *Circular statistics in R*. Oxford University Press.

Taylor, S. J. (1982). Financial returns modelled by the product of two stochastic processes-a study of the daily sugar prices 1961-75. *Time series analysis: theory and practice*, *1*, 203–226.

Wehrly, T. E., & Johnson, R. A. (1980). Bivariate models for dependence of angular observations and a related markov process. *Biometrika*, *67*, 255–256.