

# Inference for multivariate extremes via a semi-parametric angular-radial model

Callum John Rowlandson Murphy-Barltrop<sup>1,2\*</sup>, Ed Mackay<sup>3</sup>, and Philip  
Jonathan<sup>4,5</sup>

<sup>1</sup>Technische Universität Dresden, Institut Für Mathematische Stochastik, Helmholtzstraße 10,  
01069 Dresden, Germany

<sup>2</sup>Center for Scalable Data Analytics and Artificial Intelligence (ScaDS.AI), Dresden/Leipzig,  
Germany

<sup>3</sup>Department of Engineering, University of Exeter, TR10 9FE, United Kingdom

<sup>4</sup>Department of Mathematics and Statistics, Lancaster University LA1 4YF, United Kingdom

<sup>5</sup>Shell Research Limited, London SE1 7NA, United Kingdom

\*Correspondence to: callum.murphy-barltrop@tu-dresden.de

15<sup>th</sup> January, 2024

## Abstract

The modelling of multivariate extreme events is important in a wide variety of applications, including flood risk analysis, metocean engineering and financial modelling. A wide variety of statistical techniques have been proposed in the literature; however, many such methods are limited in the forms of dependence they can capture, or make strong parametric assumptions about data structures. In this article, we introduce a novel inference framework for multivariate extremes based on a semi-parametric angular-radial model. This model overcomes the limitations of many existing approaches and provides a unified paradigm for assessing joint tail behaviour.

Alongside inferential tools, we also introduce techniques for assessing uncertainty and goodness of fit. Our proposed technique is tested on simulated data sets alongside observed metocean time series', with results indicating generally good performance.

*Keywords:* Multivariate Extremes, Extremal Dependence, Generalised Additive Models, Coordinate Systems

# 1 Introduction

## 1.1 Multivariate extreme value modelling

The modelling of multivariate extremes is an active area of research, with applications spanning many domains, including meteorology (Chavez-Demoulin and Davison, 2005), metocean engineering (Jonathan and Ewans, 2013; Vanem et al., 2022), financial modelling (Castro-Camilo et al., 2018) and flood risk assessment (Diaconu et al., 2021). Typically, approaches in this research field are comprised of two steps: first, modelling the extremes of individual variables and transforming to common margins, followed by modelling of the dependence between the extremes of different variables. We refer to this dependence as the extremal dependence structure henceforth.

This article discusses inference for multivariate extremes using an angular-radial model for the *probability density function*, illustrated using examples in two dimensions. To place the proposed model in context, we first provide a brief synopsis of the existing literature for multivariate extremes. Given a random vector  $(X, Y) \in \mathbb{R}^2$  with marginal distribution functions  $F_X$  and  $F_Y$ , the strength of dependence in the upper tail of  $(X, Y)$  can be quantified in terms of the tail dependence coefficient,  $\chi \in [0, 1]$ , defined as the limiting probability

$$\chi = \lim_{u \rightarrow 1} \Pr(F_X(X) > u \mid F_Y(Y) > u),$$

when this limit exists (Joe, 1997). When  $\chi > 0$ , the components of  $(X, Y)$  are said to

be *asymptotically dependent* (AD) in the upper tail, and when  $\chi = 0$ , they are said to be *asymptotically independent* (AI). Much of the focus of recent work in multivariate extreme value theory has been related to developing a general framework for modelling joint extremes of  $(X, Y)$  which is applicable to both AD and AI cases, and can be used to evaluate joint tail behaviour in the region where at least one variable is large.

To discuss the approaches proposed to date and their associated limitations, it is helpful to categorise them in terms of whether they assume heavy- or light-tailed margins, and whether they consider the distribution or density function. Classical multivariate extreme value theory assumes heavy-tailed margins, and is based on the framework of multivariate regular variation (MRV, Resnick, 1987). It addresses the case where  $\chi > 0$ , and has been widely studied – see Beirlant et al. (2004), de Haan and Ferreira (2006) and Resnick (2007) for reviews. Under some regularity conditions, equivalent asymptotic descriptions of joint extremal behaviour can be obtained from either the density or distribution function (de Haan and Resnick, 1987).

In the MRV framework, any distribution with  $\chi = 0$  has the same asymptotic representation. To address this issue, Ledford and Tawn (1996, 1997) proposed a method to characterise joint extremes for both AI and AD distributions in the region where both variables are large. The resulting framework also assumes heavy-tailed margins and is referred to as hidden regular variation (HRV, Resnick, 2002). However, for AI distributions, a description of extremal behaviour in the region where both variables are large may not be the most useful, since extremes of both variables are unlikely to occur simultaneously. Moreover, for AI distributions with certain regularity conditions, the asymptotic representation in this framework is governed only by the properties of the distribution along the line  $y = x$  (Mackay and Jonathan, 2023). To provide a more useful representation for AI distributions, applicable in the region where either variable is large, Wadsworth and Tawn (2013) introduced an asymptotic model for the joint survivor function on standard exponential margins. In contrast to the MRV framework, the resulting model provides a useful description of AI distributions, but all AD

distributions have the same representation.

More recently, there has been interest in modelling the limiting shapes of scaled sample clouds, or *limit sets*, first introduced by Davis et al. (1988); see, for instance, Simpson and Tawn (2022); Wadsworth and Campbell (2023); Majumder et al. (2023) and Papastathopoulos et al. (2023). Nolde and Wadsworth (2022) showed that the limit set is directly linked to several representations for multivariate extremes.. For a given distribution, the limit set is obtained by evaluating the asymptotic behaviour of the joint density function on lighted tailed margins. However, the limit set itself does not provide a full description of the asymptotic joint density or distribution, so is less useful from a practical modelling perspective.

To understand the limitations of the methods discussed above, it is instructive to provide an illustration of the joint distribution and density functions on heavy- and light-tailed margins for AI and AD random vectors. All the methods discussed above have equivalent representations in angular-radial coordinates, so without loss of generality, we consider the angular-radial dependence. The first step for most methods for modelling multivariate extremes is to transform variables to common margins. Define

$$\begin{aligned}(X_P, Y_P) &= ((1 - F_X(X))^{-1}, (1 - F_Y(Y))^{-1}) \in [1, \infty)^2, \\(X_E, Y_E) &= (-\log(1 - F_X(X)), -\log(1 - F_Y(Y))) \in [0, \infty)^2,\end{aligned}$$

so that  $(X_P, Y_P)$  and  $(X_E, Y_E)$  have standard Pareto and exponential margins, respectively. Note that  $(X_E, Y_E) = (\log(X_P), \log(Y_P))$ , and that the dependence structure or copula of  $(X, Y)$  remains unchanged by the marginal transformation (Sklar, 1959). Furthermore, the joint survivor function  $\bar{F}_P(x, y) = \Pr(X_P > x, Y_P > y)$  is related to the joint survivor function of  $(X_E, Y_E)$  by  $\bar{F}_E(x, y) = \bar{F}_P(\exp(x), \exp(y))$ . Moreover, if  $(X_P, Y_P)$  has joint density function  $f_P(x, y)$ , then  $(X_E, Y_E)$  has joint density  $f_E(x, y) = \exp(-r)f_P(\exp(x), \exp(y))$ , where  $r = x + y$ .

Figure 1.1 shows the joint survivor and density functions for the AD Joe copula (see the

Supplementary Material) on standard Pareto and exponential margins. Rays of constant angle on each margin are also shown. On Pareto margins, with the axes shown on a logarithmic scale, lines of constant angle asymptote to lines with unit gradient. As such, the MRV framework provides a description of joint tail behaviour in the region close to the line  $\log(y) = \log(x)$ , i.e., where  $X_P$  and  $Y_P$  are of similar magnitudes. In this region, the contours of the joint density and survivor functions asymptote to a curve of constant shape, which describes the joint extremal behaviour in this region. In contrast, the angular-radial description appears different on exponential margins. For the joint survivor function, the contours of constant probability appear to asymptote towards the line  $\max(x, y) = c$  for some constant  $c$ . Wadsworth and Tawn (2013) showed that is the case for all AD distributions. Informally, this is because for a distribution to be AD, the probability mass must be concentrated close to the line  $y = x$ , so when the density is integrated to obtain the survivor function, the dominant contribution comes from this region. In contrast to the joint survivor function, the angular-radial description of the joint density is not the same for all AD distributions on exponential margins.

Figure 1.2 shows a similar set of plots for the AI Gaussian copula (see the Supplementary Material). In this case, contours for both the joint density and joint survivor function are curved on both sets of margins. The angular-radial model on Pareto margins describes the section of the curves close to the line  $y = x$ , which asymptote to straight lines as  $x + y \rightarrow \infty$ . Therefore, the HRV description of the asymptotic behaviour is effectively a straight line approximation to a curve, and is only applicable in the region close to the line  $y = x$ ; see Mackay and Jonathan (2023) for details. In contrast, the angular-radial description of both the density and survivor functions on exponential margins provides a more useful description of asymptotic behaviour.

In some applications it is useful to describe the extremal behaviour of a random vector for both large and small values of certain variables; see Section 1.2. In this case, it is more useful to work on symmetric two-sided margins, rather than one-sided margins. Figure 1.3 shows

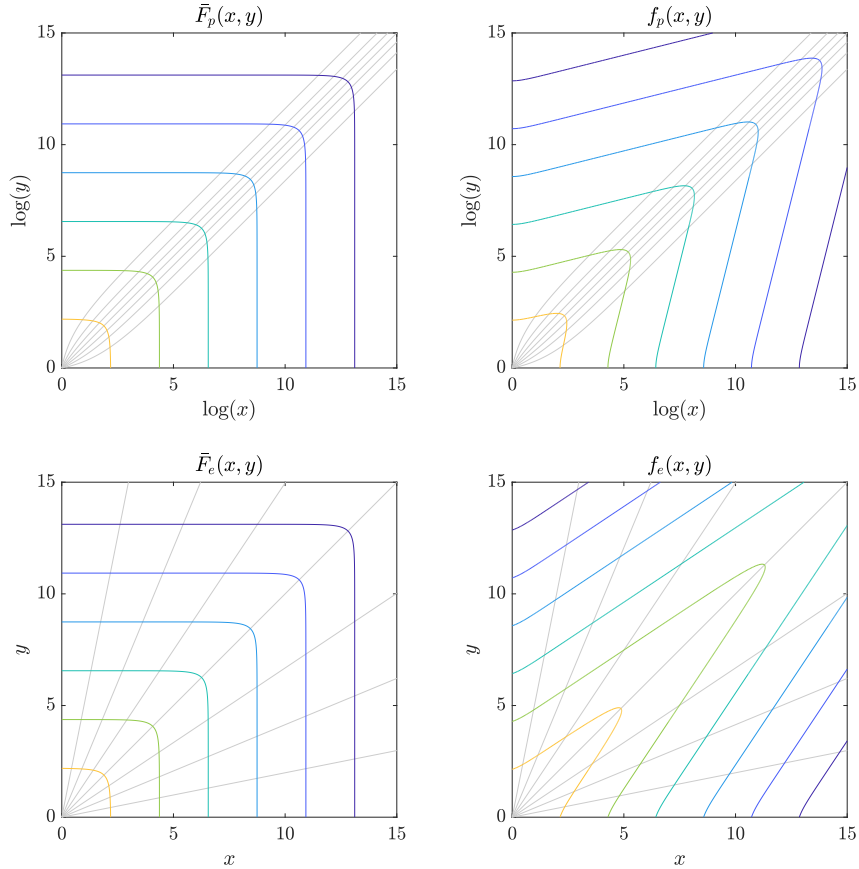


Figure 1.1: Representations of a Joe copula (with parameter  $\alpha = 3$ ) on standard Pareto (upper row) and standard exponential margins (lower row). Left plots: Contours of joint survivor function at equal logarithmic increments. Right plots: Contours of joint density function at equal logarithmic increments. Light grey lines show rays of constant angle on each margin.

the joint survivor and density functions for a Gaussian copula on standard Laplace margins. The angular-radial variation of the joint survivor function is useful in the first quadrant of the plane, but is less useful in the other quadrants. In the second and fourth quadrants, the contours of the joint survivor function asymptote to the corresponding marginal levels, providing no information about the asymptotic behaviour of the distribution in this region. In contrast, the joint density function provides useful asymptotic information in all regions of the plane.

This motivates an intuitively-appealing angular-radial description of the joint density function proposed e.g. by Mackay (2022), referred to as the semi-parametric angular-radial

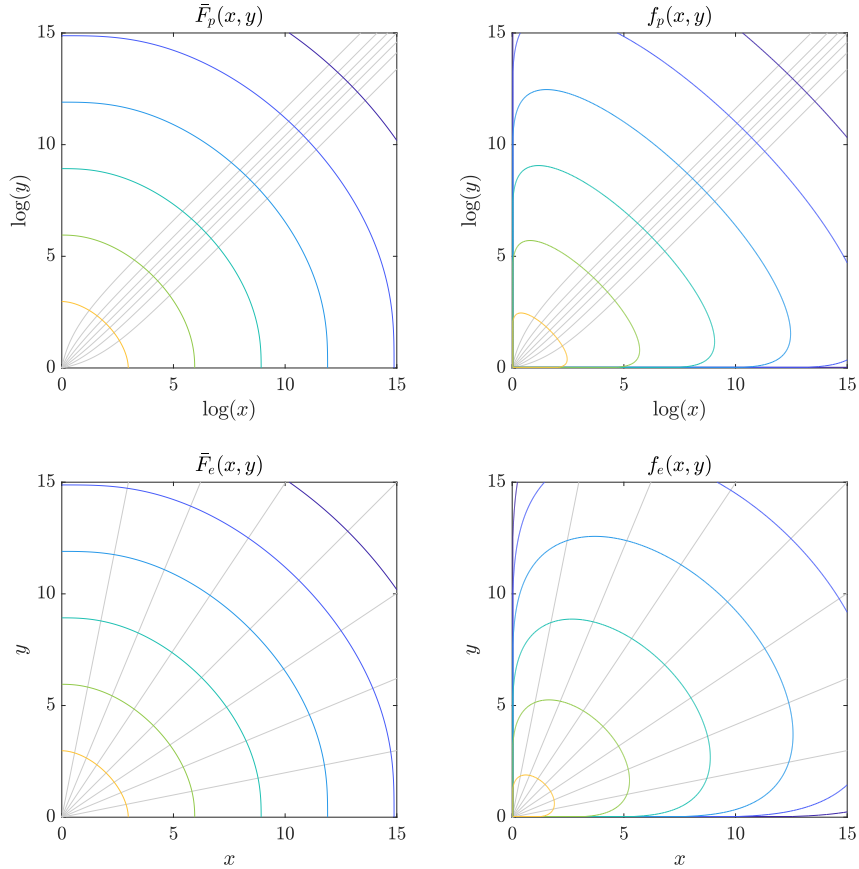


Figure 1.2: As previous figure, but for Gaussian copula with  $\rho = 0.5$ .

(SPAR) model, considered in the current work. A similar model was recently proposed by Papastathopoulos et al. (2023), although the application was only considered for standard Laplace margins. However, the SPAR framework can be applied on any type of margin. Mackay and Jonathan (2023) showed that on heavy-tailed margins, SPAR is consistent with the MRV/HRV frameworks, and on light-tailed margins, SPAR is consistent with limit set theory. However, the SPAR framework is more general than limit set theory, as it provides an explicit model for the density in extreme regions of the variable space. Moreover, there are distributions which have degenerate limit sets in some regions, for which there is still a useful SPAR representation.

In the SPAR framework, variables are transformed to angular-radial coordinates, and it is assumed that the conditional radial distribution is in the domain of attraction of an extreme value distribution. This implies the radial tail conditioned on angle can be approximated by

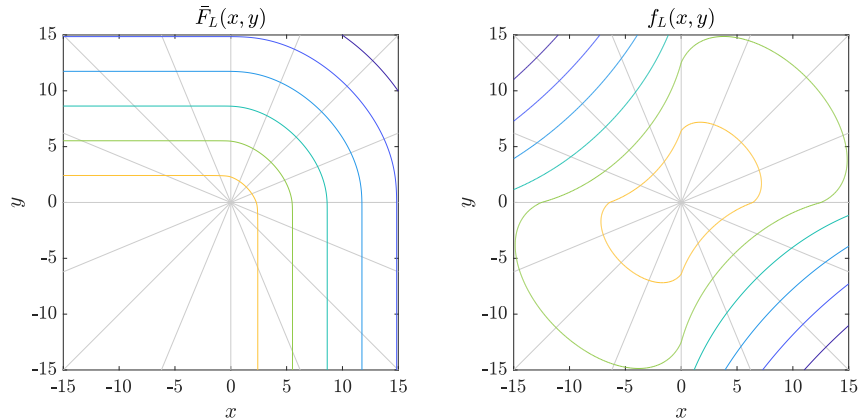


Figure 1.3: Gaussian copula on standard Laplace margins. Left: Contours of joint survivor function at equal logarithmic increments. Right: Contours of joint density function at equal logarithmic increments. Light grey lines show rays of constant angle.

a non-stationary generalised Pareto (GP) distribution. The SPAR approach generalises the model proposed by [Wadsworth et al. \(2017\)](#), in which angular and radial components are assumed to be independent. In the [Wadsworth et al. \(2017\)](#) model, the margins and angular-radial coordinate system are selected so that the assumption of independent angular and radial components is satisfied. The SPAR framework removes this requirement, providing a more flexible representation for multivariate extremes.

While a strong theoretical foundation for the SPAR model is provided in [Mackay and Jonathan \(2023\)](#), inference for this model has not yet been demonstrated. Inference via this framework would offer advantages over many existing approaches, and a fitted SPAR model could be used to estimate extreme quantities commonly applied in practice, such as risk measures ([Murphy-Barltrop et al., 2023](#)) and joint tail probabilities ([Keef et al., 2013](#)).

The SPAR model reframes multivariate extreme value modelling as non-stationary peaks over threshold (POT) modelling with angular dependence. Many approaches have been proposed for non-stationary POT inference e.g. [Randell et al. \(2016\)](#); [Youngman \(2019\)](#); [Zanini et al. \(2020\)](#). In this paper, we introduce an ‘off-the-shelf’ inference framework for the SPAR model. This framework, which utilises generalised additive models (GAMs; [Wood, 2017](#)) for capturing the relationship between radial and angular components, offers a high degree of flexibility and can capture a wide variety of extremal dependence structures, as demonstrated



in Sections 6 and 7. Our approach offers utility across a wide range of applications and provides a convenient, practical framework for performing inference on multivariate extremes. Moreover, our inference framework is ready to use by practitioners; open-source software for fitting the SPAR model is available at <https://github.com/callumbarltrop/SPAR>. For ease of discussion and illustration, we restrict attention to the bivariate setting throughout, noting that the SPAR model is not limited to this setting.

## 1.2 Motivating examples

To demonstrate the practical applicability of our proposed inference framework, we consider three bivariate metocean time series made up of zero-up-crossing period,  $T_z$ , and significant wave height,  $H_s$ , observations. We label these data sets as A, B and C, with each data set corresponding to a location off the coast of North America. Datasets A and B were previously considered in a benchmarking exercise for environmental contours (Haselsteiner et al., 2021). Observations were recorded on an hourly basis over 40, 31 and 42 year time periods for data sets A, B and C, resulting in  $n_A = 320740$ ,  $n_B = 241815$  and  $n_C = 328247$  observations, respectively, once missing observations are taken into account. Exploratory analysis indicates the joint time series are approximately stationary over the observation period. Understanding the joint extremes of metocean variables is important in the field of ocean engineering for assessing the reliability of offshore structures. Wave loading on structures is dependent on both wave height and period, and the largest loads on a structure may not necessarily occur with the largest wave heights. Resonances in a structure may result in the largest responses occurring with either short- or long-period waves, meaning it is necessary to characterise the joint distribution in both of these ranges. These data sets are illustrated in Figure 1.4.

Metocean data sets of this type can often exhibit complex dependence structures, for which many multivariate models fail to account. For example, data set B exhibits clear asymmetry in its dependence structure. Moreover, as demonstrated in Haselsteiner et al. (2021), many existing approaches for modelling metocean data sets perform poorly in prac-

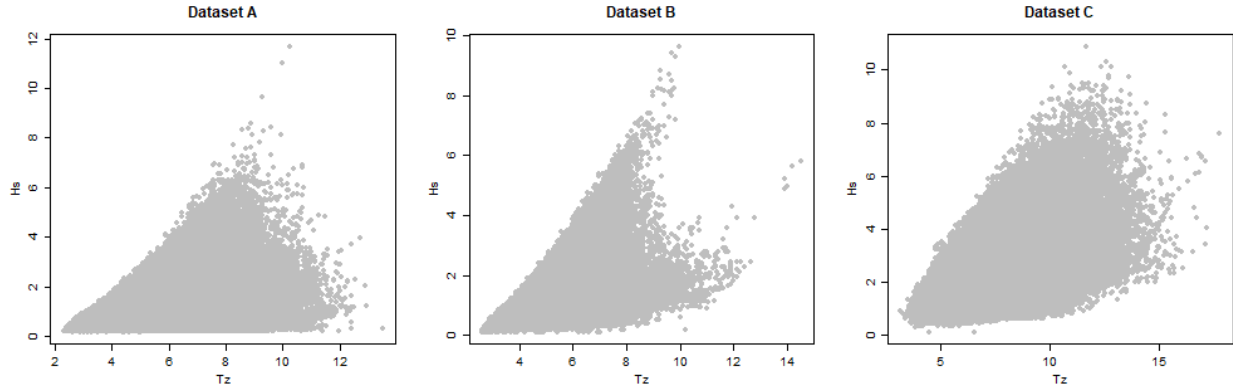


Figure 1.4: Metocean data sets A (left), B (centre) and C (right) comprised of hourly  $T_z$  and  $H_s$  observations.

tice, often misrepresenting the joint tail behaviour or not offering sufficient flexibility to capture the complex data structures. These shortcomings can have drastic consequences if fitted models are used to inform the design bases for offshore structures, as is common in practice.

This paper is structured as follows. In Section 2, we briefly introduce the SPAR model and outline our assumptions. In Section 3, we introduce a technique to estimate the density of the angular component. In Section 4, we introduce a framework for estimating the density of the radial component, conditioned on a fixed angle. In Section 5, we introduce tools for quantifying uncertainty and assessing goodness of fit when applying the SPAR model in practice. In Section 6 and 7, we apply the proposed framework to simulated and real data sets, respectively, illustrating the proposed framework can accurately capture a wide range of extremal dependence structures for both prescribed and unknown marginal distributions. We conclude in Section 8 with a discussion and outlook on future work.

## 2 The SPAR Model

### 2.1 Coordinate systems

Let  $(X, Y)$  denote a random vector in  $\mathbb{R}^2$  with continuous joint density function  $f_{X,Y}$  and simply connected support containing the point  $(0, 0)$ . The SPAR model for  $f_{X,Y}$  requires a transformation from Cartesian to polar coordinates. Polar coordinates can be defined in various ways; see Mackay and Jonathan (2023) for discussion. In this paper, we restrict attention to two particular angular-radial systems corresponding to the  $L1$  and  $L2$  norms, defined as  $\|(x, y)\|_p := (|x|^p + |y|^p)^{1/p}$ ,  $p = 1, 2$ , for  $(x, y) \in \mathbb{R}^2$ . We define  $R_p := \|(X, Y)\|_p$ ,  $p = 1, 2$ , and consider these variables as radial components of  $(X, Y)$ . Such definitions of radial variables are common in multivariate extreme value models (e.g., de Haan and de Ronde, 1998; Wadsworth et al., 2017). When using the  $L2$  norm to define the radial variable, the corresponding angular variable is usually defined as  $\Theta = \text{atan2}(X, Y)$ , where  $\text{atan2}$  is the four-quadrant inverse tan function. The map between  $(X, Y)$  and  $(R_2, \Theta)$  is bijective on  $\mathbb{R}^2 \setminus \{(0, 0)\}$ . When using the  $L1$  norm to define radii, the angular variable is typically defined as  $W := X/\|(X, Y)\|_1$  (e.g., Resnick, 1987, chapter 5). The random vector  $(R_1, W)$  has a one-to-one correspondence with  $(X, Y)$  in the upper half of the plane ( $Y \geq 0$ ), but the use of the vector  $(R_1, W)$  becomes ambiguous if we are interested in the full plane, since  $W$  contains no information about the sign of  $Y$ .

With this in mind, we follow Mackay and Jonathan (2023) and define the bijective angular functions  $\mathcal{A}_p : \mathcal{U}_p \rightarrow (-2, 2]$ , where  $\mathcal{U}_p := \{(u, v) \in \mathbb{R}^2 \mid \|(u, v)\|_p = 1\}$  is the unit circle for the  $L_p$  norm. For  $p = 1, 2$ , these are defined as

$$\begin{aligned}\mathcal{A}_1(u, v) &:= \varepsilon(v)(1 - u) \\ \mathcal{A}_2(u, v) &:= \frac{2}{\pi} \text{atan2}(u, v),\end{aligned}$$

where  $\varepsilon(v) = 1$  for  $v \geq 0$  and  $-1$  otherwise, is the generalised signum function. The functions

$\mathcal{A}_p(u, v)$  give a scaled measure of the distance along the unit circle  $\mathcal{U}_p$  from the point  $(1, 0)$  to  $(u, v)$ , measured counter-clockwise.

With angular functions established, we define the angular variables of  $(X, Y)$  to be  $Q_p := \mathcal{A}_p(X/R_p, Y/R_p)$ ,  $p = 1, 2$ . The corresponding the radial-angular mapping  $t : \mathbb{R}^2 \setminus \{(0, 0)\} \rightarrow (0, \infty) \times (-2, 2]$  given by

$$t(x, y) := \left( \|(x, y)\|_p, \mathcal{A}_p \left( \frac{x}{\|(x, y)\|_p}, \frac{y}{\|(x, y)\|_p} \right) \right),$$

is bijective for  $p = 1, 2$ . Consequently, we can recover  $(X, Y)$  from its radial and angular components, i.e.,  $(X, Y) = R_p \mathcal{A}_p^{-1}(Q_p)$  for  $p = 1, 2$ . We note that  $Q_2 = 2\Theta/\pi$ . However, we use the variable  $Q_2$  here, in preference to  $\Theta$ , so that the angular range is the same for both  $Q_1$  and  $Q_2$ . The joint density of  $(R_p, Q_p)$  can be written in terms of the joint density of  $(X, Y)$ ,

$$f_{R_1, Q_1}(r_1, q_1) = r_1 f_{X, Y}(r_1 \mathcal{A}_1^{-1}(q_1)), \quad f_{R_2, Q_2}(r_2, q_2) = \frac{\pi r_2}{2} f_{X, Y}(r_2 \mathcal{A}_2^{-1}(q_2)),$$

where the terms  $r_1$  and  $(\pi r_2)/2$  are the Jacobians of the respective transformations. For ease of notation, we henceforth drop the subscripts on the radial and angular components and simply let  $(R, Q)$  denote one of the coordinate systems, with corresponding joint density function  $f_{R, Q}$ .

Mackay and Jonathan (2023) showed that the choice of coordinate system does not affect whether the SPAR model assumptions (discussed below) are satisfied. However, the coordinates may affect the inference, so in the examples presented in Sections 6 and 7, we consider both  $L1$  and  $L2$  polar coordinates.

## 2.2 Conditional radial tail assumption

Applying Bayes theorem, the joint density  $f_{R, Q}$  can be written in the conditional form  $f_{R, Q}(r, q) = f_Q(q) f_{R_q}(r | q)$ , where  $f_Q(q)$  denotes the marginal density of  $Q$ ,  $R_q := R | (Q = q)$

and  $f_{R_q}(r | q)$  denotes the density of  $R_q$ , with corresponding distribution function  $F_{R_q}(r | q)$ . Viewed in this way, the modelling of joint extremes is reduced to the modelling of the angular density,  $f_Q$ , and the tail of the conditional density,  $f_{R_q}$ .

Given any  $\gamma \in (0, 1)$ , define  $u_\gamma : (-2, 2] \rightarrow \mathbb{R}_+$  as  $u_\gamma(q) = \inf\{r \mid F_{R_q}(r | q) \geq \gamma\}$  for all  $q \in (-2, 2]$ , implying  $\Pr(R_q \leq u_\gamma(q)) = \gamma$ . We refer to  $u_\gamma(q)$  as the threshold function henceforth. For the SPAR model, we assume that for all  $q \in (-2, 2]$ , there exists a normalising function  $c_q : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  such that

$$\Pr\left(\frac{R_q - u_\gamma(q)}{c_q(u_\gamma(q))} \leq r \mid R_q > u_\gamma(q)\right) \rightarrow 1 - \{1 + \xi(q)r\}_+^{-1/\xi(q)}, \quad r > 0, \quad (2.1)$$

as  $\gamma \rightarrow 1^-$ , with  $\xi(q) \in \mathbb{R}$ . The right hand side of equation (2.1) denotes the cumulative distribution function of a generalised Pareto (GP) distribution, and we term  $\xi(q)$  the shape parameter function. The case  $\xi(q) = 0$  can be interpreted as the limit of equation (2.1) as  $\xi(q) \rightarrow 0$ . Assumption (2.1) is equivalent to the assumption that  $R_q$  is in the domain of attraction of an extreme value distribution (Balkema and de Haan, 1974). Given the wide range of univariate distributions satisfying this assumption, it is reasonable to expect the convergence of (2.1) to hold in many cases for  $R_q$  also. Mackay and Jonathan (2023) showed that this assumption holds for a wide variety of theoretical examples.

This convergence motivates a model for the upper tail of  $R_q$ . Assuming that equation (2.1) approximately holds for some  $\gamma < 1$  close to 1, we have

$$\Pr\left(R_q - u_\gamma(q) \leq r \mid R_q > u_\gamma(q)\right) \approx F_{GP}(r \mid \tau(q), \xi(q)) := 1 - \left\{1 + \frac{\xi(q)r}{\tau(q)}\right\}_+^{-1/\xi(q)}, \quad r > 0, \quad (2.2)$$

for some  $\tau(q) \in \mathbb{R}_+$  which we refer to as the scale parameter function. The inclusion of the scale parameter removes the need to estimate the normalising function  $c_q$ , and this is equivalent to the standard peaks over threshold approximation used in univariate extreme value theory (Davison and Smith, 1990).

Given  $q \in (-2, 2]$  and  $r \geq u_\gamma(q)$ , assumption (2.2) implies that

$$\begin{aligned}\bar{F}_{R_q}(r | q) &= \Pr(R_q > u_\gamma(q)) \left[ \Pr(R_q > r \mid R_q > u_\gamma(q)) \right], \\ &\approx (1 - \gamma) \bar{F}_{GP}(r - u_\gamma(q) \mid \tau(q), \xi(q)),\end{aligned}$$

where  $\bar{F}_-(\cdot) := 1 - F_-(\cdot)$  denotes the survivor function. The joint density of  $(R, Q)$  in the region  $\mathcal{U}_\gamma := \{(r, q) \in (0, \infty) \times (-2, 2] \mid r \geq u_\gamma(q)\}$  is then given by

$$f_{R,Q}(r, q) = f_Q(q) f_{R_q}(r | q) \approx (1 - \gamma) f_Q(q) f_{GP}(r - u_\gamma(q) \mid \tau(q), \xi(q)), \quad (2.3)$$

where  $f_{GP}$  is the GP density function. Equation (2.3) implies that the SPAR model is defined within the region  $\mathcal{U}_\gamma$ .

To simplify the inference, we also assume that the functions  $f_Q(q)$ ,  $u_\gamma(q)$ ,  $\tau(q)$  and  $\xi(q)$  are finite and continuous over  $q \in (-2, 2]$  and satisfy the periodicity property  $\lim_{q \rightarrow -2^+} f(q) = f(2)$ . Such conditions are not guaranteed in general, and whether they are satisfied depends on the choice of margins (among other things). Mackay and Jonathan (2023) showed that the assumptions are valid for a wide range of copulas on Laplace margins, but using one-sided margins (e.g., exponential) or heavy-tailed margins can result in the assumptions not being satisfied for the same copulas.

### 3 Angular density estimation

In this section, we consider the angular density  $f_Q$  of equation (2.3), which we estimate using kernel density (KD) smoothing techniques. Such techniques offer many practical advantages: they are nonparametric, meaning no distributional assumptions for the underlying data are required, and they give smooth, continuous estimates of density functions. These features make KD techniques desirable for the estimation of  $f_Q$ . Note that other nonparametric smooth density estimation techniques are also available (e.g., Gu, 1993; Randell et al., 2016),

but we do not consider these here.

Unlike standard KD estimators (Chen, 2017), we require functional estimates that are periodic on the angular domain  $(-2, 2]$ , motivating the use of circular density techniques (Chaubey, 2022). Given a sample  $\{q_1, q_2, \dots, q_n\}$  from  $Q$ , the KD estimate of the density function is given by

$$\hat{f}_Q(q; h) = \frac{1}{n} \sum_{i=1}^n K_h(q, q_i),$$

where  $K_h$  denotes some circular kernel with bandwidth parameter  $h$ . The bandwidth controls the smoothness of the resulting density estimate, with the smoothness increasing as  $h \rightarrow \infty$ . The goal is typically to select  $h$  as small as possible without overfitting. Within the literature, a wide range of circular kernels have been proposed; see Chaubey (2022) for an overview. We restrict attention to one particular kernel since it is perhaps the most widely used in practice (García-Portugués, 2013). Specifically, we consider the von Mises kernel,

$$K_h(q, q_i) = \frac{1}{4I_0(1/h)} \exp \left\{ \frac{1}{h} \cos \left( (q - q_i) \frac{\pi}{2} \right) \right\}, \quad (3.4)$$

where  $I_0$  is the modified Bessel function of order zero (Taylor, 2008). Here we have modified the kernel to have support on  $(-2, 2]$ , rather than the usual support of  $(-\pi, \pi]$ .

With a kernel selected, a critical issue when applying equation (3.4) in practice is the choice of  $h$ . A variety of approaches have been proposed for automatically selecting the bandwidth parameter, including plug-in values (Taylor, 2008), cross-validation techniques (Hall et al., 1987) and bootstrapping procedures (Marzio et al., 2011).

For our modelling approach, we opt not to use automatic selection techniques for the bandwidth parameter; instead, we select  $h$  on a case-by-case basis, using the diagnostics proposed in Section 5.2 to inform selection. In unreported results, we found many of the automatic selection methods to perform poorly in practice, and it has been shown that such techniques can fail for multi-modal densities (Oliveira et al., 2012). Multi-modality is often observed within the angular density (Mackay and Jonathan, 2023), suggesting it is better

not to select  $h$  using automatic techniques.

## 4 Conditional density estimation

We now consider the conditional density of equation (2.3). For simplicity, we assume that  $\gamma \in (0, 1)$  is fixed at some high level for each  $q \in (-2, 2]$ . In practice, the choice of non-exceedance probability is non-trivial, and sensitivity analyses must be performed to ensure an appropriate value is selected; see Sections 5 and 7 for further details. Note that this is directly analogous to the threshold selection problem in univariate analyses; see Murphy et al. (2023) for a recent overview.

To apply equation (2.2), we require estimates of the threshold and GP parameter functions, denoted  $u_\gamma(q)$ ,  $\tau(q)$  and  $\xi(q)$  respectively. As noted in Section 1.1, this is equivalent to performing a non-stationary peaks over threshold analysis on the conditional radial variable  $R_q$ , with  $q$  viewed as a covariate.

Throughout this article, we let  $(\mathbf{r}, \mathbf{q}) := \{(r_i, q_i) \mid i = 1, 2, \dots, n\}$  denote a sample of size  $n \in \mathbb{N}$  from  $(R, Q)$ . In this section we introduce two methods for inference. The first approach assumes the conditional radial distribution is locally stationary over a small angular range. In the second approach, spline-based modelling techniques are used to estimate the threshold and parameter functions as smoothly-varying functions of angle. The local stationary inference is used as a precursor to the spline-based inference, providing a useful comparison and ‘sense check’ on results.

### 4.1 Local stationary inference

We compute local stationary estimates at a fixed grid of values  $\mathcal{Q}_{grid} := \{-2 + 4i/M \mid i = 1, 2, \dots, M\} \subset (-2, 2]$ , where  $M$  denotes some large positive integer, selected to ensure  $\mathcal{Q}_{grid}$  has sufficient coverage on  $(-2, 2]$ . For each  $q \in \mathcal{Q}_{grid}$ , we assume there exists a local neighbourhood  $\mathcal{Q}_q = [q - \delta, q + \delta]$ ,  $\delta > 0$ , such that the distribution of  $R_{q^*}$  is stationary for



$q^* \in \mathcal{Q}_q$ . This is true in the limit as  $\delta \rightarrow 0$ , and a reasonable approximation for small  $\delta$ .

In practice, rather than fixing the size of the interval, we select the  $N$  nearest observations in terms of the angular distance from  $q$ , defined as  $d(q)_i := \min\{|q_i - q|, 4 - |q_i - q|\}$ ,  $i = 1, \dots, n$ , for some value  $N \ll n$ . Define  $\mathcal{I}_q \subset \{1, 2, \dots, n\}$  to be the index set of the  $N$  smallest order statistics of  $d(q)_i$ . Local estimates of the threshold and parameter functions can be obtained from the corresponding radial set  $\mathcal{R}_q := \{r_i \mid i \in \mathcal{I}_q\}$ . Specifically, we define  $\hat{u}_\gamma^l(q)$  to be the  $\gamma$  empirical quantile of  $\mathcal{R}_q$ , and  $\hat{\sigma}^l(q)$  and  $\hat{\xi}^l(q)$  to be maximum likelihood estimates of the GP distribution parameters obtained from the set  $\{r_i - \hat{u}_\gamma^l(q_i) \mid i \in \mathcal{I}_q, r_i > \hat{u}_\gamma^l(q_i)\}$ . Choosing an appropriate value for  $N$  involves a bias-variance trade-off; selecting too large (small) a value too will increase the bias (variability) of the resulting pointwise threshold and parameter estimates. For our modelling procedure, this selection is not crucial, since local estimates are merely used as a means to inform the smooth estimation procedure presented in Section 4.2.

## 4.2 Smooth inference for the SPAR model

We now consider smooth estimation of the threshold and parameter functions. For this, we employ the approach of Youngman (2019), in which GAMs are used to capture covariate relationships; software for this approach is given in Youngman (2020). Our procedure is two-fold; we first estimate the threshold function  $u_\gamma(q)$  for a given  $\gamma$ , then estimate the parameter functions  $\tau(q)$  and  $\xi(q)$  using the resulting threshold exceedances.

This section is structured as follows. First, we provide a high-level overview of GAM-based modelling techniques. We then introduce procedures for estimating the threshold and parameter functions via the GAM framework. Finally, we discuss the selection of the basis dimensions required for the GAM formulations.

### 4.2.1 GAM-based procedures

GAMs are a flexible class of regression models that allow for complex, non-linear relationships between response and predictor variables. They extend the traditional linear regression

model by allowing the response to be modelled as a sum of smooth basis functions of the predictor variables. They are particularly useful when the relationship between the response and predictor variables is complex in nature and cannot be easily captured using standard parametric regression techniques.

Employing the GAM framework, the threshold and parameter functions can be represented through a sum of smooth basis functions, or splines. For an arbitrary function  $g : (-2, 2] \rightarrow \mathbb{R}$ , we write

$$g(q) = \beta_0 + \sum_{j=1}^k B_j(q)\beta_j, \quad (4.5)$$

where  $B_j$ ,  $j \in \{1, 2, \dots, k\}$  denote smooth basis functions,  $\beta_j$ ,  $j \in \{0, 1, \dots, k\}$  denote coefficients and  $k \in \mathbb{N}$  denotes the basis dimension. To apply equation (4.5) in practice, one must first select a family of basis functions  $B_j$ ,  $j \in \{1, 2, \dots, k\}$ . A wide variety of bases have been proposed in the literature; see [Perperoglou et al. \(2019\)](#) for an overview. We restrict attention to one particular type of basis function known as a cubic spline. Cubic splines are widely used in practice to capture non-linear relationships, and exhibit many desirable properties, such as optimality in various respects, continuity and smoothness ([Wood, 2017](#)). Moreover, cubic splines can be modified to ensure periodicity by imposing conditions on the coefficients, resulting in a cyclic cubic spline. In the context of the SPAR framework, these properties are desirable to ensure the estimated threshold and parameters functions are smooth and continuous, and that they satisfy periodicity on the interval  $(-2, 2]$ .

With basis functions selected, an important consideration is the basis dimension size  $k$ ; this corresponds to the number of knots of the spline function. This selection represents a trade-off, since selecting too many knots will result in higher computational burden and parameter variance, while selecting too few will not offer sufficient flexibility for capturing non-linear relationships. We consider this trade-off in detail in [Section 4.2.3](#).

Given  $k$ , the next step is to determine the knot locations; these are points where spline sections join. The knots should be more closely spaced in regions where more observations are available. In our case, we define knots at empirical quantiles of the angular variable  $Q$

corresponding to a set of equally spaced probability levels; this is typical for spline based modelling procedures.

With a GAM formulated, the final step is to estimate the spline coefficients  $\beta_j$ ,  $j \in \{0, 1, \dots, k\}$ . Various methods have been proposed for this estimation (Wood, 2017). We have opted to use the restricted maximum likelihood (REML) approach of Wood et al. (2016). For this technique, the log-likelihood function is penalised in a manner that avoids over-fitting, and cross-validation is used to automatically select the corresponding penalty parameters. Estimation via REML avoids the use of MCMC, which can be computationally expensive in practice; see Wood (2017) for further details.

#### 4.2.2 Estimation of the threshold and GP parameter functions

Estimation of the threshold function  $u_\gamma(q)$  is equivalent to estimating quantiles of  $R_q$  over  $q \in (-2, 2]$ , motivating the use of quantile regression techniques. Employing the GAM framework with  $g_{u_\gamma}$  defined as in equation 4.5, we set  $g_{u_\gamma}(q) := \log(u_\gamma(q))$ , so that  $u_\gamma(q) = \exp(g_{u_\gamma}(q)) > 0$ . We then employ the approach of Youngman (2019), whereby a misspecified asymmetric Laplace model is assumed for  $R_q$ , and REML is used to estimate the coefficients associated with  $g_{u_\gamma}$ . By altering the pinball loss function typically used in quantile regression procedures (Koenker et al., 2017), this approach avoids computational issues that can often arise within such procedures; see the Supplementary Material for further details.

Similar to  $u_\gamma$ , we define  $g_\tau(q) = \log(\tau(q))$  and  $g_\xi(q) = \xi(q)$ , with  $g_\tau$ ,  $g_\xi$  defined as in equation (4.5). Again applying the approach of Youngman (2019), we estimate the coefficients associated with  $g_\tau$  and  $g_\xi$  using REML. Further details about this estimation procedure can be found in the Supplementary Material.

#### 4.2.3 Selecting basis dimensions

An important consideration when specifying the GAM forms for both the threshold and parameter functions is the basis dimension. Selecting an appropriate dimension is essential

for ensuring accuracy and flexibility in spline modelling procedures (Wand, 2000; Perperoglou et al., 2019). Generally speaking, selecting too few knots may result in functional estimates that do not capture the underlying covariate relationships, while parameter variance increases for larger dimensions.

While some approaches have been proposed for automatic dimension selection (e.g., Kauermann and Opsomer, 2011), most available spline based modelling procedures select the dimension on a case-by-case basis using practical considerations. Moreover, as long as the basis dimension is sufficiently large enough, the resulting modelling procedure should be insensitive to the exact value, or the knot locations (Wood, 2017). This is due to the REML estimation framework, which penalises over-fitting, thus dampening the effect of adding additional knots to the spline formulation. As such, it is preferable in practice to select more knots than one believes is truly necessary to capture the covariate relationships. Therefore, we select reasonably large basis dimensions for the data sets considered in Sections 6 and 7.

## 5 Practical tools for SPAR model inference

In this section, we introduce practical tools to aid with implementation of the inference frameworks presented in Sections 3 and 4. Specifically, we introduce a tool for quantifying uncertainty in the SPAR framework, alongside diagnostic tools for assessing goodness of fit. The latter tools can also be used to inform the selection of tuning parameters, such as the non-exceedance probability  $\gamma$ , the bandwidth parameter  $h$ , and the basis dimension.

### 5.1 Evaluating uncertainty

When applying the SPAR modelling framework in practice, uncertainty will arise for each of the estimated components; namely, the angular density, threshold and parameter functions. In practice, this uncertainty is a result of sampling variability combined with model misspecification. The former arises due to finite sample sizes only partially representing the entire

population, while the latter arises from modelling frameworks not fully capturing the complex features of the data. Quantifying this uncertainty is crucial for interpreting statistical results and making informed decisions based on the inherent modelling limitations.

To quantify uncertainty in SPAR model fits, we must consider each model component in turn. For this, we take a similar approach to [Haselsteiner et al. \(2021\)](#) and [Murphy-Barltrop et al. \(2023\)](#), and consider a fixed angle  $q \in \mathcal{Q}_{grid}$ , with  $\mathcal{Q}_{grid}$  defined as in Section 4.1. We then quantify the estimation uncertainty for each model component while keeping the angle fixed. Specifically, we propose the following bootstrap procedure: for  $b = 1, \dots, B$ , where  $B \in \mathbb{N}$  denotes some large positive integer, do the following

1. Resample the original data set (with replacement) to produce a new sample of size  $n$ .
2. Compute the point estimate of the angular density at  $q$ , denoted  $\hat{f}_{Q,b}(q)$ , using the methodology described in Section 3.
3. Compute the point estimates of the threshold, scale and shape parameters at  $q$ , denoted  $\hat{u}_{\gamma,b}(q)$ ,  $\hat{\tau}_b(q)$  and  $\hat{\xi}_b(q)$  respectively, using the methodology described in Section 4.2.

We remark that the choice of resampling procedure can be adapted to incorporate data sets exhibiting temporal dependence. In this case, rather than using a standard bootstrap, one can apply a block bootstrap ([Kunsch, 1989](#)). This sampling scheme retains temporal dependence in the resampled data set, ensuring the additional uncertainty that arises due to lower effective sample sizes is accounted for ([Politis and Romano, 1994](#)). See [Keef et al. \(2013\)](#) and [Murphy-Barltrop et al. \(2023\)](#) for applications of block bootstrapping within the extremes literature.

Given a significance level  $\alpha \in (0, 1)$ , we use the outputs from the bootstrapping procedure to construct estimates of the median and  $100(1 - \alpha)\%$  confidence interval for each model component at  $q$ . Considering the angular density, for example, these quantities are computed using the set  $\{\hat{f}_{Q,b}(q) \mid b \leq k \leq B\}$ . Assuming the estimation procedure is unbiased, one

would expect  $\Pr(\hat{f}_{Q,b}^{\alpha/2}(q) \leq f_Q(q) \leq \hat{f}_{Q,b}^{1-\alpha/2}(q)) \approx (1-\alpha)$ , where  $\hat{f}_{Q,b}^{\alpha/2}(q)$  and  $\hat{f}_{Q,b}^{1-\alpha/2}(q)$  denote the empirical  $\alpha/2$  and  $1-\alpha/2$  quantile estimates from  $\{\hat{f}_{Q,b}(q) \mid 1 \leq b \leq B\}$ , respectively.

Repeating this procedure for all  $q \in \mathcal{Q}_{grid}$  allows one to evaluate uncertainty over the angular domain for each model component, thus quantifying the SPAR model uncertainty. This in turn allows us to evaluate uncertainty in quantities computed from the SPAR model, such as isodensity contours or return level sets; see Sections 6 and 7.

## 5.2 Evaluating goodness of fit

To assess goodness of fit for our inference framework, we present a range of diagnostic tools. These proposed diagnostics allow one to assess the reliability and sensitivity of fitted SPAR models. In practice, these diagnostics aid with the selection of tuning parameters, including the threshold probability  $\gamma \in (0, 1)$ , which are essential for accurately modelling joint tail behaviour. Example applications of these tools can be found in Section 7.

For the first diagnostic tool, consider a fixed  $\gamma \in (0, 1)$  and let  $\hat{u}_\gamma(q)$ ,  $\hat{\tau}_\gamma(q)$  and  $\hat{\xi}_\gamma(q)$ ,  $q \in (-2, 2]$ , denote estimates of the threshold, scale and shape functions, respectively. Furthermore, let  $\mathcal{U}_\gamma^o := \{(r_i, q_i) \mid i \in \{1, \dots, n\}, r_i > \hat{u}_\gamma(q_i)\}$  be the set of observed threshold exceedances. Applying the probability integral transform, the sample

$$\mathcal{R}_\gamma^{exp} := \left\{ -\log \left\{ 1 - F_{GP}(r - \hat{u}_\gamma(q) \mid \hat{\tau}_\gamma(q), \hat{\xi}_\gamma(q)) \right\} \mid (r, q) \in \mathcal{U}_\gamma^o \right\}$$

should be approximately standard exponentially distributed. However, this result will only hold for an appropriate choices of tuning parameters; in particular,  $\gamma \in (0, 1)$  needs to have been selected high enough such that equation (2.1) holds approximately. Letting  $r_\gamma^{(j)}$ ,  $j = 1, \dots, m$ , denote the  $j/(m+1)$  empirical quantiles of  $\mathcal{R}_\gamma^{exp}$ , with  $m := |\mathcal{R}_\gamma^{exp}|$ , we consider the following pairs on the exponential scale:

$$\left\{ \left[ r_\gamma^{(j)}, -\log \left( 1 - \frac{j}{m+1} \right) \right] \mid j = 1, \dots, m \right\}. \quad (5.6)$$

Comparing these values via a quantile-quantile (QQ) plot gives us a visual means to assess the goodness of fit for the overall model. For an appropriate choice of tuning parameters, one would expect good agreement between model and theoretical quantiles, i.e., points lying near to the  $y = x$  line in the QQ plot. Moreover, uncertainty in the QQ plot can be quantified via nonparametric bootstrapping. This tool is similar to the diagnostic proposed in Youngman (2019), and comparing QQ plots over different values of  $\gamma$  and  $k$  gives us a means to compare SPAR model fits, and thus select tuning parameters in practice.

While comparing the quantiles from equation (5.6) provides an overall summary of the SPAR model fit, this diagnostic provides no information about the relative quality in different angular regions. Specifically, we do not know for which angles  $q \in (-2, 2]$  the SPAR model approximation holds well. Therefore, we also present a localised diagnostic to assess the relative performance of the SPAR model fits in different regions. For this diagnostic, consider  $q \in \{-1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2\}$ , corresponding to a variety of regions in  $\mathbb{R}^2$ . For each  $q$ , take the local radial window  $\mathcal{R}_q$  as defined in Section 4.1. Treating  $\mathcal{R}_q$  as a sample from  $R_q$ , we compare the observed quantiles with the fitted SPAR model quantiles, resulting in a localised QQ plot. Similar to before, uncertainty can be quantified via nonparametric bootstrapping. Comparing the resulting QQ plots over different angles, and different values of  $\gamma$  and  $k$ , provides another means to assess model performance.

Finally, we propose comparing the estimated angular density, obtained using the methodology of Section 3, with the corresponding density function computed from the histogram. This comparison allows one to assess whether the choice of bandwidth parameter,  $h$ , is appropriate for a given data structure.

## 6 Simulation study

### 6.1 Study set up

In this section, we evaluate the performance of the smooth inference framework introduced in Section 4.2 via simulation. We do not consider the local estimation approach of Section 4.1 here, since our proposed local estimates are only meant as a means of assessing smooth SPAR estimates when the true values are unknown.

We consider four copulas on standard Laplace margins; Gaussian, Frank, t and Joe, as defined in the Supplementary Material. These distributions represent a range of dependence structures. Note that analogous dependence coefficients to  $\chi$  can be defined to quantify the strength of extremal dependence in other regions of the plane (see Mackay and Jonathan, 2023). In the following, we denote the four quadrants of  $\mathbb{R}^2$  as  $\mathbb{Q}_1, \dots, \mathbb{Q}_4$ . For  $\rho > 0$ , the Gaussian copula has intermediate dependence in  $\mathbb{Q}_1$  and  $\mathbb{Q}_3$  (Hua and Joe, 2011), and negative dependence in  $\mathbb{Q}_2$  and  $\mathbb{Q}_4$ . The Frank copula is AI in all quadrants, whereas the t copula is AD in all quadrants. Finally, the Joe copula is AD in  $\mathbb{Q}_1$ , negatively dependent in  $\mathbb{Q}_2$  and  $\mathbb{Q}_4$ , and AI in  $\mathbb{Q}_3$ . Samples from each copula are shown in Figure 6.1, together with the corresponding values of the copula parameters used in the simulation studies. In each case, the sample size is  $n = 10,000$ . One can observe the variety in dependence structures, as evidenced by the shape of data clouds. For the distributions considered here, the asymptotic shape parameter function is  $\xi(q) = 0$ ,  $q \in (-2, 2]$ , and the asymptotic scale parameter functions can be derived analytically (see Mackay and Jonathan, 2023). The true values of the threshold functions  $u_\gamma(q)$  and angular density functions  $f_Q(q)$  can be calculated using numerical integration. Hence, in all cases, the target values for the SPAR model parameters are known.

To evaluate performance, we simulate 500 samples from each distribution and for every sample, apply the methods outlined in Section 3 and 4.2 to estimate all SPAR model components. Using these estimates, we compute isodensity contours, defined as  $\{(x, y) \in \mathbb{R}^2 \mid$



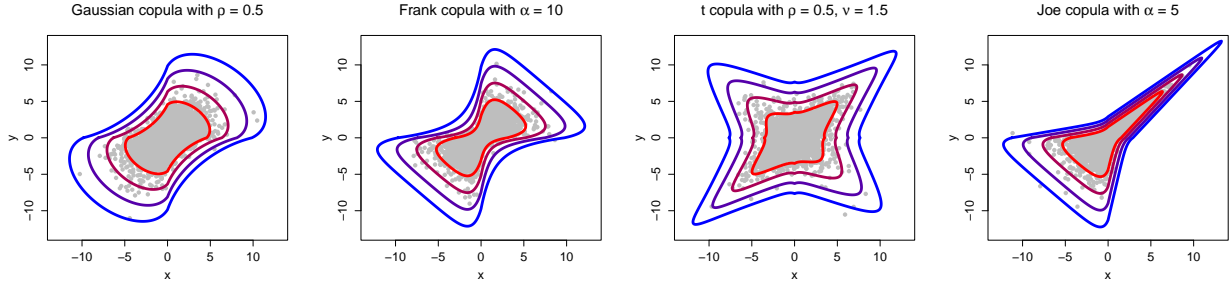


Figure 6.1: Example data sets of size  $n = 10,000$  and isodensity contours from each copula on standard Laplace margins. The red to blue lines in each plot represent the joint density levels  $p \in \{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$ .

$f_{X,Y}(x, y) = p\}$  for some  $p$ . In particular, we consider  $p \in \{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$ ; the corresponding true contours for each distribution are given in Figure 6.1. These density values represent regions of low probability mass, corresponding to joint extremal behaviour. Moreover, estimates of the joint density are appropriate for evaluating the overall performance of the SPAR model, since in practice, capturing the joint density is crucial for ensuring one can accurately extrapolate into the joint tail.

Alongside isodensity contours, we also compare the estimated GP scale parameter functions and angular density functions to their corresponding target values. For the former, we remark that for each copula, the conditional radial distribution  $f_{R_q}(r|q)$  only converges to a GP distribution in limit as  $r \rightarrow \infty$ , implying we are unlikely to accurately estimate the asymptotic GP parameter functions for finite samples; see Mackay and Jonathan (2023). We remark that even with this caveat, we still obtain high quality estimates for the isodensity contours at extreme levels.

Uncertainty in the estimation procedure is quantified by adapting the procedure of Section 5.1 across the 500 simulated samples. This allows us to compute median estimates and confidence intervals for isodensity contours, scale functions and angular density functions.

Although the choice of coordinate system does not affect whether the SPAR model assumptions are satisfied, the asymptotic SPAR parameter functions do depend on the coordinate system. Since smooth, continuous splines are used to represent the GP parameter

functions, the choice of coordinate system may affect the quality of model fits. The simulation study is therefore conducted using both  $L1$  and  $L2$  polar coordinates.

## 6.2 Tuning parameters

We first consider the tuning parameters required for the smooth inference procedure, as outlined in Section 4.2.3. For each copula, the threshold and scale parameter functions appeared to vary in a similar manner over angle. Furthermore, we fix a constant value of the shape parameter with angle, i.e.,  $\xi(q) = \xi \in \mathbb{R}$  for all  $q \in (-2, 2]$ , since for each copula, the tail behaviour remains constant over angles (Mackay and Jonathan, 2023). As discussed in Section 7, this is not true in the general case, so fixing  $\xi(q)$  to be constant is an additional constraint imposed on the model.

As noted previously, it is better to select a basis dimension  $k$  that is larger than one would expect to be necessary. We considered a range of dimensions in the interval  $[5, 50]$ , and compared the resulting model fits across each of the four copulas. From this analysis, we found that setting  $k = 25$  was sufficiently flexible for capturing the angular dependence for both the threshold and scale functions.

We are also required to select a non-exceedance probability  $\gamma \in (0, 1)$ . As observed in Mackay and Jonathan (2023), each of the four copulas exhibits a different rate of convergence to the asymptotic form. Therefore, different values of  $\gamma$  may be appropriate for these different dependence structures. However, we instead opt to keep  $\gamma$  fixed across all copulas. This is for consistency in the estimation framework, as well as to show that even in the case when the model is misspecified, the corresponding inference framework is still robust enough to approximate the true model. We considered a range of  $\gamma$  values, restricting our attention to the interval  $[0.5, 0.95]$ , and compared the resulting model fits. As expected, the performance for each copula varied non-homogeneously across  $\gamma$  values. Ultimately, we found that setting  $\gamma = 0.8$  appeared sufficient for approximating the conditional radial tails for each dependence structure. In particular, this value appeared high enough to give approximate convergence

to a GP distribution model, without giving a large degree of variability in model estimates.

Finally, for estimation of the angular density, we fix the bandwidth parameter at  $h = 1/50$  for each copula. Our results show that for these copula examples, this bandwidth is sufficiently flexible to approximate the true angular density functions across the majority of angles.

### 6.3 Results

Figure 6.2 compares the median estimates of isodensity contours, obtained using the  $L1$  coordinate system, to the true contours at a range of low density levels; the corresponding plots for the  $L2$  coordinate system are given in the Supplementary Material. For both coordinate systems, one can observe generally good agreement between the sets of contours, suggesting the modelling framework is, on average, capturing the dependence structure of each copula. Furthermore, plots comparing the median estimates from both coordinate systems can also be found in the Supplementary Material. These plots show a similar overall performance for both systems, with perhaps a slight preference for the  $L1$  estimates.

Plots comparing the estimated GP scale parameter functions, and associated confidence intervals, to the known asymptotic functions are given in the Supplementary Material. In some angular regions, the estimated isodensity contours and scale functions from the SPAR model do not agree with the known values; for example, in the region around  $q = 0.5$  for the Joe copula. In this case, there is a sharp cusp in the asymptotic GP scale parameter function. Similarly, there is a sharp cusp in the true GP scale parameter for the t copula at  $q = \pm 0.5, \pm 1.5$ . As the inference framework assumes that the scale is a smooth function of angle, this behaviour is not properly captured. Despite the GAM representation not being able to capture these cusps, the overall performance of the estimated SPAR model is still reasonable in these regions. Furthermore, there is poor agreement between the estimated and asymptotic scale functions for the Frank copula. This is likely due to the relatively slow convergence of this distribution to its asymptotic form, and hence the poorer approximation

by the GP distribution.

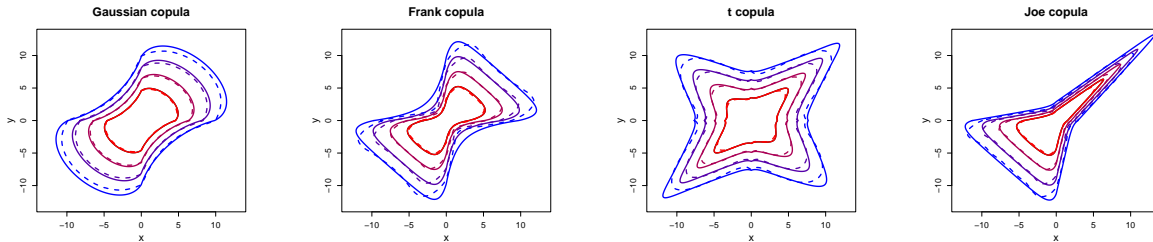


Figure 6.2: Comparison of true (thick lines) and median estimated (dashed lines) isodensity contours under the  $L1$  coordinate system. In each plot, the red to blue lines represent the joint density levels  $p \in \{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$ .

Next, we consider the uncertainty of the isodensity contours for  $p \in \{10^{-3}, 10^{-6}\}$  in the radial-angular space. Figure 6.3 shows the median contour estimates, along with estimated 95% confidence intervals, obtained under the  $L1$  coordinate system; the corresponding plots for  $L2$  coordinates are given in the Supplementary Material. One can observe that the true contours are generally well captured within the estimated uncertainty regions. The exception in the  $10^{-6}$  density contour for the Frank copula in  $\mathbb{Q}_2$  and  $\mathbb{Q}_4$ , owing to the aforementioned slow rate of convergence to the asymptotic form for this copula.

Finally, we compare median estimates of the angular density functions, alongside estimated confidence regions, to the corresponding true density functions for each copula. These results are given in the Supplementary Material. Overall, we observe close agreement between the estimated and true functions at the majority of angles. However, we note that the KD estimation framework appeared unable to fully capture the modal regions for the Frank,  $t$  and Joe copulas; see Section 8 for further discussion.

Overall, when compared to the truth, the SPAR model estimates perform well for each copula. This observation suggests that our proposed inference framework, with appropriate tuning parameters, can capture the extremal dependence structure across a range of copulas with differing dependence classes. This illustrates both the flexibility and robustness of the SPAR approach, and its advantages over many alternative multivariate models.

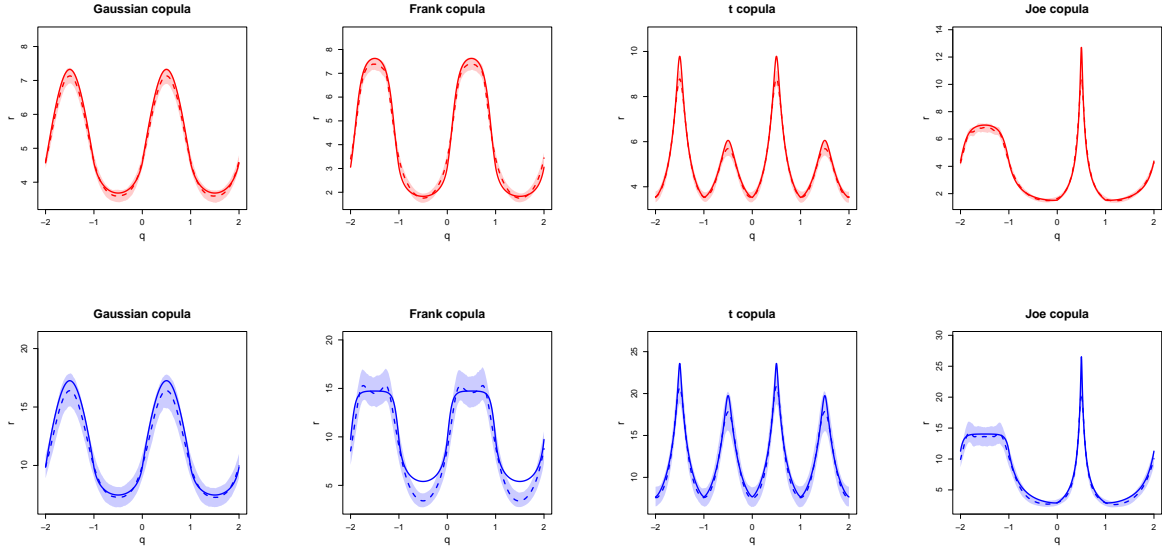


Figure 6.3: Comparison of median estimated isodensity contours (dashed lines), with 95% confidence intervals (shaded region), to true contours (solid lines) for joint density level  $p = 10^{-3}$  (top row) and  $p = 10^{-6}$  (bottom row), with estimates obtained using  $L1$  polar coordinates.

## 7 Case Study

In this section, we apply the techniques introduced in Sections 3 and 4 to the data sets A, B and C introduced in Section 1.2. We show that the resulting model fits are physically plausible and capture the complex dependence features of each data set. We also apply the tools introduced in Section 5 to quantify uncertainty and assess goodness of fit; the resulting diagnostics indicate generally good performance.

### 7.1 Pre-processing

The simulation study in Section 6 considered data on standard Laplace margins. This is because the SPAR model assumptions are satisfied for a wide range of copulas on Laplace margins, and the resulting asymptotic representations are relatively simple (Mackay and Jonathan, 2023). However, the SPAR framework does not pre-suppose any particular choice of margins, and Mackay and Jonathan (2023) also showed that SPAR representations arise for random vectors with bounded and heavy tailed margins.

For the case of the metocean datasets, the margins are unknown. In practice, estimation of the marginal distributions, as is common in many extreme value analyses, introduces a high degree of additional modelling uncertainty, and poor marginal estimates affect the quality of the resulting multivariate inference (Towe et al., 2023). Therefore, we opt not to transform the margins of the metocean time series, and to instead fit the SPAR model on the original scale of the data. With suitable selections of tuning parameters, we demonstrate below that our inference framework is flexible enough to capture the observed extremal dependence structures for the metocean data sets without the need for marginal transformation.

An important consideration for the SPAR model is where to place the origin of the polar coordinate system. When using Laplace margins, a natural choice is to locate the origin at  $(x, y) = (0, 0)$ . When working on the original scale of the data, the choice is less clear. One option would be to place the polar origin at  $(x, y) = (0, 0)$ . However, this would restrict the range of angles for which the SPAR model offers a useful representation, since both variables we are considering here take only positive values. To account for this, we normalise the data to have zero mean and unit variance, and select the polar origin at  $(x, y) = (0, 0)$  in the normalised variable space. Define normalised variables  $(\tilde{T}_z, \tilde{H}_s) := ((T_z - \mu_{T_z})/\sigma_{T_z}, (H_s - \mu_{H_s})/\sigma_{H_s})$ , where  $(\mu_{T_z}, \mu_{H_s})$  and  $(\sigma_{T_z}, \sigma_{H_s})$  denote the estimated means and standard deviations of  $(T_z, H_s)$ , respectively.

We henceforth assume that for each data set, the normalised joint density function,  $f_{\tilde{T}_z, \tilde{H}_s}$ , satisfies the assumptions of the SPAR model, and apply the statistical techniques introduced in Sections 3, 4 and 5; these results are presented in Section 7.3. Throughout this section, we present all results for the  $L1$  coordinate system; the corresponding results for  $L2$  coordinates are given in the Supplementary Material, with both systems resulting in similar model fits.

We remark that each metocean time series exhibits non-negligible temporal dependence. Following Section 5.1, we apply block bootstrapping throughout this section whenever quantifying uncertainty, with the block size set to 4 days. This block size appeared appropriate to account for the observed dependence in each time series. Note that temporal dependence

could alternatively be accounted for by ‘declustering’ the data and only modelling peak values. However, in multivariate applications, what constitutes a ‘peak value’ is ambiguous since the extremes of each variable do not necessarily occur simultaneously; see Mackay et al. (2021) and Mackay and de Hauteclocque (2023) for further discussion.

## 7.2 Tuning parameters

Prior to inference, we must first select each of the relevant tuning parameters for the methodologies discussed in Sections 3 and 4. Since the true dependence features are unknown, we use local estimates of the SPAR model, obtained using the framework of Section 4.1, to inform the choice of basis dimensions for the smooth inference procedure of Section 4.2.

To obtain local estimates, we are required to specify the number of reference angles  $M$ , the number of order statistics  $N$ , and the non-exceedance probability  $\gamma$ . For the first two values, we set  $M = 200$  and  $N = 500$ ; we found these values to be adequate to ensure a high degree of coverage over the angular interval  $(-2, 2]$ , and to give angular windows that appeared approximately stationary. For selecting  $\gamma$ , we tested a range of probabilities in the interval  $[0.5, 0.95]$ . Through this testing, we set  $\gamma = 0.7$ , since this value appeared to give approximate convergence to a GP tail across the majority of local windows. The same non-exceedance probability is also used for the smooth SPAR model estimates.

With local estimates obtained, we then consider smooth estimation of the SPAR components. Notably, for each of the time series, we observe clear trends in the locally estimated shape parameter function; it would therefore not be appropriate to specify this parameter as constant. This makes sense when one considers the shapes of the data clouds illustrated in Figure 1.4; the radial behaviour varies significantly over angles. Furthermore, both metocean variables are bounded below by 0; therefore, we would expect shorter tails in angular directions that intersect the axes. A range of basis dimensions were tested, and from this analysis, we fixed  $k = 35$  for the threshold and scale functions, and  $k = 12$  for the shape function. These values appeared to offer adequate flexibility for capturing the trends observed over the

angular variable.

Finally, for angular density estimation, we follow Section 6.2 and set the bandwidth  $h = 1/50$ . This value offered sufficient flexibility to capture the observed angular distributions.

### 7.3 Results

Figure 7.1 compares the threshold and parameter function estimates for data set B from the local and smooth inference procedures; the corresponding plots for data sets A and C are given in the Supplementary Material. The shaded regions in this figure denote the 95% bootstrapped confidence intervals for the smooth model fits. One can observe generally good agreement for each component of the SPAR model. We remark that the local estimates appear unstable, and hence unreliable, for certain angles; this is not surprising, given the small sample size of the angular window. However, the general overall agreement suggests the smooth SPAR estimates are accurately capturing the observed dependence features for each data set, providing evidence that the chosen tuning parameters are appropriate.

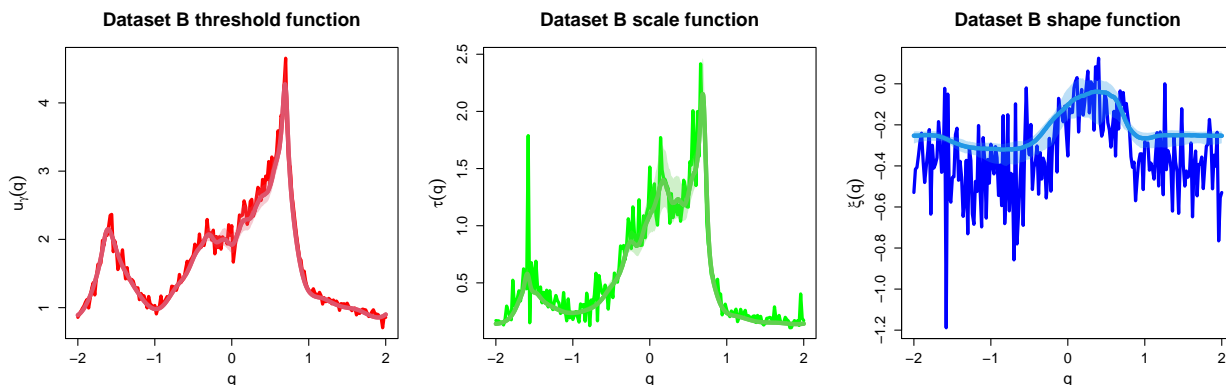


Figure 7.1: Comparison of estimated local (rough lines) and smooth threshold (red, left), scale (green, middle) and shape (blue, right) functions for data set B with the  $L_1$  coordinate system, with shaded regions denoting 95% confidence intervals.

Following Section 5.2, we compare the median angular density functions from the KD estimation technique with empirical histograms. These comparisons are given in Figure 7.2 for each data set, and one can observe good agreement between the estimated quantities.



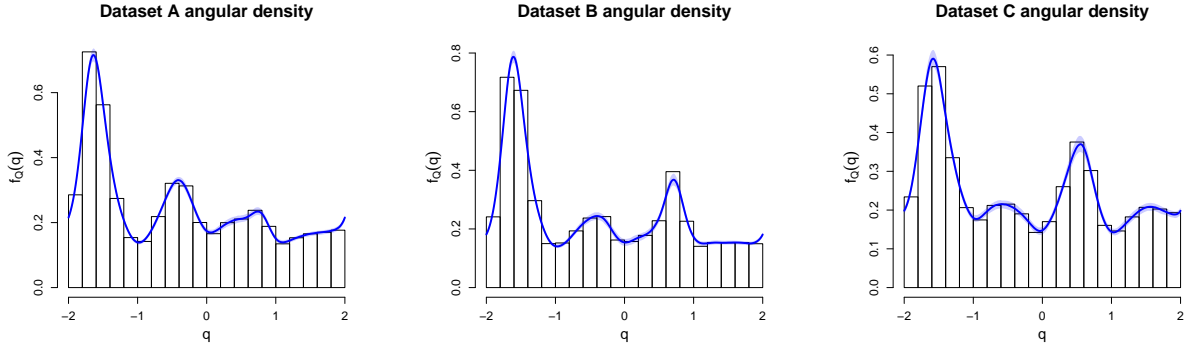


Figure 7.2: Comparison of estimated median (blue lines) angular density functions to histograms for data sets A (left), B (centre), and C (right) with the  $L_1$  coordinate system. The shaded regions in each plot denote the estimated 95% confidence intervals.

Estimates of isodensity contours are shown in Figure 7.3. These joint density contours are given on the original scale of the data, rather than on the normalised scale, and we consider the joint density levels  $p \in \{10^{-3}, 10^{-6}\}$ , corresponding to regions of low probability mass. The estimated isodensity contours appear to capture the shape and structure of each data cloud well. Furthermore, we note that the SPAR model appears able to capture the observed asymmetric dependence structures, illustrating the flexibility and robustness of this modelling framework.

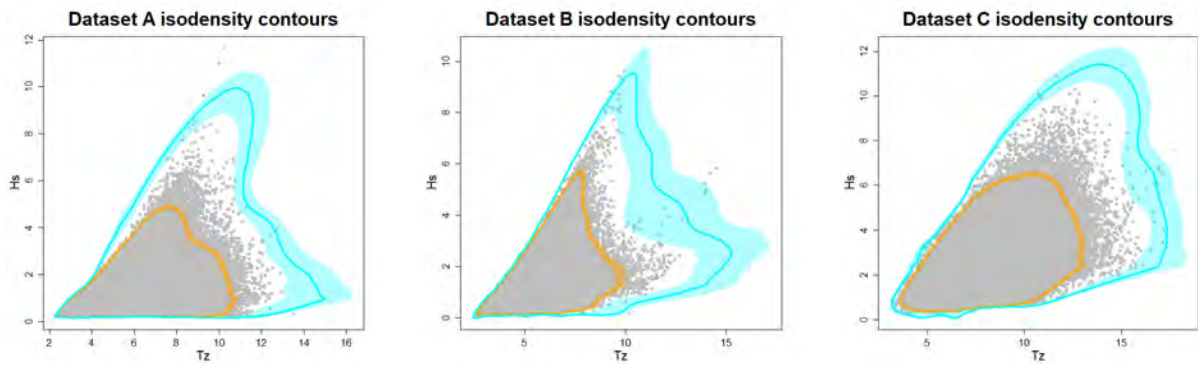


Figure 7.3: Estimated median isodensity contours at  $p = 10^{-3}$  (orange lines) and  $p = 10^{-6}$  (cyan lines) for data sets A (left), B (centre), and C (right) with the  $L_1$  coordinate system. The shaded region for each contour denote the 95% bootstrapped confidence intervals.

To further demonstrate the utility of the SPAR framework, we also use the fitted model to obtain return level sets for each of the data sets. Return level sets are commonly used in

ocean engineering for the design of offshore structures. They can be defined in various ways, but are generally defined in terms of marginal probabilities under various rotations of the coordinate axes, or in terms of the probability of an observation falling anywhere outside the set (the so-called ‘total exceedance probability’ of a contour); see Mackay and Haselsteiner (2021). Papastathopoulos et al. (2023) noted that SPAR-type models offer a natural way for total exceedance probability contours to be constructed. For contour exceedance probability  $1 - \gamma > a \in [0, 1]$ , the radius of the contour at angle  $q$  is the  $(1 - a - \gamma)/(1 - \gamma)$  quantile of the GP distribution with parameter vector  $(u_\gamma(q), \xi(q), \tau(q))$ . For any angle  $q \in (-2, 2]$ , the probability of an observation exceeding this radius is equal to  $a$ ; consequently, the probability of observing data outside of the resulting contour set is equal to  $a$ . When observations are independent and the distribution is stationary, we can define such sets in terms of return periods; given a number of years  $K \in \mathbb{N}$ , the  $K$ -year return level set is the set corresponding to the probability  $a := 1/n_y K$ , where  $n_y$  denotes the number of observations per year. One would expect to observe data points outside of the return level set once, on average, every  $K$  years. Given the temporal dependence observed within the metocean data sets, we note that such an interpretation is not possible due to clustering of extreme events, and as such these estimates are conservative (Mackay et al., 2021). However, the resulting return level sets can still provide a useful summary of joint extreme behaviour.

Plots of estimated median 10 year return level sets for each data set are given in Figure 7.4, along with 95% bootstrapped confidence intervals. These sets, obtained by computing GP distribution quantiles from the fitted model, appear sensible in shape and structure when compared to the data cloud. Moreover, given the lengths of observation windows of each data set, we would not expect to observe many datapoints outside of the return level set; this is clearly true in every case. Furthermore, a comparison of return level sets from the two coordinate systems is given in the Supplementary Material, where one can observe generally good agreement between the estimated sets

We note that simulation from the SPAR model is straightforward; a simulation scheme is

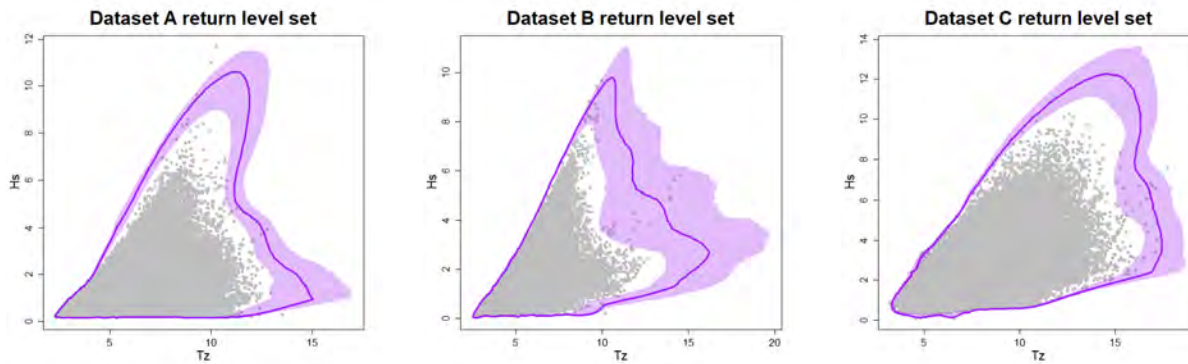


Figure 7.4: Estimated median 10 year return level sets (purple lines) for data sets A (left), B (centre), and C (right) with the  $L1$  coordinate system. The shaded region for each return level set denotes the 95% bootstrapped confidence region.

given in the Supplementary Material, alongside examples for each of the metocean data sets.

Finally, to assess goodness of fit, we apply the remaining diagnostic tools presented in Section 5.2. The resulting QQ plots are given in the Supplementary Material. Note that the ‘overall’ QQ plots (i.e., for all threshold exceedances) indicate generally good performance, albeit with some overestimation for the most extreme exceedances. Considering the local window QQ plots, it is clear that the performance of the fitted SPAR model varies across different angular regions. Although we observe generally good agreement between quantiles, there is clearly better agreement for certain angles, suggesting rates of convergence to the GP tail model may vary over angle for these data sets.

## 8 Discussion

In this paper, we have introduced a novel inference framework for the SPAR model of Mackay and Jonathan (2023). We have explored the properties of this framework, and introduced practical tools for quantifying uncertainty and assessing goodness of fit. Furthermore, we have applied this framework to simulated and real data sets in Sections 6 and 7, with results indicating that the proposed framework captures joint tail behaviour across a wide range of data structures. Moreover, this framework has been recently applied in Mackay et al. (2024), where the authors show the SPAR model can accurately capture joint extremes of wind speeds

and wave heights, and extreme response distributions for a variety of metocean data sets. Our proposed modelling framework is one of the first multivariate extreme value modelling techniques that can be applied without marginal transformation, offering an advantage over competing approaches and removing a significant degree of model variability.

Noting that the SPAR model has only been developed recently, this work is the first attempt to apply this modelling framework in practice, and it is likely that other inference approaches will follow. While our proposed framework performs well in general, we acknowledge there exist some shortcomings that could provide the motivation for future work.

The results from Section 6 indicate that the proposed angular density estimation framework from Section 3 performs poorly for some copulas in regions around the angular mode(s). However, we note that even with this caveat, the KD estimation framework appeared adequate for capturing the angular distribution for the observed data sets in Section 7. Future work could explore whether using alternative angular density estimation approaches (e.g., Gu, 1993; Randell et al., 2016) could further improve performance.

Observe that for the AD copulas considered in Section 6, the true isodensity functions exhibit clear cusps, where the underlying GP scale function is non-differentiable. Such sections cannot be captured under the current framework, since the use of cyclic cubic splines for smooth estimation imposes differentiability at all angles. Future work could explore how such behaviour could be captured in the inference framework. For example, one could use a spline representation that allows for superimposed knots. Combined with a more general spline inference procedure, this alternative representation could allow for optimal estimation of both the number and locations of knots, while simultaneously giving cusps in the estimated SPAR functions (Hastie et al., 2009).

From Section 7, one can observe that for the estimated isodensity contours and return levels obtained using the  $L1$  coordinate system, there exist distinct cusps at certain angles; these arise due to the square shape of  $\mathcal{U}_1$ . We acknowledge that such cusps are not realistic for practical applications, and consequently, estimates from the  $L2$  coordinate system may

be preferable in such settings.

When estimating the shape parameter functions in Section 7, we did not impose any functional constraints, even though the variables considered must have finite lower and upper bounds, and hence cannot be in the domain of attraction of a GP distribution with a non-negative shape parameter. However, even without bounding the shape parameter, we note that across all of our model fits, the estimated shape functions were almost always homogeneously negative, indicating the proposed framework is flexible enough to detect the form of tail behaviour directly from the data. Future work could explore whether imposing physical constraints on the shape function improves the quality of model fits.

Following on from Section 7.3, it appears that having one non-exceedance probability for all angles may not be optimal for fitting the SPAR model in practice due to different rates of convergence at different angles. Exploring techniques for selecting and estimating threshold functions with varying rates of exceedance (e.g., Northrop and Jonathan, 2011) remains an open area for future work.

We have restricted attention to the bivariate setting throughout this work. This decision was motivated by parsimony, as well as the fact many of the examples given Mackay and Jonathan (2023) are for bivariate vectors. A natural avenue for future work would therefore be expanding the proposed inference framework to the general  $d$ -dimensional setting.

Finally, we note that in Mackay and Jonathan (2023), the authors also derive a link between the SPAR model and the limit set representation for multivariate extremes. Specifically, the radius of the limit set at a fixed angle is given by the asymptotic shape parameter of the SPAR representation. We believe the inference approach we have proposed could be adapted for the estimation of limit sets, though additional care will be required given estimates obtained from finite sample sizes seldom equal limiting asymptotic quantities in practice.

## Supplementary Material

- **Supplementary Material for “Inference for multivariate extremes via a semi-parametric angular-radial model”**: File containing supporting figures and additional information about the REML procedures and simulation study. (.pdf file)

## Declarations

### Ethical Approval

Not Applicable

### Availability of supporting data

The data sets analysed in Section 7 are freely available online at <https://github.com/EC-BENCHMARK-ORGANIZERS/EC-BENCHMARK>.

### Competing interests

The authors have no relevant financial or non-financial interests to disclose.

### Funding

This work was supported by EPSRC grant numbers EP/L015692/1 and EP/Y016297/1.

### Authors' contributions

All authors contributed to the study conception and design. Model development was performed by Ed Mackay and Phil Jonathan. Material preparation and initial analysis were performed by Callum Murphy-Barltrop. The first draft of the manuscript was written by Callum Murphy-Barltrop and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

## Acknowledgments

This paper is based on work partly completed while Callum Murphy-Barltrop was part of the EPSRC funded STOR-i centre for doctoral training (EP/L015692/1). Ed Mackay was funded by the EPSRC Supergen Offshore Renewable Energy Hub, United Kingdom (EP/Y016297/1). We would like to thank Ben Youngman for his assistance with the `evgam` package in the R computing language.

## References

- Balkema, A. A. and de Haan, L. (1974). Residual Life Time at Great Age. *The Annals of Probability*, 2:792–804.
- Beirlant, J., Goegebeur, Y., Teugels, J., and Segers, J. (2004). *Statistics of Extremes*. Wiley.
- Castro-Camilo, D., de Carvalho, M., and Wadsworth, J. (2018). Time-varying extreme value dependence with application to leading European stock markets. *Annals of Applied Statistics*, 12:283–309.
- Chan, Y. and Li, H. (2008). Tail dependence for multivariate t-copulas and its monotonicity. *Insurance: Mathematics and Economics*, 42(2):763–770.
- Chaubey, Y. P. (2022). *Directional Statistics for Innovative Applications*, pages 351–378. Springer.
- Chavez-Demoulin, V. and Davison, A. C. (2005). Generalized additive modelling of sample extremes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54:207–222.
- Chen, Y.-C. (2017). A tutorial on kernel density estimation and recent advances. *Biostatistics & Epidemiology*, 1:161–187.

- Davis, R. A., Mulrow, E., and Resnick, S. I. (1988). Almost sure limit sets of random samples in  $\mathbb{R}^d$ . *Advances in Applied Probability*, 20(3):573–599.
- Davison, A. C. and Smith, R. L. (1990). Models for Exceedances Over High Thresholds. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 52:393–425.
- de Haan, L. and de Ronde, J. (1998). Sea and Wind: Multivariate Extremes at Work. *Extremes*, 1:7–45.
- de Haan, L. and Ferreira, A. (2006). *Extreme value theory: an introduction*, volume 3. Springer.
- de Haan, L. and Resnick, S. I. (1987). On regular variation of probability densities. *Stochastic processes and their applications*, 25:83–93.
- Diaconu, D. C., Costache, R., and Popa, M. C. (2021). An Overview of Flood Risk Analysis Methods. *Water*, 13(4).
- García-Portugués, E. (2013). Exact risk improvement of bandwidth selectors for kernel density estimation with directional data. *Electronic Journal of Statistics*, 7:1655–1685.
- Geraci, M. and Bottai, M. (2007). Quantile regression for longitudinal data using the asymmetric Laplace distribution. *Biostatistics*, 8:140–154.
- Gu, C. (1993). Smoothing Spline Density Estimation: A Dimensionless Automatic Algorithm. *Journal of the American Statistical Association*, 88:495–504.
- Hall, P., Watson, G. S., and Cabrera, J. (1987). Kernel Density Estimation with Spherical Data. *Biometrika*, 74:751.
- Haselsteiner, A. F., Coe, R. G., Manuel, L., Chai, W., Leira, B., Clarindo, G., Soares, C. G., Ásta Hannesdóttir, Dimitrov, N., Sander, A., Ohlendorf, J. H., Thoben, K. D., de Hauteclocque, G., Mackay, E., Jonathan, P., Qiao, C., Myers, A., Rode, A., Hildebrandt,



- A., Schmidt, B., Vanem, E., and Huseby, A. B. (2021). A benchmarking exercise for environmental contours. *Ocean Engineering*, 236:1–29.
- Hastie, T., Tibshirani, R., Friedman, J. H., and Friedman, J. H. (2009). *The elements of statistical learning: data mining, inference, and prediction*, volume 2. Springer.
- Hosking, J. R. and Wallis, J. R. (1987). Parameter and quantile estimation for the generalized Pareto distribution. *Technometrics*, 29:339–349.
- Hua, L. and Joe, H. (2011). Tail order and intermediate tail dependence of multivariate copulas. *Journal of Multivariate Analysis*, 102(10):1454–1471.
- Joe, H. (1997). *Multivariate Models and Multivariate Dependence Concepts*. Chapman and Hall/CRC.
- Jonathan, P. and Ewans, K. (2013). Statistical modelling of extreme ocean environments for marine design: A review. *Ocean Engineering*, 62:91–109.
- Kauermann, G. and Opsomer, J. D. (2011). Data-driven selection of the spline dimension in penalized spline regression. *Biometrika*, 98:225–230.
- Keef, C., Tawn, J. A., and Lamb, R. (2013). Estimating the probability of widespread flood events. *Environmetrics*, 24:13–21.
- Koenker, R., Chernozhukov, V., He, X., and Peng, L. (2017). *Handbook of Quantile Regression*. Chapman and Hall/CRC.
- Kunsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *The Annals of Statistics*, 17:1217–1241.
- Ledford, A. W. and Tawn, J. A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187.

- Ledford, A. W. and Tawn, J. A. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 59:475–499.
- Mackay, E. (2022). Improved Models for Multivariate Metocean Extremes. Technical report, Supergen ORE Hub.
- Mackay, E. and de Hauteclocque, G. (2023). Model-free environmental contours in higher dimensions. *Ocean Engineering*, 273:113959.
- Mackay, E., de Hauteclocque, G., Vanem, E., and Jonathan, P. (2021). The effect of serial correlation in environmental conditions on estimates of extreme events. *Ocean Engineering*, 242:110092.
- Mackay, E. and Haselsteiner, A. F. (2021). Marginal and total exceedance probabilities of environmental contours. *Marine Structures*, 75:1–24.
- Mackay, E. and Jonathan, P. (2023). Modelling multivariate extremes through angular-radial decomposition of the density function. *arXiv*, 2310.12711.
- Mackay, E., Murphy-Barltrop, C., and Jonathan, P. (2024). The SPAR model: a new paradigm for multivariate extremes. Application to joint distributions of metocean variables. In *43rd International Conference on Ocean, Offshore & Arctic Engineering*, page OMAE2024/130932, Singapore.
- Majumder, R., Shaby, B. A., Reich, B. J., and Cooley, D. (2023). Semiparametric Estimation of the Shape of the Limiting Multivariate Point Cloud. *arXiv*, 2306.13257.
- Marzio, M. D., Panzera, A., and Taylor, C. C. (2011). Kernel density estimation on the torus. *Journal of Statistical Planning and Inference*, 141:2156–2173.
- Murphy, C., Tawn, J. A., and Varty, Z. (2023). Automated threshold selection and associated inference uncertainty for univariate extremes. *arXiv*, 2310.17999.

- Murphy-Barltrop, C. J. R., Wadsworth, J. L., and Eastoe, E. F. (2023). New estimation methods for extremal bivariate return curves. *Environmetrics*, e2797:1–22.
- Nolde, N. and Wadsworth, J. L. (2022). Linking representations for multivariate extremes via a limit set. *Advances in Applied Probability*, 54:688–717.
- Northrop, P. J. and Jonathan, P. (2011). Threshold modelling of spatially dependent non-stationary extremes with application to hurricane-induced wave heights. *Environmetrics*, 22:799–809.
- Oh, H.-S., Lee, T. C. M., and Nychka, D. W. (2011). Fast Nonparametric Quantile Regression With Arbitrary Smoothing Methods. *Journal of Computational and Graphical Statistics*, 20:510–526.
- Oliveira, M., Crujeiras, R., and Rodríguez-Casal, A. (2012). A plug-in rule for bandwidth selection in circular density estimation. *Computational Statistics & Data Analysis*, 56:3898–3908.
- Papastathopoulos, I., de Monte, L., Campbell, R., and Rue, H. (2023). Statistical inference for radially-stable generalized Pareto distributions and return level-sets in geometric extremes. *arXiv*, 2310.06130.
- Perperoglou, A., Sauerbrei, W., Abrahamowicz, M., and Schmid, M. (2019). A review of spline function procedures in R. *BMC Medical Research Methodology*, 19:1–16.
- Politis, D. N. and Romano, J. P. (1994). The Stationary Bootstrap. *Journal of the American Statistical Association*, 89:1303–1313.
- Randell, D., Turnbull, K., Ewans, K., and Jonathan, P. (2016). Bayesian inference for nonstationary marginal extremes. *Environmetrics*, 27:439–450.
- Resnick, S. (2002). Hidden Regular Variation, Second Order Regular Variation and Asymptotic Independence. *Extremes*, 5:303–336.

- Resnick, S. I. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer New York.
- Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer.
- Simpson, E. S. and Tawn, J. A. (2022). Estimating the limiting shape of bivariate scaled sample clouds for self-consistent inference of extremal dependence properties. *arXiv*, 2207.02626.
- Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. *Publ. Inst. Statist. Univ. Paris*, 8:229–231.
- Taylor, C. C. (2008). Automatic bandwidth selection for circular density estimation. *Computational Statistics and Data Analysis*, 52:3493–3500.
- Towe, R., Randell, D., Kensler, J., Feld, G., and Jonathan, P. (2023). Estimation of associated values from conditional extreme value models. *Ocean Engineering*, 272:113808.
- Vanem, E., Zhu, T., and Babanin, A. (2022). Statistical modelling of the ocean environment – A review of recent developments in theory and applications. *Marine Structures*, 86.
- Wadsworth, J. and Campbell, R. (2023). Statistical inference for multivariate extremes via a geometric approach. *arXiv*, 2208.14951.
- Wadsworth, J. L. and Tawn, J. A. (2013). A new representation for multivariate tail probabilities. *Bernoulli*, 19:2689–2714.
- Wadsworth, J. L., Tawn, J. A., Davison, A. C., and Elton, D. M. (2017). Modelling across extremal dependence classes. *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, 79:149–175.
- Wand, M. P. (2000). A Comparison of Regression Spline Smoothing Procedures. *Computational Statistics*, 15:443–462.

- Wood, S. N. (2017). *Generalized Additive Models*. Chapman and Hall/CRC.
- Wood, S. N., Pya, N., and Säfken, B. (2016). Smoothing Parameter and Model Selection for General Smooth Models. *Journal of the American Statistical Association*, 111:1548–1563.
- Youngman, B. (2020). evgam: Generalised Additive Extreme Value Models. *R Package*.
- Youngman, B. D. (2019). Generalized Additive Models for Exceedances of High Thresholds With an Application to Return Level Estimation for U.S. Wind Gusts. *Journal of the American Statistical Association*, 114:1865–1879.
- Yu, K. and Moyeed, R. A. (2001). Bayesian quantile regression. *Statistics & Probability Letters*, 54:437–447.
- Zanini, E., Eastoe, E., Jones, M. J., Randell, D., and Jonathan, P. (2020). Flexible covariate representations for extremes. *Environmetrics*, 31:1–28.

# Supplementary Material to ‘Inference for multivariate extremes via a semi-parametric angular-radial model’

## S1 Additional information for Section 4.4.2

In this section, we provide further details about the REML schemes used to estimate the GAM parameters associated with the threshold and parameter functions. We first consider the GAM formulation of the threshold function  $\log(u_\gamma(q)) = g_{u_\gamma}(q)$ , which we estimate via quantile regression techniques. For finite sample sizes, conditional quantile regression requires us to compute

$$\hat{\beta}_{u_\gamma} := \arg \min_{\beta_{u_\gamma} \in \mathbb{R}^{k+1}} \sum_{i=1}^n (\rho_\gamma(r_i - \exp(g_{u_\gamma}(q_i)))) , \quad (\text{S1})$$

where  $\beta_{u_\gamma}$  denotes the spline coefficients associated with  $g_{u_\gamma}(q)$ , and  $\rho_\gamma(x) = (\gamma - 1)x$  for  $x < 0$  and  $\gamma x$  for  $x \geq 0$  (Koenker et al., 2017). However, optimisation of (S1) is non-trivial and computational issues often result due to the non-differentiability of  $\rho_\gamma$  at zero.

To overcome these computational issues, Youngman (2019) proposed the following misspecified model

$$(R \mid \exp(g_{u_\gamma}(q)), \sigma(q), \gamma, c, Q = q) \sim \text{ALD}(\exp(g_{u_\gamma}(q)), \sigma(q), \gamma, c),$$

with corresponding density function

$$f_{\text{ALD}}(r) = \frac{\gamma(1-\gamma)}{\sigma(q)} \exp \left\{ -\rho_{\gamma,c} \left( \frac{r - \exp(g_{u_\gamma}(q))}{\sigma(q)} \right) \right\}, \quad r \in \mathbb{R},$$

where  $\sigma(q) > 0$  and  $\rho_{\gamma,c}$  denotes the modified check function of Oh et al. (2011). This check

function is defined as

$$\rho_{\gamma,c}(x) = \begin{cases} (\gamma - 1)(2x + c) & \text{for } x < -c \\ (1 - \gamma)x^2/c & \text{for } -c \leq x < 0 \\ \gamma x^2/c & \text{for } 0 \leq x < c \\ \gamma(2x - c) & \text{for } c \leq x \end{cases}$$

for  $c > 0$ , where  $c$  is chosen close to zero. Unlike the standard loss function  $\rho_\gamma$ ,  $\rho_{\gamma,c}$  is differentiable at zero, offering significant advantages for inference in terms of speed and computational efficiency; see Oh et al. (2011) for further details. For our framework, we set  $c = 0.5$ ; this is the default value suggested in Youngman (2020).

Setting  $\log(\sigma(q)) = g_\sigma(q)$ , where  $g_\sigma(q)$  denotes a GAM formulation from equation (4.5) of the main article, and letting  $\boldsymbol{\beta}_\sigma$  denote the associated coefficient vector, the log-likelihood associated with the mis-specified model given of equation (S1) is given by

$$\ell(\boldsymbol{\beta}_{u_\gamma}, \boldsymbol{\beta}_\sigma; \mathbf{r}, \mathbf{q}, \gamma, c) \propto - \sum_{i=1}^n g_\sigma(q_i) - \sum_{i=1}^n \rho_{\gamma,c} \left( \frac{r_i - \exp(g_{u_\gamma}(q_i))}{\exp(g_\sigma(q_i))} \right). \quad (\text{S2})$$

Treating  $\sigma(q)$  as a nuisance parameter function, maximisation of equation (S2) with respect to  $\boldsymbol{\beta}_{u_\gamma}$  is equivalent to minimisation of equation (S1). Consequently, the resulting estimate for  $u_\gamma(q)$  gives an estimate of the conditional quantile function for  $R_q$ . For further details, along with detailed examples, see Yu and Moyeed (2001) and Geraci and Bottai (2007). Note that the same spline basis dimensions are used for both  $u_\gamma$  and  $\sigma$ .

Given an estimate of  $u_\gamma(q)$ , let  $\mathcal{I}_\gamma^o := \{i \in \{1, 2, \dots, n\} \mid r_i \geq u_\gamma(q_i)\}$  denote the observed indices of threshold exceedances. Recalling the GAM formulations for the GP scale and shape,  $\log(\tau(q)) = g_\tau(q)$  and  $\xi(q) = g_\xi(q)$ , the log-likelihood function of the GP tail model is

given by

$$\ell(\boldsymbol{\beta}_\tau, \boldsymbol{\beta}_\xi; \mathbf{r}_{\mathcal{I}_\gamma^o}, \mathbf{q}_{\mathcal{I}_\gamma^o}) = - \sum_{i \in \mathcal{I}_\gamma^o} g_\tau(q_i) + \sum_{i \in \mathcal{I}_\gamma^o} \left( -\frac{1}{g_\xi(q_i)} - 1 \right) \log \left( 1 + \frac{g_\xi(q_i)(r_i - u_\gamma(q_i))}{\exp(g_\tau(q_i))} \right), \quad (\text{S3})$$

where  $\mathbf{r}_{\mathcal{I}_\gamma^o} := \{r_i \mid i \in \mathcal{I}_\gamma^o\}$ ,  $\mathbf{q}_{\mathcal{I}_\gamma^o} := \{q_i \mid i \in \mathcal{I}_\gamma^o\}$ , and  $\boldsymbol{\beta}_\tau, \boldsymbol{\beta}_\xi$  denote the coefficient vectors associated with  $g_\tau$  and  $g_\xi$ , respectively. Minimising equation (S3) with respect to  $\boldsymbol{\beta}_\tau$  and  $\boldsymbol{\beta}_\xi$  results in estimates of the parameter functions.

Minimisation of equations (S2) and (S3) is achieved via REML, with the corresponding penalty parameters estimated via cross validation. We refer to Wood et al. (2016) and Wood (2017) for further details.

## S2 Additional information and plots for Section 6

In Section 6 of the main article, we consider four copula examples. The first is the Gaussian copula; given  $\rho \in [-1, 1]$ , termed the Pearson correlation coefficient, this is given by

$$C(u_1, u_2; \rho) = \boldsymbol{\Phi}(\Phi^{-1}(u_1), \Phi^{-1}(u_2); \Sigma), \quad \Sigma := \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$$

where  $\boldsymbol{\Phi}$  is the bivariate Gaussian cumulative distribution function with correlation matrix  $\Sigma$  and  $\Phi^{-1}$  is the inverse of the standard univariate Gaussian cumulative distribution function.

The parameter  $\rho$  controls the form and strength of dependence.

Secondly, for any  $\alpha \neq 0$ , the Frank copula is defined as

$$C(u_1, u_2; \alpha) = -\frac{1}{\alpha} \log \left( 1 + \frac{\prod_{i=1}^2 (e^{-\alpha u_i} - 1)}{e^{-\alpha} - 1} \right),$$

with the parameter  $\alpha$  controlling the strength and form of dependence.



Next, for any  $\rho \in [-1, 1]$  and  $\nu > 0$ , the t copula is given by

$$C(u_1, u_2; \rho, \nu) = T(t_\nu^{-1}(u_1), t_\nu^{-1}(u_2); \Sigma, \nu)$$

where  $T$  is the bivariate t cumulative distribution function with correlation matrix  $\Sigma$  and  $\nu$  degrees of freedom, and  $t_\nu^{-1}$  is the inverse of the univariate t cumulative distribution function with  $\nu$  degrees of freedom. The strength of dependence in the tails is controlled by both  $\nu$  and  $\rho$  (Chan and Li, 2008).

Finally, given any  $\alpha > 0$ , the Joe copula is given by

$$C(u_1, u_2; \alpha) = 1 - \left(1 - \sum_{i=1}^2 u_i^\alpha\right)^{\frac{1}{\alpha}}$$

where  $\alpha > 0$  controlling the strength of dependence.

For each copula, we simulate data on standard Laplace margins, for which the marginal distribution is given by

$$F_L(x) := \frac{1}{2} (1 + \text{sgn}(x) \cdot (1 - e^{-|x|})), \quad x \in \mathbb{R}. \quad (\text{S4})$$

To achieve this, we first simulate data from each copula on standard uniform margins using the `copula` package in the R computing language. We then transform this data to standard Laplace using the inverse of equation (S4).

Figure S1 compares the median estimates of isodensity contours, obtained using the  $L2$  coordinate system, to the true contours at a range of low density levels.

Figure S2 compares the median isodensity contours estimates from the two coordinate systems for two density levels  $p \in \{10^{-3}, 10^{-6}\}$ . One can observe that the differences between the sets of median estimates are negligible.

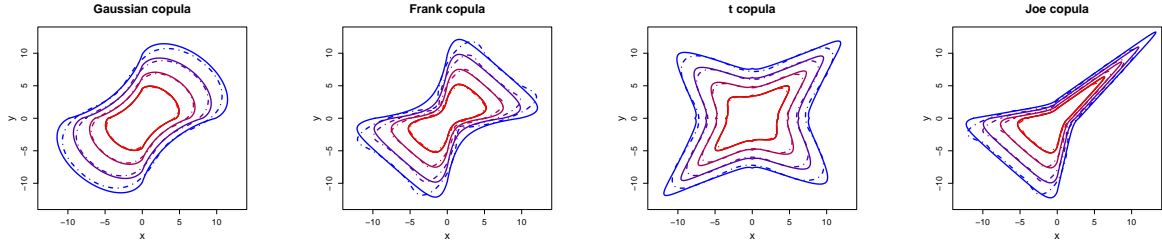


Figure S1: Comparison of true (thick lines) and median estimated (dot-dashed lines) isodensity contours under the  $L_2$  coordinate system. In each plot, the red to blue lines represent the joint density levels  $p \in \{10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}\}$ .

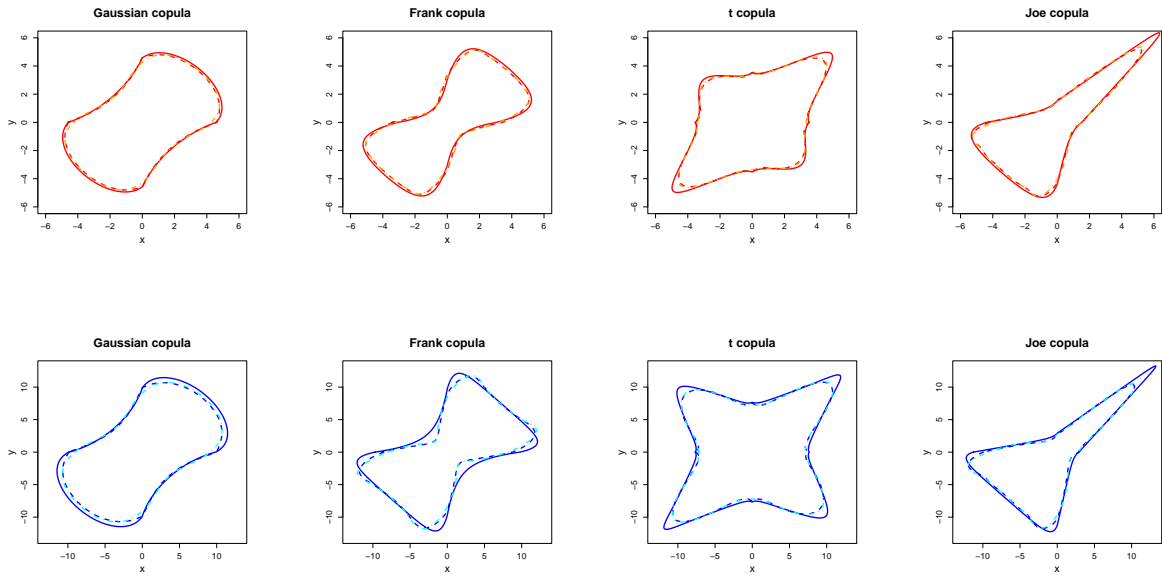


Figure S2: Comparison of estimated median isodensity contours to truth (solid lines) for the  $L_1$  (dashed lines) and  $L_2$  (dot-dashed lines) coordinate systems. The top and bottom rows give the comparisons for  $p = 10^{-3}$  and  $p = 10^{-6}$ , respectively.

Figures S3 and S4 illustrate the median scale parameter estimates, with confidence intervals, for the  $L_1$  and  $L_2$  coordinate systems, respectively. Considering the fact estimates were computed using finite sample sizes, these plots illustrate reasonable agreement between the estimated and asymptotic scale parameter functions in most cases. One notable exception is the Frank copula, for which the estimated scale functions perform poorly. This is likely due to the relatively slow convergence of this distribution to its asymptotic form, as discussed in Mackay and Jonathan (2023). Furthermore, the majority of the  $\xi$  estimates obtained over the simulated data sets are slightly negative. This discrepancy from the asymptotic value

( $\xi = 0$ ) partly explains why the estimated scale functions estimates are biased high in most cases, since the scale and shape parameters are negatively correlated under the maximum likelihood framework (Hosking and Wallis, 1987).

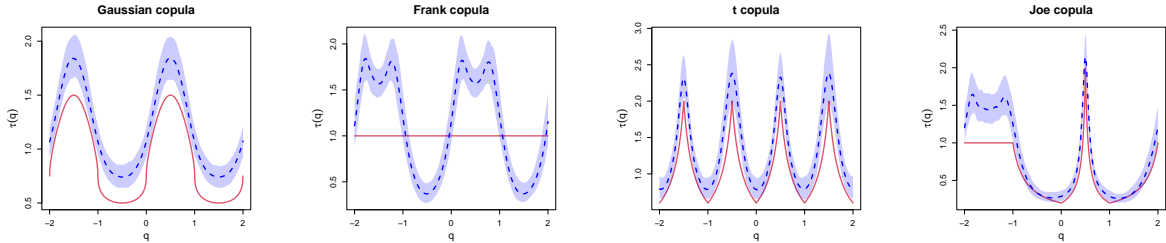


Figure S3: Comparison of true (solid red lines) and median estimated (dashed blue lines) scale parameter functions under the  $L1$  coordinate system. In each plot, the shaded regions give the estimated 95% confidence intervals.

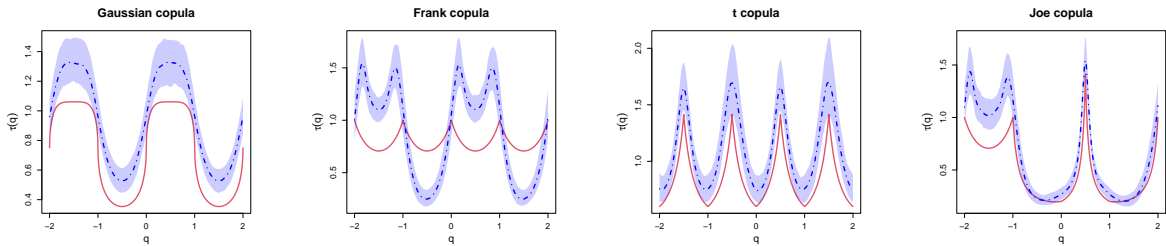


Figure S4: Comparison of true (solid red lines) and median estimated (dot-dashed blue lines) scale parameter functions under the  $L2$  coordinate system. In each plot, the shaded regions give the estimated 95% confidence intervals.

Figure S5 illustrates the median contour estimates for  $p \in \{10^{-3}, 10^{-6}\}$  on the radial-angular scale for the  $L2$  coordinate system, along with estimated 95% confidence intervals. As with the  $L1$  coordinates, the estimated confidence intervals capture the true contours in most cases.

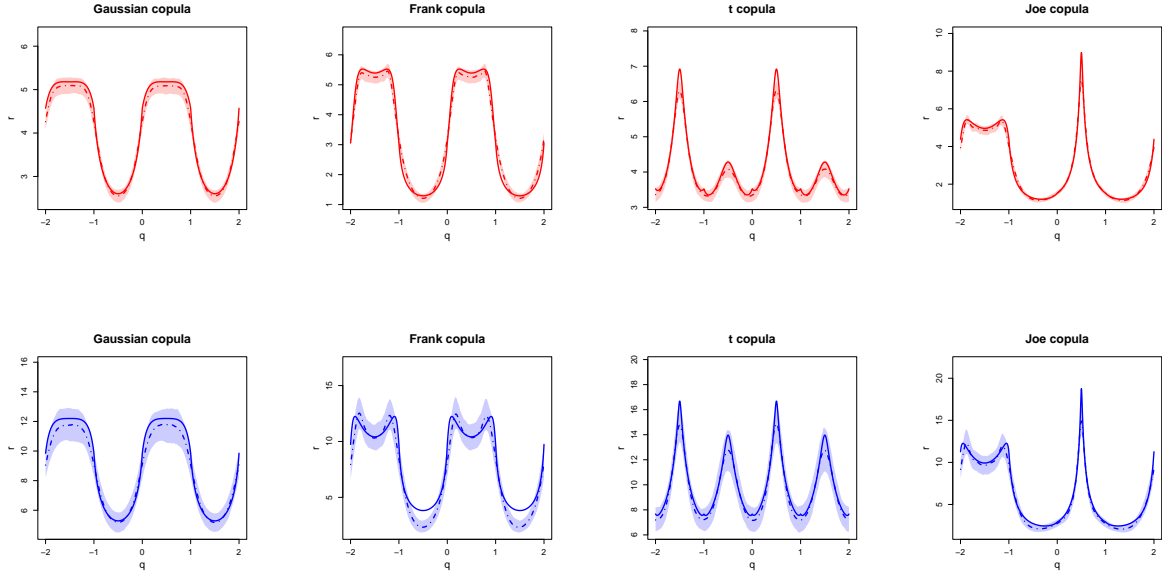


Figure S5: Comparison of median estimated isodensity contours (dashed lines), with 95% confidence intervals (shaded region), to true contours (solid lines) for joint density level  $p = 10^{-3}$  (top row) and  $p = 10^{-6}$  (bottom row), with estimates obtained using  $L2$  polar coordinates.

Figures S6 and S7 compare the true and estimated angular density functions for the  $L1$  and  $L2$  coordinate systems, respectively. One can observe generally good agreement for both coordinate systems. We note that the KD estimation framework appears unable to fully capture the modal regions for the Frank,  $t$  and Joe copulas. We observed a marginal improvement in performance for these regions when the bandwidth parameter,  $h$ , was decreased; however, this significantly increased the variability in the resulting angular density estimates, and consequently, we opted to keep  $h$  fixed at  $1/50$ . Moreover, we note that even for extremely small bandwidth parameters, the proposed KD framework was unable to approximate the modal behaviour of the Joe copula, suggesting the sample size ( $n = 10,000$ ) is not large enough to fully capture the true density function for this particular example.

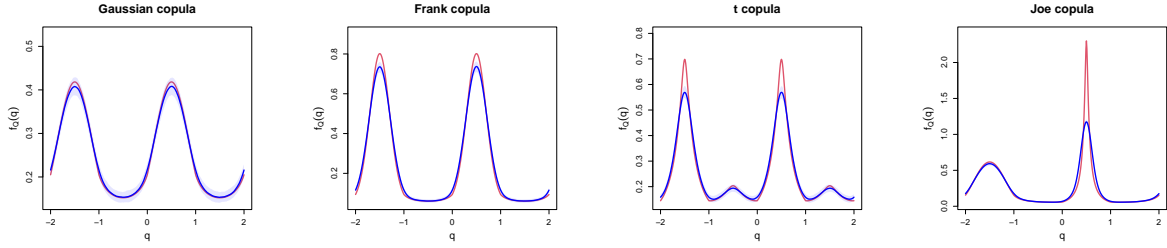


Figure S6: Comparison of true (red) and estimated median (blue) angular density functions, alongside 95% confidence intervals, for the  $L1$  coordinate system. In each plot, the shaded regions illustrate the estimated confidence intervals.

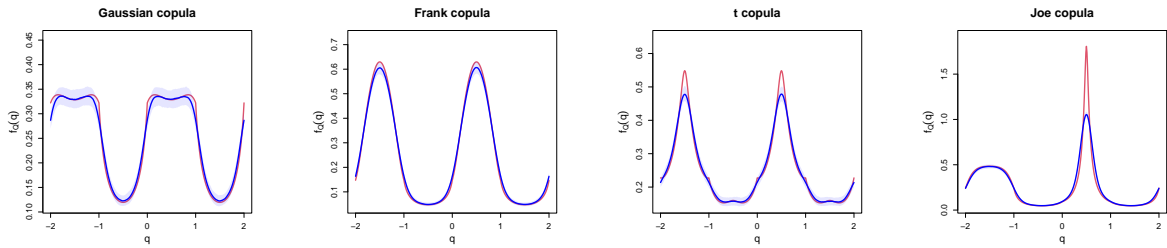


Figure S7: Comparison of true (red) and estimated median (blue) angular density functions, alongside 95% confidence intervals, for the  $L2$  coordinate system. In each plot, the shaded regions illustrate the estimated confidence intervals.

### S3 Additional plots for Section 7

Figures S1 and S2 compare the smoothly and locally estimated threshold and parameters functions for data sets A and C, respectively, under the  $L1$  coordinate system. One can observe generally good agreement for each component of the SPAR model.

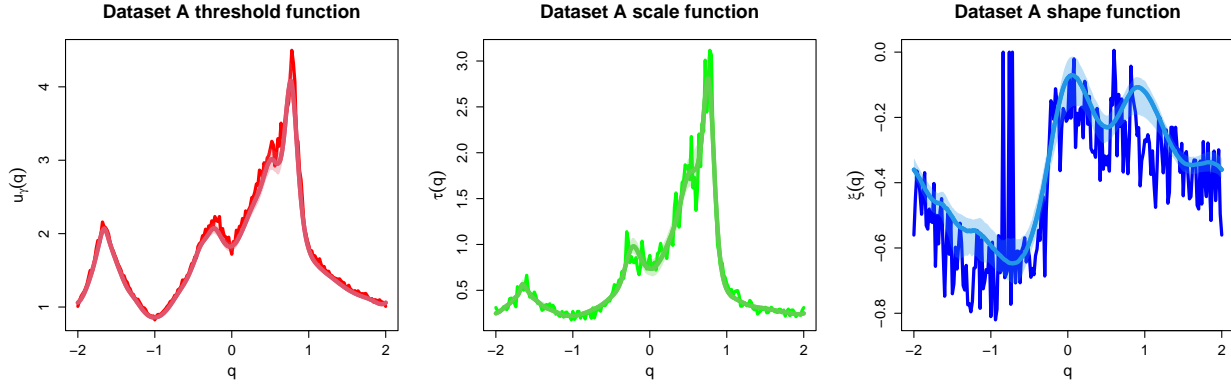


Figure S1: Comparison of estimated local (rough lines) and smooth threshold (red, left), scale (green, middle) and shape (blue, right) functions for data set A with the  $L_1$  coordinate system, with shaded regions denoting 95% confidence intervals.

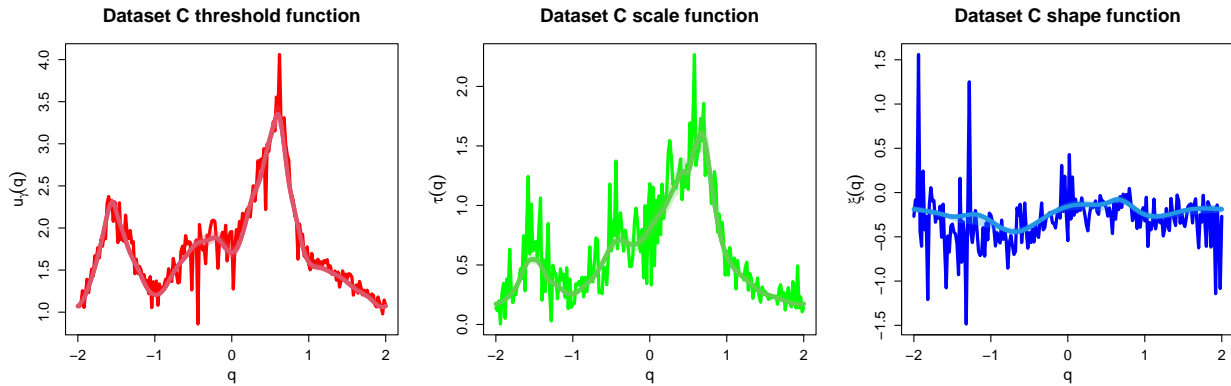


Figure S2: Comparison of estimated local (rough lines) and smooth threshold (red, left), scale (green, middle) and shape (blue, right) functions for data set C with the  $L_1$  coordinate system, with shaded regions denoting 95% confidence intervals.

Figures S3, S4 and S5 compare the smoothly and locally estimated threshold and parameters functions for data sets A, B and C, respectively, under the  $L_2$  coordinate system. As with the  $L_1$  coordinates, we obtain good agreement for each model component.

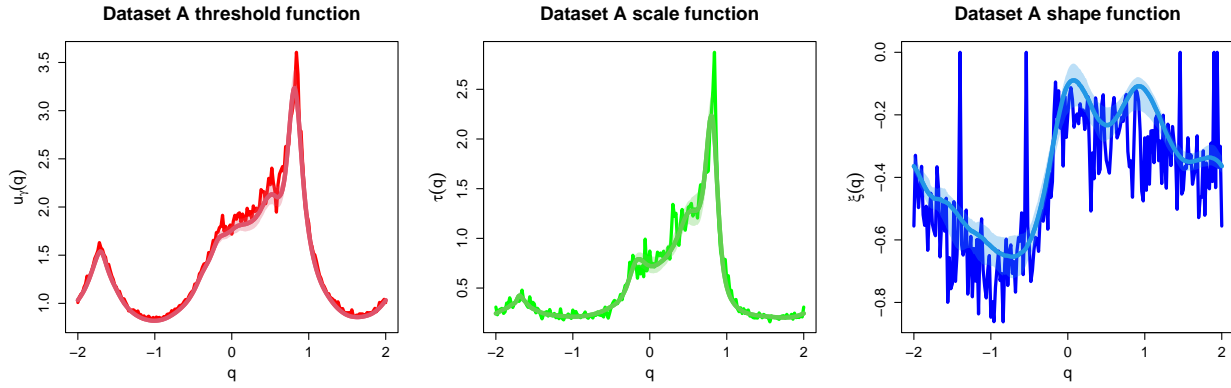


Figure S3: Comparison of estimated local (rough lines) and smooth threshold (red, left), scale (green, middle) and shape (blue, right) functions for data set A with the  $L_2$  coordinate system, with shaded regions denoting 95% confidence intervals.

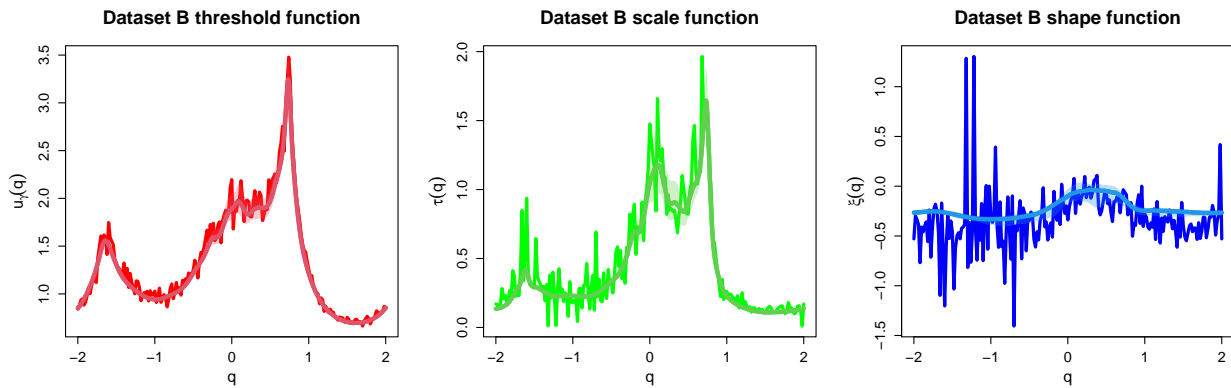


Figure S4: Comparison of estimated local (rough lines) and smooth threshold (red, left), scale (green, middle) and shape (blue, right) functions for data set B with the  $L_2$  coordinate system, with shaded regions denoting 95% confidence intervals.

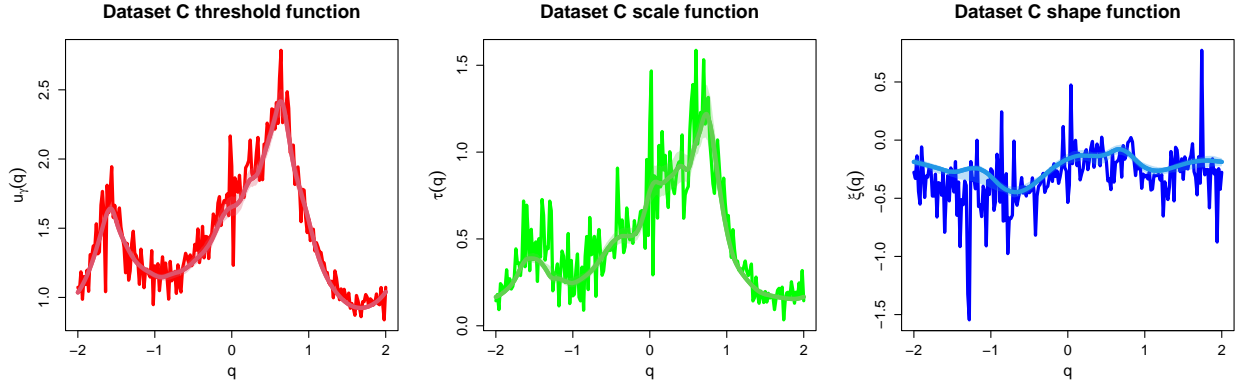


Figure S5: Comparison of estimated local (rough lines) and smooth threshold (red, left), scale (green, middle) and shape (blue, right) functions for data set C with the  $L_2$  coordinate system, with shaded regions denoting 95% confidence intervals.

Figure S6 compares the estimated median angular density functions, and corresponding confidence intervals, with the histograms for each data set under the  $L_2$  coordinate system. The two sets of estimates are in good agreement, providing evidence that the chosen bandwidth parameter is appropriate.

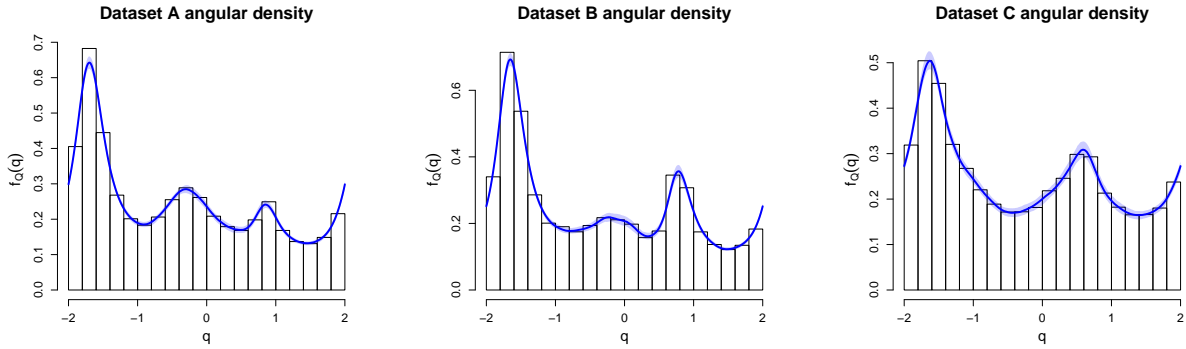


Figure S6: Comparison of estimated median (blue lines) angular density functions to histograms for data sets A (left), B (centre), and C (right) with the  $L_2$  coordinate system. The shaded regions in each plot denote the estimated 95% confidence intervals.

Figure S7 illustrates the estimated median isodensity contours, and corresponding uncertainty regions, for the density levels  $p \in \{10^{-3}, 10^{-6}\}$  under the  $L_2$  coordinate system. As with the  $L_1$  coordinates, the estimated contours capture the shape and structure of each data set.



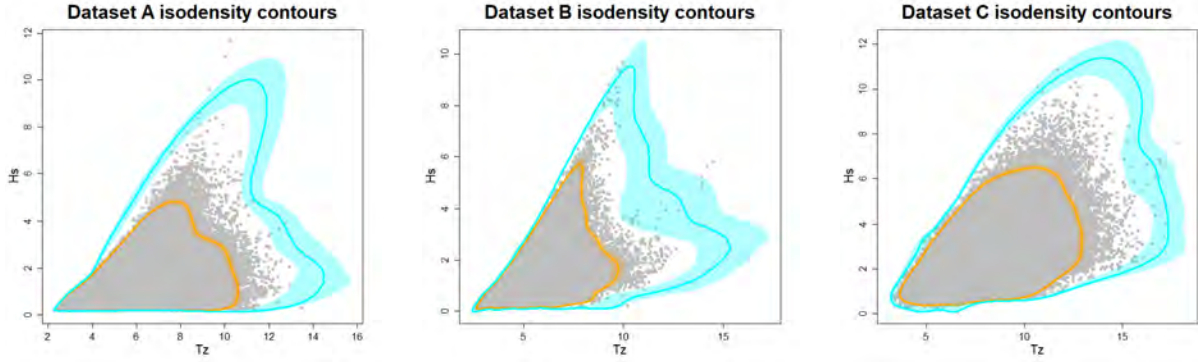


Figure S7: Estimated median isodensity contours at  $p = 10^{-3}$  (orange lines) and  $p = 10^{-6}$  (cyan lines) for data sets A (left), B (centre), and C (right) with the  $L2$  coordinate system. The shaded region for each contour denote the 95% bootstrapped confidence intervals.

Figure S8 illustrates the estimated median 10 year return level sets for each data set, along with 95% bootstrapped confidence intervals, obtained using the  $L2$  coordinate system. The estimated sets appear to capture the features of the observed data, and we observe only a handful of observations outside of the return level set for each data set.

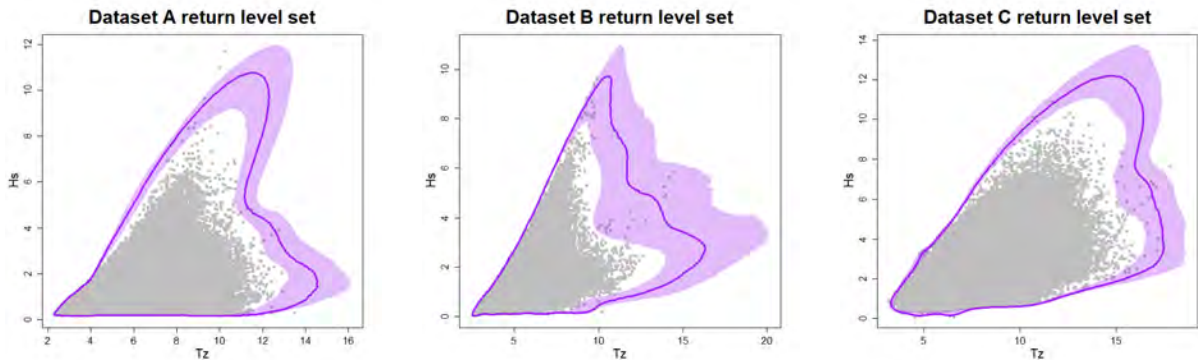


Figure S8: Estimated median 10 year return level sets (purple lines) for data sets A (left), B (centre), and C (right) with the  $L2$  coordinate system. The shaded region for each return level set denotes the 95% bootstrapped confidence region.

Figure S9 compares the estimated 10 year return level sets from the two coordinate systems for each of the data sets. One can observe generally good agreement between the two systems, although we note some small differences in the estimates for data set A at certain angles. The overall agreement, however, provides evidence of consistency between the two modelling approaches.

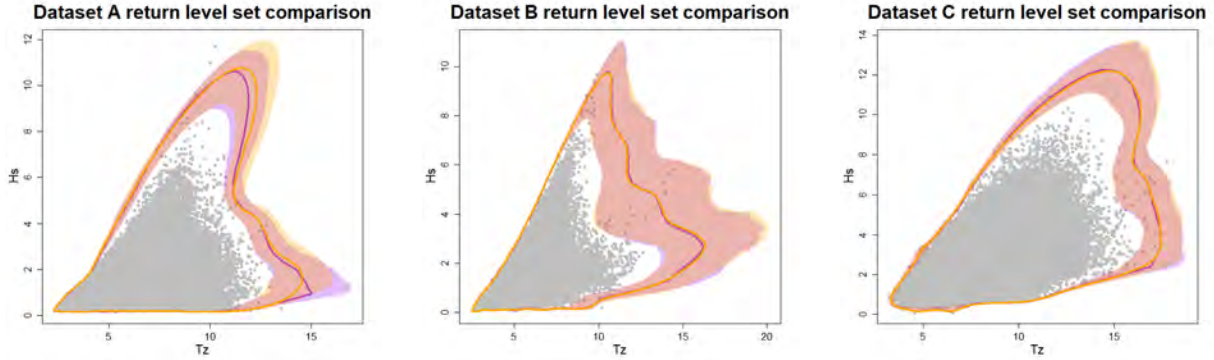


Figure S9: Comparison of  $L1$  (purple) and  $L2$  (orange) coordinate system return level sets for data sets A (left), B (centre), and C (right). The shaded region for each return level set denotes the 95% bootstrapped confidence intervals.

As noted in the main article, fitted SPAR models can be used to simulate new observations. This simulation is straightforward. We start by generating a random number  $u^*$ , uniformly distributed in  $[0, 1]$ . A random angle  $q \in (-2, 2]$  can then be calculated by applying the probability integral transform so that  $q = F_Q^{-1}(u^*)$ , where  $F_Q$  denotes the estimated distribution function of  $Q$ . A corresponding radial value  $r$ , is then simulated as a random value from the GP distribution with parameter vector  $(u_\gamma(q), \xi(q), \tau(q))$ . The resulting pair  $(r, q)$  is then a random sample from the SPAR model. Figure S10 and S11 illustrate simulated data points for each metocean data set, overlaid on top of the observed time series, for the  $L1$  and  $L2$  coordinate systems, respectively. For each data set, we simulated the same number of new observations as the original sample size. One can observe that the simulated data sets closely resemble the threshold exceeding observations. Moreover, these plots illustrate the regions  $\mathcal{U}_\gamma$  for which the SPAR model is valid.

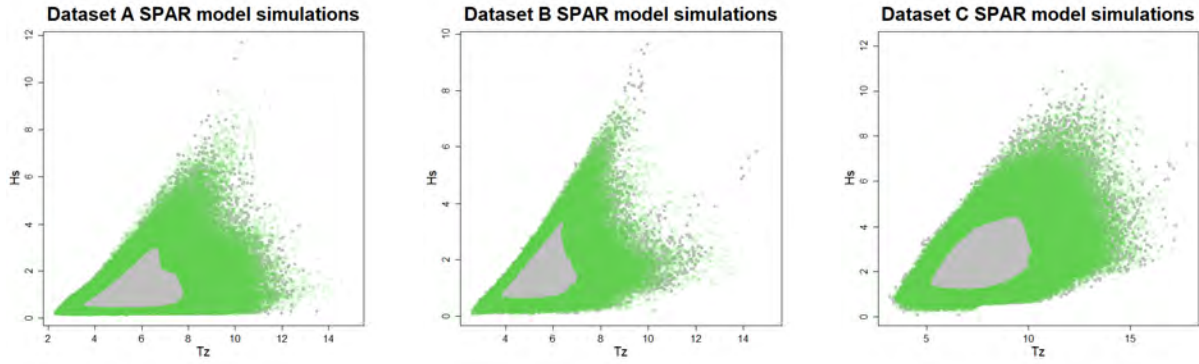


Figure S10: Model simulations (green) against observed time series (grey) for data sets A (left), B (centre), and C (right) with the  $L1$  coordinate system.

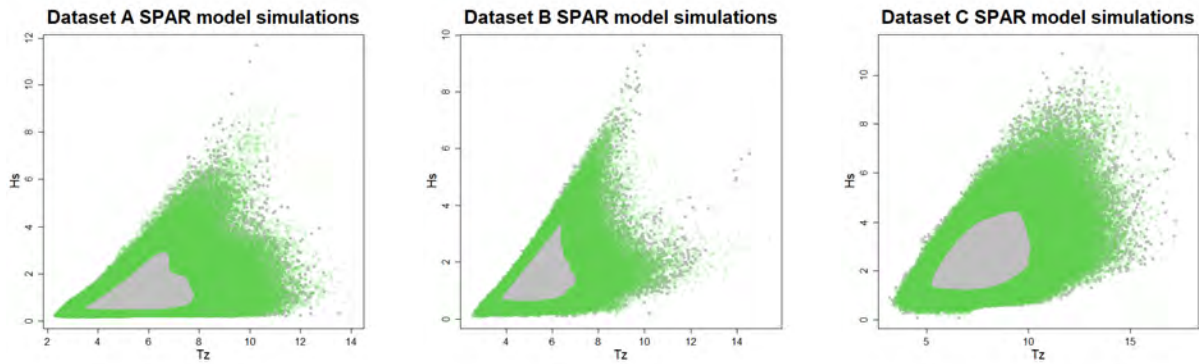


Figure S11: Model simulations (green) against observed time series (grey) for data sets A (left), B (centre), and C (right) with the  $L2$  coordinate system.

Figures S12 and S13 illustrate QQ plots on the standard exponential scale obtained from the SPAR model fits on each data set using the  $L1$  and  $L2$  coordinate systems, respectively. There is good overall agreement, albeit with some slight overestimation from the model at the most extreme radial values.

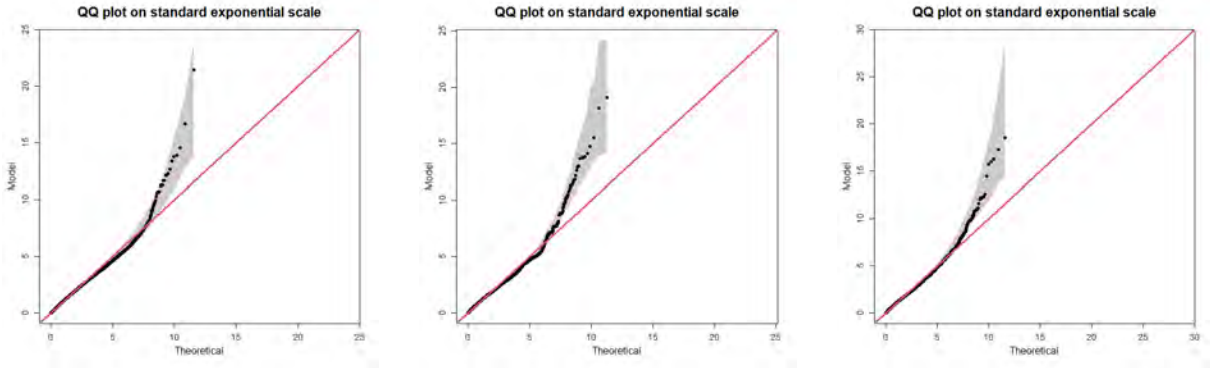


Figure S12: QQ plot on the standard exponential scale obtained from the fitted SPAR models for data sets A (left), B (centre), and C (right) with the  $L1$  coordinate system. The shaded region for each plot denotes the empirical 95% confidence region.

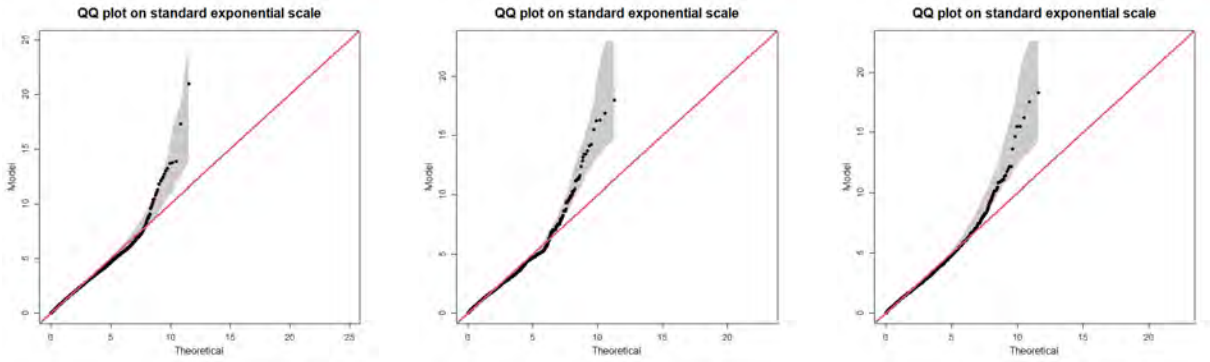


Figure S13: QQ plot on the standard exponential scale obtained from the fitted SPAR models for data sets A (left), B (centre), and C (right) with the  $L2$  coordinate system. The shaded region for each plot denotes the empirical 95% confidence region.

Figures S14, S15 and S16 give the local window QQ plots for the SPAR model fits on data sets A, B, and C, respectively, under the  $L1$  coordinate system. The corresponding plots for the  $L2$  coordinate system are given in Figures S17, S18 and S19. We observe generally good agreement between the estimated model and observed quantiles. Note that the step-like behaviour observed in some of the plotted quantiles occurs due to repeated observations in the time series'; these are likely a result of rounding errors in measurement equipment.

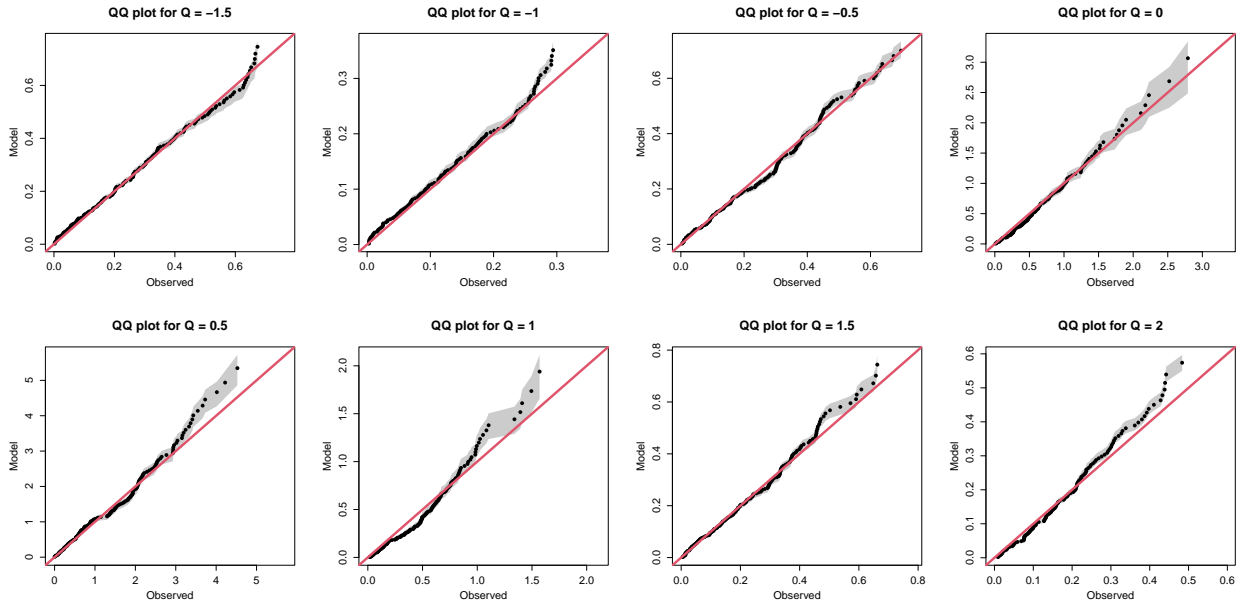


Figure S14: Local QQ plots for the fitted SPAR model on data set A with the  $L_1$  coordinate system. The shaded region for each plot denotes the empirical 95% confidence region.

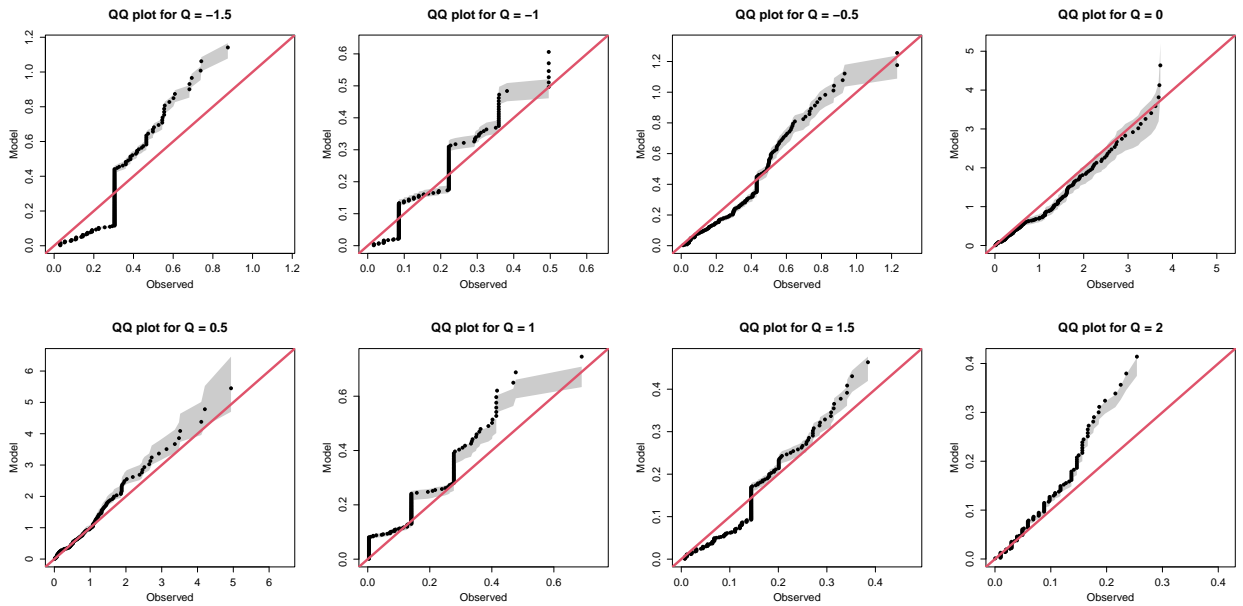


Figure S15: Local QQ plots for the fitted SPAR model on data set B with the  $L_1$  coordinate system. The shaded region for each plot denotes the empirical 95% confidence region.

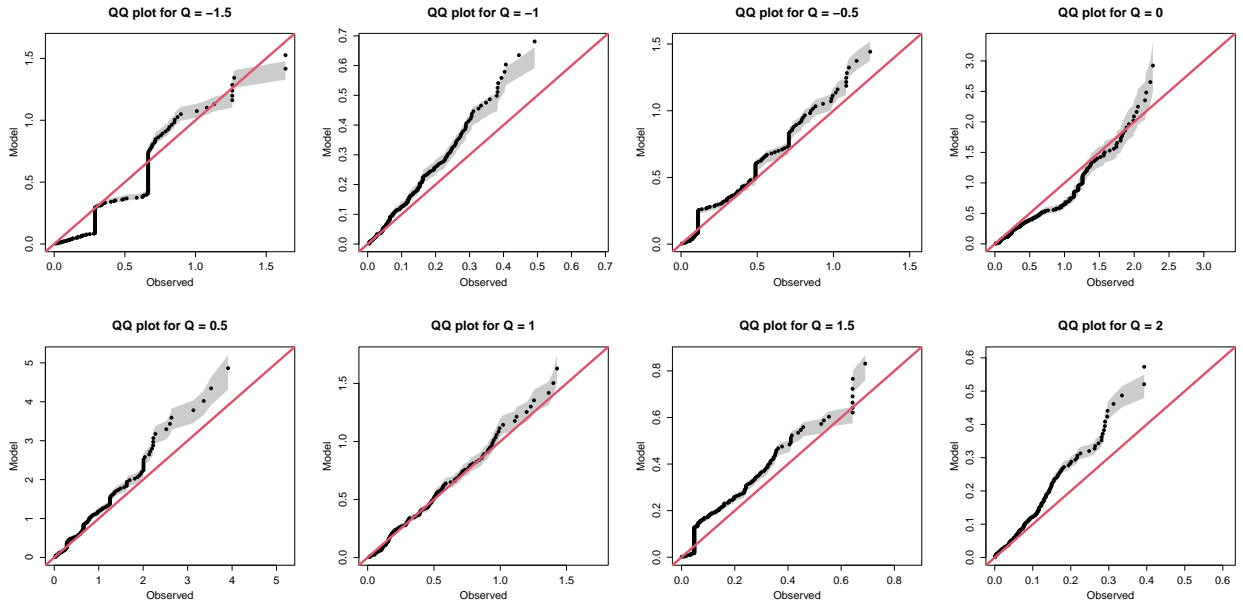


Figure S16: Local QQ plots for the fitted SPAR model on data set C with the  $L_1$  coordinate system. The shaded region for each plot denotes the empirical 95% confidence region.

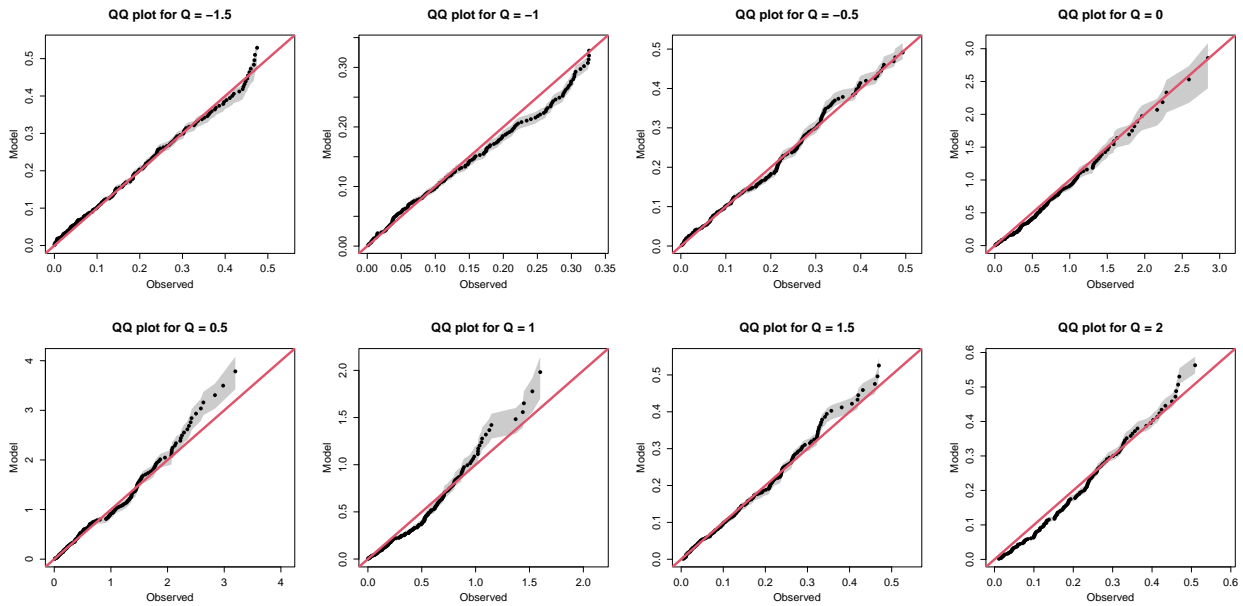


Figure S17: Local QQ plots for the fitted SPAR model on data set A with the  $L_2$  coordinate system. The shaded region for each plot denotes the empirical 95% confidence region.

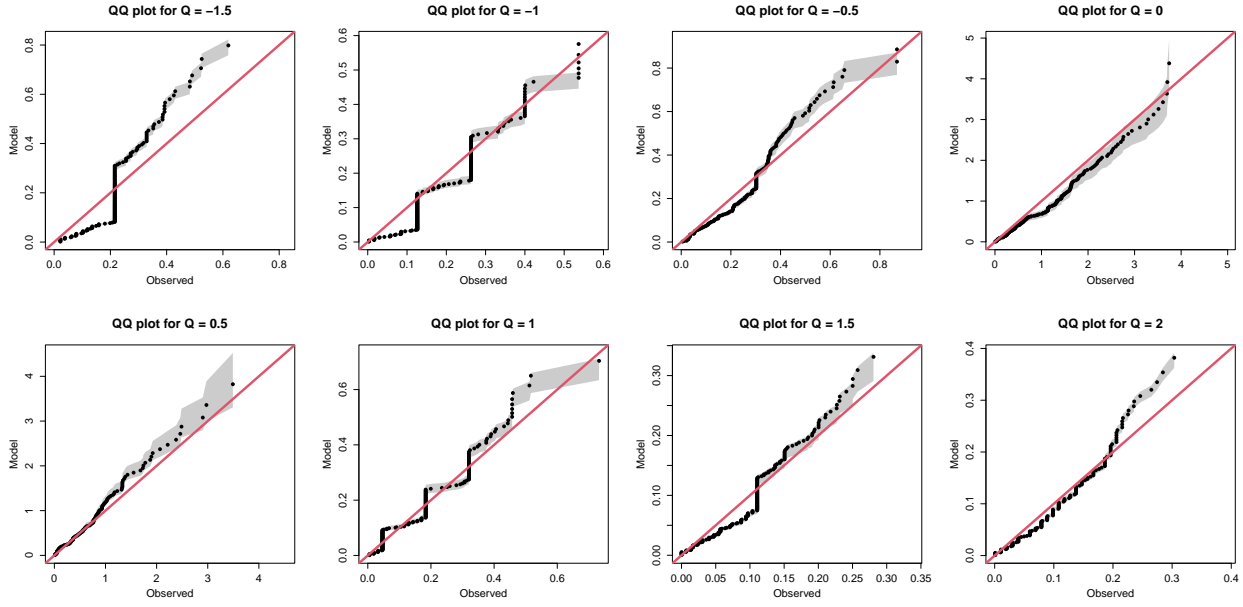


Figure S18: Local QQ plots for the fitted SPAR model on data set B with the  $L_2$  coordinate system. The shaded region for each plot denotes the empirical 95% confidence region.

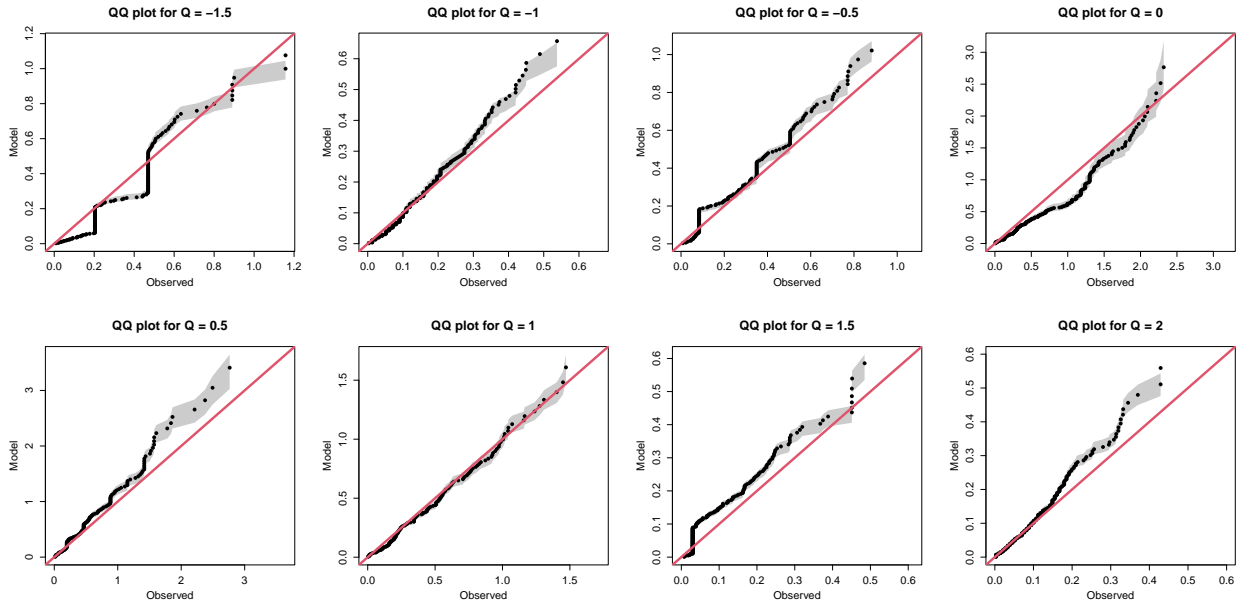


Figure S19: Local QQ plots for the fitted SPAR model on data set C with the  $L_2$  coordinate system. The shaded region for each plot denotes the empirical 95% confidence region.