

# A tractable, parsimonious and flexible model for cylindrical data, with applications

Toshihiro Abe\* and Christophe Ley\*\*<sup>1</sup>

\* Nanzan University, Nagoya, Japan

E-mail: abetosh@ss.nanzan-u.ac.jp

\*\* Ghent University, Gent, Belgium

E-mail: christophe.ley@ugent.be

---

## Abstract

New cylindrical distributions are proposed by combining the sine-skewed von Mises distribution (circular part) with the Weibull distribution (linear part). This new model, the WeiSSVM, enjoys numerous advantages: simple normalizing constant and hence very tractable density, parameter-parsimony and interpretability, good circular-linear dependence structure, easy random number generation thanks to known marginal/conditional distributions, and flexibility illustrated via excellent fitting abilities. Inferential issues, such as independence testing, circular-linear respectively linear-circular regression, can easily be tackled with the new model, which is applied on two real data sets.

*Keywords:* Circular-linear data, circular-linear regression, distributions on the cylinder, sine-skewed von Mises distribution, Weibull distribution

---

## 1. Introduction

Cylindrical data are observations that consist of a directional part (a set of angles), which is often of a circular nature (a single angle), and a linear part (mostly a positive real number). This explains the alternative terminology of directional-linear or circular-linear data. Such data occur frequently in natural sciences; typical examples are wind direction and another climatological variable such as wind speed or air temperature, the direction an animal moves and the distance moved, or wave direction and wave height. Recent studies of cylindrical data include the exploration of wind direction and SO<sub>2</sub> concentration ([1]), the analysis of Japanese earthquakes ([2]), the link between wildfire orientation and burnt area ([3]), and space-time modeling of sea currents in the Adriatic Sea ([4], [5]).

A non-trivial yet fundamental problem is the joint modeling of the directional/circular and linear variables via the construction of cylindrical probability distributions. The best known examples stem from Mardia and Sutton (1978) [6], conditioning from a trivariate normal distribution, and Johnson and Wehrly (1978) [7], invoking maximum entropy principles. The latter also provide in their paper a general way, based on copulas, to construct circular-linear distributions with specified marginals.

What desirable properties should a “good” cylindrical distribution possess? It should be able to model diverse shapes, in other words present good fitting aptitudes, yet it should ideally

---

<sup>1</sup>Corresponding author. Address: Ghent University, Department of Applied Mathematics, Computer Science and Statistics, Krijgslaan 281, S9, Campus Sterre, 9000 Gent, Belgium. Phone number: + 32 9 264 49 08.

remain of a tractable form (this is crucial for stochastic properties, estimation purposes, and circular-linear regression) and be parsimonious in terms of parameters at play. The marginal and conditional distributions should optimally be well-known and flexible (e.g., there is no reason for the circular component to be always symmetric), whilst the dependence structure has to take care of a reasonable joint behavior. Indeed, numerous examples of cylindrical data require that the circular concentration tends to increase with the linear component, as identified in the seminal paper [8].

All these conditions are well fulfilled by the new model we propose in the present paper. Its probability density function (pdf) is of the form

$$(\theta, x) \mapsto \frac{\alpha\beta^\alpha}{2\pi \cosh(\kappa)} (1 + \lambda \sin(\theta - \mu)) x^{\alpha-1} \exp[-(\beta x)^\alpha (1 - \tanh(\kappa) \cos(\theta - \mu))], \quad (1)$$

where  $(x, \theta) \in [0, \infty) \times [-\pi, \pi)$ ,  $\alpha, \beta > 0$ ,  $-\pi \leq \mu < \pi$ ,  $\kappa \geq 0$  and  $-1 \leq \lambda \leq 1$ . The roles of the distinct parameters will be explained in Section 2, as well as the construction underpinning (1). Stochastic properties such as marginal and conditional distributions, random number generation, moment and correlation calculations are presented in Section 3. We will in particular stress the capacity of our new density to model cylindrical data with length-increasing circular concentration. Maximum likelihood estimation and the ensuing efficient likelihood ratio tests (including tests for circular-linear independence) are discussed in Section 4, as well as circular-linear and linear-circular regression. A Monte Carlo simulation study (Section 5) reveals a good behavior of maximum likelihood estimates. In order to give the reader a better idea of the strengths of our new model, we review in Section 6 the main competitor cylindrical distributions from the literature, and compare them to our model on basis of objective criteria. The excellent modeling capacities of our new model are illustrated by means of two real data sets in Section 7. We conclude the paper by some final comments in Section 8, where we refer to two very recent papers that have already used our model as important building block.

## 2. A new model for circular-linear data: the WeiSSVM

Johnson and Wehrly proposed in [7] a very simple distribution able to fit cylindrical data where the circular concentration increases with the length of the linear part. Their density reads

$$(\theta, x) \mapsto \frac{\beta}{2\pi \cosh(\kappa)} \exp[-\beta x (1 - \tanh(\kappa) \cos(\theta - \mu))], \quad (2)$$

with  $-\pi \leq \mu < \pi$ ,  $\beta > 0$  and  $\kappa \geq 0$ . The linear conditional density is the (negative) exponential, while the circular conditional density given  $X = x$  is of the form

$$\theta \mapsto \frac{1}{2\pi I_0(x\beta \tanh(\kappa))} \exp[\beta x \tanh(\kappa) \cos(\theta - \mu)], \quad (3)$$

where  $I_0(\kappa)$  is the modified Bessel function of the first kind and order zero. The mapping (3) is the popular von Mises density with location  $\mu$  and concentration  $\beta x \tanh(\kappa)$ , often considered as the circular analogue of the normal distribution. We attract the reader's attention to the fact that we have slightly reparameterized the original Johnson-Wehrly parameterization which would correspond to using  $\beta$  and  $\kappa_1 = \beta \tanh(\kappa)$  instead of  $\beta$  and  $\kappa$ , and hence adding the condition that  $\kappa_1 < \beta$  in view of  $\beta / \cosh(\kappa) = (\beta^2 - \kappa_1^2)^{1/2}$ . With our parameterization we avoid this condition, which is an advantage for numerical maximization methods.

A drawback of the Johnson-Wehrly model (2) is its lack of flexibility. Both its conditional and marginal circular densities are symmetric (see Section 3.2 for details), when  $\kappa = 0$  the circular

contribution in (2) boils down to the uniform law on  $[-\pi, \pi)$ , and the circular concentration varies linearly with  $x$  (see (3)). In order to overcome these limitations, we have applied two separate transformations to the Johnson-Wehrly density: a power transformation  $x \mapsto x^{1/\alpha}$  for  $\alpha > 0$  to the linear part, and a perturbation of the circular part via multiplication with  $\theta \mapsto (1 + \lambda \sin(\theta - \mu))$  for  $\lambda \in [-1, 1]$ . The former is the classical way to turn an exponential distribution on  $\mathbb{R}^+$  to the Weibull distribution with pdf  $x \mapsto \alpha \beta x^{\alpha-1} \exp[-\beta x^\alpha]$ , which is a very popular distribution to model diverse natural phenomena. The effect of the perturbation is known in circular statistics as “sine-skewing” a reflectively symmetric distribution, see [9]. Whenever  $\lambda \neq 0$ , the resulting density becomes skewed, whereas symmetry is retrieved for  $\lambda = 0$ ; moreover, the perturbation leaves the normalizing constant untouched. The combined effect of both transformations (plus the change from  $\beta$  to  $\beta^\alpha$  mainly for aesthetic reasons) thus yields the pdf

$$\frac{\alpha \beta^\alpha}{2\pi \cosh(\kappa)} (1 + \lambda \sin(\theta - \mu)) x^{\alpha-1} \exp[-(\beta x)^\alpha (1 - \tanh(\kappa) \cos(\theta - \mu))],$$

which we term *WeiSSVM* for the interplay between the linear Weibull part and the circular sine-skewed von Mises part. 2D contour plots of the density (1) are given in Figure 1 and show the versatility of our new model.

Parameter interpretation becomes now clear:  $\mu$  and  $\lambda$  respectively endorse the role of circular location and skewness parameters, while  $\beta$  and  $\alpha$  are linear scale and shape parameters. The parameter  $\kappa$  plays, as in the original Johnson-Wehrly model, the role of circular concentration and circular-linear dependence parameter. Independence is attained when  $\kappa = 0$ , in which case the density (1) becomes the product of the linear Weibull and the circular cardioid distribution with location  $\mu + \pi/2$  and concentration  $\lambda$ , see the first row of Figure 1.

### 3. Properties of the WeiSSVM

#### 3.1. The normalizing constant

As can be seen from (1), the normalizing constant is very simple, which is a strong asset of our proposal. Indeed, it is not rare to encounter normalizing constants defined in terms of infinite series of functions, as for the Kato-Shimizu model [10], see also (8).

#### 3.2. Marginal and conditional distributions

The marginal density of the circular component  $\Theta$  from pdf (1) is given by

$$\begin{aligned} f(\theta) &= \frac{1}{2\pi \cosh(\kappa)} (1 + \lambda \sin(\theta - \mu)) \int_0^\infty \alpha \beta^\alpha x^{\alpha-1} \exp[-(\beta x)^\alpha (1 - \tanh(\kappa) \cos(\theta - \mu))] dx \\ &= \frac{1}{2\pi \cosh(\kappa)} \frac{1 + \lambda \sin(\theta - \mu)}{1 - \tanh(\kappa) \cos(\theta - \mu)} \\ &= \frac{1 - \tanh^2(\kappa/2)}{2\pi} \frac{1 + \lambda \sin(\theta - \mu)}{1 + \tanh^2(\kappa/2) - 2 \tanh(\kappa/2) \cos(\theta - \mu)}, \end{aligned}$$

which is the sine-skewed wrapped Cauchy distribution [9], a flexible extension of the symmetric wrapped Cauchy distribution.

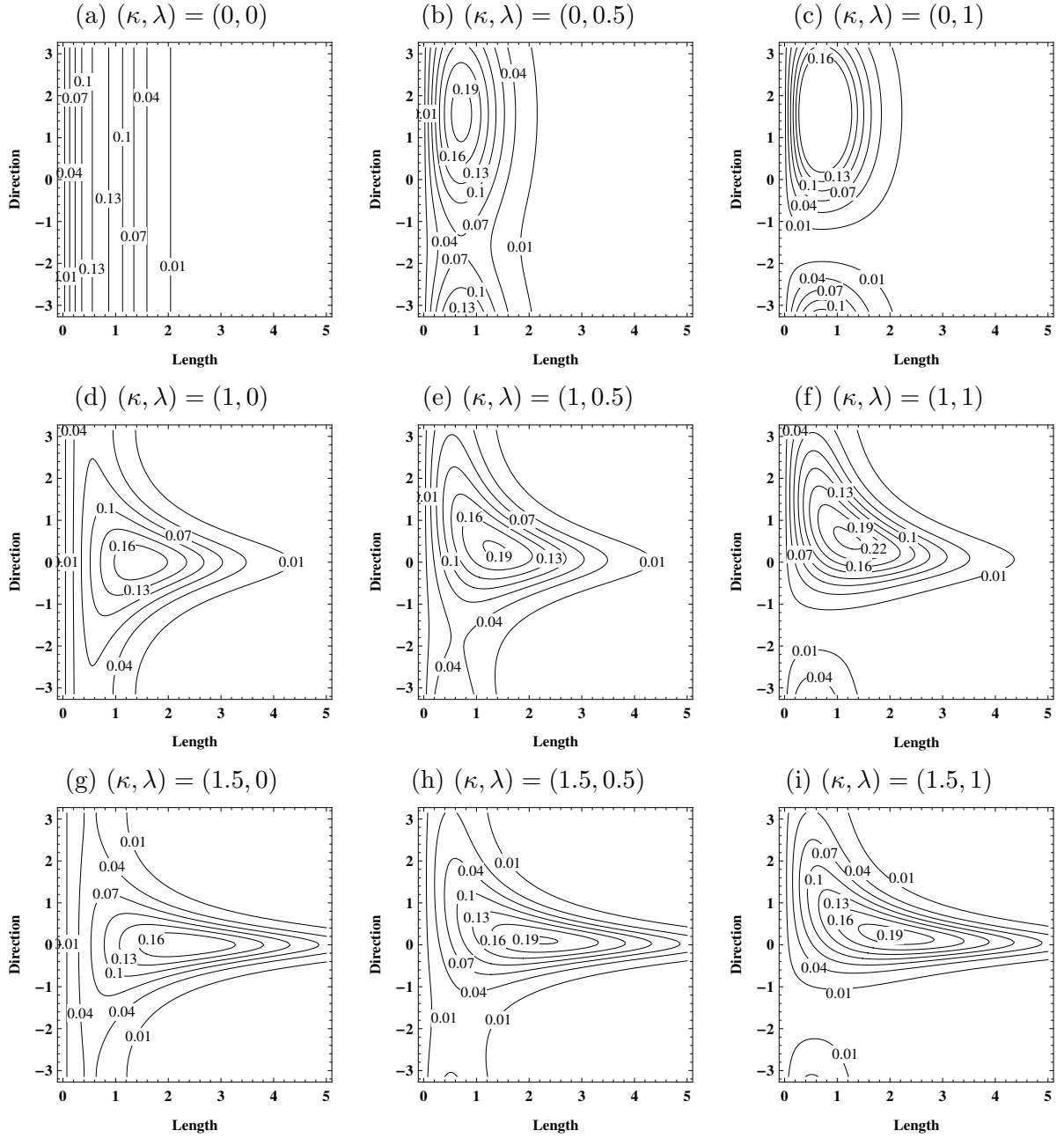


Figure 1: Contour plots of the WeiSSVM density (1) over  $[0, 5] \times [-\pi, \pi]$  for  $(\alpha, \beta, \mu) = (2, 1, 0)$  with (a)  $(\kappa, \lambda) = (0, 0)$  (Weibull and uniform) (b)  $(\kappa, \lambda) = (0, 0.5)$  (Weibull and cardioid), (c)  $(\kappa, \lambda) = (0, 1)$  (Weibull and cardioid), (d)  $(\kappa, \lambda) = (1, 0)$ , (e)  $(\kappa, \lambda) = (1, 0.5)$ , (f)  $(\kappa, \lambda) = (1, 1)$ , (g)  $(\kappa, \lambda) = (1.5, 0)$ , (h)  $(\kappa, \lambda) = (1.5, 0.5)$  and (i)  $(\kappa, \lambda) = (1.5, 1)$ .

The marginal density of the linear component  $X$  from pdf (1) in turn corresponds to

$$\begin{aligned}
f(x) &= \frac{1}{2\pi \cosh(\kappa)} \alpha \beta^\alpha x^{\alpha-1} \int_{-\pi}^{\pi} (1 + \lambda \sin(\theta - \mu)) \exp[-(\beta x)^\alpha (1 - \tanh(\kappa) \cos(\theta - \mu))] d\theta \\
&= \frac{1}{2\pi \cosh(\kappa)} \alpha \beta^\alpha x^{\alpha-1} \exp[-(\beta x)^\alpha] \int_{-\pi}^{\pi} \exp[(\beta x)^\alpha \tanh(\kappa) \cos(\theta - \mu)] d\theta \\
&= \frac{I_0(x^\alpha \beta^\alpha \tanh(\kappa))}{\cosh(\kappa)} \alpha \beta^\alpha x^{\alpha-1} \exp[-(\beta x)^\alpha].
\end{aligned}$$

This is an extended version of the marginal density of (2) given in [7]; as already noticed, it simplifies to the Weibull when  $\kappa = 0$ .

The conditional densities from (1) are now readily given by

$$f(\theta|x) = \frac{1}{2\pi I_0(x^\alpha \beta^\alpha \tanh(\kappa))} (1 + \lambda \sin(\theta - \mu)) \exp[(\beta x)^\alpha \tanh(\kappa) \cos(\theta - \mu)] \quad (4)$$

and

$$f(x|\theta) = \alpha \left[ \beta (1 - \tanh(\kappa) \cos(\theta - \mu))^{1/\alpha} \right]^\alpha x^{\alpha-1} \exp \left[ - \left( \beta (1 - \tanh(\kappa) \cos(\theta - \mu))^{1/\alpha} x \right)^\alpha \right]. \quad (5)$$

Both densities are quite common: (4) is the sine-skewed von Mises distribution with concentration  $(\beta x)^\alpha \tanh(\kappa)$  (note how the concentration now varies with  $x^\alpha$  instead of simply  $x$ ) whereas (5) is the Weibull with shape parameter  $\beta (1 - \tanh(\kappa) \cos(\theta - \mu))^{1/\alpha}$ . In general the conditional density (4) is unimodal, but it can become bimodal when the absolute value of the skewness parameter  $\lambda$  is large (from our experiment,  $|\lambda| > 0.918$ ) because of the structure of the sine-skewed von Mises distribution, see [9].

### 3.3. Random number generation

Thanks to the results of the previous section, we can describe a simple random number generation algorithm by decomposing  $f(\theta, x)$  into  $f(x|\theta)f(\theta)$ , in other words, by first generating  $\Theta \sim f(\theta)$  and then  $X|\Theta = \theta \sim f(x|\theta)$ . The algorithm goes as follows.

Step 1: Generate a random variable  $\Theta_1$  following a (symmetric) wrapped Cauchy law with location  $\mu$  and concentration  $\tanh(\kappa/2)$ , and generate independently  $U \sim Unif[0, 1]$ .

Step 2: Define  $\Theta$  as

$$\begin{cases} \Theta_1 & \text{if } U < (1 + \lambda \sin(\Theta_1 - \mu))/2 \\ -\Theta_1 & \text{if } U \geq (1 + \lambda \sin(\Theta_1 - \mu))/2; \end{cases}$$

$\Theta$  then follows the sine-skewed wrapped Cauchy distribution.

Step 3: Generate  $X$  from a Weibull with shape parameter  $\beta (1 - \tanh(\kappa) \cos(\Theta - \mu))^{1/\alpha}$ .

Random number generation from sine-skewed distributions follows from general skew-symmetric theory on  $\mathbb{R}^k$ ; see [11].

### 3.4. Moment expressions

The moments of the Weibull distribution and trigonometric moments of the sine-skewed von Mises distribution are given explicitly. These nice properties are inherited to our model.

For  $n = 1, 2, \dots$  and  $m = 1, 2, \dots$ , we have

$$\begin{aligned}
& \mathbb{E}[X^n \cos(m\Theta)] \\
&= \frac{\alpha\beta^\alpha}{2\pi \cosh(\kappa)} \int_0^\infty \int_{-\pi}^\pi x^n \cos(m\theta) (1 + \lambda \sin(\theta)) x^{\alpha-1} \exp[-(\beta x)^\alpha (1 - \tanh(\kappa) \cos(\theta))] d\theta dx \\
&= \frac{1}{2\pi \cosh(\kappa)} \int_{-\pi}^\pi \cos(m\theta) \int_0^\infty \alpha\beta^\alpha x^n x^{\alpha-1} \exp[-(\beta x)^\alpha (1 - \tanh(\kappa) \cos(\theta))] dx d\theta \\
&= \frac{1}{2\pi \cosh(\kappa)} \int_{-\pi}^\pi \cos(m\theta) \frac{\Gamma(1 + n/\alpha)}{\beta^n (1 - \tanh(\kappa) \cos(\theta))^{n/\alpha+1}} d\theta \\
&= \frac{\Gamma(n/\alpha + 1)(\cosh(\kappa))^{n/\alpha+1}}{\cosh(\kappa)\beta^n} \int_{-\pi}^\pi \frac{1}{2\pi} \cos(m\theta) \frac{1}{(\cosh(\kappa) - \sinh(\kappa) \cos(\theta))^{n/\alpha+1}} d\theta \\
&= \frac{\Gamma(n/\alpha + 1)(\cosh(\kappa))^{n/\alpha} \Gamma(n/\alpha + 1 - m) P_{n/\alpha}^m(\cosh(\kappa))}{\beta^n \Gamma(n/\alpha + 1)} \\
&= \frac{(\cosh(\kappa))^{n/\alpha} \Gamma(n/\alpha + 1 - m)}{\beta^n} P_{n/\alpha}^m(\cosh(\kappa)),
\end{aligned}$$

where  $P_\nu^m(z)$  is the associated Legendre function of the first kind of degree  $\nu$  and order  $m$  given by (equation 8.711.2 of [12], p. 969)

$$P_\nu^m(z) = \frac{(-\nu)_m}{\pi} \int_0^\pi \frac{\cos mt}{(z + \sqrt{z^2 - 1} \cos t)^{\nu+1}} dt = \frac{\Gamma(\nu + 1)}{\pi \Gamma(\nu - m + 1)} \int_{-\pi}^0 \frac{\cos mt}{(z - \sqrt{z^2 - 1} \cos t)^{\nu+1}} dt.$$

Here, we used the relation

$$(-\nu)_m = \frac{\Gamma(m - \nu)}{\Gamma(-\nu)} = (-1)^m \frac{\Gamma(\nu + 1)}{\Gamma(\nu - m + 1)}.$$

Similarly,

$$\begin{aligned}
& \mathbb{E}[X^n \sin(m\Theta)] \\
&= \frac{\alpha\beta^\alpha}{2\pi \cosh(\kappa)} \int_0^\infty \int_{-\pi}^\pi x^n \sin(m\theta) (1 + \lambda \sin(\theta)) x^{\alpha-1} \exp[-(\beta x)^\alpha (1 - \tanh(\kappa) \cos(\theta))] d\theta dx \\
&= \frac{\lambda}{2\pi \cosh(\kappa)} \int_{-\pi}^\pi \sin(m\theta) \sin(\theta) \int_0^\infty \alpha\beta^\alpha x^n x^{\alpha-1} \exp[-(\beta x)^\alpha (1 - \tanh(\kappa) \cos(\theta))] dx d\theta \\
&= \frac{\lambda}{2\pi \cosh(\kappa)} \int_{-\pi}^\pi \sin(m\theta) \sin(\theta) \frac{\Gamma(n/\alpha + 1)}{\beta^n (1 - \tanh(\kappa) \cos(\theta))^{n/\alpha+1}} d\theta \\
&= \frac{\lambda \Gamma(n/\alpha + 1)(\cosh(\kappa))^{(n/\alpha+1)}}{\cosh(\kappa)\beta^n} \int_{-\pi}^\pi \frac{1}{2\pi} \frac{(\cos((m-1)\theta) - \cos((m+1)\theta))}{2(\cosh(\kappa) - \sinh(\kappa) \cos(\theta))^{n/\alpha+1}} d\theta \\
&= \frac{\lambda \Gamma(n/\alpha + 1)(\cosh(\kappa))^{n/\alpha}}{\beta^n} \frac{1}{2} \left( \frac{\Gamma(n/\alpha + 2 - m)}{\Gamma(n/\alpha + 1)} P_{n/\alpha}^{m-1}(\cosh(\kappa)) - \frac{\Gamma(n/\alpha - m)}{\Gamma(n/\alpha + 1)} P_{n/\alpha}^{m+1}(\cosh(\kappa)) \right) \\
&= \frac{\lambda(\cosh(\kappa))^{n/\alpha}}{\beta^n} \frac{\left( \Gamma(n/\alpha + 2 - m) P_{n/\alpha}^{m-1}(\cosh(\kappa)) - \Gamma(n/\alpha - m) P_{n/\alpha}^{m+1}(\cosh(\kappa)) \right)}{2}.
\end{aligned}$$

Specifying choices for  $m$  and  $n$ , and noting that the marginal of the circular part is the sine-skewed wrapped Cauchy density, we obtain the following simple moment expressions (we write  $P_\nu(z)$  for  $P_\nu^0(z)$ )

$$\begin{aligned}
\mathbb{E}[X] &= \frac{(\cosh(\kappa))^{1/\alpha} \Gamma\left(\frac{1}{\alpha} + 1\right)}{\beta} P_{1/\alpha}(\cosh(\kappa)), \\
\mathbb{E}[X^2] &= \frac{(\cosh(\kappa))^{2/\alpha} \Gamma\left(\frac{2}{\alpha} + 1\right)}{\beta^2} P_{2/\alpha}(\cosh(\kappa)), \\
\mathbb{E}[\cos(\Theta)] &= \tanh\left(\frac{\kappa}{2}\right), \quad \mathbb{E}[\cos^2(\Theta)] = \frac{1}{2} \left(1 + \tanh^2\left(\frac{\kappa}{2}\right)\right), \\
\mathbb{E}[\sin(\Theta)] &= \frac{\lambda}{2 \cosh^2\left(\frac{\kappa}{2}\right)}, \quad \mathbb{E}[\sin^2(\Theta)] = \frac{1}{2 \cosh^2\left(\frac{\kappa}{2}\right)}, \\
\mathbb{E}[X \cos(\Theta)] &= \frac{(\cosh(\kappa))^{1/\alpha} \Gamma\left(\frac{1}{\alpha}\right)}{\beta} P_{1/\alpha}^1(\cosh(\kappa)), \\
\mathbb{E}[X \sin(\Theta)] &= \frac{\lambda(\cosh(\kappa))^{1/\alpha}}{\beta} \frac{\left(\Gamma\left(\frac{1}{\alpha} + 1\right) P_{1/\alpha}(\cosh(\kappa)) - \Gamma\left(\frac{1}{\alpha} - 1\right) P_{1/\alpha}^2(\cosh(\kappa))\right)}{2}, \\
\mathbb{E}[\cos(\Theta) \sin(\Theta)] &= \frac{\lambda \tanh\left(\frac{\kappa}{2}\right)}{4 \cosh^2\left(\frac{\kappa}{2}\right)}.
\end{aligned}$$

### 3.5. Circular-linear correlation

From the moment expressions of the previous section we readily derive the following quantities:

$$\begin{aligned}
\text{Var}[X] &= \frac{(\cosh(\kappa))^{2/\alpha} \left(\Gamma\left(\frac{2}{\alpha} + 1\right) P_{2/\alpha}(\cosh(\kappa)) - \Gamma\left(\frac{1}{\alpha} + 1\right)^2 P_{1/\alpha}(\cosh(\kappa))^2\right)}{\beta^2}, \\
\text{Var}[\cos(\Theta)] &= \frac{1}{2 \cosh^2\left(\frac{\kappa}{2}\right)}, \\
\text{Var}[\sin(\Theta)] &= \frac{1}{2 \cosh^2\left(\frac{\kappa}{2}\right)} \left(1 - \frac{\lambda^2}{2 \cosh^2\left(\frac{\kappa}{2}\right)}\right), \\
\text{Cov}(X, \cos(\Theta)) &= \frac{(\cosh(\kappa))^{1/\alpha} \left(\Gamma\left(\frac{1}{\alpha}\right) P_{1/\alpha}^1(\cosh(\kappa)) - \Gamma\left(\frac{1}{\alpha} + 1\right) P_{1/\alpha}(\cosh(\kappa)) \tanh\left(\frac{\kappa}{2}\right)\right)}{\beta}, \\
\text{Cov}(X, \sin(\Theta)) &= \frac{\lambda(\cosh(\kappa))^{1/\alpha} \left(\Gamma\left(\frac{1}{\alpha} + 1\right) \tanh^2\left(\frac{\kappa}{2}\right) P_{1/\alpha}(\cosh(\kappa)) - \Gamma\left(\frac{1}{\alpha} - 1\right) P_{1/\alpha}^2(\cosh(\kappa))\right)}{2\beta}, \\
\text{Cov}(\cos(\Theta), \sin(\Theta)) &= -\frac{\lambda \tanh\left(\frac{\kappa}{2}\right)}{4 \cosh^2\left(\frac{\kappa}{2}\right)}.
\end{aligned}$$

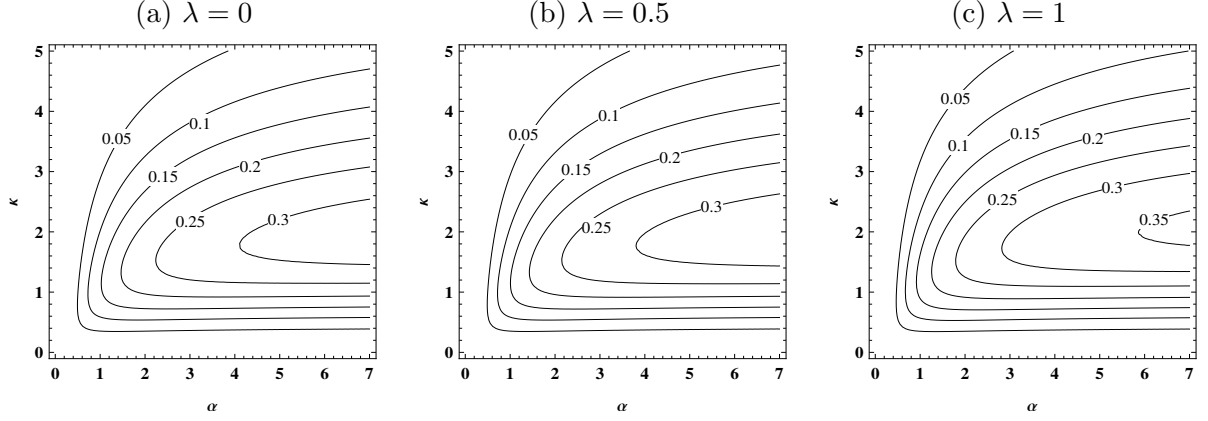


Figure 2: Contour plots of the circular-linear correlation  $R_{x\theta}^2$  as a function of  $(\alpha, \kappa)$  over  $[0, 7] \times [0, 5]$  for (a)  $\lambda = 0$ , (b)  $\lambda = 0.5$ , (c)  $\lambda = 1$ .

Using these expressions, the correlations are given by

$$\begin{aligned}
r_{xc} &= \text{Corr}(X, \cos(\Theta)) \\
&= \frac{\sqrt{2} \cosh\left(\frac{\kappa}{2}\right) \left( \Gamma\left(\frac{1}{\alpha}\right) P_{1/\alpha}^1(\cosh(\kappa)) - \Gamma\left(\frac{1}{\alpha} + 1\right) P_{1/\alpha}(\cosh(\kappa)) \tanh\left(\frac{\kappa}{2}\right) \right)}{\sqrt{\Gamma\left(\frac{2}{\alpha} + 1\right) P_{2/\alpha}(\cosh(\kappa)) - \Gamma\left(\frac{1}{\alpha} + 1\right)^2 P_{1/\alpha}(\cosh(\kappa))^2}}, \\
r_{xs} &= \text{Corr}(X, \sin(\Theta)) \\
&= \frac{\lambda \cosh\left(\frac{\kappa}{2}\right) \left( \Gamma\left(\frac{1}{\alpha} + 1\right) \tanh^2\left(\frac{\kappa}{2}\right) P_{1/\alpha}(\cosh(\kappa)) - \Gamma\left(\frac{1}{\alpha} - 1\right) P_{1/\alpha}^2(\cosh(\kappa)) \right)}{\sqrt{2} \sqrt{\Gamma\left(\frac{2}{\alpha} + 1\right) P_{2/\alpha}(\cosh(\kappa)) - \Gamma\left(\frac{1}{\alpha} + 1\right)^2 P_{1/\alpha}(\cosh(\kappa))^2} \sqrt{1 - \frac{\lambda^2}{2 \cosh^2\left(\frac{\kappa}{2}\right)}}}, \\
r_{cs} &= \text{Corr}(\cos(\Theta), \sin(\Theta)) \\
&= -\frac{\lambda \tanh\left(\frac{\kappa}{2}\right)}{2 \sqrt{1 - \frac{\lambda^2}{2 \cosh^2\left(\frac{\kappa}{2}\right)}}}.
\end{aligned}$$

The circular-linear correlation, which was proposed by [13] and [14], can then be obtained via the formula

$$R_{x\theta}^2 = \frac{r_{xc}^2 + r_{xs}^2 - 2r_{cs}r_{xc}r_{xs}}{1 - r_{cs}^2}.$$

We see from the above expressions that the circular-linear correlation neither depends on the parameter  $\beta$  nor on the sign of  $\lambda$ . In particular, when  $\lambda = 0$ , it simplifies to  $R_{x\theta}^2 = r_{xc}^2$ . The influence of  $\alpha$ ,  $\kappa$  and  $|\lambda|$  are shown via contour plots in Figure 2. We observe that, for fixed  $\kappa$  and  $\lambda$ , the correlation increases with  $\alpha$ . The influence of  $\kappa$  goes as follows: at fixed  $\alpha$  and  $\lambda$ , the correlation increases for small values of  $\kappa$  until it reaches its maximum, and then decreases. This phenomenon is directly inherited from the Johnson-Wehrly construction, as we show in Figure 3.

### 3.6. A generalization of the WeiSSVM

A natural generalization of our WeiSSVM model consists in replacing the linear Weibull part with the *Generalized Gamma distribution* [15], resulting in the Generalized Gamma sine-skewed

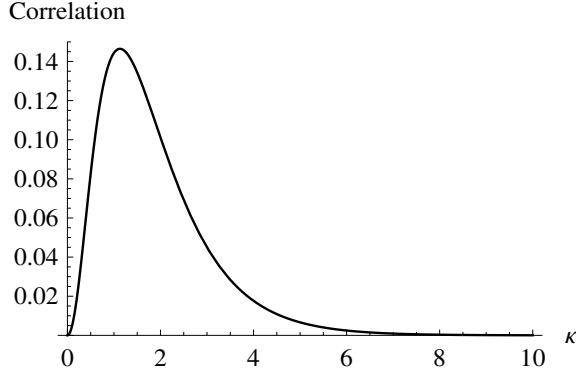


Figure 3: Circular-linear correlation  $R_{x\theta}^2$  for the Johnson-Wehrly model ( $\alpha = 1$  and  $\lambda = 0$ ) as a function of  $\kappa$ .

von Mises (GGSSVM) density

$$(\theta, x) \mapsto C (1 + \lambda \sin(\theta - \mu)) x^{\alpha-1} \exp[-(\beta x)^\gamma (1 - \tanh(\kappa) \cos(\theta - \mu))], \quad (6)$$

with  $\alpha, \gamma, \beta > 0$ ,  $\kappa \geq 0$ ,  $-1 \leq \lambda \leq 1$  and  $-\pi \leq \mu < \pi$ . The normalizing constant is calculated as follows:

$$\begin{aligned} & \int_{-\pi}^{\pi} \int_0^{\infty} (1 + \lambda \sin(\theta - \mu)) x^{\alpha-1} \exp[-(\beta x)^\gamma (1 - \tanh(\kappa) \cos(\theta - \mu))] dx d\theta \\ &= \frac{\Gamma(\alpha/\gamma)}{\gamma \beta^\alpha} \int_{-\pi}^{\pi} \frac{1}{(1 - \tanh(\kappa) \cos(\Theta))^{\alpha/\gamma}} d\Theta \\ &= \frac{2\pi \Gamma(\alpha/\gamma) (\cosh(\kappa))^{\alpha/\gamma} P_{\alpha/\gamma-1}(\cosh(\kappa))}{\gamma \beta^\alpha}. \end{aligned}$$

The WeiSSVM clearly corresponds to  $\gamma = \alpha$  in (6). All properties of the GGSSVM are obtained along the same lines as our developments in the previous sections, albeit with more involved calculations. It is to be noted that the circular marginal distribution for the GGSSVM is the sine-skewed Jones–Pewsey distribution (see [9, 16]).

We prefer the WeiSSVM over the GGSSVM because of its simplicity, parameter parsimony, higher tractability and its neat link with the Johnson-Wehrly distribution. This explains why the main focus of the present paper lies on the WeiSSVM and why we only briefly mention the density (6).

## 4. Statistical inference

### 4.1. Parameter estimation

Let  $(\theta_1, x_1), \dots, (\theta_n, x_n)$  be a sample of  $n$  independent and identically distributed couples of angular and linear observations drawn from the distribution with density (1). Then the log-likelihood function can be expressed as

$$\begin{aligned} \ell(\alpha, \beta, \mu, \kappa, \lambda) &= (\alpha - 1) \sum_{i=1}^n \log x_i - \beta^\alpha \sum_{i=1}^n x_i^\alpha (1 - \tanh(\kappa) \cos(\theta_i - \mu)) \\ &+ \sum_{i=1}^n \log(1 + \lambda \sin(\theta_i - \mu)) + n(\alpha \log \beta + \log \alpha - \log(2\pi \cosh(\kappa))). \end{aligned} \quad (7)$$

The elements of the score vector are just the first-order partial derivatives of (7) with respect to each of the parameters:

$$\begin{aligned}\frac{\partial \ell}{\partial \alpha} &= \sum_{i=1}^n \log x_i - \beta^\alpha \sum_{i=1}^n \log(\beta x_i) x_i^\alpha (1 - \tanh(\kappa) \cos(\theta_i - \mu)) + n \left( \log \beta + \frac{1}{\alpha} \right), \\ \frac{\partial \ell}{\partial \beta} &= -\alpha \beta^{\alpha-1} \sum_{i=1}^n x_i^\alpha (1 - \tanh(\kappa) \cos(\theta_i - \mu)) + \frac{n\alpha}{\beta}, \\ \frac{\partial \ell}{\partial \mu} &= \beta^\alpha \tanh(\kappa) \sum_{i=1}^n x_i^\alpha \sin(\theta_i - \mu) - \lambda \sum_{i=1}^n \frac{\cos(\theta_i - \mu)}{1 + \lambda \sin(\theta_i - \mu)}, \\ \frac{\partial \ell}{\partial \kappa} &= \frac{\beta^\alpha}{(\cosh(\kappa))^2} \sum_{i=1}^n x_i^\alpha \cos(\theta_i - \mu) - n \tanh(\kappa), \\ \frac{\partial \ell}{\partial \lambda} &= \sum_{i=1}^n \frac{\sin(\theta_i - \mu)}{1 + \lambda \sin(\theta_i - \mu)}.\end{aligned}$$

It is difficult to give closed-form expressions for the maximum likelihood estimates (MLEs), hence numerical methods should be used to find the solutions. We used the function **NMaximize** in *Mathematica* where the default numerical maximization algorithm is *Nelder-Mead*, and added the respective constraints on the parameters (e.g.,  $\mu$  lying between  $-\pi$  and  $\pi$ ). We encountered no problems in the optimization procedure.

#### 4.2. Submodel and independence testing

Testing for submodels of the WeiSSVM model is straightforward via likelihood ratio tests. For each parameter  $\psi \in \{\alpha, \beta, \mu, \kappa, \lambda\}$ , we denote  $\hat{\psi}$  the unconstrained maximum likelihood estimate and  $\hat{\psi}_0$  the maximum likelihood estimate under the respective null hypotheses. Two particular instances are of interest. On the one hand, testing for the Johnson-Wehrly submodel, which is taken care of by the test statistic

$$T_{\text{JW}} = -2(\log \ell(1, \hat{\beta}_0, \hat{\mu}_0, \hat{\kappa}_0, 0) - \log \ell(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\kappa}, \hat{\lambda})),$$

rejecting  $\mathcal{H}_0 : (\alpha = 1) \cap (\lambda = 0)$  at asymptotic level  $\eta$  whenever  $T_{\text{JW}}$  exceeds  $\chi_{2,1-\eta}^2$ , the  $\eta$ -upper quantile of the chi-square distribution with 2 degrees of freedom. On the other hand, we are interested in testing for circular-linear independence via the test statistic

$$T_{\text{Indep}} = -2(\log \ell(\hat{\alpha}_0, \hat{\beta}_0, \hat{\mu}_0, 0, \hat{\lambda}_0) - \log \ell(\hat{\alpha}, \hat{\beta}, \hat{\mu}, \hat{\kappa}, \hat{\lambda})),$$

rejecting  $\mathcal{H}_0 : \kappa = 0$  in favor of  $\mathcal{H}_1 : \kappa > 0$  at asymptotic level  $\eta$  whenever  $T_{\text{Indep}}$  exceeds  $\chi_{1,1-\eta}^2$ . Such tests have a long-standing history in the statistical literature; see [3] for a recent proposal, based on directional-linear kernel density estimation, and for references.

#### 4.3. Circular-linear and linear-circular regression

Since the conditional distributions take a very simple form, the WeiSSVM model lends itself for circular-linear as well as linear-circular regression, similarly as in [7]. The mean and variance of  $X$  given  $\Theta = \theta$  correspond to

$$\mathbb{E}[X|\Theta = \theta] = \frac{1}{\beta (1 - \tanh(\kappa) \cos(\theta - \mu))^{1/\alpha}} \Gamma\left(\frac{1}{\alpha} + 1\right)$$

and

$$\text{Var}[X|\Theta = \theta] = \frac{1}{\beta^2 (1 - \tanh(\kappa) \cos(\theta - \mu))^{2/\alpha}} \left( \Gamma\left(\frac{2}{\alpha} + 1\right) - \Gamma\left(\frac{1}{\alpha} + 1\right)^2 \right),$$

respectively. The first mean direction and mean resultant length of  $\Theta$  given  $X = x$  are provided by

$$\mu_{1|X=x} = \arg((\beta x)^\alpha \tanh(\kappa) + i\lambda),$$

with  $\arg$  denoting the argument of a complex number, and

$$\rho_{1|X=x} = \frac{I_1((\beta x)^\alpha \tanh(\kappa))}{(\beta x)^\alpha \tanh(\kappa) I_0((\beta x)^\alpha \tanh(\kappa))} \sqrt{(\beta x)^{2\alpha} \tanh^2(\kappa) + \lambda^2},$$

respectively. The parameters in each regression model are readily estimated via maximum likelihood, see Section 4.1.

## 5. Monte Carlo simulation study

In order to provide further insight into the statistical properties of the WeiSSVM, we investigated the performance of ML estimation on basis of a Monte Carlo simulation study with sample sizes  $n = 30, 50$  and  $100$ . For each sample size, we generated  $M = 3000$  independent samples of size  $n$  using the random number generation described in Section 3.3. We did this for various combinations of parameter values  $\alpha, \beta, \mu, \kappa$  and  $\lambda$ . At each replication we computed the biases and mean squared errors of the parameter estimates. The averaged results and choices of the parameters are shown in Table 1 and Table 2.

From both tables, the bias and MSE of all estimators decrease as  $n$  increases, as one could expect. It is also not surprising that the bias for  $\lambda$  when the true value lies on the boundary ( $-1$  and  $1$ ) is higher than in the non-boundary cases; the situation is inverted for the MSE. Overall no particular anomaly could be observed, which further indicates that the parameters in the model are well identified.

## 6. A brief review of, and comparison with, existing cylindrical models from the literature

For the sake of consistency with the original proposals in the literature, we shall use in what follows the same parameters as the authors of the diverse proposals. This entails, of course, that some of our parameters (e.g., the skewness parameter  $\lambda$ ) will endorse different roles in the following models; this should however not raise any concerns, as we explain in detail the parameters for each model.

### 6.1. The Mardia-Sutton and Kato-Shimizu models

Kato and Shimizu (2008) [10] propose a cylindrical distribution as an extension of the distribution by Mardia and Sutton (1978) [6]. Their model has as density

$$f_{KS}(\theta, x) = C \exp \left[ -\frac{\{x - \mu(\theta)\}^2}{2\sigma^2} + \kappa_1 \cos(\theta - \mu_1) + \kappa_2 \cos\{2(\theta - \mu_2)\} \right], \quad (8)$$

Table 1: Bias and MSE of the maximum likelihood estimates of the parameters of the WeiSSVM density, obtained on basis of 3000 simulated samples of size  $n$  from the WeiSSVM model with  $\alpha = 2$ ,  $\beta = 0.1$ ,  $\mu = 0$ ,  $\kappa = 1$  and  $\lambda = -1, -0.5, 0, 0.5, 1$ .

$\lambda$	$n$	$\alpha$	$\beta$	$\mu$	$\kappa$	$\lambda$
-1	30	0.1323 (0.1124)	0.0045 (0.0003)	-0.0381 (0.0283)	0.0638 (0.0558)	0.0584 (0.0268)
	50	0.0748 (0.0557)	0.0028 (0.0001)	-0.0239 (0.0145)	0.0353 (0.0312)	0.0361 (0.0096)
	100	0.0388 (0.0243)	0.0015 (0.0001)	-0.0163 (0.0068)	0.0220 (0.0149)	0.0204 (0.0028)
-0.5	30	0.1248 (0.1069)	0.0024 (0.0002)	-0.0129 (0.0591)	0.0366 (0.0543)	0.0248 (0.1374)
	50	0.0697 (0.0553)	0.0012 (0.0001)	0.0009 (0.0302)	0.0164 (0.0291)	-0.0024 (0.0750)
	100	0.0380 (0.0245)	0.0007 (0.0001)	0.0003 (0.0114)	0.0116 (0.0124)	-0.0004 (0.0335)
0	30	0.1229 (0.1099)	0.0017 (0.0002)	-0.0072 (0.0692)	0.0207 (0.0517)	0.0080 (0.1670)
	50	0.0792 (0.0547)	0.0013 (0.0000)	-0.0019 (0.0316)	0.0223 (0.0260)	-0.0034 (0.0913)
	100	0.0383 (0.0250)	0.0005 (0.0000)	-0.0030 (0.0129)	0.0090 (0.0118)	0.0005 (0.0452)
0.5	30	0.1206 (0.1068)	0.0024 (0.0002)	0.0032 (0.0551)	0.0301 (0.0524)	-0.0053 (0.1248)
	50	0.0806 (0.0544)	0.0011 (0.0001)	0.0032 (0.0275)	0.0191 (0.0295)	-0.0109 (0.0743)
	100	0.0340 (0.0250)	0.0006 (0.0001)	-0.0008 (0.0127)	0.0058 (0.0127)	-0.0088 (0.0366)
1	30	0.1277 (0.1100)	0.0048 (0.0003)	0.0407 (0.0302)	0.0648 (0.0600)	-0.0558 (0.0237)
	50	0.0749 (0.0572)	0.0029 (0.0001)	0.0243 (0.0151)	0.0395 (0.0313)	-0.0361 (0.0099)
	100	0.0351 (0.0255)	0.0017 (0.0001)	0.0120 (0.0067)	0.0228 (0.0153)	-0.0205 (0.0026)

Table 2: Bias and MSE of the maximum likelihood estimates of the parameters of the WeiSSVM density, obtained on basis of 3000 simulated samples of size  $n$  from the WeiSSVM model with  $\alpha = 2$ ,  $\beta = 1$ ,  $\mu = 0$ ,  $\kappa = 1$  and  $\lambda = -1, -0.5, 0, 0.5, 1$ .

$\lambda$	$n$	$\alpha$	$\beta$	$\mu$	$\kappa$	$\lambda$
-1	30	0.1262 (0.1118)	0.0479 (0.0292)	-0.0396 (0.0285)	0.0661 (0.0541)	0.0592 (0.0249)
	50	0.0735 (0.0536)	0.0285 (0.0150)	-0.0229 (0.0153)	0.0382 (0.0317)	0.0398 (0.0105)
	100	0.0354 (0.0240)	0.0153 (0.0064)	-0.0141 (0.0068)	0.0231 (0.0149)	0.0219 (0.0029)
-0.5	30	0.1320 (0.1138)	0.0247 (0.0233)	-0.0054 (0.0605)	0.0333 (0.0543)	0.0124 (0.1309)
	50	0.0775 (0.0572)	0.0096 (0.0114)	-0.0024 (0.0325)	0.0185 (0.0301)	0.0062 (0.0773)
	100	0.0365 (0.0254)	0.0069 (0.0057)	-0.0015 (0.0121)	0.0115 (0.0135)	0.0069 (0.0346)
0	30	0.1170 (0.1086)	0.0163 (0.0201)	-0.0152 (0.0786)	0.0174 (0.0513)	0.0204 (0.1692)
	50	0.0674 (0.0566)	0.0115 (0.0105)	-0.0033 (0.0346)	0.0167 (0.0279)	0.0092 (0.0940)
	100	0.0406 (0.0254)	0.0034 (0.0048)	-0.0030 (0.0131)	0.0096 (0.0119)	-0.0004 (0.0453)
0.5	30	0.1207 (0.1079)	0.0262 (0.0222)	0.0156 (0.0674)	0.0312 (0.0561)	-0.0195 (0.1343)
	50	0.0763 (0.0570)	0.0121 (0.0120)	0.0051 (0.0282)	0.0202 (0.0291)	-0.0094 (0.0746)
	100	0.0389 (0.0234)	0.0077 (0.0058)	0.0021 (0.0131)	0.0091 (0.0130)	-0.0024 (0.0349)
1	30	0.1291 (0.1113)	0.0466 (0.0259)	0.0427 (0.0273)	0.0662 (0.0572)	-0.0602 (0.0242)
	50	0.0751 (0.0577)	0.0293 (0.0144)	0.0252 (0.0154)	0.0412 (0.0331)	-0.0391 (0.0109)
	100	0.0364 (0.0245)	0.0156 (0.0064)	0.0136 (0.0064)	0.0190 (0.0149)	-0.0190 (0.0025)

where  $-\pi \leq \theta < \pi$ ,  $-\infty < x < \infty$ ,  $\sigma > 0$ ,  $\kappa_1, \kappa_2 > 0$ ,  $-\pi \leq \mu_1 < \pi$ ,  $-\pi/2 \leq \mu_2 < \pi/2$ ,  $\mu(\theta) = \mu' + \lambda \cos(\theta - \nu)$ ,  $-\infty < \mu' < \infty$ ,  $\lambda > 0$ ,  $-\pi \leq \nu < \pi$  and its normalizing constant  $C$  is provided by

$$C^{-1} = (2\pi)^{3/2} \sigma \left\{ I_0(\kappa_1) I_0(\kappa_2) + 2 \sum_{j=1}^{\infty} I_j(\kappa_2) I_{2j}(\kappa_1) \cos\{2j(\mu_1 - \mu_2)\} \right\}.$$

The conditional distribution of  $X$  given  $\Theta = \theta$  is a normal distribution and the marginal distribution of  $\Theta$  is the generalized von Mises distribution [17]. The conditional distribution of  $\Theta$  given  $X = x$  is also the generalized von Mises distribution, and the marginal distribution of  $X$  does not admit a simple form; see [10] for details. The dependence is obviously regulated via their parameter  $\lambda$ , independence occurring for  $\lambda = 0$ , leading to the product of a normal and the generalized von Mises. A clear drawback of the Kato-Shimizu model is that the density involves an infinite sum in the normalizing constant which, in practice, must be approximated using a finite sum of central terms.

The Mardia-Sutton model is obtained by setting  $\kappa_2 = 0$  in (8). The infinite sum in the normalizing constant then vanishes, resulting in a simpler density. All properties from above of course are the same, except that the generalized von Mises is replaced with the von Mises.

## 6.2. The Johnson-Wehrly-2 and Johnson-Wehrly-3 models

Besides what we may now call ExpVM model, [7] have also proposed the density

$$f_{\text{JW2}}(\theta, x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\kappa^2/(4\sigma^2)} C \exp \left\{ -\frac{(x - \lambda)^2}{2\sigma^2} + \frac{\kappa x}{\sigma^2} \cos(\theta - \mu) \right\},$$

where  $-\pi \leq \theta < \pi$ ,  $-\infty < x < \infty$ ,  $-\infty < \lambda < \infty$ ,  $\kappa, \sigma > 0$  and  $-\pi \leq \mu < \pi$ , and with normalizing constant

$$C^{-1} = 2\pi \left\{ I_0 \left( \frac{\kappa\lambda}{\sigma^2} \right) I_0 \left( \frac{\kappa^2}{4\sigma^2} \right) + 2 \sum_{j=1}^{\infty} I_j \left( \frac{\kappa^2}{4\sigma^2} \right) I_{2j} \left( \frac{\kappa\lambda}{\sigma^2} \right) \right\}.$$

As in the previous model, the conditional distribution of  $X$  given  $\Theta = \theta$  is a normal distribution and the marginal distribution of  $\Theta$  is the generalized von Mises, whereas the conditional distribution of  $\Theta$  given  $X = x$  is the von Mises distribution, and the marginal distribution of  $X$  is proportional to  $\exp\{-(x - \lambda)^2/(2\sigma^2)\} I_0(\kappa x/\sigma^2)$ ; see [10], who have studied the Johnson-Wehrly-2 model, for more details. [7] have noticed as drawback that, in case of independence (here,  $\kappa = 0$ ), the circular component is forced to be uniform.

In order to overcome the latter limitation, Johnson and Wehrly have further proposed the density

$$f_{\text{JW3}}(\theta, x) = C \exp \{ -\lambda x + \kappa x \cos(\theta - \mu_1) + \nu \cos(\theta - \mu_2) \}, \quad (9)$$

where  $-\pi \leq \theta < \pi$ ,  $0 < x < \infty$ ,  $\lambda > \kappa > 0$  and  $-\pi \leq \mu_1, \mu_2 < \pi$ , and with normalizing constant

$$C^{-1} = \frac{2\pi}{\sqrt{\lambda^2 - \kappa^2}} \left[ I_0(\nu) + 2 \sum_{j=1}^{\infty} \frac{\kappa^j I_j(\nu) \cos\{j(\mu_1 - \mu_2)\}}{(\lambda + \sqrt{\lambda^2 - \kappa^2})^j} \right].$$

This density has as conditional circular distribution the von Mises and as conditional linear distribution an exponential; the circular marginal density is proportional to  $\exp\{\nu \cos(\theta - \mu_2)\}/\{\lambda - \kappa \cos(\theta - \mu_1)\}$ , while the linear marginal is of a complicated form. Independence is attained at  $\kappa = 0$ , with (9) becoming the product of an exponential and the von Mises density.

### 6.3. The Fernández-Durán model

[7] have further proposed a general, copula-like way of defining a cylindrical density, namely via the expression

$$(\theta, x) \mapsto 2\pi g\{2\pi(F_{\Theta}(\theta) + F_X(x))\}f_{\Theta}(\theta)f_X(x), \quad (10)$$

where  $g$  and  $f_{\Theta}$  are circular densities,  $f_X$  is a linear density, and  $F_{\Theta}$  and  $F_X$  stand for the corresponding cumulative distribution functions. As established in Theorem 5 of [7], such a formulation ensures the marginal densities are  $f_{\Theta}$  and  $f_X$ , respectively. When  $g$  is uniform, (10) becomes the simple product of both marginals. This nice construction, which does *not* underpin our model (1), has been put to use by Fernández-Durán (2007) [18] with the Weibull as linear component  $f_X$  and both  $g$  and  $f_{\Theta}$  circular densities based on nonnegative trigonometric sums, of the form  $\frac{1}{2\pi} + \frac{1}{\pi} \sum_{j=1}^n (a_j \cos(j\theta) + b_j \sin(j\theta))$  with  $a_j - ib_j = 2\pi \sum_{\nu=0}^{n-j} c_{\nu+j} \bar{c}_{\nu}$  for complex numbers  $c_j$  such that  $\sum_{j=0}^n |c_j|^2 = (2\pi)^{-1}$ ; see [18] for details. The number  $n$  of terms in the sum is not fixed, hence figures as an additional parameter ( $n = 0$  is the uniform,  $n = 1$  the cardioid). Conditional densities are given by standard copula theory, but their forms are usually not known and of a complicated form for  $n > 1$ .

### 6.4. Comparison

Cylindrical distributions can be compared in terms of the following criteria: (i) is the density expressed in simple terms, hence tractable, (ii) is the (in)dependence structure “good” in the sense of [7], and (iii) does the model give rise to reasonable (in the sense of well-known) marginal and conditional distributions.

We hope to have convincingly demonstrated in Section 3 that our new model, the WeiSSVM, does satisfy these criteria. The classical Mardia-Sutton model enjoys the same properties, except that the marginal linear distribution is of a more complicated form. Its extension by Kato and Shimizu suffers from the same flaw, plus has a more involved density due to the infinite sum in the normalizing constant. As already mentioned earlier in the paper, the Johnson-Wehrly 1 model fails to satisfy (ii), idem for the model JW2 that moreover also has no simple density. The latter comment also applies to JW3, a model that however does enjoy a good (in)dependence structure, but in turn gives rise to less tractable marginal laws. This is exactly the opposite of the Fernández-Durán model: the construction precisely allows to control the marginal laws, at the expense of very intricate conditional laws.

Thus, on basis of this comparison, it appears that the WeiSSVM and the Mardia-Sutton model are the best choices if one wishes to satisfy the above-mentioned criteria. A further important criterion, namely the capacity of a model to produce a good fit to various data sets, is investigated in the next section.

## 7. Fitting two circular-linear real data sets

In this section we shall illustrate the good fitting behavior of the WeiSSVM by analyzing two popular data sets from the literature. While the first data set reflects exactly the characteristics of data the WeiSSVM is tailor-made for, namely concentration increasing with length and circular skewness, these attributes are much less marked in the second data set. This allows for a meaningful assessment of the modeling capacities of our new model. In each case, we will compare the WeiSSVM with the Johnson-Wehrly distribution, the independence model, the GGSSVM of Section 3.6 and the Mardia-Sutton (MS) and Kato-Shimizu (KS) models. Our means of comparison shall be the Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), and we will apply the tests of Section 4.2. We draw the reader’s attention to

Table 3: Maximum likelihood estimates, maximized log-likelihood (MLL), Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for the Weibull sine-skewed von Mises (WeiSSVM) and its competitor models, the Generalized Gamma sine-skewed von Mises (GGSSVM), the Johnson-Wehrly (JW), the independence (Indep.), Mardia-Sutton (MS) and Kato-Shimizu (KS) models, fitted to the blue periwinkle data.

Distributions	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\mu}$	$\hat{\kappa}$	$\hat{\lambda}$		MLL	AIC	BIC	
WeiSSVM	2.01	0.05		-1.90	1.68	1.00		-168.57	347.13	354.30	
GGSSVM	2.00	0.05	2.04	-1.90	1.70	1.00		-168.56	349.13	357.73	
JW		0.10		-1.70	1.44			-182.93	371.86	376.16	
Indep.	1.54	0.02		-2.97		1.00		-187.25	382.50	388.24	
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\nu}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	MLL	AIC	BIC
MS	28.58	24.43	29.63	-2.11	-1.52		2.59		-176.88	365.75	374.36
KS	28.58	24.43	29.63	-2.11	-0.97	0.73	8.16	3.46	-168.46	352.93	364.40

the fact that both the Mardia-Sutton and Kato-Shimizu models are defined over  $\mathbb{R} \times \mathcal{C}_1$  with  $\mathcal{C}_1$  the unit circle in  $\mathbb{R}^2$ , whereas the above-mentioned densities are defined over  $\mathbb{R}^+ \times \mathcal{C}_1$ .

### 7.1. Periwinkle data

We give an analysis of  $n = 31$  observations which consist of the movements of blue periwinkles after they had been transplanted downshore from the height at which they normally live. The data set was taken from Table 1 of [8]; see that paper for details about the experience.

A visual inspection of the data points in Figure 4 reveals that the concentration of the circular part tends to increase with length, which is precisely one of the features that the WeiSSVM model can well incorporate. Moreover, [10] have shown that, on basis of the Pewsey test of symmetry (see [19]), the circular part of the data is asymmetric.

Table 3 presents the maximum likelihood estimates, maximized log-likelihood, Akaike and Bayesian Information Criterion values obtained from all models under investigation. As we can see from Table 3, the location parameters of the GGSSVM and its submodels are close (note that the location of the Independence model is  $-2.97 + \pi/2 = -1.40$ , as explained at the end of Section 2) and the WeiSSVM has the lowest AIC and BIC values. It clearly improves on Johnson-Wehrly and Mardia-Sutton, and even on the flexible Kato-Shimizu model. It is quite remarkable to notice the tiny difference in the maximized log-likelihood between WeiSSVM and the embedding model, the GGSSVM. The likelihood ratio test for the Johnson-Wehrly submodel (w.r.t. the WeiSSVM) takes value  $T_{JW} = -2(-182.93 + 168.57) = 28.72$ , with  $p$ -value = 0.00, which emphatically rejects the Johnson-Wehrly model. Even stronger, the independence test yields  $T_{\text{Indep}} = -2(-187.25 + 168.57) = 37.36$ , stressing the dependence between the angular and the linear part.

As a conclusion, our WeiSSVM model (with 5 parameters) is a good-fitting and parsimonious model for the periwinkle data set. For visual impression, we have superimposed the contour plot of the fitted WeiSSVM model on a scatter plot of the data in the panel making up Figure 4.

### 7.2. Wind direction and temperature data

As second example, we consider the original data set from [6], consisting of 28 measurements of wind direction and temperature at Kew during the period 1956-1960. The data are taken from Table 1 in [6] and illustrated in Figure 5. Although the effect noticed for the periwinkle data, namely high concentration for high linear values, is less marked here, Mardia and Sutton have noted (and established) a strong dependence between the circular and the linear component.

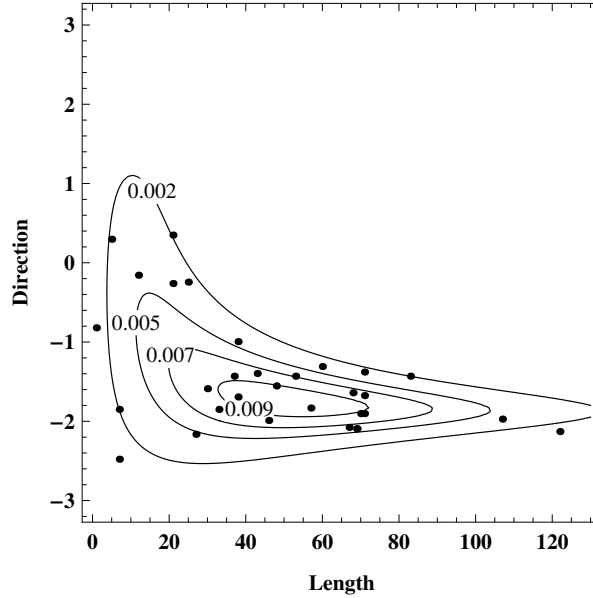


Figure 4: Contour plot of the blue periwinkle data (in lengths and radians), together with the fitted WeiSSVM density. The data are plotted over  $[0, 125] \times [-\pi, \pi)$ , with the distance indicated in cm.

It has been shown in [10] that the Mardia-Sutton model is extremely good for this data set; it is therefore very interesting to compare it with the WeiSSVM. Table 4 contains the maximum likelihood estimates, maximized log-likelihood, Akaike and Bayesian information criterion values. The (circular) location parameters of our proposed models are almost the same (again, the location of the Independence model is  $-0.62 + \pi/2 = 0.95$ ). We see that our WeiSSVM model best incorporates the non-trivial behavior of this data set (its AIC and BIC values are clearly below that of the MS model), and again it is much better than the Johnson-Wehrly model (which is clearly rejected as submodel). A contour plot of the fitted WeiSSVM model with a scatter plot of the data is provided in Figure 5. We finally note that the independence test of course heavily rejects ( $p$ -value = 0.00) the null of independence, hereby agreeing with [6].

Table 4: Maximum likelihood estimates, maximized log-likelihood (MLL), Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) values for the Weibull sine-skewed von Mises (WeiSSVM) and its competitor models, the Generalized Gamma sine-skewed von Mises (GGSSVM), the Johnson-Wehrly (JW), the independence (Indep.), Mardia-Sutton (MS) and Kato-Shimizu (KS) models, fitted to the wind-temperature data.

Distributions	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\gamma}$	$\hat{\mu}$	$\hat{\kappa}$	$\hat{\lambda}$	MLL	AIC	BIC		
WeiSSVM	10.72	0.02		0.54	1.10	0.49	-125.70	261.39	268.05		
GGSSVM	10.78	0.02	10.64	0.54	1.09	0.49	-125.69	263.39	271.38		
JW		0.03		0.83	0.60		-180.77	367.55	371.54		
Indep.	8.90	0.02		-0.62		0.78	-134.32	276.64	281.97		
	$\hat{\mu}$	$\hat{\sigma}$	$\hat{\lambda}$	$\hat{\nu}$	$\hat{\mu}_1$	$\hat{\mu}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	MLL	AIC	BIC
MS	42.07	4.86	5.01	0.36	0.88		1.14		-128.10	268.19	276.19
KS	42.07	4.86	5.01	0.36	1.09	0.48	1.02	0.53	-126.66	269.32	279.98

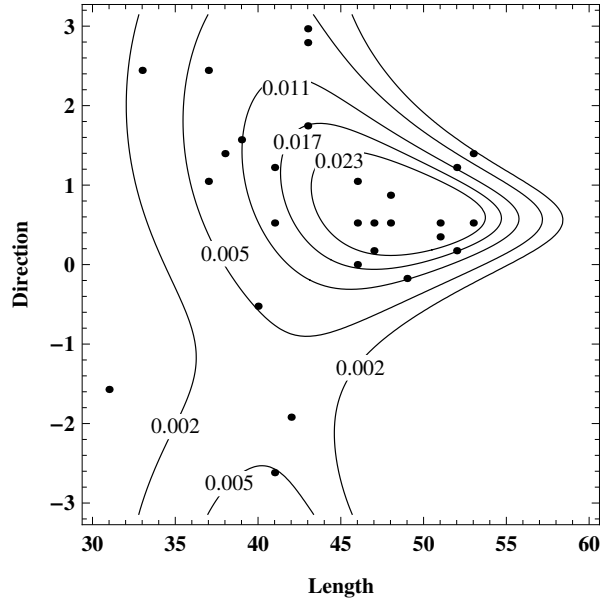


Figure 5: Contour plot of the wind and temperature data (in lengths and radians), together with the fitted WeiSSVM density. The data are plotted over  $[30, 60] \times [-\pi, \pi]$ , with temperature indicated in Fahrenheit.

## 8. Conclusion

We have introduced in this paper a new distribution for cylindrical data, the WeiSSVM. We have described its various good properties: very simple density expression, clear parameter interpretation, good (in)dependence structure, well-known conditional and marginal distributions, immediate and intuitive random number generation, tractable moment expressions and hence circular-linear correlation, simple parameter estimation properties and excellent fitting capacities. Moreover, the WeiSSVM lends itself well to linear-circular or circular-linear regression purposes. A potential drawback could be that the parameter  $\kappa$  regulates at the same time the dependence structure and the circular concentration; on the other hand, this parameter parsimony may precisely be perceived as advantage by some researchers.

As shown in Sections 6 and 7, the WeiSSVM thus improves on other models from the literature, mainly thanks to its simplicity, good interpretability and fitting capacities. Quoting one of the referees, “I strongly believe that the proposed distribution could be applied on several real data”. This applicability of the WeiSSVM is underlined by the fact that it has already been used as important building block in two environmental studies (based on a previous arXiv version of the present paper). [20] use a mixture of WeiSSVM (or Abe-Ley, as they call it) densities as data distribution in a hidden Markov model designed to account for features of cylindrical time series. This allows them to analyze the dynamics of waves in the Adriatic Sea. [21] propose a cylindrical spatial model to get further insights into marine currents in the Adriatic Sea; in their model they combine a mixture of WeiSSVM densities with a Potts-based segmentation of the area interest.

## Acknowledgements

Toshihiro Abe was supported in part by JSPS KAKENHI Grant Number 15K17593 and Nanzan University of Pache Research Subsidy I-A-2 for the 2015 academic year. Christophe

Ley was supported in part by the Fonds National de la Recherche Scientifique, Communauté française de Belgique, via a Mandat de Chargé de Recherche. Both authors would like to thank three anonymous referees for helpful comments that led to an improvement of the present paper.

## References

- [1] E. García-Portugués, R. M. Crujeiras, W. González-Manteiga, Exploring wind direction and SO<sub>2</sub> concentration by circular-linear density estimation, *Stochastic Environmental Research and Risk Assessment* 27 (2013) 1055–1067.
- [2] M.-Z. Wang, K. Shimizu, K. Uesu, An analysis of earthquakes latitude, longitude and magnitude data by use of directional statistics, *Japanese Journal of Applied Statistics* (2013) 29–44 (in Japanese).
- [3] E. García-Portugués, A. M. G. Barros, R. M. Crujeiras, W. González-Manteiga, J. M. C. Pereira, A test for directional-linear independence, with applications to wildfire orientation and size, *Stochastic Environmental Research and Risk Assessment* 28 (2014) 1261–1275.
- [4] F. Wang, A. E. Gelfand, G. Jona-Lasinio, Joint spatio-temporal analysis of a linear and a directional variable: space-time modeling of wave heights and wave directions in the Adriatic Sea, *Statistica Sinica* 25 (2015) 25–39.
- [5] F. Lagona, M. Picone, A. Maruotti, S. Cosoli, A hidden Markov approach to the analysis of space-time environmental data with linear and circular components, *Stochastic Environmental Research and Risk Assessment* 29 (2015) 397–409.
- [6] K. V. Mardia, T. W. Sutton, A model for cylindrical variables with applications, *Journal of the Royal Statistical Society Series B* 40 (1978) 229–233.
- [7] R. A. Johnson, T. E. Wehrly, Some angular-linear distributions and related regression models, *Journal of the American Statistical Association* 73 (1978) 602–606.
- [8] N. I. Fisher, A. J. Lee, Regression models for an angular response, *Biometrics* 48 (1992) 665–677.
- [9] T. Abe, A. Pewsey, Sine-skewed circular distributions, *Statistical Papers* 52 (2011) 683–707.
- [10] S. Kato, K. Shimizu, Dependent models for observations which include angular ones, *Journal of Statistical Planning and Inference* 138 (2008) 3538–3549.
- [11] J. Wang, J. Boyer, M. G. Genton, A skew-symmetric representation of multivariate distribution, *Statistica Sinica* 14 (2004) 1259–1270.
- [12] I. S. Gradshteyn, I. M. Ryzhik, *Tables of integrals, series, and products*, 8th Edn., London: Academic Press, 2015.
- [13] K. V. Mardia, Linear-circular correlation coefficients and rhythmometry, *Biometrika* 63 (1976) 403–405.
- [14] R. A. Johnson, T. E. Wehrly, Measures and models for angular correlation and angular-linear correlation, *Journal of the Royal Statistical Society Series B* 39 (1977) 222–229.
- [15] E. W. Stacy, A generalization of the Gamma distribution, *Annals of Mathematical Statistics* 33 (1962) 1187–1192.

- [16] M. C. Jones, A. Pewsey, A family of symmetric distributions on the circle, *Journal of the American Statistical Association* 100 (2005) 1422–1428.
- [17] V. M. Maksimov, Necessary and sufficient statistics for a family of shifts of probability distributions on continuous bicomact groups, *Rossiskaya Akademiya Nauk. Teor. Veroyatnost. i Primenen.* 12 (1967) 307–321 (in Russian), English Translation: *Theory of Probability and its Applications* **12**, 267–280.
- [18] J. J. Fernández-Durán, Models for circular-linear and circular-circular data constructed from circular distributions based on nonnegative trigonometric sums, *Biometrics* 63 (2007) 579–585.
- [19] A. Pewsey, Testing circular symmetry, *Canadian Journal of Statistics* 30 (2002) 591–600.
- [20] F. Lagona, M. Picone, A. Maruotti, A hidden Markov model for the analysis of cylindrical time series, *Environmetrics* 26 (2015) 534–544.
- [21] F. Lagona, M. Picone, Model-based segmentation of spatial cylindrical data, *Journal of Statistical Computation and Simulation*.