

# Dynamic Modelling of Heavy-Tailed Cylindrical Time Series

Chris Toumping Fotso\*, Yeliz Özer<sup>†</sup>, Dario Palumbo<sup>‡</sup>, Philipp Sibbertsen<sup>§</sup>

March 5, 2026

## Abstract

A dynamic modelling for heavy-tailed cylindrical time series is developed by combining score-driven models with a generalised Pareto-type cylindrical distribution. The proposed specification extends existing cylindrical models by allowing location, scale, concentration, and crucially, the tail index of the linear component through the conditional distribution of speed to vary according to its score. Whereas the Weibull-von Mises model, whose linear component exhibits exponentially decaying tails, the GPar specification admits polynomial tail decay. An explicit expression for the time-varying circular-linear dependence measure is also derived. The methodology is applied to high-frequency data from two onshore wind turbines in Germany. The empirical results indicate that allowing time-varying tail thickness leads to overall improvements compared to the Weibull-von Mises model. The proposed model provides a flexible and computationally tractable framework for analysing heavy-tailed cylindrical time series in environmental and energy applications.

**Keywords:** cylindrical distributions, dynamic correlation, generalised Pareto, score-driven models, Weibull-von Mises, wind energy.

**JEL codes:** C13, C18, C22, C46, Q42.

---

\*Corresponding Author. Leibniz University Hannover, Institute of Statistics, Königsworther Platz 1, D-30167 Hannover, Germany. E-Mail: [toumping@statistik.uni-hannover.de](mailto:toumping@statistik.uni-hannover.de). Phone: +49-511-762-3783

<sup>†</sup>Leibniz University Hannover. E-Mail: [oezer@statistik.uni-hannover.de](mailto:oezer@statistik.uni-hannover.de)

<sup>‡</sup>Ca' Foscari University of Venice, E-Mail: [dario.palumbo@unive.it](mailto:dario.palumbo@unive.it)

<sup>§</sup>Leibniz University Hannover. E-Mail: [sibbertsen@statistik.uni-hannover.de](mailto:sibbertsen@statistik.uni-hannover.de).

# 1 Introduction

Directional data plays a central role in environmental and physical systems, where angular measurements describe movement or flow and magnitudes quantify intensity. A canonical example is wind direction and wind speed, whose joint behaviour is of increasing relevance in the context of climate change and renewable energy integration. High-frequency observations of such variables reveal rapid regime changes and short-lived extreme events that cannot be adequately captured by models that ignore non-linear dynamics or tail behaviour.

Statistically, wind direction and wind speed form cylindrical data, combining a circular component with a non-negative linear magnitude that is typically skewed and may exhibit heavy tails. Modelling these components separately neglects their structural dependence and may lead to distorted inference, particularly during periods of extreme variability. This has motivated the development of joint circular–linear distributions. Early constructions include the regression-based framework of [Fisher and Lee \(1992\)](#), copula-based approaches such as [Johnson and Wehrly \(1978\)](#), and the generalized von Mises–normal model of [Kato and Shimizu \(2008\)](#). A tractable and interpretable parametric specification is the Weibull Sine Skewed–von Mises (WeiSSVM) distribution of [Abe and Ley \(2017\)](#).

While the WeiSSVM model accommodates skewness, its linear component remains of Weibull type and therefore exhibits exponentially decaying tails. In applications where extreme events occur more frequently, such behaviour may be restrictive. To address this limitation, [Imoto et al. \(2019\)](#) introduced the Generalized Pareto-type cylindrical (GPar) distribution, which extends the Abe–Ley model through a generalized Gamma mixture representation. The resulting linear component exhibits polynomially decaying tails, allowing for substantially higher probabilities of extreme observations while preserving tractability and interpretability. Despite these advantages, the GPar distribution has so far been studied only in a static framework.

Many practical applications, however, involve time series in which distributional features change over time. Score-driven models ([Creal et al., 2013](#); [Harvey, 2013](#)) provide a natural mechanism for introducing such dynamics, allowing parameters to vary according to the score of the one-step-ahead predictive density. These updates are likelihood-based, automatically adapt to the underlying distribution, and remain computationally attractive because the likelihood is available in closed form. Recent contributions by ([Harvey and Palumbo, 2023](#); [Harvey et al., 2024](#)) demonstrate that circular and cylindrical distributions can be embedded within this framework, yielding dynamic specifications for wind direction and speed based on the Weibull–von Mises (WeiVM) distribution.

The present paper extends this literature by developing a score-driven specification based on the cylindrical GPar distribution. Our contribution is threefold. First, we embed the GPar model within an observation-driven framework, allowing for time variation in the circular location and linear scale parameters. Second, we introduce a dynamic tail mechanism for the linear component by allowing the tail

index to be time-varying. This feature enables the model to capture changes in extremal behaviour directly within the cylindrical distribution, rather than through exogenous regime-switching mechanisms. Third, we derive the implied time-varying circular–linear correlation coefficient and show how it depends on the evolving distributional parameters, providing a coherent dynamic measure of dependence consistent with the cylindrical structure.

The dynamic tail specification is particularly relevant in high-frequency wind applications. At very short timescales, yaw alignment and control systems rely on accurate modelling of rapid directional shifts and gust-induced variability. Misalignment reduces efficiency and increases structural loads. At coarser timescales, such as 15 minutes, wind direction and speed are directly linked to electricity market operations, where these intervals constitute trading units in several European markets ([EPEX SPOT SE, 2025](#)). In both contexts, allowing the degree of tail thickness to vary over time provides a flexible way to capture episodic bursts of variability and extreme behaviour.

Empirically, we apply the proposed methodology to Supervisory Control and Data Acquisition (SCADA) wind turbine data observed at high frequency. We compare the dynamic GPar specification with the WeiVM benchmark and show that allowing for polynomial tails and time-varying tail thickness yields substantial improvements in likelihood-based criteria and forecasting performance. The estimated tail dynamics capture episodes of increased extremal activity, while the implied correlation measure reveals meaningful time variation in directional–speed dependence.

The remainder of the paper is organised as follows. Section 2 describes the data. Section 3 introduces the dynamic cylindrical GPar model and its score-driven specification, including the derivation of the time-varying correlation measure. Section 4 presents the empirical findings. Section 5 concludes.

## 2 Data

For the empirical analysis, we use SCADA data from two wind turbines located at two different operational onshore wind farms in Germany. Wind farm 1 is located in an inland region, whereas wind farm 2 is located near the North Sea coast. The two locations belong to different wind zones, with the inland wind farm characterised by lower expected wind speeds and the coastal farm associated with substantially stronger wind conditions. In addition to the SCADA measurements, we have access to selected technical information (e.g., hub height).

Since the focus of this study is not on modelling entire wind farms, we restrict the analysis to one turbine per site, namely the southernmost turbine at each wind farm. Wind conditions are measured by the nacelle-mounted sensor located behind the rotor plane. By selecting the outermost turbine on the south/south-west edge of each wind farm, we mitigate potential wake effects from turbines further inside the wind farm and from neighbouring wind farms under the prevailing inflow directions. This

consideration is particularly relevant for the coastal wind farm, where the density of wind farms is higher.

The SCADA data we have access to are recorded at a 10-second frequency and the timestamp is provided in Coordinated Universal Time, so daylight-saving time adjustments caused by one-hour duplicate observations are not required. To evaluate the performance of the proposed model across different sampling frequencies, the two turbines are analysed at different temporal resolutions. As many analyses in the literature are conducted at lower sampling frequencies, we subsample the time series to 15-minute intervals. Although we do not model wind power production, we further motivate the 15-minute sampling by noting that 15-minute products are commonly used in the German intraday market and, since October 2025, also in the day-ahead spot market ([EPEX SPOT SE \(2025\)](#)). Further, we need to decide on the time period on which to base our empirical analysis. To avoid an entirely arbitrary choice, we focus on autumn and summer, as these seasons provide sufficiently different wind regimes and allow us to assess the model under heterogeneous conditions, including periods with pronounced variability and potential heavy-tailed behaviour. In general, winter months are known for higher wind speeds, whereas summer months are typically less productive in terms of wind power generation due to lower wind speeds. For autumn 2024, wind speeds were slightly lower overall than in the reference period 2004–2023 ([anemos Gesellschaft für Umweltmeteorologie mbH, Feb. 12, 2025](#); [Bundesnetzagentur and Bundeskartellamt, Nov. 26, 2025](#)). Our decision to focus on autumn is further motivated by its pronounced short-term variability: autumn represents a transition season between summer and winter and is associated with changing large-scale weather patterns. Such changes can include, for instance, periods influenced by frontal systems as well as calmer episodes, leading to intermittent short-term variability. We therefore use autumn 2024 to evaluate whether the model can reproduce ramp-like rapid dynamics, i.e., pronounced changes over short time scales ranging from minutes to a few hours, which are known to be a key challenge in wind power prediction ([Lochmann et al., 2023](#)). For the coastal turbine, we instead focus on a short summer window at the original 10-second resolution. This choice allows us to examine the model’s behaviour under high-frequency directional variability in a typically low-wind summer regime, thereby providing a useful contrast to the more dynamic autumn conditions while keeping the sample size broadly comparable for benchmarking purposes. Preliminary analyses indicated that single-day summer samples provided insufficient variation in wind direction, whereas the selected three-day window retains adequate dynamics and exhibits the heavy-tailed features that motivate our modelling framework. For the turbine in the inland region, we subsample to 15-minute intervals (Turbine 1), whereas for the turbine at the coast we use the original 10-second frequency (Turbine 2). For Turbine 1, we select the full meteorological autumn season of 2024 (September–November), yielding 8,734 observations. Two values are missing; we treat these as negligible and exclude them from the analysis. For Turbine 2, we restrict the sample to a three-day summer period (10–12 July 2024), yielding 25,920 observations without any missing values. Tab. 1 summarises the key characteristics of

the resulting data sets.

Table 1: Overview of the two turbine data sets.

	Turbine 1	Turbine 2
Geographical location in Germany	Inland	Coastal (North Sea)
Temporal coverage	1 Sep.–30 Nov. 2024	10–12 July 2024
Temporal resolution	15 minutes	10 seconds
Number of observed days	91	3
Number of expected observations	8,736	25,920
Missing values	2	0
Expected wind conditions	Moderate	High
Nacelle height (relative)	Higher	Lower

The variables of interest are wind speed (m/s) and wind direction ( $0, 2\pi$ ). Wind direction follows the meteorological convention commonly used in wind energy applications: it indicates the direction from which the wind is coming, with 0 corresponding to North and angles increasing clockwise. Although the time series plot (see Fig. 1) may suggest the presence of calm winds, we do not observe any wind speed equal to  $0\text{ m/s}$ . From the wind rose plots (Fig. 3) we identify the prevailing incoming wind directions during autumn 2024 for Turbine 1 (Panel A), and for 10–12 July 2024 for Turbine 2 (Panel B). Turbine 1 exhibits dominant wind directions from the west and south, whereas Turbine 2 shows a more pronounced inflow from the southwest. We also note that Turbine 2 has a lower hub height compared to Turbine 1, still the measured average wind speed at Turbine 2 is higher. As a technical reference point, wind turbines typically require wind speeds of roughly  $4\text{ m/s}$  to start operating and reach near-rated performance at higher wind speeds around  $10\text{ m/s}$ . Wind direction from Turbine 1 shows a broader directional distribution, with most observations in the  $4\text{--}10\text{ m/s}$  class; the largest frequencies are observed in the western and southern sectors (each around 15%). Wind direction from Turbine 2 reveals a more concentrated directional pattern, mainly from south-western (between 20–25%) to north-western (nearly 20%). Wind speeds higher than  $10\text{ m/s}$  (between 5–10%) come from north to north-east.

The histograms in Fig. 2 complement the wind roses in Fig. 3 by summarising the marginal distribution of wind speed for the two turbines. For both series, the wind speed is positively skewed, with most observations concentrated at low to moderate wind speeds and a right tail extending to higher values. The wind roses additionally reveal how the wind speed classes are distributed across wind direction

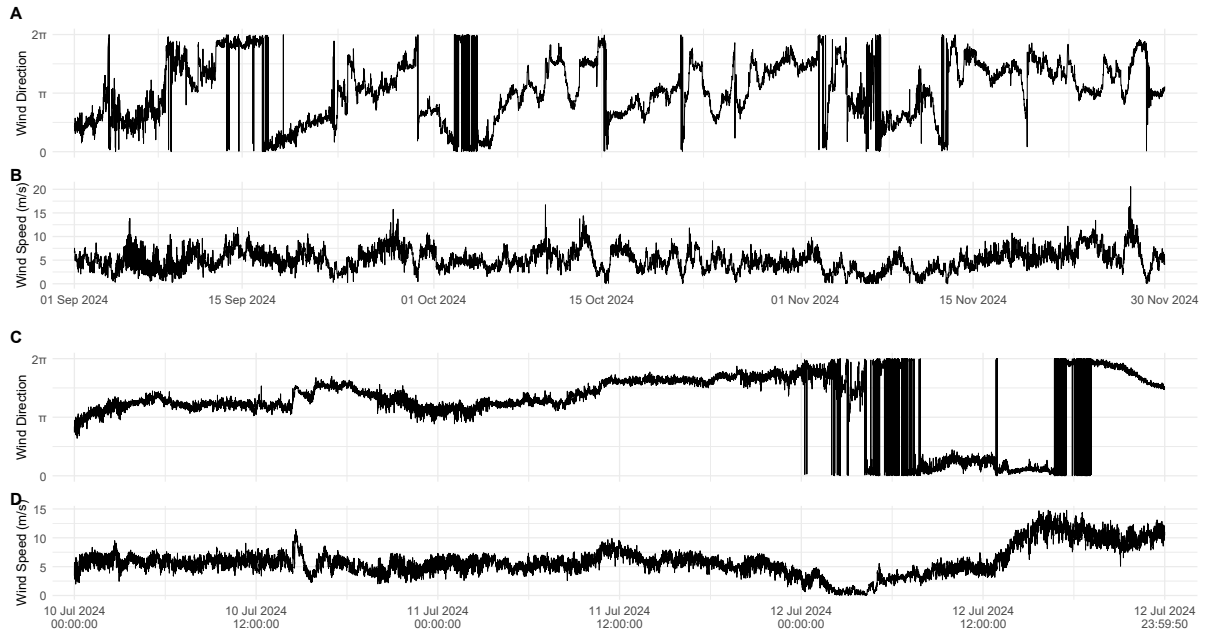


Figure 1: Wind direction and wind speed for study turbines. (A) Turbine 1 (inland site), based on 15-minute SCADA data from meteorological autumn 2024 (September–November). (B) Turbine 2 (coastal North Sea), based on the selected summer period (10–12 July 2024).

sectors. In particular, the coastal 10-second series (Turbine 2, Panel B) shows a more concentrated directional pattern in the wind roses, consistent with a clearer sectoral wind structure than in the inland 15-minute series (Turbine 1, Panel A). Both histograms display right-skewed unconditional distributions, with most observations concentrated at low to moderate wind speeds and a visible extension in the upper range. Relative to the fitted lognormal curves, the empirical distributions suggest the presence of heavier upper tails, particularly for Turbine 1, while Turbine 2 exhibits a more concentrated centre but still a noticeable mass at higher wind speeds. These patterns indicate that wind speed may exhibit non-negligible upper tails, providing empirical support for modelling the linear component with a flexible distribution that accommodates heavy-tailed behaviour. This motivates the generalized Pareto-type specification introduced in the next section.

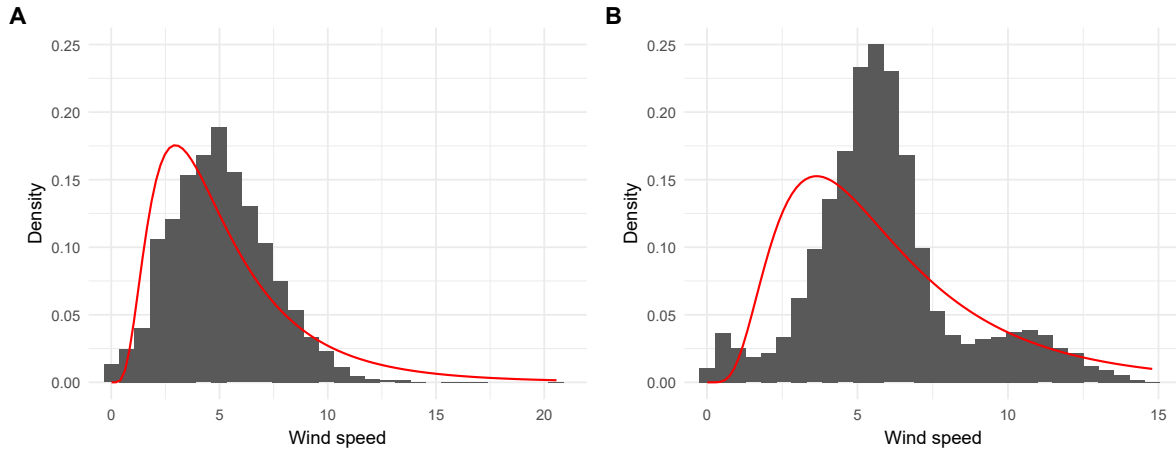


Figure 2: Histograms for wind speed of the two study turbines with fitted lognormal density curves (red). (A) Turbine 1 (inland site), based on 15-minute SCADA data from meteorological autumn 2024 (September–November). (B) Turbine 2 (coastal North Sea), based on based on 10-second SCADA data on the selected summer period (10–12 July 2024).

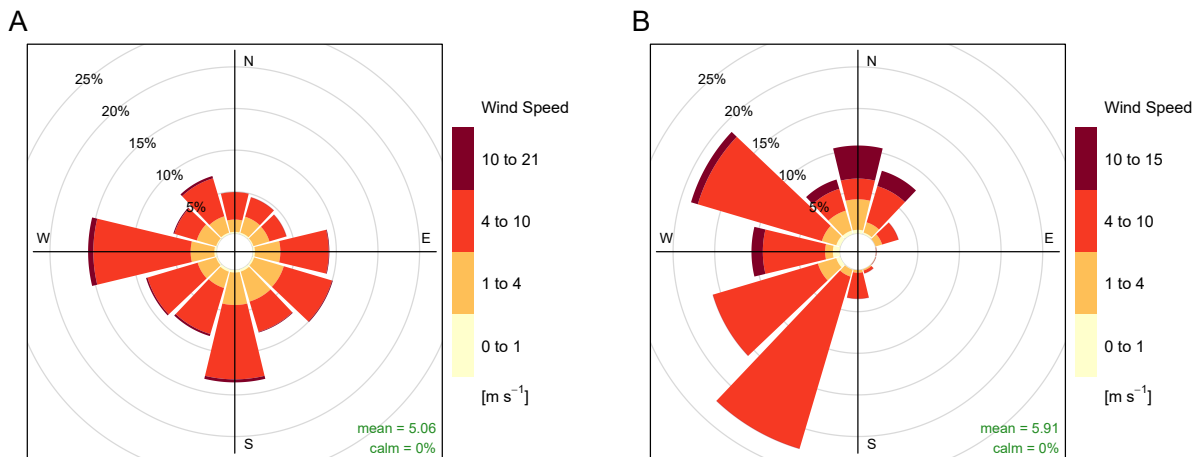


Figure 3: Wind roses for the two study turbines. Radial bar length indicates the relative frequency of observations by wind direction (%), and colours indicate wind speed classes (m/s). The classes are defined to reflect broad turbine operating regimes, including near cut-in conditions (around 4 m/s), the approximate rated-operation range (around 10 m/s), and high-wind conditions approaching the cut-out range (approximately 21–25 m/s, where turbines are shut down for safety). (A) Turbine 1 (inland site), based on 15-minute SCADA data from meteorological autumn 2024 (September–November). (B) Turbine 2 (coastal North Sea), based on 10-second SCADA data on the selected summer period (10–12 July 2024).

### 3 The Dynamic Cylindrical GPar Model

A bivariate distribution for a circular and a linear variable takes the form of a cylinder. This section shows how a dynamic modelling framework based on such a distribution can be constructed.

Imoto et al. (2019) introduced a joint circular and linear distribution with heavy-tailed linear variable: the Generalized Pareto-type cylindrical distribution (GPar) with density,

$$f_{GPar}(y_t, x_t) = \frac{\sqrt{1 - \kappa^2}}{2\pi\sigma\delta} \left(\frac{x_t}{\sigma}\right)^{\frac{1}{\delta}-1} \left[1 + \frac{\tau}{\delta} \left(\frac{x_t}{\sigma}\right)^{\frac{1}{\delta}} (1 - \kappa \cos(y_t - \mu))\right]^{-\frac{\delta}{\tau}-1}, \text{ for } -\pi \leq y_t < \pi \text{ and } x_t \geq 0,$$

where  $-\pi \leq \mu < \pi$  is the location for the circular variable  $y_t$  and  $\sigma > 0$  is the scale of the linear variable  $x_t$ . The remaining shape parameters satisfy  $\tau, \delta > 0$  and  $0 < \kappa < 1$ .

Following a useful reparameterisation, where  $\ln \sigma = \lambda$ ,  $1/\delta = \alpha$ ,  $\delta/\tau = \zeta$  and  $\kappa = \tanh(v)$ , the density can be rewritten as

$$f_{Y,X}(y_t, x_t) = \frac{\alpha}{2\pi e^\lambda \cosh(v)} (x_t e^{-\lambda})^{\alpha-1} \left[1 + \frac{1}{\zeta} (x_t e^{-\lambda})^\alpha (1 - \tanh v \cos(y_t - \mu))\right]^{-(\zeta+1)},$$

with  $-\pi \leq \mu < \pi$ ,  $\lambda \in \mathbb{R}$ ,  $\zeta, \alpha, v > 0$ . In this reparameterisation, as  $\zeta \rightarrow \infty$  the distribution tends to the Weibull-Von Mises (WeiVM) distribution of Abe and Ley (2017). It follows that  $v$  serves as the concentration parameter for the circular variable.

**Remark:** In Imoto et al. (2019) the GPar distribution is stated for  $x > 0$ . However, it is naturally supported on  $[0, \infty)$ . For  $\alpha > 0$ , the density behaves as

$$f_{Y,X}(y, x) \sim C(y) x^{\alpha-1} \text{ as } x \rightarrow 0,$$

which is locally integrable. Hence, for every  $\varepsilon > 0$ ,

$$\int_0^\varepsilon f_{Y,X}(y, x) dx < \infty,$$

and therefore  $P(X = 0) = 0$ . The point  $x = 0$  is a boundary of measure zero and does not require a separate zero-mass component. This is analogous to the classical Weibull distribution when the shape parameter is smaller than one: the density may diverge at the origin but remains integrable.

As in the WeiVM distribution, the marginal distribution of the circular distribution is distributed Wrapped Cauchy with density

$$f_Y(y_t) = \frac{\sqrt{1 - \tanh^2(v)}}{2\pi (1 - \tanh(v) \cos(y_t - \mu))}.$$

On the other hand, if we consider the conditional distribution of the linear variable  $x_t$  conditional on the circular variable  $y_t$ , after applying the transformation

$$\varphi(y_t) = e^\lambda \left( \frac{\zeta}{1 - \tanh(v) \cos(y_t - \mu)} \right)^{\frac{1}{\alpha}},$$

the conditional density becomes

$$f_{X|Y}(x_t|y_t) = \frac{\alpha \zeta}{\varphi(y_t)} \left( \frac{x_t}{\varphi(y_t)} \right)^{\alpha-1} \left[ 1 + \left( \frac{x_t}{\varphi(y_t)} \right)^\alpha \right]^{-(\zeta+1)}.$$

which shows exactly that  $x_t|y_t \sim \text{Burr}(\varphi(y_t), \alpha, \zeta)$ , where  $\varphi(y_t)$  is the scale parameter, while  $\alpha$  and  $\zeta$  are shape parameters determining the tail behaviour, since the tail index is defined as  $\eta = \alpha \zeta$ . The Burr is a special case of the three parameters Generalised Beta of the Second Kind (GB2), with shape parameters  $\alpha$ ,  $\zeta$  unrestricted and  $\xi = 1$ , see [McDonald and Xu \(1995\)](#) and [Harvey \(2013\)](#). When  $\alpha = 1$  we obtain the generalized Pareto distribution. When, in addition,  $\zeta = 1$ , the log-logistic is obtained.  $F$  is obtained as a special case in which  $\alpha = 1$  and  $\zeta = \nu_1/2 = \nu_2/2$ . Finally, if  $\zeta \rightarrow \infty$  the Weibull is obtained as a limiting case.

The family of distributions generated by the Burr is very flexible and has been extensively studied in the score-driven literature; see, among others [Harvey \(2013\)](#), [Harvey and Palumbo \(2023a\)](#). It has been particularly used in the modelling of extreme positive financial variables, such as realised volatility and range, due to the fact that the score-driven filters generated by these conditional distributions are robust to the presence of outliers if the data display features of heavy tail, see [Harvey \(2022\)](#).

Overall, the corresponding log density of the joint cylindrical distribution is as follows.

$$\ln f(y_t, x_t) = -\ln 2\pi + \ln \alpha - \lambda - \ln \cosh(v) + (\alpha - 1)(\ln x_t - \lambda) - (\zeta + 1) \ln \left[ 1 + \frac{1}{\zeta} (x_t e^{-\lambda})^\alpha (1 - \tanh(v) \cos(y_t - \mu)) \right]. \quad (1)$$

In what follows, we first introduce the baseline specification of the model allows for time-varying circular location and linear scale in [3.1](#). Extensions permitting time-varying concentration and tail thickness are introduced in [3.2](#) and [3.3](#).

### 3.1 Score-driven dynamics for location and scale

Data generated by a real-valued time-series model, with support  $-\infty < z_t < \infty$ , can be mapped into wrapped circular observations taking values in  $[-\pi, \pi)$  by defining

$$y_t = z_t \bmod(2\pi) - \pi, \quad t = 1, \dots, T, \quad (2)$$

see [Breckling \(1989\)](#) and [Fisher and Lee \(1994\)](#). A modelling specification for circular data is then given by

$$z_t = \mu_{t|t-1} + \varepsilon_t, \quad -\pi \leq \varepsilon_t < \pi, \quad t = 1, \dots, T, \quad (3)$$

where the innovations  $\varepsilon_t$  are independent and identically distributed (IID) draws from a circular distribution with zero location, and  $\mu_{t|t-1}$  denotes the time- $t$  filtered location based on information available at

time  $t - 1$ . Following [Harvey et al. \(2024\)](#), a stationary first-order observation-driven updating equation for the location filter is

$$\mu_{t+1|t} = (1 - \phi)\omega + \phi\mu_{t|t-1} + \kappa u_t, \quad |\phi| < 1, \quad t = 1, \dots, T, \quad (4)$$

with  $\mu_{1|0} = \omega$  as the unconditional location of  $\mu_{t|t-1}$ . The innovation variable  $u_t$  is taken to be a function of the circular observations  $y_t$ .

A continuous circular density  $f(\cdot; \boldsymbol{\theta})$  satisfies the periodicity condition

$$f(z \pm 2\pi k; \boldsymbol{\theta}) = f(z; \boldsymbol{\theta}),$$

for any integer  $k$ , where  $\boldsymbol{\theta}$  collects the parameters. Provided that the derivatives of the log-density with respect to the elements of  $\boldsymbol{\theta}$  are continuous, they inherit the same periodicity property. Consequently, the path of  $\mu_{t|t-1}$  is identical whether the innovation variable  $u_t$  is constructed from  $z_t$  or from its wrapped counterpart  $y_t$ . It follows that, in the model defined by (3), (4) and (2), the conditional distribution of  $y_t$  coincides with that of  $z_t$ , and hence the likelihood function for the observed wrapped series  $\{y_t\}_{t=1}^T$  is the same as that for the latent unwrapped variables  $\{z_t\}_{t=1}^T$ .

Within this framework, [Harvey et al. \(2024\)](#) study score-driven filters in the case where the innovation  $u_t$  is proportional to the conditional score, as in [Creal et al. \(2013\)](#) and [Harvey \(2013\)](#). For the location parameter, this corresponds to

$$u_t = \frac{\partial \ln f(y_t; \mu_{t|t-1}, v)}{\partial \mu_{t|t-1}}.$$

They show that, for circular distributions,  $u_t$  is invariant to wrapping and remains IID. Moreover, for the stationary von Mises model they derive the asymptotic distribution of the maximum likelihood estimators of  $\phi$ ,  $\kappa$  and  $\omega$  in (4).

In [Harvey and Palumbo \(2023a\)](#) they extended the dynamic framework to a cylindrical distribution based on WeiVM defined in [Abe and Ley \(2017\)](#), constructing a filter for both the location parameter of the circular variable and the scale parameter of the linear variable. The present work is a direct extension of the bivariate framework of [Harvey and Palumbo \(2023a\)](#).

To model the variation in the circular variable, we introduce a score-driven filter for the circular location parameter  $\mu$  from its conditional score obtained from the log-density (Eq. 1).

$$\frac{\partial \ln f}{\partial \mu} = (1 + \varsigma) \frac{\tanh(v) \sin(y_t - \mu)}{1 - \tanh(v) \cos(y_t - \mu)} b_t,$$

where

$$b_t = \frac{\frac{1}{\varsigma} (x_t e^{-\lambda})^\alpha (1 - \tanh(v) \cos(y_t - \mu))}{1 + \frac{1}{\varsigma} (x_t e^{-\lambda})^\alpha (1 - \tanh(v) \cos(y_t - \mu))}.$$

The quantity  $b_t \in (0, 1)$ . Given that the linear variable  $x_t$  conditionally on  $y_t$  is Burr distributed, is easy to show that  $b_t$  is conditionally distributed  $b_t | y_t \sim \text{beta}(1, \zeta)$ , see [Harvey \(2013\)](#) and [Harvey and Palumbo \(2023a\)](#).

As in the cylindrical score-driven WeiVM model, the linear variable  $x_t$  enters the score with respect to the circular location parameter. However, rather than linearly, in this case, it does so through the variable  $b_t$ , and its impact is limited by its bounds. In addition, the information quantity with respect to the location parameter  $\mu$  is

$$\mathcal{I}_{\mu\mu} = \frac{(1 + \zeta)}{(2 + \zeta)} \sinh(v)^2.$$

As a consequence, the filter for the location of the circular variable can be constructed as

$$\mu_{t+1|t} = \omega_\mu(1 - \phi_\mu) + \phi_\mu \mu_{t|t-1} + \kappa_\mu u_{t\mu}, \quad u_{t\mu} = \frac{\partial \ln f}{\partial \mu} \mathcal{I}_{\mu\mu}^{-1}, \quad t = 1, \dots, T,$$

where  $\omega_\mu \in \mathbb{R}$ ,  $|\phi_\mu| < 1$  and  $\kappa_\mu > 0$ .

**Remark:** As for the model in [Harvey and Palumbo \(2023a\)](#) based on the WeiVM distribution, given the dependence of  $u_{t\mu}$  on the variable  $b_t$ , whenever  $x_t = 0$ , also  $u_{t\mu} = 0$ . This means that the model is robust to the presence of zeros in the linear variable, which empirically occurs when the wind speed is 0 and, as a consequence, there is no direction variable  $y_t$ . Due to this feature, this model is also robust to the missing in wind direction caused by the absence of wind, and the filter provides a one-step-ahead forecast naturally.

At the same time, a model for the dynamics of the scale parameter of the linear variable  $\lambda$  can be constructed in a similar way.

$$\frac{\partial \ln f}{\partial \lambda} = \alpha [(1 + \zeta) b_t - 1],$$

which is consistent with the location/scale filter for the Burr model; see [Harvey \(2013\)](#) and [Harvey and Palumbo \(2023a\)](#). In this case, we can also see that as  $x_t$  increases, since  $0 < b_t < 1$  the score is bounded between  $-\alpha$  and  $\alpha\zeta$ . As  $\alpha$  falls, the tail index falls, and therefore the filter windsorises high values of  $x_t$  in the presence of heavy-tailed observations, so it is robust to the presence of outliers.

The information quantity with respect to the location parameter  $\lambda$  is

$$\mathcal{I}_{\lambda\lambda} = \frac{\alpha^2 \zeta}{2 + \zeta}.$$

As a consequence a filter for the scale of the linear variable can be constructed as

$$\lambda_{t+1|t} = \omega_\lambda(1 - \phi_\lambda) + \phi_\lambda \lambda_{t|t-1} + \kappa_\lambda u_{t\lambda}, \quad u_{t\lambda} = \frac{\partial \ln f}{\partial \lambda} \mathcal{I}_{\lambda\lambda}^{-1}, \quad t = 1, \dots, T,$$

where  $\omega_\lambda \in \mathbb{R}$ ,  $|\phi_\lambda| < 1$  and  $\kappa_\lambda > 0$ .

Then the parameters of the model  $\theta = (\omega_\mu, \phi_\mu, \kappa_\mu, \omega_\lambda, \phi_\lambda, \kappa_\lambda, v, \alpha, \zeta)'$  can be estimated through maxi-

mum likelihood, maximising the log-likelihood function obtained from 1,

$$\ln L(\boldsymbol{\theta}) = \sum_{t=1}^T \ln f(y_t, x_t) = -T \ln 2\pi + T \ln \alpha - \sum_{t=1}^T \lambda_{t|t-1} - T \ln \cosh(v) + (\alpha - 1) \sum_{t=1}^T (\ln x_t - \lambda_{t|t-1}) \quad (5)$$

$$- (\zeta + 1) \sum_{t=1}^T \ln \left[ 1 + \frac{1}{\zeta} (x_t e^{-\lambda_{t|t-1}})^\alpha (1 - \tanh(v) \cos(y_t - \mu_{t|t-1})) \right]. \quad (6)$$

while the filters are initialised at their unconditional mean, i.e.  $\mu_{1|0} = \omega_\mu$  and  $\lambda_{1|0} = \omega_\lambda$ .

## 3.2 Time-varying concentration

The model can be easily extended to include a time-varying concentration parameter  $v$ . This feature has been shown in [Harvey et al. \(2024\)](#) that can capture heteroscedasticity in the circular variable  $y_t$ . To do so, we apply the link function  $\vartheta = \ln v$  and then construct a filter for the parameter  $\vartheta_{t|t-1}$  based on its conditional score.

$$\frac{\partial \ln f}{\partial \vartheta} = v \frac{(1 + \zeta)}{\cosh(v)^2} \frac{\cos(y - \mu)}{(1 - \tanh(v) \cos(y_t - \mu))} b_t - \tanh(v),$$

Then the information quantity with respect to  $\vartheta$  is

$$\mathcal{I}_{\vartheta\vartheta} = v^2 \frac{(1 + \zeta) + \zeta \tanh(v)^2}{(2 + \zeta)},$$

then the resulting filter is

$$\vartheta_{t+1|t} = \omega_\vartheta (1 - \phi_\vartheta) + \phi_\vartheta \vartheta_{t|t-1} + \kappa_\vartheta u_{t\vartheta}, \quad u_{t\vartheta} = \frac{\partial \ln f}{\partial \vartheta} \mathcal{I}_{\vartheta\vartheta}^{-1}, \quad t = 1, \dots, T,$$

where  $\omega_\vartheta \in \mathbb{R}$ ,  $|\phi_\vartheta| < 1$  and  $\kappa_\vartheta > 0$ . The additional parameters can also be estimated by ML through the log-likelihood function in 6 while the filter can be initialised at its unconditional mean, i.e.  $\vartheta_{1|0} = \omega_\vartheta$ .

## 3.3 Time-varying tail index for the linear variable.

While [Harvey \(2013\)](#) studied the properties of filters derived by the GB2 family of distributions, which include the Burr. [Caivano and Harvey \(2014\)](#) have studied its logarithm, the Exponential Generalised beta of the Second Kind (EGB2) family. It contains the exponential Burr (EBurr) distribution for  $\xi = 1$  which tends to a logistic if also  $\zeta = 1$ ; see [McDonald and Xu \(1995\)](#) and [Caivano and Harvey \(2014\)](#). In particular, [Harvey and Palumbo \(2023a\)](#) have identified that, once the logarithm of a standardised Burr variable  $\varepsilon_t = (x/\zeta)^\alpha$  is taken, we have that:

$$\ln \varepsilon_t = \alpha (\ln x - \ln \varphi),$$

where  $\ln \varphi$  becomes the location of the EBurr variable, while  $\alpha$ , from driving the tail index in the level variable, becomes the inverse of the scale parameter for the logged EBurr variable. [Harvey and Palumbo \(2023b\)](#) have constructed a score-driven filter for the  $\alpha$  parameter of the Burr to capture the time heteroscedasticity in the log realized volatility ( $RV$ ) variables, which reflected a time varying tail index in the  $RV$  in levels.

In practice, a time varying  $\alpha$  would affect the way the score  $u_{t\lambda}$  reacts to the presence of outliers in the linear variable  $x_t$ , windsorising *more heavily* as the distribution become more heavy-tailed.

To construct a filter, we apply the link function  $\chi = \ln \alpha$  and then construct a filter for the parameter  $\chi_{t|t-1}$  based on its conditional score.

$$\frac{\partial \ln f}{\partial \chi} = 1 + \ln(x_t e^{-\lambda})^\alpha [1 - (1 + \varsigma) b_t],$$

Then the information quantity with respect to  $\chi$  is

$$\mathcal{I}_{\chi\chi} = F_0 + F_1 A_1 + F_2 \left[ A_1^2 + 2 \text{Li}_2 \left( \tanh^2 \left( \frac{v}{2} \right) \right) \right],$$

where,

$$A_1 = \log \varsigma + \log \cosh(v) + 2 \log \cosh \left( \frac{v}{2} \right),$$

and  $F_0, F_1, F_2$  defined in (11) (in the appendix E) are constants that depends on  $\varsigma$ , the digamma  $\psi(\cdot)$  and trigamma  $\psi'(\cdot)$  functions, while  $\text{Li}_2(\cdot)$  is the dilogarithm.

Then the resulting filter is

$$\chi_{t+1|t} = \omega_\chi (1 - \phi_\chi) + \phi_\chi \chi_{t|t-1} + \kappa_\chi u_{t\chi} \quad u_{t\chi} = \frac{\partial \ln f}{\partial \chi} \mathcal{I}_{\chi\chi}^{-1}, \quad t = 1, \dots, T,$$

where  $\omega_\chi \in \mathbb{R}$ ,  $|\phi_\chi| < 1$  and  $\kappa_\chi > 0$ . The additional parameters can also be estimated by ML through the log-likelihood function in 6 while the filter can be initialised at its unconditional mean, i.e.  $\chi_{1|0} = \omega_\chi$ .

### 3.4 Diagnosing the Presence of Dynamics in Parameters

In the score-driven literature the Lagrange multiplier, or score, test against serial correlation in location parameters is based on the portmanteau or Box-Ljung statistic constructed from the autocorrelations of the scores; see [Harvey \(2013\)](#), section 2.5. For time varying location in models based on the conditional von Mises distribution with concentration  $v > 0$ , the scores under the null hypothesis of constant location are proportional to  $\sin(y_t - \bar{y}_d)$ . and the sample autocorrelations correspond to the circular autocorrelations (CACFs) in [Jammalamadaka and Sengupta \(2001\)](#), p. 176–9.

If there are more fixed or dynamic shape parameters in the model, the limiting distribution of the Box-

Ljung statistic applied on fitted scores under the null of no dynamics is a chi-square mixture and can be approximated through block bootstrap. Nevertheless, looking at the results of the statistics under the standard Box-Ljung chi-square limiting distribution is still informative of the presence of dynamics and it has been successfully used in the literature, see [Palumbo \(2021a\)](#), [Harvey and Palumbo \(2023b\)](#) and [Harvey and Palumbo \(2023a\)](#). For these reasons, we will use the Box-Ljung test as a diagnostic tool to detect the presence of dynamics in the parameters of the distribution. Moreover, if is applied on the fitted scores after fitting a dynamic model, the Box-Ljung test can still be a usefull tool to detect residual correlation in the dynamics of the parameters and can be used as a measure of goodness of fit.

### 3.5 Dynamic Circular-Linear Correlation

[Imoto et al. \(2019\)](#) introduce for the GPar distribution a coefficient of correlation between the linear and the circular variable. This concept has also been explored by [Abe and Ley \(2017\)](#) in their WeiVM distribution. The derivation of the coefficient is based on the results of [Johnson and Wehrly \(1977\)](#) and [Mardia \(1976\)](#), which are direct extensions of the Spearman correlation coefficient in the case of linear and circular variables. This type of correlation has been extensively used in empirical works concerning environmental data, see, among others, [Wang et al. \(2015\)](#), [Beyene et al. \(2018\)](#), [Collins and Norton \(2024\)](#), and more recently also in economics, see [Lourenço and Rua \(2023\)](#).

In our case, the coefficient can be expressed as

$$R_{X\Theta}^2 = \frac{r_{XC}^2 - 2r_{CS}r_{XC}r_{XS}}{1 - r_{CS}^2},$$

where  $r_{CS} = \text{Corr}(\cos \Theta, \sin \Theta)$ ,  $r_{XC} = \text{Corr}(X, \cos \Theta)$ , and  $r_{XS} = \text{Corr}(X, \sin \Theta)$ . For both the distributions GPar and WeiVM, the authors showed that if the variables  $X$  and  $\Theta$  were standardised, the resulting correlation coefficient depends only on the remaining shape parameters of the distributions. Considering that for our distribution  $r_{CS}$  and  $r_{XS}$  are 0, deriving the correlation coefficient from  $r_{XC}$  gives us,

$$R_{X\Theta}^2 = r_{XC}^2 = 2\alpha\zeta \cosh(v)^{2-\frac{2}{\alpha}} \frac{\Gamma\left(\frac{1}{\alpha}\right)^2 \Gamma\left(\zeta - \frac{1}{\alpha}\right)^2 \left[ P_{1/\alpha}^1(\cosh(v)) - \frac{\tanh(v/2)}{\alpha} P_{1/\alpha}^0(\cosh(v)) \right]^2}{\Gamma(\zeta) \left[ 2\Gamma\left(\frac{2}{\alpha}\right) \Gamma\left(\zeta - \frac{2}{\alpha}\right) P_{1/\alpha}^2(\cosh(v)) - \frac{\zeta\Gamma\left(\frac{1}{\alpha}\right)^2 \Gamma\left(\zeta - \frac{1}{\alpha}\right)^2}{\alpha\Gamma(\zeta)} P_{1/\alpha}^0(\cosh(v)) \right]},$$

This quantity is constrained  $0 < R_{X\Theta}^2 < 1$  due to the nature of the two marginal distributions. If in our case the shape parameters are dynamic, in particular  $v$  and  $\alpha$  which model heteroscedasticity in the circular and in the logarithm of the conditional linear variable, the resulting correlation coefficient will also be time varying.

## 4 Empirical analysis

In this section, we evaluate the empirical performance of the proposed model using high-frequency wind data. To facilitate the comparison, we benchmark our specification against the cylindrical score-driven model of [Harvey and Palumbo \(2023a\)](#), which provides a natural reference case for comparison. Both models are estimated within the same score-driven framework so that differences in performance can be attributed to the distributional assumptions rather than to the updating mechanism.

### 4.1 Benchmark model and computational setup

The WeiVM density is given by

$$f(y, x) = \frac{\alpha}{2\pi \cosh v} \exp(-\alpha\lambda) x^{\alpha-1} \exp\left[-\left(\frac{x}{\exp(\lambda)}\right)^\alpha (1 - \tanh(v) \cos(y - \mu))\right], \quad (7)$$

for  $-\pi \leq y, \mu < \pi, x \geq 0, \alpha > 0, v \geq 0$ , where  $\exp(\lambda)$  is the scale parameter  $\varphi$  of the linear variable and  $v$  controls the concentration of the circular component. The score-driven filters are specified in the same way for both models. The information quantities required for scaling the score are available in [Harvey and Palumbo \(2023a\)](#) and are therefore omitted here.

Estimation is carried out in R<sup>1</sup>. Numerical optimisation relies on the COBYLA algorithm, implemented through the `nloptr`<sup>2</sup> package ([Ypma et al., 2018](#)). COBYLA (Constrained Optimization BY Linear Approximations) is a derivative-free optimisation routine designed to handle nonlinear objective functions and parameter constraints without requiring gradient information. The `nloptr` package provides an interface in R to the NLOpt library, which contains a collection of state-of-the-art numerical optimisation algorithms for nonlinear problems. A derivative-free method is convenient because the likelihood is highly nonlinear, making gradients difficult to compute reliably. COBYLA also allows parameter constraints to be enforced directly and provides stable convergence.

Tab. 2 describes all model specifications considered in the empirical application, stating the corresponding conditional distributions GPar (Model 1–5) and WeiVM (Model 6–8) and indicating which parameters are either static or time-varying in each model specification.

### 4.2 Performance evaluation measures

To assess the empirical performance of competing models, we compute a set of complementary goodness-of-fit measures in line with [Harvey et al. \(2024\)](#) and [Harvey and Palumbo \(2023a\)](#) in the case of modelling with a dynamic Burr distribution.

---

<sup>1</sup>Version 4.4.3

<sup>2</sup>Version 2.2.1

Table 2: Overview of empirical model specifications.

Model	Dist.	Static parameters	Time-varying parameters	Feature
(1)	GPar	$\mu, \lambda, \vartheta, \chi, \varsigma$	—	Static GPar model (no dynamics)
(2)	GPar	$\vartheta, \chi, \varsigma$	$\mu_{t t-1}, \lambda_{t t-1}$	Adds dynamics in direction, location, and speed scale
(3)	GPar	$\chi, \varsigma$	$\mu_{t t-1}, \lambda_{t t-1}, \vartheta_{t t-1}$	Adds time-varying concentration
(4)	GPar	$\vartheta, \varsigma$	$\mu_{t t-1}, \lambda_{t t-1}, \chi_{t t-1}$	Adds time-varying tail thickness
(5)	GPar	$\varsigma$	$\mu_{t t-1}, \lambda_{t t-1}, \vartheta_{t t-1}, \chi_{t t-1}$	Full GPar: dynamics in location, scale, concentration, and tail index
(6)	WeiVM	$\mu, \lambda, \vartheta, \chi$	—	Static WeiVM model (no dynamics)
(7)	WeiVM	$\vartheta, \chi$	$\mu_{t t-1}, \lambda_{t t-1}$	Benchmark with dynamic location and scale
(8)	WeiVM	$\chi$	$\mu_{t t-1}, \lambda_{t t-1}, \vartheta_{t t-1}$	Benchmark with dynamic location, scale, and concentration

Note:  $\mu_{t|t-1}$  is the filtered circular location,  $\lambda_{t|t-1}$  the filtered log-scale of wind speed,  $\vartheta_{t|t-1} = \log v_{t|t-1}$  the filtered log-concentration, and  $\chi_{t|t-1} = \log \alpha_{t|t-1}$  the filtered log-shape governing tail thickness in GPar.

Likelihood-based model comparison is conducted using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).

For wind direction, we report the dispersion measure following [Harvey et al. \(2024\)](#) adapted to our specification as

$$D = 1 - \frac{1}{T} \sum_{t=1}^T \cos(y_t - \mu_{t|t-1}), \quad 0 \leq D \leq 1,$$

where  $y_t$  denotes the observed wind direction and  $\mu_{t|t-1}$  the filtered circular location. Lower values of  $D$  indicate a tighter directional fit.

Together with the relative improvement statistic

$$A_\delta = 1 - \frac{D}{D_\delta},$$

with

$$D_\delta = 1 - \frac{1}{T-1} \sum_{t=2}^T \cos(y_t - y_{t-1}),$$

where  $D_\delta$  denotes the dispersion of the random walk benchmark. Positive values of  $A_\delta$  indicate an improvement over the benchmark.

For wind speed, model fit is evaluated using the mean squared error (MSE) which provides a conventional measure of level accuracy in wind energy applications.

$$\text{MSE}(x_t, \hat{x}_t) = \frac{1}{T} \sum_{t=1}^T (x_t - \hat{x}_t)^2.$$

While widely used, the MSE is known to be sensitive to extreme values, which may impact model comparisons. We therefore additionally report the QLIKE loss function as proposed by [Patton \(2011\)](#)

and defined as:

$$\text{QLIKE}(x_t, \hat{x}_t) = \frac{1}{T} \sum_{t=1}^T \left( \log \hat{x}_t + \frac{x_t}{\hat{x}_t} \right),$$

where  $x_t$  denotes the observed wind speed and  $\hat{x}_t$  is the expectation of the Burr distribution defined as:

$$\hat{x}_t = \mathbb{E}[X | Y = y_t, \mathcal{F}_{t-1}] = \varphi(y_t) \frac{\Gamma\left(1 + \frac{1}{\alpha_{t|t-1}}\right) \Gamma\left(\zeta - \frac{1}{\alpha_{t|t-1}}\right)}{\Gamma(\zeta)},$$

where  $\varphi(y_t)$  is defined as:

$$\varphi_t(y_t) = e^{\lambda_{t|t-1}} \left( \frac{\zeta}{1 - \kappa_{t|t-1} \cos(\theta_t - \mu_{t|t-1})} \right)^{1/\alpha_{t|t-1}}.$$

### 4.3 Results

Tab. 3 and 4 summarise the comparative performance of the dynamic specifications for both the WeiVM and GPar models. Across turbines and sampling frequencies, increasing flexibility systematically improves likelihood-based criteria. However, the gains are markedly stronger under the GPar specification, particularly when the tail index is allowed to evolve. This indicates that accommodating time-varying tail thickness provides additional explanatory power beyond that achieved by dynamic scale and concentration alone.

The goodness-of-fit measures for wind direction,  $D$  and  $A_\delta$ , reveal a more nuanced pattern. For the GPar specification, directional performance remains stable as flexibility increases, and improvements in directional fit do not come at the expense of speed modelling. For instance, in Turbine 1 (Autumn, 15-minute frequency), the most flexible specification (Model 5) delivers the best directional fit according to both  $A_\delta$  and  $D$ , while maintaining essentially unchanged QLIKE values relative to simpler specifications. A similar behaviour is observed for Turbine 2 at the 10-second frequency: although Model 2 slightly outperforms Model 5 in some directional and MSE metrics, the overall speed-related performance, as measured by QLIKE, remains comparable across specifications.

In contrast, for the WeiVM model, additional flexibility tends to improve directional goodness-of-fit but at the cost of deteriorating slightly speed-related performance. Moving from Model 5 to Model 8 leads to lower values of  $D$  and  $A_\delta$ , yet slightly higher MSE and QLIKE, indicating a trade-off between directional and linear fit under the lighter-tailed structure. This difference highlights the role of the heavier-tailed GPar specification in jointly capturing direction and speed without compromising predictive accuracy.

Fig. 4 illustrates the filtered wind direction and speed for Model 5 at the 15-minute frequency during autumn 2024 (Turbine 1). Both GPar and WeiVM track the main directional movements closely, consistent with their similar  $D$  values. However, Model 5 achieves a moderate improvement in  $A_\delta$  and a clearer advantage in speed-related metrics. The lower MSE and improved QLIKE indicate that the presence

of dynamic tail index enhances the fit during periods of increased variability. A comparable pattern emerges in the 10-second data for Turbine 2, where the GPar specification consistently outperforms the WeiVM benchmark across the main performance measures.

The drivers of the GPar model's improved fit become apparent in Fig. 5. Panel A shows  $1/\alpha_{t|t-1}$ , the scale parameter of the logarithm of speed, which drives the tail behaviour of the conditional distribution of speed. Periods of high values of  $\alpha_{t|t-1}$  are associated with lighter tails, while lower values correspond to heavier tail behaviour. This is reflected in Panel D, which displays the exact time-varying tail index  $\alpha \cdot \zeta$  for the conditional distribution of speed. With a mean of to a mean of 7.621, over the sample the tail index ranges from periods with a maximum of 11.997, to periods with a minimum of 1.685 where fewer finite moments exist. These values indicate substantial variation in tail thickness over time, with periods with lower values corresponding to periods of low speed where the logarithm of speed becomes more volatile and speed is more prone to extreme movements. In terms of the filter, a lower tail index implies that the score with respect to  $\lambda_{t|t-1}$  windsorises the observations more heavily, making the filter less sensitive to large deviations and outliers, with more moderate innovations. Panel B shows the inverse of the concentration parameter  $1/v_{t|t-1}$ , which is the dispersion of the direction variable. The periods when the tail index is low also coincides with periods when the concentration parameter is also low and dispersion of the direction variable increases becoming more uncertain. This confirms the fact that periods with low and extreme wind speed are also associated with greater directional uncertainty.

The last panel displays dynamic correlation. Here, higher values of  $R_{X\Theta}^2$  indicate that certain arcs of the circle tend to be associated with particular wind-speed ranges. Panel D shows that  $R_{X\Theta}^2$  throughout the sample remains mostly at low to moderate levels, up to approximately 0.17. Occasionally, it falls towards 0 for short periods of time. As we can observe, these drops correspond to periods in which wind direction or wind speed are more uncertain, identified also by drops in concentration or in the tail index. The relative magnitude of these two combined effects determines the size of the movements in  $R_{X\Theta}^2$  over time.

For the 10-second data, we observe a similar qualitative behaviour in the inverse scale, inverse concentration, and tail index. However, the estimated values of  $\alpha \cdot \zeta$  are substantially higher, ranging from 48.390 to 1.773 with an average of 21.177. This indicates predominantly lighter tails than in the 15-minute data, but still periods of very extreme behaviour. This is because the shorter time horizon captures more locally persistent wind conditions. As illustrated in Fig. 1 Panel C, and D, wind direction remains largely stable over the period, with only a single pronounced shift first towards 0 and then towards higher values. This smoother evolution reduces regime mixing and results in a higher average tail index. The tail index indeed falls significantly as the series approaches 0 between the second and the third day. However, we can still see that the tail index falls in the first part of the second day, this time not because the series is merely approaching 0 but because it effectively becomes more volatile. Overall, this behaviour indicates a stronger and more stable circular-linear dependence over the three-day period which can be seen also

in Panel C, where the correlation coefficient is persistently high. Furthermore, this is consistent with the wind rose plot (Panel B of Fig. 3) and the raw time series (Panels C and D of Fig. 1), which show a clearer sector-specific direction–speed structure. The main reduction in dependence occurs around 03:00 on 12 July 2024, when lower wind speed levels coincide with a marked directional transition across wind sectors.

Table 3: Estimated parameters using GPar and WeiVM for Turbine 1. Autumn 2024, 15-minute frequency.

<b>GPar</b>																				
Model	AIC	BIC	Logl	MSE	$A_0$	$D$	$QLike$	$\omega_\mu$	$\phi_\mu$	$\kappa_\mu$	$\omega_\lambda$	$\phi_\lambda$	$\kappa_\lambda$	$\omega_s$	$\phi_s$	$\kappa_s$	$\omega_\chi$	$\phi_\chi$	$\kappa_\chi$	$\zeta$
(1)	71,168.51	71,203.89	-35,579.26	5.2757	-14.808	0.817	2.622	3.816	-	-	5.020	-	-	0.581	-	-	2.595	-	-	4.504
(2)	37,009.41	37,073.09	-18,495.71	1.2977	-2.379	0.174	2.528	2.65846	0.99626	0.11011	0.54003	0.96888	0.13840	2.038	-	-	5.277	-	-	1.669
(3)	34,222.56	34,300.39	-17,100.28	1.3818	-1.088	0.107	2.530	1.82810	0.99846	0.11091	1.79511	1.00000	0.13103	0.17888	0.99668	0.01193	5.655	-	-	1.463
(4)	36,707.94	36,785.77	-18,342.97	1.3123	-0.896	0.098	2.530	1.98914	0.99534	0.16665	1.27825	1.00000	0.08790	1.532	-	-	2.02467	0.93638	0.10219	0.876
(5)	30,149.01	30,240.99	-15,061.51	1.4006	-0.857	0.096	2.529	1.92064	0.99998	0.09710	1.13277	0.95855	0.09343	0.10863	0.98033	0.09125	2.01565	0.96757	0.04485	1.005
<b>WeiVM</b>																				
(6)	70,437.416	70,465.716	-35,214.708	5.0824	-14.837	0.819	2.619	3.847	-	-	0.182	-	-	0.296	-	-	2.351	-	-	-
(7)	42,436.139	42,492.739	-21,210.070	1.8054	-1.959	0.153	2.539	1.690	1.000	0.196	0.713	0.995	0.116	1.701	-	-	3.833	-	-	-
(8)	38,920.487	38,991.236	-19,450.243	3.2117	-0.955	0.101	2.541	1.972	1.000	0.145	0.361	0.978	0.116	0.092	0.992	0.106	3.333	-	-	-

Note: (1)&(6): static model, further  $\omega_\lambda$ ,  $\omega_s$ , and  $\omega_\chi$  are passed through the exponential link function; (2)&(7): time-varying location, and scale; (3)&(8): time-varying location, scale, and concentration; (4): time-varying location, scale, and tail index; (5): time-varying location, scale, concentration, and tail index. Whenever a parameter is specified as static (i.e. no  $\phi$  and  $\kappa$  are reported), the corresponding entry in the  $\omega$  column represents the estimated static parameter after transformation through the link function.

Table 4: Estimated parameters across all models GPar and WeiVM for Turbine 2. 10–12 July 2024, 10-second frequency.

GPar																					
Model	AIC	BIC	Logl	MSE	$A_\delta$	$D$	$QLike$	$\omega_\mu$	$\phi_\mu$	$\kappa_\mu$	$\omega_\lambda$	$\phi_\lambda$	$\kappa_\lambda$	$\omega_\vartheta$	$\phi_\vartheta$	$\kappa_\vartheta$	$\omega_\chi$	$\phi_\chi$	$\kappa_\chi$	$\zeta$	
(1)	209,388.17	209,456.87	-104,689.09	7.1229	-61.817	0.511	2.787	4.990	-	-	5.268	-	-	0.737	-	-	2.727	-	-	-	3.713
(2)	24,375.12	24,498.77	-12,178.56	0.3053	-1.752	0.022	2.660	4.70589	0.99920	0.05060	1.21155	0.99998	0.07569	2.993	-	-	12.915	-	-	-	2.640
(3)	18,472.58	18,623.71	-9,225.29	0.2757	-0.399	0.011	2.660	4.03325	0.99552	0.04717	1.82507	0.99959	0.06995	-0.40534	0.97879	0.11764	14.606	-	-	-	1.408
(4)	16,052.15	16,203.28	-8,015.07	0.3466	-1.539	0.020	2.661	4.91086	0.99915	0.07589	3.22992	0.99847	0.07447	2.927	-	-	2.02225	0.99831	0.01869	1.463	-
(5)	7,205.07	7,383.68	-3,589.54	0.2849	-0.828	0.014	2.660	5.08658	1.00000	0.03365	1.29345	0.99809	0.05551	-0.02566	0.99528	0.03928	0.92935	0.99980	0.04957	1.040	-

**WeiVM**

(6)	209,490.850	209,523.501	-104,741.425	6.6010	-68.488	0.566	2.780	5.297	-	-	0.169	-	-	0.553	-	-	2.496	-	-	-	-
(7)	42,161.307	42,226.609	-21,072.653	0.5399	-2.060	0.025	2.667	3.875	1.000	0.245	1.057	0.990	0.224	2.830	-	-	7.537	-	-	-	-
(8)	40,469.272	40,550.900	-20,224.636	0.5745	-0.867	0.015	2.669	4.995	1.000	0.156	0.952	0.990	0.135	-0.874	1.000	0.148	7.192	-	-	-	-

Note: (1)&(6): static model, further  $\omega_\lambda$ ,  $\omega_\vartheta$ , and  $\omega_\chi$  are passed through the exponential link function; (2)&(7): time-varying location, and scale; (3)&(8): time-varying location, scale, and concentration; (4): time-varying location, scale, and tail index; (5): time-varying location, scale, concentration, and tail index. Whenever a parameter is specified as static (i.e. no  $\phi$  and  $\kappa$  are reported), the corresponding entry in the  $\omega$  column represents the estimated static parameter after transformation through the link function.

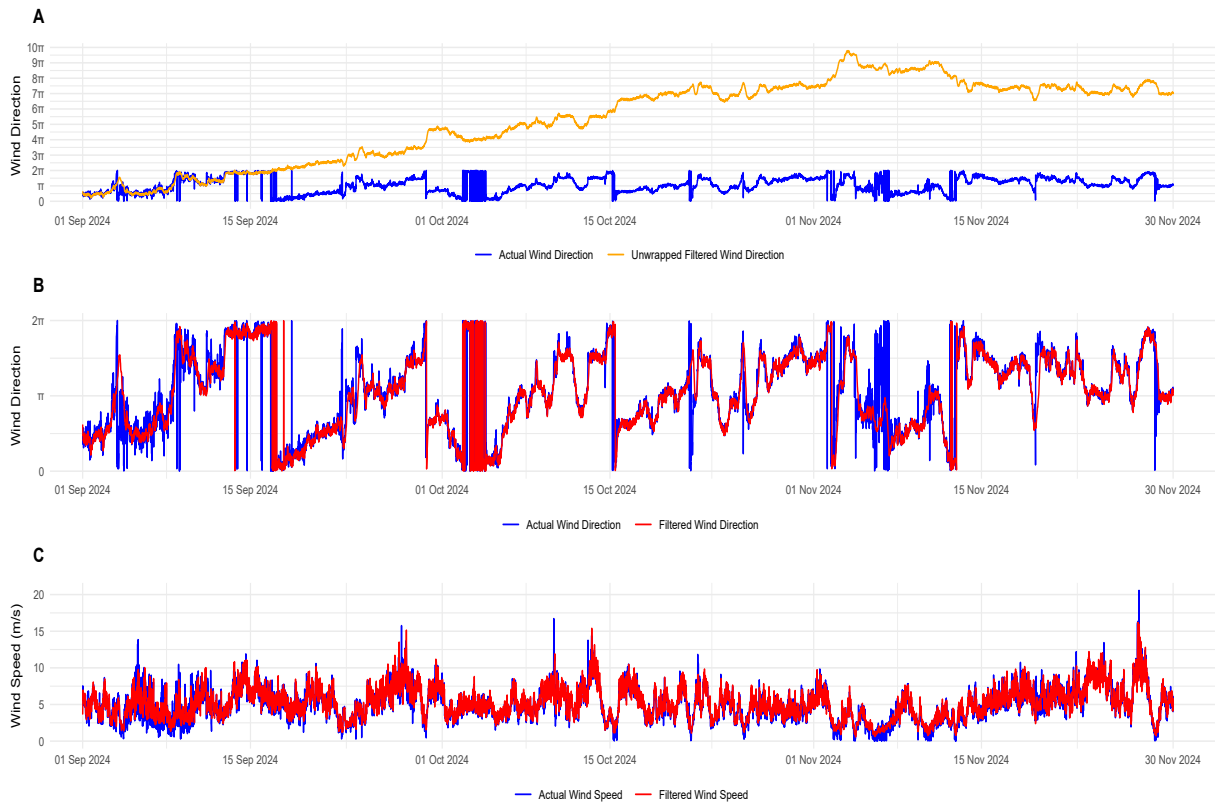


Figure 4: Filtered wind direction and wind speed using GPar for Turbine 1. Autumn 2024, 15-minute frequency.

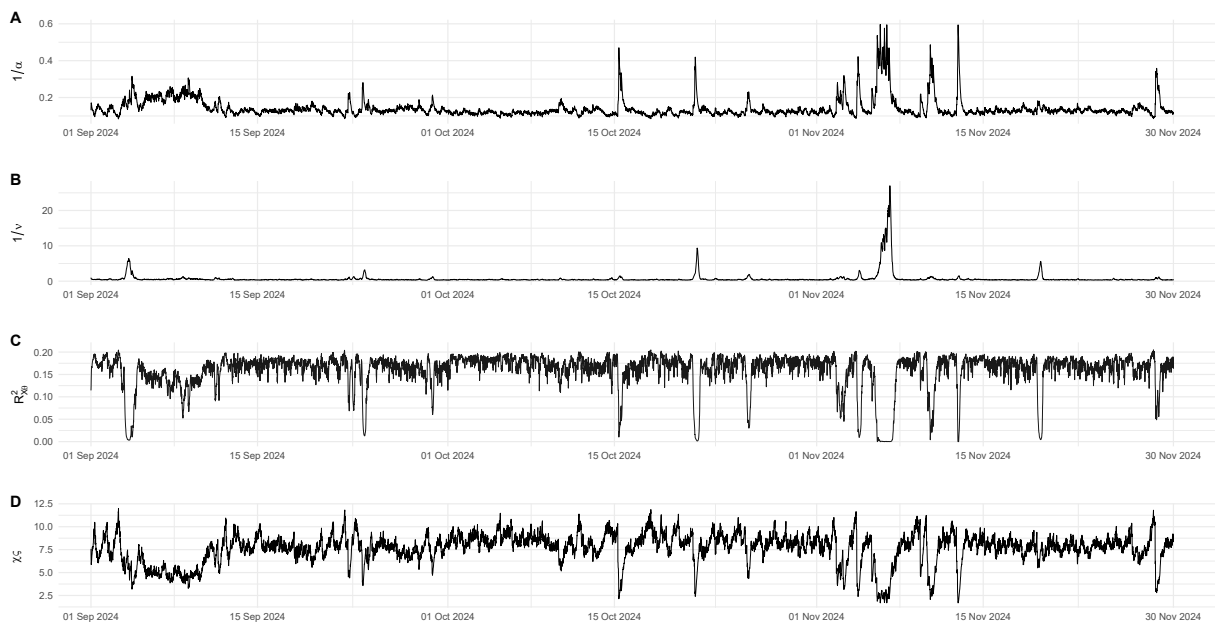


Figure 5: Time-varying scale, concentration, correlation, and tail index using GPar for Turbine 1. Autumn 2024, 15-minute frequency.

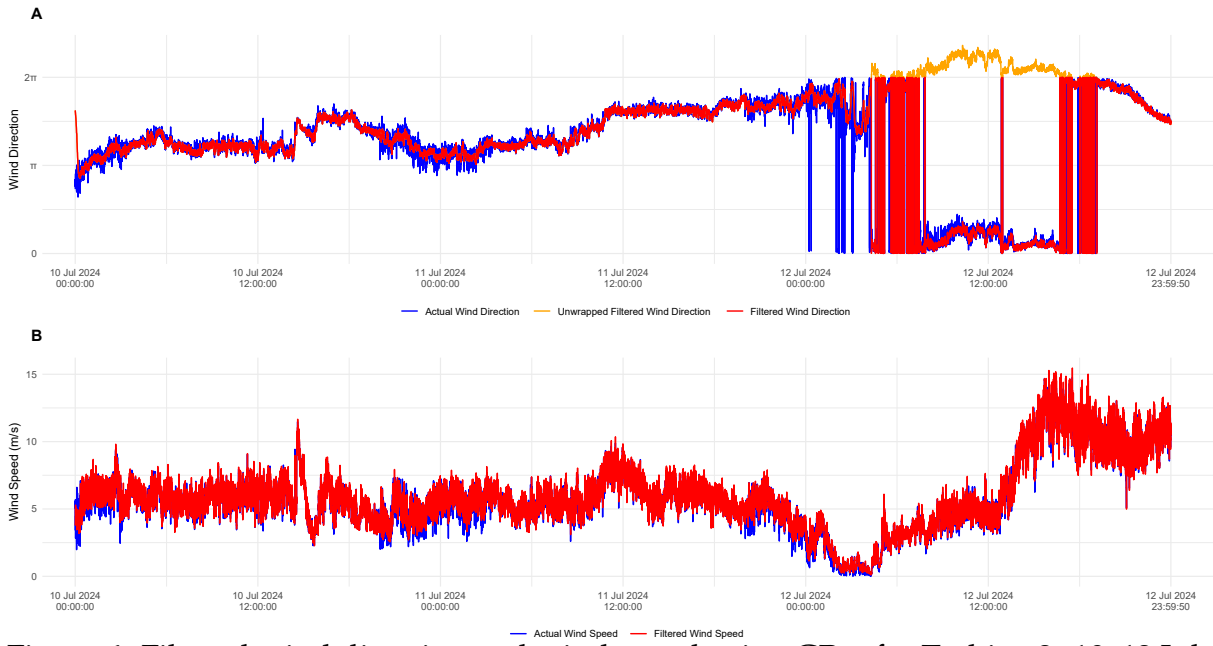


Figure 6: Filtered wind direction and wind speed using GPar for Turbine 2. 10–12 July 2024, 10-second frequency.

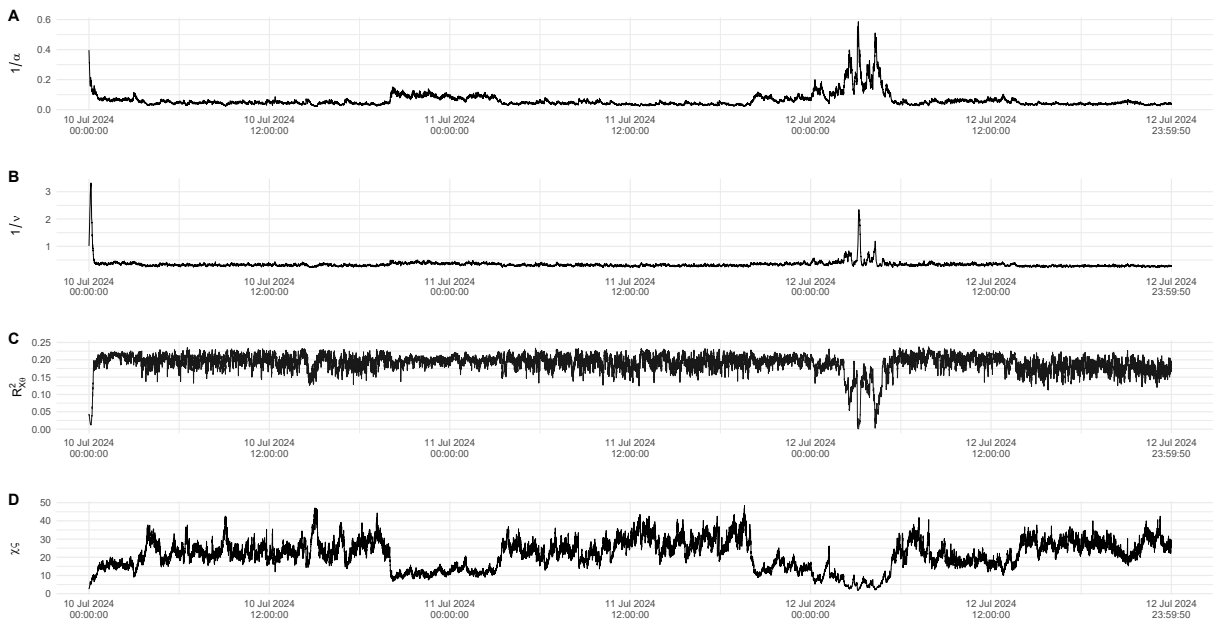


Figure 7: Time-varying scale, concentration, correlation, and tail index using GPar for Turbine 2. 10–12 July 2024, 10-second frequency.

Tab. 5 presents Ljung–Box tests applied to the fitted score residuals for Turbine 1 (autumn, 15-minute frequency). For the GPar specification, increasing flexibility substantially reduces residual autocorrelation across components. In particular, in Model 5, the statistic associated with the parameter  $\chi_{t|t-1}$  becomes insignificant at lag 1, and the remaining statistics are markedly smaller than in simpler specifications, indicating that the dynamic tail mechanism effectively absorbs short-term dependence. Although some longer-run dependence remains at lag 24, its magnitude is considerably reduced relative to the baseline models. In contrast, the WeiVM specification continues to exhibit large and highly significant Ljung–Box statistics across direction, scale, and concentration, even in the most flexible specification. This suggests that the Weibull-based structure leaves substantial serial dependence unmodelled.

Tab. 6 reports Ljung–Box tests for Turbine 2 (10-second frequency). The results broadly mirror those obtained for Turbine 1. For the GPar specification, increasing flexibility reduces residual autocorrelation, particularly at lag 1, where the statistics associated with the parameters  $\lambda_{t|t-1}$  and  $\chi_{t|t-1}$  become small or insignificant in the more flexible specifications. At lag 24, some dependence remains, which is expected given the higher temporal resolution of the data. In contrast, the WeiVM specification continues to display large and significant Ljung–Box statistics across components, indicating that residual serial dependence persists.

Overall, the GPar model again provides a better dynamic description of the data than the lighter tailed WeiVM model.

Table 5: Ljung-Box tests on fitted score residuals for Turbine 1. Autumn 2024, 15-minute frequency.

lag = 1								
GPar					WeiVM			
Model	$\mu$	$\lambda$	$\vartheta$	$\chi$	Model	$\mu$	$\lambda$	$\vartheta$
(1)	7420.001***	5726.283***	7135.826***	4563.935***	(6)	6773.136***	5062.736***	6304.436***
(2)	36.956***	15.566***	896.304***	88.675***	(7)	23.335***	216.379***	1460.315***
(3)	53.251***	1.823	414.743***	420.962***	(8)	224.093***	40.935***	766.003***
(4)	72.843***	1.730	197.258***	16.689***				
(5)	5.503**	64.171***	3.895**	0.327				
lag = 24								
GPar					WeiVM			
(1)	131613.856***	89606.962***	133900.759***	46810.198***	(6)	6773.136***	5062.736***	6304.436***
(2)	174.343***	91.524***	14199.509***	1680.408***	(7)	23.335***	216.379***	1460.315***
(3)	160.590***	115.328***	2665.882***	5611.211***	(8)	224.093***	40.935***	766.003***
(4)	231.239***	26.676	3049.760***	108.876***				
(5)	132.910***	150.381***	154.310***	79.971***				

Note: Significance: 0.01 (\*\*\*), 0.05 (\*\*), 0.1 (\*).

Table 6: Ljung-Box tests on fitted score residuals for Turbine 2. 10–12 July 2024, 10-second frequency.

lag = 1								
GPar					WeiVM			
Model	$\mu$	$\lambda$	$\vartheta$	$\chi$	Model	$\mu$	$\lambda$	$\vartheta$
(1)	25340.712***	24707.385***	24967.714***	24705.020***	(6)	25284.039***	24719.346***	24766.610***
(2)	1.924	44.711***	2739.227***	201.920***	(7)	359.389***	423.234***	9100.340***
(3)	155.436***	3.847**	32.292***	2158.663***	(8)	390.462***	978.908***	7768.074***
(4)	476.464***	3.252*	968.571***	1.757				
(5)	400.480***	2.104	360.287***	2.765*				
lag = 24								
GPar					WeiVM			
(1)	579118.379***	531194.476***	539529.113***	518563.948***	(6)	565572.049***	518892.162***	515969.353***
(2)	691.542***	1288.459***	9911.790***	1741.511***	(7)	680.293***	502.290***	35467.815***
(3)	696.986***	932.965***	1204.897***	20567.796***	(8)	576.437***	1045.957***	10891.294***
(4)	1164.446***	1276.613***	6312.007***	683.462***				
(5)	733.417***	1043.772***	816.566***	640.303***				

Note: Significance: 0.01 (\*\*\*), 0.05 (\*\*), 0.1 (\*).

## 5 Conclusion

This paper develops a dynamic cylindrical modelling framework based on the GPar distribution for circular–linear time series. The proposed specification embeds the heavy-tailed cylindrical distribution of Imoto et al. (2019) within an observation-driven structure, allowing key distributional parameters to vary according to the score of the predictive density. In doing so, the model extends existing cylindrical score-driven approaches by introducing polynomial tail behaviour and, crucially, by permitting the tail index of the linear component to vary dynamically.

The methodological contribution is threefold. First, we provide a tractable dynamic specification for the cylindrical GPar distribution, allowing for time variation in circular location and linear scale. Second, we introduce a dynamic tail mechanism that captures changes in extremal behaviour directly within the conditional distribution, rather than through exogenous regime-switching devices. This feature enhances robustness to extreme observations while preserving likelihood-based tractability. Third, we derive the implied time-varying circular–linear correlation coefficient and show how evolving tail thickness and concentration jointly shape the strength of directional–speed dependence over time.

The empirical application to high-frequency SCADA wind turbine data demonstrates the practical relevance of these extensions. Relative to the WeiVM benchmark, the dynamic GPar model delivers systematic improvements in likelihood-based criteria and forecasting performance, particularly when

the tail index is allowed to vary. The estimated tail dynamics reveal substantial temporal variation in tail thickness, with heavier tails coinciding with periods of increased speed volatility and directional dispersion. In addition, the model provides a coherent dynamic measure of circular–linear dependence, highlighting episodes in which direction and speed become more weakly or strongly coupled.

From a modelling perspective, the results illustrate that allowing for heavier and time-varying tail thickness materially improves the joint description of direction and magnitude in environments characterised by extreme variability. Importantly, these gains are achieved within a fully likelihood-based and computationally tractable framework, which makes the approach suitable for high-frequency applications.

The proposed methodology is not limited to wind data. Many environmental and engineering applications involve cylindrical time series in which heavy-tailed magnitudes and evolving dependence structures are present. Future research may explore multivariate extensions, incorporation of exogenous covariates, or the integration of spatial dependence across multiple turbines or geographical locations.

A further natural extension is to assess the proposed framework using substantially higher-frequency data, such as 50 Hz measurements commonly available in operational and test turbine monitoring systems. At such control-relevant time scales, additional short-run dynamics may become important, potentially requiring richer persistence structures, higher-order score updates, or additional time-varying components in the tail or concentration parameters. Evaluating the adequacy of the cylindrical heavy-tailed specification under these conditions would provide insight into the limits of observation-driven dynamics at very high sampling frequencies and inform potential refinements for real-time control applications. The framework developed here provides a flexible foundation for such developments.

## **Conflicts of interest**

The authors declare that they have no competing interests.

## **Funding**

The authors received no funding for the research.

## **Data availability**

The data were provided for research purposes only by enercity AG and enercity erneuerbare and are not publicly available. The implementation code will be made available at <https://github.com/chriscmr>.

## Author contributions statement

D.P. proposed the implementation of the GPar cylindrical distribution framework and outlined the implementation of the filters. C.T.F. implemented the models (WeiVM and GPar) in software and conducted all computer experiments. C.T.F., D.P., P.S., and Y.Ö. developed the methodological framework. Y.Ö. performed data collection, preprocessing, and visualizations. C.T.F. and D.P. derived the theoretical results, in particular the information matrix. All authors contributed to writing, the interpretation of the results, and reviewed the manuscript. D.P. and P.S. supervised the research.

## 6 Acknowledgments

We are grateful to Andrew Harvey for valuable suggestions and comments. We also thank Shogo Kato, Francesco Lagona, Christoph Ley, Arthur Pewsey, and Ashis Sengupta for insightful discussions on cylindrical data and distributions, which contributed to the development of this paper. We are indebted to Leopoldo Catania and Alessandra Luati for early suggestions on the modelling of cylindrical time series and for inviting us to present an early stage of this work, which ultimately evolved into the present paper, at the Department of Statistics "Paolo Fortunati" at the University of Bologna. We further thank the participants of the ADISTA 2025 workshop at the Université du Luxembourg for helpful feedback. Finally, we acknowledge enercity AG and enercity erneuerbare for providing the high-frequency wind turbine SCADA data used in this study.

## References

- Abe, Toshihiro and Christophe Ley (2017). "A tractable, parsimonious and flexible model for cylindrical data, with applications". In: *Econometrics and Statistics* 4, pp. 91–104.
- anemos Gesellschaft für Umweltmeteorologie mbH (Feb. 12, 2025). *Wind- und Ertragsindex Report 2024*. *Wind- und Ertragsindex Report für das Windjahr 2024*. Reppenstedt. URL: <https://www.anemos.de/Windreport.pdf> (visited on 02/08/2026).
- Beyene, Mussie T, Shaleen Jain, and Ramesh C Gupta (2018). "Linear-circular statistical modeling of Lake Ice-out dates". In: *Water Resources Research* 54(10), pp. 7841–7858.
- Breckling, Jens (1989). "Application to Series of Residual Wind Directions". In: *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*. Springer, pp. 207–215.
- Bundesnetzagentur and Bundeskartellamt (Nov. 26, 2025). *Monitoringbericht 2025*. Bonn. URL: <https://data.bundesnetzagentur.de/Bundesnetzagentur/SharedDocs/Mediathek/Monitoringberichte/MonitoringberichtEnergie2025.pdf> (visited on 02/08/2026).
- Caivano, Michele and Andrew Harvey (2014). "Time-series models with an EGB2 conditional distribution". In: *Journal of Time Series Analysis* 35(6), pp. 558–571.
- Collins, Scott F and Andrea Norton (2024). "Prevailing wind patterns influence the distribution of plastics in small urban lakes". In: *Scientific Reports* 14(1), p. 17741.
- Creal, Drew, Siem Jan Koopman, and André Lucas (2013). "Generalized autoregressive score models with applications". In: *Journal of Applied Econometrics* 28(5), pp. 777–795.
- EPEX SPOT SE (2025). *Introduction of 15-minute MTU contracts for Day-Ahead - SDAC*. URL: <https://www.epexspot.com/en/15-minute-products-market-coupling> (visited on 02/08/2026).
- Fisher, N.I. and A.J. Lee (1994). "Time series analysis of circular data." In: *Journal of the Royal Statistical Society B*(70), pp. 327–332.
- Fisher, Nicholas I and Alan J Lee (1992). "Regression models for an angular response". In: *Biometrics*, pp. 665–677.
- Gradshteyn, IS and IM Ryzhik (2015). *Errata for Table of Integrals, Series, and Products*.
- Harvey, Andrew, Stan Hurn, Dario Palumbo, and Stephen Thiele (2024). "Modelling circular time series". In: *Journal of Econometrics* 239(1), p. 105450.
- Harvey, Andrew and Dario Palumbo (2023a). "Regime switching models for circular and linear time series". In: *Journal of Time Series Analysis* 44(4), pp. 374–392.
- Harvey, Andrew and Dario Palumbo (2023b). "Score-driven models for realized volatility". In: *Journal of Econometrics* 237(2), p. 105448.
- Harvey, Andrew C (2013). *Dynamic models for volatility and heavy tails: with applications to financial and economic time series*. Econometric Society Monographs 52. Cambridge University Press.

- Harvey, Andrew C. (2022). "Score-Driven Time Series Models". In: *Annual Review of Statistics and Its Application* 9, pp. 321–342.
- Imoto, Tomoaki, Kunio Shimizu, and Toshihiro Abe (2019). "A cylindrical distribution with heavy-tailed linear part". In: *Japanese Journal of Statistics and Data Science* 2(1), pp. 129–154.
- Jammalamadaka, S Rao and Ambar Sengupta (2001). *Topics in circular statistics*. Vol. 5. World Scientific Publishing Co.: New Jersey.
- Johnson, Richard A and Thomas Wehrly (1977). "Measures and models for angular correlation and angular-linear correlation". In: *Journal of the Royal Statistical Society Series B: Statistical Methodology* 39(2), pp. 222–229.
- Johnson, Richard A and Thomas E Wehrly (1978). "Some angular-linear distributions and related regression models". In: *Journal of the American Statistical Association* 73(363), pp. 602–606.
- Kato, Shogo and Kunio Shimizu (2008). "Dependent models for observations which include angular ones". In: *Journal of Statistical Planning and Inference* 138(11), pp. 3538–3549.
- Lochmann, Moritz, Heike Kalesse-Los, Michael Schäfer, Ingrid Heinrich, and Ronny Leinweber (2023). "Analysing wind power ramp events and improving very short-term wind power predictions by including wind speed observations". In: *Wind Energy* 26(6), pp. 573–588.
- Lourenço, Nuno and António Rua (2023). "Business cycle clocks: Time to get circular". In: *Empirical Economics* 65(4), pp. 1513–1541.
- Mardia, KV (1976). "Linear-circular correlation coefficients and rhythmometry". In: *Biometrika*, pp. 403–405.
- McDonald, James B. and Yexiao J. Xu (1995). "A generalization of the beta distribution with applications". In: *Journal of Econometrics* 66(1), pp. 133–152.
- Palumbo, D. (2021a). "Testing and Modelling Time Series with Time Varying Tails". In: *Cambridge Working Paper Series*.
- Palumbo, Dario (2021b). *Testing and Modelling Time Series with Time Varying Tails*. CWPE 2111. Faculty of Economics, University of Cambridge. URL: <https://www.econ.cam.ac.uk/sites/default/files/publication-cwpe-pdfs/cwpe2111.pdf>.
- Patton, Andrew J (2011). "Volatility forecast comparison using imperfect volatility proxies". In: *Journal of Econometrics* 160(1), pp. 246–256.
- Wang, Fangpo, Alan E. Gelfand, and Giovanna Jona-Lasinio (2015). "JOINT SPATIO-TEMPORAL ANALYSIS OF A LINEAR AND A DIRECTIONAL VARIABLE: SPACE-TIME MODELING OF WAVE HEIGHTS AND WAVE DIRECTIONS IN THE ADRIATIC SEA". In: *Statistica Sinica* 25(1), pp. 25–39.
- Ypma, Jelmer, Hans W Borchers, Dirk Eddelbuettel, and Maintainer Jelmer Ypma (2018). "Package 'nloptr'". In: *R package version* 1(1).

## A Expectations of functions of Beta functions

Following Lemma 1 in [Harvey \(2013\)](#), pg 23, a random variable  $b$  distributed with a  $beta(\alpha, \beta)$ <sup>3</sup> while  $w(b)$  is a function of a  $b$  with finite expectation,

$$E \left[ b^h (1-b)^k w(b) \right] = \frac{B(\alpha+h, \beta+k)}{B(\alpha, \beta)} E[w(b)], \quad h > -\alpha, k > -\beta$$

where  $B(\alpha, \beta)$  is a beta function and now the expectation on the right-hand side is now understood to be with respect to a  $beta(\alpha+h, \beta+k)$  distribution. Then as, also derived in [Palumbo \(2021b\)](#)

$$E \left[ \ln^h(b) \ln^k(1-b) \right] = \frac{1}{B(\alpha, \beta)} \frac{\partial^h}{\partial \alpha^h} \frac{\partial^k}{\partial \beta^k} B(\alpha, \beta)$$

where  $\frac{\partial}{\partial \alpha} B(\alpha, \beta) = B(\alpha, \beta) [\psi(\alpha) - \psi(\alpha + \beta)]$  and  $\frac{\partial^{(l)}}{\partial \alpha^{(l)}} \psi(\alpha) = \psi^{(l)}(\alpha)$  is the multigamma function for various values of  $l = 0, 1, \dots$ . As a consequence,

$$\begin{aligned} E \left[ b^h (1-b)^k \ln^i(b) \ln^j(1-b) \right] &= E \left[ b^h (1-b)^k \right] \frac{1}{B(\alpha+h, \beta+k)} \frac{\partial^i}{\partial \alpha^{hi}} \frac{\partial^j}{\partial \beta^{kj}} B(\alpha+h, \beta+k) \\ &= \frac{1}{B(\alpha, \beta)} \frac{\partial^i}{\partial \alpha^{hi}} \frac{\partial^j}{\partial \beta^{kj}} B(\alpha+h, \beta+k) \end{aligned}$$

## B Information Matrix for $\mu$

Starting off with the score with respect to the location parameter  $\mu$ ,

$$\frac{\partial \ln f}{\partial \mu} = (1 + \zeta) \frac{\tanh(v) \sin(y_t - \mu)}{1 - \tanh(v) \cos(y_t - \mu)} b_t,$$

where

$$b_t = \frac{\frac{1}{\zeta} (x_t e^{-\lambda})^\alpha (1 - \tanh(v) \cos(y_t - \mu))}{1 + \frac{1}{\zeta} (x_t e^{-\lambda})^\alpha (1 - \tanh(v) \cos(y_t - \mu))},$$

which is conditionally distributed as  $b_t | y_t \sim beta(1, \zeta)$ , see [Harvey and Palumbo \(2023a\)](#). Then we have that its information matrix is computed in the following way

---

<sup>3</sup>this means that  $1-b$  is distributed with a  $beta(\beta, \alpha)$

$$\begin{aligned}
E \left[ \left( \frac{\partial \ln f}{\partial \mu} \right)^2 \right] &= E \left[ (1 + \varsigma)^2 \frac{\tanh^2(v) \sin^2(y_t - \mu)}{(1 - \tanh(v) \cos(y_t - \mu))^2} b_t^2 \right] \\
&= (1 + \varsigma)^2 E \left[ \frac{\tanh^2(v) \sin^2(y_t - \mu)}{(1 - \tanh(v) \cos(y_t - \mu))^2} E(b_t^2 | y_t) \right] \\
&= \frac{2(1 + \varsigma)}{(2 + \varsigma)} E \left[ \frac{\tanh^2(v) \sin^2(y_t - \mu)}{(1 - \tanh(v) \cos(y_t - \mu))^2} \right]
\end{aligned}$$

given that the conditional expectation of  $E(b_t^2 | y_t) = 2/(1 + \varsigma)(2 + \varsigma)$ . Then we will make use of the relationship  $\sin^2 \theta = \frac{1 - \cos 2\theta}{2}$  and the integral definition of the associated Legendre function of the first kind with degree  $n$  and order  $m$ , as described in Eq. 8.711.2 of [Gradshteyn and Ryzhik \(2015\)](#), p. 969, as used in [Abe and Ley \(2017\)](#)

$$P_n^m(z) = \frac{\Gamma(n+1)}{\pi \Gamma(n-m+1)} \int_{-\pi}^0 \frac{\cos mt}{(z - \sqrt{z^2 - 1} \cos t)^{n+1}} dt$$

Then using the pdf of the marginal wrapped cachy distribution for  $y_t$

$$f_Y(y_t) = \frac{\sqrt{1 - \tanh^2(v)}}{2\pi (1 - \tanh(v) \cos(y_t - \mu))}$$

we have that

$$\begin{aligned}
&= \frac{2(1 + \varsigma)}{(2 + \varsigma)} E \left[ \frac{\tanh^2(v)}{2(1 - \tanh(v) \cos(y_t - \mu))^2} - \frac{\tanh^2(v) \cos 2(y_t - \mu)}{2(1 - \tanh(v) \cos(y_t - \mu))^2} \right] \\
&= \frac{(1 + \varsigma)}{(2 + \varsigma)} \tanh^2(v) \cosh^3(v) \sqrt{1 - \tanh^2(v)} \left( \frac{1}{\pi} \int_{-\pi}^0 \frac{d\theta}{(\cosh(v) - \sinh(v) \cos(\theta))^3} \right. \\
&\quad \left. - \frac{1}{\pi} \int_{-\pi}^0 \frac{\cos(2\theta)}{(\cosh(v) - \sinh(v) \cos(\theta))^3} d\theta \right) \\
&= \frac{(1 + \varsigma)}{(2 + \varsigma)} \sinh^2(v) \left( P_2^0(\cosh(v)) - \frac{1}{2} P_2^2(\cosh(v)) \right)
\end{aligned}$$

Given the tabulated results for the associated Legendre function of the first kind, such as  $P_2^0(x) = (3x^2 - 1)/2$  and  $P_2^2(x) = 3(x^2 - 1)$  we have that

$$\begin{aligned}
&= \frac{(1 + \varsigma)}{(2 + \varsigma)} \frac{\sinh^2(v)}{2} \left( (3 \cosh^2(v) - 1) - 3(\cosh^2(v) - 1) \right) \\
&= \frac{(1 + \varsigma)}{(2 + \varsigma)} \sinh^2(v) = \mathcal{I}_{\mu\mu}
\end{aligned}$$

where we used the relations  $1 - \cosh^2(v) = \sinh^2(v)$

## C Information Matrix for $\lambda$

Starting with the score with respect to the parameter driving the scale of the conditional linear distribution,  $\lambda$ , we have that

$$\frac{\partial \ln f}{\partial \lambda} = \alpha [(1 + \zeta) b_t - 1]$$

where  $b_t$  is conditionally distributed  $b_t|y_t \sim \text{beta}(1, \zeta)$ , which is consistent with [Harvey \(2013\)](#) and [Harvey and Palumbo \(2023b\)](#).

Then the information quantity with respect to the parameter  $\lambda$  can be obtained by

$$\begin{aligned} E \left[ \frac{\partial^2 \ln f}{\partial^2 \lambda} \right] &= \alpha (1 + \zeta) E \left[ \frac{\partial b_t}{\partial \lambda} \right] \\ &= -\alpha^2 (1 + \zeta) E [E (b_t (1 - b_t) | y_t)] \\ &= -\frac{\alpha^2 \zeta}{2 + \zeta} \end{aligned}$$

therefore

$$\mathcal{I}_{\lambda\lambda} = -E \left[ \frac{\partial^2 \ln f}{\partial^2 \lambda} \right] = \frac{\alpha^2 \zeta}{2 + \zeta},$$

## D Information Matrix for $v$

Starting off with the score with respect to the concentration parameter  $v$ ,

$$\frac{\partial \ln f}{\partial v} = \frac{1 + \zeta}{\cosh(v)^2} \frac{\cos(y - \mu)}{(1 - \tanh(v) \cos(y - \mu))} b_t - \tanh(v),$$

Then the information quantity with respect to the parameter  $v$  can be obtained as

$$\begin{aligned}
E \left[ \left( \frac{\partial \ln f}{\partial v} \right)^2 \right] &= E \left[ \left( \frac{1 + \zeta}{\cosh(v)^2} \frac{\cos(y - \mu)}{(1 - \tanh(v) \cos(y - \mu))} b_t - \tanh(v) \right)^2 \right] \\
&= \frac{(1 + \zeta)^2}{\cosh(v)^4} E \left[ \frac{\cos(y - \mu)^2}{(1 - \tanh(v) \cos(y - \mu))^2} E(b_t^2 | y_t) \right] + \tanh(v)^2 \\
&\quad - 2 \tanh(v) \frac{(1 + \zeta)}{\cosh(v)^2} E \left[ \frac{\cos(y - \mu)}{(1 - \tanh(v) \cos(y - \mu))} E(b_t | y_t) \right] \\
&= \frac{(1 + \zeta)}{(2 + \zeta) \cosh(v)^4} E \left[ \frac{1 + \cos 2(y - \mu)}{(1 - \tanh(v) \cos(y - \mu))^2} \right] + \tanh(v)^2 \\
&\quad - 2 \frac{\tanh(v)}{\cosh(v)^2} E \left[ \frac{\cos(y - \mu)}{(1 - \tanh(v) \cos(y - \mu))} \right] \\
&= \frac{(1 + \zeta)}{(2 + \zeta) \cosh(v)^2} \left( \frac{1}{\pi} \int_{-\pi}^0 \frac{1}{(\cosh(v) - \sinh(v) \cos(y - \mu))^3} \right. \\
&\quad \left. + \frac{1}{\pi} \int_{-\pi}^0 \frac{\cos 2(y - \mu)}{(\cosh(v) - \sinh(v) \cos(y - \mu))^3} \right) + \\
&\quad + \tanh(v)^2 - 2 \frac{\tanh(v)}{\cosh(v)} \frac{1}{\pi} \int_{-\pi}^0 \frac{\cos(y - \mu)}{(\cosh(v) - \sinh(v) \cos(y - \mu))^2} \\
&= \frac{(1 + \zeta)}{(2 + \zeta) \cosh(v)^2} \left( P_2^0(\cosh(v)) + \frac{1}{2} P_2^2(\cosh(v)) \right) + \tanh(v)^2 - 2 \frac{\tanh(v)}{\cosh(v)} P_1^1(\cosh(v)) \\
&= \frac{(1 + \zeta)}{(2 + \zeta) \cosh(v)^2} \left( \frac{1}{2} (3 \cosh(v)^2 - 1) + \frac{3}{2} (\cosh(v)^2 - 1) \right) + \tanh(v)^2 \\
&\quad - 2 \frac{\tanh(v)}{\cosh(v)} \sqrt{\cosh(v)^2 - 1} \\
&= \frac{(1 + \zeta)}{(2 + \zeta) \cosh(v)^2} (3 \cosh(v)^2 - 2) + \tanh(v)^2 - 2 \tanh(v) \\
&= \frac{(1 + \zeta)}{(2 + \zeta)} \left( 3 - \frac{2}{\cosh(v)^2} \right) - \left( 1 - \frac{1}{\cosh(v)^2} \right) \\
&= \frac{(1 + \zeta) + \zeta \tanh(v)^2}{(2 + \zeta)} = \mathcal{I}_{vv}
\end{aligned}$$

where  $P_1^1(x) = \sqrt{x^2 - 1}$  for  $x > 1$ .

## E Information for the tail index $\chi$

We define  $\kappa = \tanh(v)$  and  $\delta_t = y_t - \mu$ .

The score for  $\chi$  is

$$s_{\chi,t} = 1 + \alpha \log(x_t e^{-\lambda}) \left[ 1 - (1 + \varsigma)b_t \right],$$

and given that

$$\alpha \log(x_t e^{-\lambda}) = \log \varsigma - \log(1 - \kappa \cos \delta_t) + \log \frac{b_t}{1 - b_t},$$

define

$$a(\delta) = \log \varsigma - \log(1 - \kappa \cos \delta), \quad \ell_t = \log \frac{b_t}{1 - b_t}.$$

Then the score can be written

$$s_{\chi,t} = 1 + (a + \ell)(1 + \beta b_t), \quad \beta = -(1 + \varsigma).$$

Let  $Y = (a + \ell)(1 + \beta b)$  so that

$$s_{\chi,t}^2 = 1 + 2Y + Y^2, \quad Y^2 = (a + \ell)^2(1 + 2\beta b + \beta^2 b^2).$$

We first take expectation over the Beta component since  $b \sim \text{Beta}(1, \varsigma)$ . Let  $\ell = \log \frac{b}{1-b}$  and define the Beta-logit moments

$$r_0 = \mathbb{E}[b], \quad r_1 = \mathbb{E}[b^2], \quad m_0 = \mathbb{E}[\ell], \quad m_1 = \mathbb{E}[\ell^2], \quad q_0 = \mathbb{E}[b\ell], \quad q_1 = \mathbb{E}[b^2\ell], \quad p_0 = \mathbb{E}[b\ell^2], \quad p_1 = \mathbb{E}[b^2\ell^2],$$

all taken under  $b \sim \text{Beta}(1, \varsigma)$ .

These admit closed forms in terms of the digamma  $\psi(\cdot)$  and trigamma  $\psi_1(\cdot)$  functions. For Beta(1,  $\varsigma$ ) one obtains

$$r_0 = \frac{1}{1 + \varsigma}, \quad r_1 = \frac{2}{(1 + \varsigma)(2 + \varsigma)}, \quad m_0 = \psi(1) - \psi(\varsigma), \quad m_1 = \psi_1(1) + \psi_1(\varsigma) + (\psi(1) - \psi(\varsigma))^2 \quad (8)$$

$$q_0 = r_0(\psi(2) - \psi(\varsigma)), \quad q_1 = r_1(\psi(3) - \psi(\varsigma)), \quad p_0 = r_0(\psi_1(2) + \psi_1(\varsigma) + (\psi(2) - \psi(\varsigma))^2), \quad (9)$$

$$p_1 = r_1(\psi_1(3) + \psi_1(\varsigma) + (\psi(3) - \psi(\varsigma))^2). \quad (10)$$

After expanding  $s_{\chi,t}^2$  and collecting powers of  $a(\delta)$  one obtains

$$\mathbb{E}[s_{\chi,t}^2 \mid \delta] = F_0 + F_1 a(\delta) + F_2 a(\delta)^2,$$

where

$$F_2 = 1 + 2\beta r_0 + \beta^2 r_1, \quad F_1 = 2(1 + \beta r_0)m_0 + 2\beta q_0 + 2\beta^2 q_1, \quad F_0 = 1 + 2(m_0 + \beta q_0) + m_1 + 2\beta p_0 + \beta^2 p_1. \quad (11)$$

Taking expectation over  $\delta$  which follows a wrapped-Cauchy therefore yields

$$\mathcal{I}_{X\lambda} = F_0 + F_1 A_1 + F_2 A_2, \quad (12)$$

where  $A_1 = \mathbb{E}[a(\delta)]$ ,  $A_2 = \mathbb{E}[a(\delta)^2]$ . We now compute these wrapped-Cauchy expectations. Remember that the wrapped-Cauchy density is given by

$$f(\delta) = \frac{\sqrt{1 - \kappa^2}}{2\pi(1 - \kappa \cos \delta)}, \quad 0 < \kappa < 1.$$

Introduce the Poisson parameter

$$\rho = \frac{1 - \sqrt{1 - \kappa^2}}{\kappa}, \quad 0 < \rho < 1,$$

so that

$$1 - \kappa \cos \delta = \frac{1}{1 + \rho^2} (1 - 2\rho \cos \delta + \rho^2), \quad \sqrt{1 - \kappa^2} = \frac{1 - \rho^2}{1 + \rho^2}.$$

Hence

$$\log(1 - \kappa \cos \delta) = -\log(1 + \rho^2) + L(\delta),$$

with  $L(\delta) = \log(1 - 2\rho \cos \delta + \rho^2)$ .

Using the factorization

$$1 - 2\rho \cos \delta + \rho^2 = (1 - \rho e^{i\delta})(1 - \rho e^{-i\delta}),$$

and the Taylor expansion of  $\log(1 - w) = -\sum_{n=1}^{\infty} w^n/n$  for  $|w| < 1$  gives the Fourier series

$$L(\delta) = -2 \sum_{n=1}^{\infty} \frac{\rho^n}{n} \cos(n\delta).$$

For the wrapped Cauchy distribution, one has

$$\mathbb{E}[\cos(n\delta)] = \rho^n,$$

so that

$$\mathbb{E}[L(\delta)] = -2 \sum_{n \geq 1} \frac{\rho^{2n}}{n} = 2 \log(1 - \rho^2).$$

To obtain the  $\mathbb{E}[L^2(\delta)]$ , write

$$L(\delta) = \sum_{n \geq 1} \ell_n \cos(n\delta), \quad \ell_n = -\frac{2\rho^n}{n},$$

and from the Parseval identity we have:

$$\mathbb{E}[L(\delta)^2] = \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \rho^k, \quad (13)$$

with  $a_0$  and  $a_k$  define as:

$$\frac{a_0}{2} = \frac{1}{2\pi} \int_0^{2\pi} L(\delta)^2 d\delta, \quad a_k = \frac{1}{\pi} \int_0^{2\pi} L(\delta)^2 \cos(k\delta) d\delta, \quad k \geq 1. \quad (14)$$

The cosine Fourier expansion of  $L(\delta)^2$  using the discrete convolution gives

$$L(\delta)^2 = \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \ell_n \ell_m \cos(n\delta) \cos(m\delta). \quad (15)$$

Substituting (15) in (14) and using cosine identities shows that the integral is non-zero only if  $k = |n - m|$  or  $k = n + m$ , yielding the convolution representation

$$a_k = \frac{1}{2} \sum_{n,m \geq 1} \ell_n \ell_m \left( \mathbf{1}\{k = |n - m|\} + \mathbf{1}\{k = n + m\} \right).$$

Reindexing and substituting  $\ell_n = -2\rho^n/n$  yields

$$\begin{aligned} a_k &= \frac{1}{2} \sum_{n=1}^{k-1} \ell_n \ell_{k-n} + \sum_{n \geq 1} \ell_n \ell_{n+k} \\ &= 2\rho^k \sum_{n=1}^{k-1} \frac{1}{n(k-n)} + 4 \sum_{n=1}^{\infty} \frac{\rho^{2n+k}}{n(n+k)} \\ &= 4\rho^k \frac{H_{k-1}}{k} + 4 \sum_{n=1}^{\infty} \frac{\rho^{2n+k}}{n(n+k)}, \quad H_{k-1} = \sum_{j=1}^{k-1} \frac{1}{j} \end{aligned}$$

Using the identities  $4 \sum_{k=1}^{\infty} \frac{H_{k-1}}{k} \rho^{2k} = 2 \log^2(1 - \rho^2)$  and  $\frac{1}{n+k} = \int_0^1 u^{n+k-1} du$ ,  $n, k \geq 1$  lead to

$$\begin{aligned} \sum_{k=1}^{\infty} a_k \rho^k &= 4 \sum_{k=1}^{\infty} \frac{H_{k-1}}{k} \rho^{2k} + 4 \sum_{n=1}^{\infty} \sum_{k=1}^{\infty} \frac{\rho^{2n+2k}}{n(n+k)} \\ &= \int_0^1 \frac{1}{u} \left( \sum_{n=1}^{\infty} \frac{(\rho^2 u)^n}{n} \right) \left( \sum_{k=1}^{\infty} (\rho^2 u)^k \right) du \\ &= \int_0^1 \frac{\rho^2 [-\log(1 - \rho^2 u)]}{1 - \rho^2 u} du \\ &= 4 \log^2(1 - \rho^2) \end{aligned} \quad (16)$$

For  $a_0$  we use the definition of the dilogarithm  $\text{Li}_2$  to get:

$$\frac{a_0}{2} = \frac{1}{2\pi} \int_0^{2\pi} L(\delta)^2 d\delta = 2 \sum_{n=1}^{\infty} \frac{\rho^{2n}}{n^2} = 2 \text{Li}_2(\rho^2). \quad (17)$$

Hence it follows from (16) and (17) that:

$$\mathbb{E}[L(\delta)^2] = 2 \text{Li}_2(\rho^2) + 4 \log^2(1 - \rho^2).$$

Finally, using standard identities for hyperbolic functions, we have that

$$\rho = \tanh\left(\frac{v}{2}\right), \quad 1 - \rho^2 = \text{sech}^2\left(\frac{v}{2}\right), \quad 1 + \rho^2 = \frac{\cosh(v)}{\cosh^2(v/2)}$$

Therefore

$$A_1 = -\mathbb{E} \left[ \log \left( \frac{1}{\zeta} (1 - \tanh(v) \cos \delta) \right) \right] = \log \zeta + \log \cosh(v) + 2 \log \cosh\left(\frac{v}{2}\right),$$

$$A_2 = \mathbb{E} \left[ \log^2 \left( \frac{1}{\zeta} (1 - \tanh(v) \cos \delta) \right) \right] = A_1^2 + 2 \text{Li}_2\left(\tanh^2\left(\frac{v}{2}\right)\right).$$

Substituting into (12) gives the final explicit form

$$\mathcal{I}_{\chi\chi} = F_0 + F_1 A_1 + F_2 \left[ A_1^2 + 2 \text{Li}_2\left(\tanh^2\left(\frac{v}{2}\right)\right) \right],$$

with  $F_0, F_1, F_2$  defined in (11).

## F Derivation of the Correlation Coefficient

We start from defining,

$$\text{Cov}(x_t e^{-\lambda}, \cos(y_t - \mu)) = \zeta^{1+\frac{1}{\alpha}} \frac{\Gamma\left(\frac{1}{\alpha}\right) \Gamma\left(\zeta - \frac{1}{\alpha}\right)}{\Gamma(\zeta)} \left[ P_{1/\alpha}^1(\cosh(v)) - \frac{\tanh(v/2)}{\alpha} P_{1/\alpha}^0(\cosh(v)) \right],$$

$$\text{Cov}(x_t e^{-\lambda}, \sin(y_t - \mu)) = 0,$$

$$\text{Cov}(\cos(y_t - \mu), \sin(y_t - \mu)) = 0,$$

where  $P_b^a(x)$  is the Legendre function of the first kind. Then, given that

$$Var(x_t e^{-\lambda}) = \frac{\zeta^{1+\frac{1}{\alpha}} \cosh(v)^{2/\alpha}}{\alpha \Gamma(\zeta)} \left[ 2\Gamma\left(\frac{2}{\alpha}\right) \Gamma\left(\zeta - \frac{2}{\alpha}\right) P_{1/\alpha}^2(\cosh(v)) - \frac{\zeta \Gamma\left(\frac{1}{\alpha}\right)^1 \Gamma\left(\zeta - \frac{1}{\alpha}\right)^2}{\alpha \Gamma(\zeta)} P_{1/\alpha}^0(\cosh(v)) \right],$$

$$Var(\cos(y_t - \mu)) = \frac{1}{2 \cosh(v)^2},$$

as a result we have that

$$R_{X\Theta}^2 = r_{XC}^2 = 2\alpha\zeta \cosh(v)^{2-\frac{2}{\alpha}} \frac{\Gamma\left(\frac{1}{\alpha}\right)^2 \Gamma\left(\zeta - \frac{1}{\alpha}\right)^2 \left[ P_{1/\alpha}^1(\cosh(v)) - \frac{\tanh(v/2)}{\alpha} P_{1/\alpha}^0(\cosh(v)) \right]^2}{\Gamma(\zeta) \left[ 2\Gamma\left(\frac{2}{\alpha}\right) \Gamma\left(\zeta - \frac{2}{\alpha}\right) P_{1/\alpha}^2(\cosh(v)) - \frac{\zeta \Gamma\left(\frac{1}{\alpha}\right)^1 \Gamma\left(\zeta - \frac{1}{\alpha}\right)^2}{\alpha \Gamma(\zeta)} P_{1/\alpha}^0(\cosh(v)) \right]},$$