Supplement to "Peaks over thresholds modelling with multivariate generalized Pareto distributions"

Anna Kiriliouk   Johan Segers
Université catholique de Louvain
Institut de Statistique, Biostatistique et Sciences Actuarielles
Voie du Roman Pays 20, B-1348 Louvain-la-Neuve, Belgium.
E-mail: anna.kiriliouk@uclouvain.be, johan.segers@uclouvain.be

Holger Rootzén
Chalmers University of Technology
Department of Mathematical Sciences
SE-412 96 Gothenburg, Sweden.
E-mail: hrootzen@chalmers.se

Jennifer L. Wadsworth
Lancaster University
Department of Mathematics and Statistics
Fylde College LA1 4YF, Lancaster, England.
E-mail: j.wadsworth@lancaster.ac.uk

A Censored likelihoods

Here we detail forms of censored likelihoods for the models proposed in Section 4. For simplicity they are presented in standardized ($\boldsymbol{\sigma} = \mathbf{1}$, $\boldsymbol{\gamma} = \mathbf{0}$) form, i.e.,

$$h_C(\boldsymbol{x}_{D \setminus C}, \boldsymbol{v}_C; \mathbf{1}, \mathbf{0}) = \int_{\times_{j \in C}(-\infty, v_j]} h(\boldsymbol{x}; \mathbf{1}, \mathbf{0}) \, \mathrm{d}\boldsymbol{x}_C, \tag{A.1}$$

for $v_j \leq 0$ and h corresponding to either $h_{\boldsymbol{T}}$ or $h_{\boldsymbol{U}}$. The generalized form of a censored likelihood is easily obtained from (A.1) as

$$h_C(\boldsymbol{x}_{D \setminus C}, \boldsymbol{v}_C; \boldsymbol{\sigma}, \boldsymbol{\gamma}) = h_C\left(\tfrac{1}{\boldsymbol{\gamma}} \log(1 + \boldsymbol{\gamma}\boldsymbol{x}_{D \setminus C}/\boldsymbol{\sigma}), \tfrac{1}{\boldsymbol{\gamma}} \log(1 + \boldsymbol{\gamma}\boldsymbol{v}_C/\boldsymbol{\sigma}); \mathbf{1}, \mathbf{0}\right)$$

$$\times \prod_{j \in D\setminus C} \frac{1}{\sigma_j + \gamma_j x_j}.$$

The support for each density is $\{\boldsymbol{x} \in \mathbb{R}^d : \boldsymbol{x} \not\leq \boldsymbol{0}\}$, and we let $|C|$ denote the cardinality of the set C.

Generators with independent Gumbel components

Case $f_{\boldsymbol{T}} = f_{\boldsymbol{V}}$.

$$h_C(\boldsymbol{x}_{D\setminus C}, \boldsymbol{v}_C; \boldsymbol{1}, \boldsymbol{0}) = e^{-\max(\boldsymbol{x})} \times \int_0^\infty t^{-1} \prod_{j \in C} e^{-(te^{v_j - \beta_j})^{-\alpha_j}} \prod_{j \in D\setminus C} \alpha_j \left(te^{x_j - \beta_j}\right)^{-\alpha_j} e^{-(te^{x_j-\beta_j})^{-\alpha_j}} \,\mathrm{d}t.$$

If all α_j are equal to α:

$$h_C(\boldsymbol{x}_{D\setminus C}, \boldsymbol{v}_C; \boldsymbol{1}, \boldsymbol{0}) = e^{-\max(\boldsymbol{x})} \frac{\alpha^{d-|C|-1}\Gamma(d-|C|) \prod_{j\in D\setminus C} e^{-\alpha(x_j - \beta_j)}}{\left(\sum_{j\in C} e^{-\alpha(v_j-\beta_j)} + \sum_{j \in D\setminus C} e^{-\alpha(x_j-\beta_j)}\right)^{d-|C|}}.$$

Case $f_{\boldsymbol{U}} = f_{\boldsymbol{V}}$.

$$h_C(\boldsymbol{x}_{D\setminus C}, \boldsymbol{v}_C; \boldsymbol{1}, \boldsymbol{0}) = \frac{\int_0^\infty \prod_{j\in C} e^{-(te^{v_j-\beta_j})^{-\alpha_j}} \prod_{j\in D\setminus C} \alpha_j \left(te^{x_j-\beta_j}\right)^{-\alpha_j} e^{-(te^{x_j-\beta_j})^{-\alpha_j}} \,\mathrm{d}t}{\Gamma(1-1/\alpha)\left(\sum_{j=1}^d e^{\beta_j \alpha}\right)^{1/\alpha}}.$$

If all α_j are equal to α:

$$h_C(\boldsymbol{x}_{D\setminus C}, \boldsymbol{v}_C; \boldsymbol{1}, \boldsymbol{0}) = \frac{\alpha^{d-|C|-1}\Gamma(d-|C|-1/\alpha)\left(\sum_{j=1}^d e^{\beta_j\alpha}\right)^{-1/\alpha} \prod_{j\in D\setminus C} e^{-\alpha(x_j-\beta_j)}}{\Gamma(1-1/\alpha)\left(\sum_{j\in C} e^{-\alpha(v_j-\beta_j)} + \sum_{j\in D\setminus C} e^{-\alpha(x_j-\beta_j)}\right)^{d-|C|-1/\alpha}}.$$

Generators with independent reverse exponential components

Case $f_{\boldsymbol{T}} = f_{\boldsymbol{V}}$.

$$h_C(\boldsymbol{x}_{D\setminus C}, \boldsymbol{v}_C; \boldsymbol{1}, \boldsymbol{0}) = e^{-\max(\boldsymbol{x})} \times$$

$$\int_0^{e^{-\max_{j\in D\setminus C}(x_j+\beta_j)}} t^{-1}\prod_{j\in C}\min(te^{v_j+\beta_j},1)^{1/\alpha_j}\prod_{j\in D\setminus C}\frac{1}{\alpha_j}(te^{x_j+\beta_j})^{1/\alpha_j}\,\mathrm{d}t \quad \text{(A.2)}$$

To evaluate this, consider two cases: (i) $\max_{j\in C}(v_j+\beta_j) < \max_{j\in D\setminus C}(x_j+\beta_j)$; and (ii) let $v_{(1)}+\beta_{(1)} > \ldots > v_{(k)}+\beta_{(k)} > \max_{j\in D\setminus C}(x_j+\beta_j) > v_{(k+1)}+\beta_{(k+1)} > \ldots$ for $j\in C$ and $k\le |C|$. In case (i), we have

$$h_C(\boldsymbol{x}_{D\setminus C},\boldsymbol{v}_C;\mathbf{1},\mathbf{0}) = e^{-\max(\boldsymbol{x})}\frac{\prod_{j\in C}e^{(v_j+\beta_j)/\alpha_j}\prod_{j\in D\setminus C}(1/\alpha_j)e^{(\beta_j+x_j)/\alpha_j}}{\left(\sum_{j=1}^d 1/\alpha_j\right)\left(e^{\max_{j\in D\setminus C}(x_j+\beta_j)}\right)^{\sum_{j=1}^d 1/\alpha_j}},$$

since on the range $t\in(0,e^{-\max_{j\in D\setminus C}(x_j+\beta_j)})$ the term $\prod_{j\in C}\min(te^{v_j+\beta_j},1)^{1/\alpha_j}$ in (A.2) is equal to $\prod_{j\in C}(te^{v_j+\beta_j})^{1/\alpha_j}$. In case (ii) this term will vary over that range, and one needs to split the integral as follows:

$$\int_0^{e^{-(v_{(1)}+\beta_{(1)})}} + \int_{e^{-(v_{(1)}+\beta_{(1)})}}^{e^{-(v_{(2)}+\beta_{(2)})}} + \cdots + \int_{e^{-(v_{(k)}+\beta_{(k)})}}^{e^{-\max_{j\in D\setminus C}(x_j+\beta_j)}} .$$

An evaluation of each integral yields that $e^{\max(\boldsymbol{x})}h_C(\boldsymbol{x}_{D\setminus C},\boldsymbol{v}_C;\mathbf{1},\mathbf{0})$ is equal to

$$\begin{aligned}
&\frac{\prod_{j\in C}e^{(v_j+\beta_j)/\alpha_j}\prod_{j\in D\setminus C}(1/\alpha_j)e^{(x_j+\beta_j)/\alpha_j}}{\left(\sum_{j=1}^d 1/\alpha_j\right)\left(e^{v_{(1)}+\beta_{(1)}}\right)^{\sum_{j=1}^d 1/\alpha_j}}\\
&\quad+\sum_{i=1}^{k-1}\Bigg\{\frac{\prod_{j\in C_{(i)}}e^{(v_j+\beta_j)/\alpha_j}\prod_{j\in D\setminus C}(1/\alpha_j)e^{(x_j+\beta_j)/\alpha_j}}{\sum_{j\in C_{(i)}}1/\alpha_j+\sum_{j\in D\setminus C}1/\alpha_j}\\
&\qquad\times\left[\left(e^{v_{(i+1)}+\beta_{(i+1)}}\right)^{-\sum_{j\in C_{(i)}}1/\alpha_j-\sum_{j\in D\setminus C}1/\alpha_j}-\left(e^{v_{(i)}+\beta_{(i)}}\right)^{-\sum_{j\in C_{(i)}}1/\alpha_j-\sum_{j\in D\setminus C}1/\alpha_j}\right]\Bigg\}\\
&\quad+\frac{\prod_{j\in C_{(k)}}e^{(v_j+\beta_j)/\alpha_j}\prod_{j\in D\setminus C}(1/\alpha_j)e^{(x_j+\beta_j)/\alpha_j}}{\sum_{j\in C_{(k)}}1/\alpha_j+\sum_{j\in D\setminus C}1/\alpha_j}\\
&\qquad\times\left[\left(e^{\max_{j\in D\setminus C}(x_j+\beta_j)}\right)^{-\sum_{j\in C_{(k)}}1/\alpha_j-\sum_{j\in D\setminus C}1/\alpha_j}-\left(e^{v_{(k)}+\beta_{(k)}}\right)^{-\sum_{j\in C_{(k)}}1/\alpha_j-\sum_{j\in D\setminus C}1/\alpha_j}\right]
\end{aligned}$$

with $C_{(i)} = C\setminus\{(1),\ldots,(i)\}$, i.e., with the indices corresponding to the i largest $v_j+\beta_j$ removed.

Case $f_{\boldsymbol{U}} = f_{\boldsymbol{V}}$. Is found similarly by noting the relation between these two approaches.

Generators with independent log-gamma components

Case $f_{\boldsymbol{T}} = f_{\boldsymbol{V}}$. Let F_j denote the cumulative distribution function of a Gamma$(\alpha_j, 1)$ random variable. Then

$$h_C(\boldsymbol{x}_{D\setminus C}, \boldsymbol{v}_C; \mathbf{1}, \mathbf{0}) = e^{-\max(\boldsymbol{x})} \prod_{j \in D\setminus C} \frac{e^{\alpha_j x_j}}{\Gamma(\alpha_j)} \int_0^\infty t^{-1} \left(\prod_{j\in D\setminus C} t^{\alpha_j} e^{-te^{x_j}} \right) \left(\prod_{j\in C} F_j(te^{v_j}) \right) \mathrm{d}t.$$

Case $f_{\boldsymbol{U}} = f_{\boldsymbol{V}}$. Defining $C_d = \int_{\Delta_{d-1}} \max(u_1, \ldots, u_d) \prod_{j=1}^d u_j^{\alpha_j - 1} \,\mathrm{d}u_1 \cdots \mathrm{d}u_{d-1}$, we have

$$\begin{aligned} h_C(\boldsymbol{x}_{D\setminus C}, \boldsymbol{v}_C; \mathbf{1}, \mathbf{0}) = {} & \frac{C_d^{-1}}{\Gamma\left(\sum_{j=1}^d \alpha_j + 1\right)} \prod_{j\in D\setminus C} e^{\alpha_j x_j} \prod_{j\in C} \Gamma(\alpha_j) \\ & \times \int_0^\infty \left(\prod_{j\in D\setminus C} t^{\alpha_j} e^{-te^{x_j}} \right) \left(\prod_{j\in C} F_j(te^{v_j}) \right) \mathrm{d}t. \end{aligned}$$

Generators with multivariate Gaussian components

Case $f_{\boldsymbol{T}} = f_{\boldsymbol{V}}$. For the Gaussian model, using abbreviated notation, the key observation is

$$\int_{\times_{j\in C}(-\infty, v_j]} h(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x}_C = h_{D\setminus C}(\boldsymbol{x}_{D\setminus C}) \int_{\times_{j\in C}(-\infty, v_j]} \frac{h(\boldsymbol{x})}{h_{D\setminus C}(\boldsymbol{x}_{D\setminus C})} \,\mathrm{d}\boldsymbol{x}_C, \tag{A.3}$$

and the ratio in the second integral can be written as a proper Gaussian density function (with parameters that depend on $\boldsymbol{x}_{D\setminus C}$). The integrand is

$$\begin{aligned} \frac{h(\boldsymbol{x})}{h_{D\setminus C}(\boldsymbol{x}_{D\setminus C})} = {} & \frac{e^{\max_{j\in D\setminus C} x_j}}{e^{\max(\boldsymbol{x})}} \frac{(\mathbf{1}^T \Sigma_{D\setminus C}^{-1} \mathbf{1})^{1/2}}{(\mathbf{1}^T \Sigma^{-1} \mathbf{1})^{1/2}} \frac{|\Sigma_{D\setminus C}|^{1/2}}{|\Sigma|^{1/2}} \frac{(2\pi)^{(d-|C|-1)/2}}{(2\pi)^{(d-1)/2}} \\ & \times \exp\left\{ -\frac{1}{2} \left[(\boldsymbol{x} - \boldsymbol{\beta})^T A (\boldsymbol{x} - \boldsymbol{\beta}) - (\boldsymbol{x}_{D\setminus C} - \boldsymbol{\beta}_{D\setminus C})^T A_{D\setminus C} (\boldsymbol{x}_{D\setminus C} - \boldsymbol{\beta}_{D\setminus C}) \right] \right\} \end{aligned} \tag{A.4}$$

with

$$A_{D\setminus C} = \Sigma_{D\setminus C}^{-1} - \frac{\Sigma_{D\setminus C}^{-1} \mathbf{1}\mathbf{1}^T \Sigma_{D\setminus C}^{-1}}{\mathbf{1}^T \Sigma_{D\setminus C}^{-1} \mathbf{1}}.$$

Firstly note that

$$e^{\max_{j\in D\setminus C} x_j} = e^{\max(\boldsymbol{x})}$$

as the maximum will not occur among the censored components. By a completion of the square it can be shown that expression (A.4) is in fact equal to

$$\frac{(2\pi)^{(|C|-d)/2}}{|\Gamma|^{1/2}} \exp\left\{-\frac{1}{2}(\boldsymbol{x}_C - \boldsymbol{\mu})^T \Gamma^{-1} (\boldsymbol{x}_C - \boldsymbol{\mu})\right\}$$

with

$$\boldsymbol{\mu} = -(K_C^T A K_C)^{-1} K_C A K_{D\setminus C}(\boldsymbol{x}_{D\setminus C} - \boldsymbol{\beta}_{D\setminus C})$$

and

$$\Gamma = (K_C^T A K_C)^{-1},$$

where K_C (respectively $K_{D\setminus C}$) is a $d \times |C|$ [respectively $d \times (d - |C|)$] matrix of 0s with 1s in the $(C_k, l)^{\text{th}}$ position, for C_k the kth index in C and $k = 1, \dots, |C|$, $l = 1, \dots, |C|$ (similarly for $K_{D\setminus C}$). Therefore equation (A.3) resolves as

$$h_{D\setminus C}(\boldsymbol{x}_{D\setminus C})\, \Phi_{|C|}(\boldsymbol{v}_C - \boldsymbol{\beta}_C; \boldsymbol{\mu}, \Gamma)$$

with $\Phi_{|C|}(\,\cdot\,; \boldsymbol{\mu}, \Gamma)$ the cdf of a $|C|$-variate multivariate Gaussian distribution with location vector $\boldsymbol{\mu}$ and covariance matrix Γ.

Case $f_{\boldsymbol{U}} = f_{\boldsymbol{V}}$. Again this can be found similarly to the above noting the relation between these two forms; see also Wadsworth and Tawn (2014).

Generators with structured components

Recall that since this is a model on the random vector $\boldsymbol{R}$, we need to differentiate between $\boldsymbol{\gamma} = \mathbf{0}$ and $\boldsymbol{\gamma} > \mathbf{0}$. We present the case $\boldsymbol{\gamma} = \mathbf{0}$ only, since the case $\boldsymbol{\gamma} > \mathbf{0}$ is very similar. Moreover, we set $\boldsymbol{v} = v\mathbf{1}$ as in Section 6.2.

Case $\boldsymbol{\gamma} = \mathbf{0}$. The censored likelihood has an analytical expression but is tedious to write down. Note that, since the density $h(\boldsymbol{x}; \mathbf{1}, \mathbf{0})$ is non-zero only for $x_1 < \dots < x_d$, we censor in $|C| = k$ components if $x_1 < \dots < x_k < v < x_{k+1} < \dots < x_d$. If $k = 1$, then for $\mathbb{1}(v < x_2 < \dots < x_d)$ and $\mathbb{1}(x_d > 0)$,

$$h_C(x_{2:d}, v; \mathbf{1}, \mathbf{0}) = \frac{d! \prod_{j=1}^d \lambda_j}{\sum_{j=1}^d \lambda_j^{-1}} \int_{-\infty}^{v} \frac{\prod_{j=1}^d e^{x_j}}{\left(\sum_{j=1}^d (\lambda_j - \lambda_{j+1}) e^{x_j}\right)^{d+1}}\, \mathrm{d}x_1$$

$$= \frac{(d-1)! e^{\sum_{j=2}^{d} x_j} \prod_{j=1}^{d} \lambda_j}{(\lambda_1 - \lambda_2) \sum_{j=1}^{d} \lambda_j^{-1}} \Bigg\{ \left(\sum_{j=2}^{d} (\lambda_j - \lambda_{j+1}) e^{x_j} \right)^{-d} - \left((\lambda_1 - \lambda_2) e^{v} + \sum_{j=2}^{d} (\lambda_j - \lambda_{j+1}) e^{x_j} \right)^{-d} \Bigg\},$$

where $x_{2:d} = (x_2, \ldots, x_d)$. Expressions for $k > 1$ follow naturally by repeated integration of the above result.

B Supporting information for Section 6

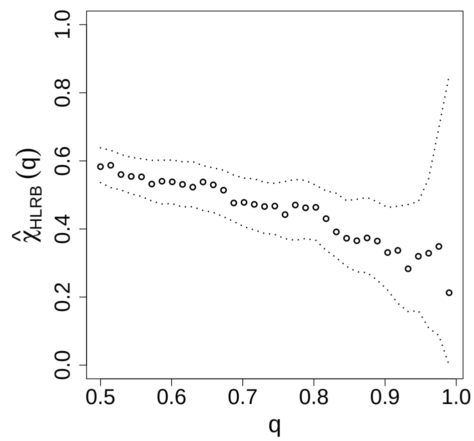

Figure 1: Negative UK bank returns: estimate of $\chi_{HLRB}(q)$ with HSBC (H), Lloyds (L), RBS (R) and Barclays (B). Approximate 95% pointwise confidence intervals are obtained by bootstrapping from $\{\boldsymbol{Y}_t : t = 1, \ldots, n\}$.

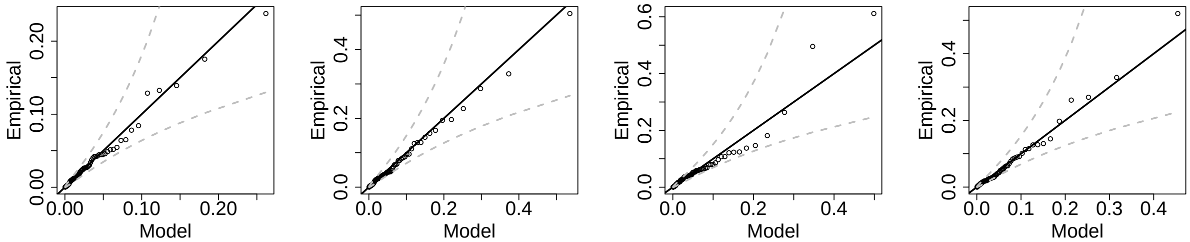

Figure 2: Negative UK bank returns: marginal QQ-plots using the fitted GP distribution. From left to right: HSBC, Lloyds, RBS and Barclays. The 95% pointwise confidence intervals are obtained by a transformation of the beta distributed order statistics of a uniform distribution.

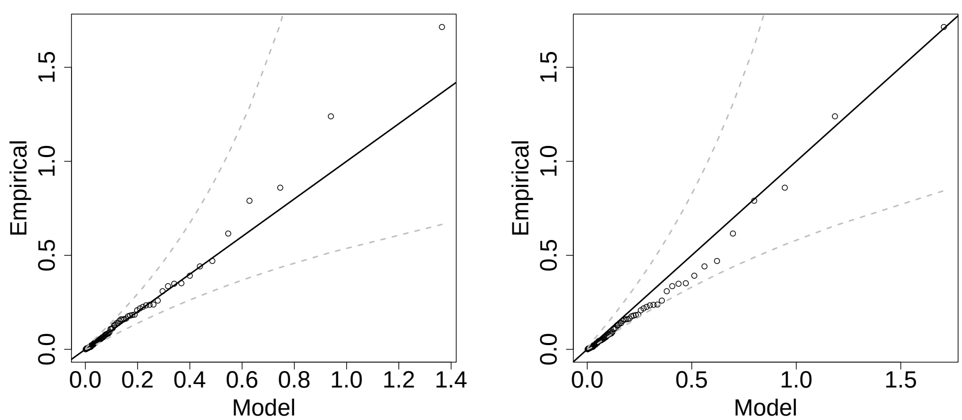

Figure 3: Negative UK bank returns: QQ-plots for GP distribution fitted by maximum likelihood to (6.1) (left) and for GP distribution with scale and shape parameter determined by the multivariate fit and Proposition 5.7 of Rootzén et al. (2016) (right). The 95% pointwise confidence intervals are obtained by a transformation of the beta distributed order statistics of a uniform distribution.

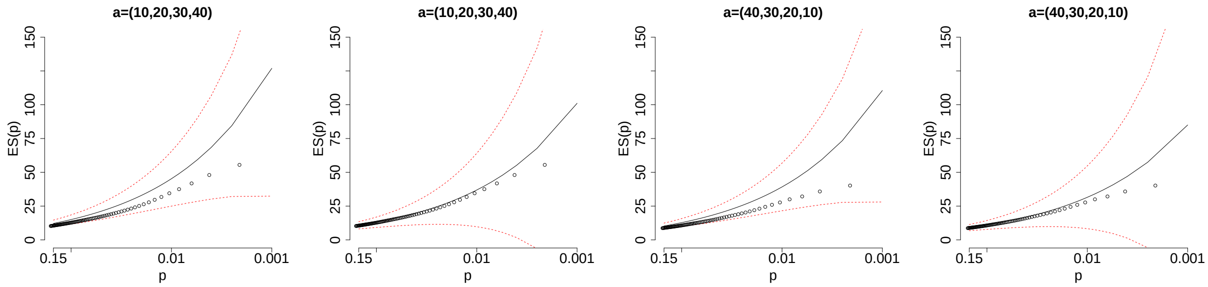

Figure 4: ES estimates and pointwise 95% delta-method confidence intervals for portfolio losses based on the weights given as the figure title. Estimates based on the multivariate GP fit are on the left; estimates based on the univariate fit are on the right.

C Supporting information for Section 7

C.1 Time trend and marginal QQ plots

We investigate the question whether there is a trend in the daily, two-day or three-day rainfall amounts by fitting a univariate GP distribution with a fixed shape parameter γ but a loglinear trend for the scale parameter, $\log \sigma(t) = a + bt$ for $t \in (0, 1]$, to the marginal components of the series $(\boldsymbol{Y}_i)_{i=1}^n$. To this end, we need to select marginal thresholds above which we fit the univariate GP distributions. For the first component, we take $u_1 = 12$ as found previously; for the second and third components, we take $u_2 = 13.5$ and $u_3 = 14$ respectively, based on inspection of parameter stability plots. For the first component, the time t corresponds to the indices $i \in \{1, \ldots, N\}$ for which $P_i > u_1$; for the second and third component, we use the time corresponding to $\max(P_i, P_{i+1})$ and $\max(P_i, P_{i+1}, P_{i+2})$.

In Table 1, we report the parameter estimates for the univariate GP fit above these thresholds. The final line shows the deviance, i.e., -2 times the difference in log-likelihood with respect to a model with $\sigma(t) \equiv \sigma$. We compare to the 95% quantile of a χ_1^2 distribution, given by 3.84. Likelihood ratio tests show that the absence of a linear trend in the logarithm of the scale parameter cannot be rejected.

Table 1: Precipitation data in Abisko: estimates of the parameters of marginal GP models with $\log \sigma(t) = a + bt$ and shape γ for thresholds $u = 12$, $u = 13.5$ and $u = 14$ respectively; standard errors in parentheses.

	$\boldsymbol{Y}_{i1}$	$\boldsymbol{Y}_{i2}$	$\boldsymbol{Y}_{i3}$
$\widehat{\gamma}$	-0.09 (0.06)	-0.05 (0.06)	-0.03 (0.06)
$\widehat{a}$	2.01 (0.12)	2.13 (0.11)	2.21 (0.11)
$\widehat{b}$	0.24 (0.21)	0.24 (0.19)	0.21 (0.17)
deviance	1.17	1.49	1.46

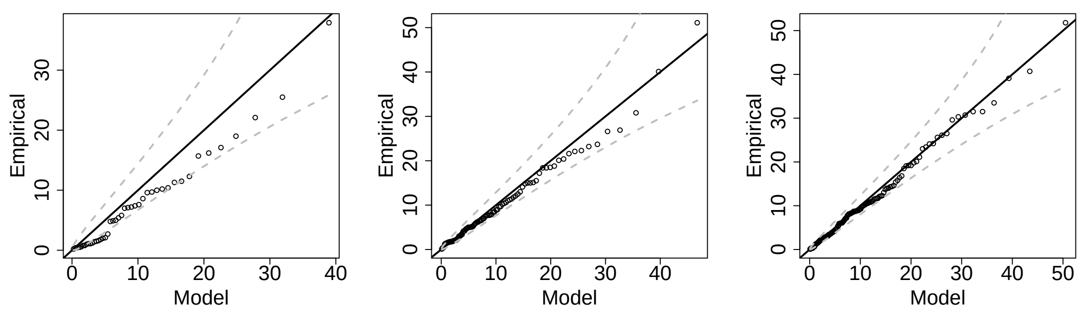

Figure 5: Precipitation data in Abisko: QQ-plots for the univariate GP distributions of the three variables Y_{i1}, Y_{i2} and Y_{i3} (left to right) for the model with $\gamma = 0$ with parameters implied by Table 4. The 95% pointwise confidence intervals are obtained by a transformation of the beta distributed order statistics of a uniform distribution.

C.2 Pairwise and trivariate χ

For the three-dimensional structured components model fitted in Section 6.2, the dependence measures χ_{12}, χ_{13}, χ_{23} and χ_{123} are

$$\chi_{12} = 1 - \frac{\lambda_1}{2(\lambda_1 + \lambda_2)}$$

$$\chi_{13} = 1 - \frac{\lambda_1(\lambda_2 + \lambda_3)^3}{(\lambda_3 + 2\lambda_2)(\lambda_2 + 2\lambda_3)(\lambda_2\lambda_3 + \lambda_1\lambda_3 + \lambda_1\lambda_2)},$$

$$\chi_{23} = 1 - \frac{\lambda_1\lambda_2(\lambda_1 + \lambda_2)^2}{(\lambda_1 + 2\lambda_2)(\lambda_2 + 2\lambda_1)(\lambda_2\lambda_3 + \lambda_1\lambda_3 + \lambda_1\lambda_2)},$$

$$\chi_{123} = 1 - \frac{\lambda_1}{2(\lambda_1 + \lambda_2)} - \frac{\lambda_1\lambda_2(4\lambda_1\lambda_2 + \lambda_1\lambda_3 + 3\lambda_2^2 + \lambda_2\lambda_3)}{3(2\lambda_1 + \lambda_2)(2\lambda_2 + \lambda_3)(\lambda_1\lambda_2 + \lambda_1\lambda_3 + \lambda_2\lambda_3)}.$$

Some properties of the structured components model can be inferred from the expression for χ_{12}; when $\lambda_1 = \lambda_2$, then $\chi_{12} = 0.75$ regardless of the value of the parameter. If $\lambda_1 \gg \lambda_2$, then $\chi_{12} \to 0.5$; if $\lambda_2 \gg \lambda_1$, then $\chi_{12} \to 1$. It is natural that this model cannot approach asymptotic independence, since it is based on cumulative sums.

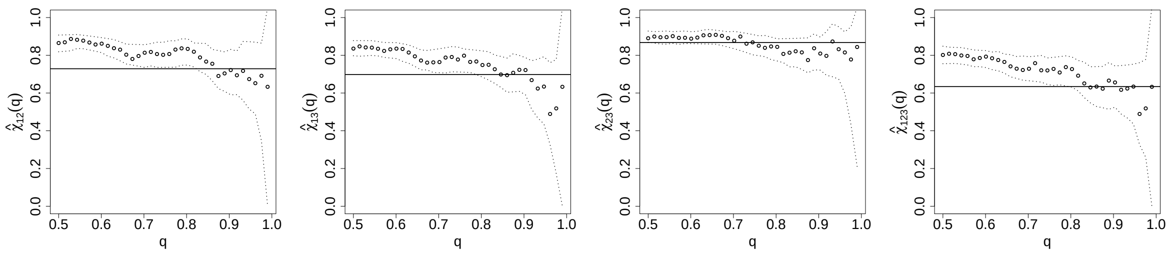

Figure 6: Precipitation data in Abisko: pairwise and trivariate $\widehat{\chi}(q)$ (dots) and model-based limiting χ (horizontal lines) for $u = 24$, with parameters implied by Table 4 for $\boldsymbol{\gamma} = \mathbf{0}$. Approximate 95% pointwise confidence intervals are obtained by bootstrapping from $\{\boldsymbol{Y}_i : i = 1, \ldots, \boldsymbol{Y}_n\}$.

References

Rootzén, H., Segers, J., and Wadsworth, J. L. (2016). Multivariate peaks over thresholds models. Available at http://arxiv.org/abs/1603.06619.

Wadsworth, J. L. and Tawn, J. A. (2014). Efficient inference for spatial extreme-value processes associated to log-Gaussian random functions. *Biometrika*, 101(1):1–15.