Multivariate Threshold Models with Applications to Wind Speed Data

Ph.D. Thesis Summary

Pál Rakonczai

Supervisor:

András Zempléni, Associate Professor, CSc

DOCTORAL SCHOOL OF MATHEMATICS

Director: Professor Miklós Laczkovich, member of HAS

DOCTORAL PROGRAMME OF APPLIED MATHEMATICS

Director: Professor György Michaletzky, DSc

EÖTVÖS LORÁND UNIVERSITY

Faculty of Science, Institute of Mathematics

Department of Probability Theory and Statistics

Budapest, 2012

1 Introduction

Subject. In this thesis I investigate and develop methods for joint modeling of extremely high values (extremes) of multivariate observations. The thesis consists of a short theoretical overview of the multivariate extreme value theory and a detailed presentation of my own scientific contribution during the recent 5 years.

Motivation. The initial motivation for the research was established within the framework of an applied statistical project[1] called "Applied stochastic models for ocean engineering, climate and safe transportation", where modeling simultaneously appearing high wind speeds - monthly maxima or exceedances over high thresholds - at different sites was one of the main research objective. The scope of the presented methods is much wider though, see e.g. the M.Sc. thesis[2] of Krusper (2011) for actuarial and financial applications or the presentation of Zempléni and Rakonczai (2011) for hydrological applications.

Impact. The most important impact of this thesis is in statistical inference and applications, providing useful material for practitioners of various disciplines working on extreme values, but simultaneously I present some theoretical considerations about certain model properties and conditions of applicability.

Programming. I devoted substantial work to the development of a new `R` software package called `mgpd` which has been first published in Rakonczai (2011) for modeling bivariate exceedances. It is freely available at the `Comprehensive R Archive Network` as a part of the `Distributions` task. It is now widely used by researchers of the field, and as the maintainer of the package I develop it following the demands of the users. The practical applications have been made by this package as well and all of them are easy to reproduce by running the cited code parts.

Outline. I outline the main probabilistic results providing the basis of modeling multivariate extremes in section 2. In section 3, I propose an universal method of constructing families of asymmetric dependence structures illustrated by some examples. At the end of this section I also introduce a novel statistical tool called *autocopula*. In section 4 goodness-of-fit and simulation methods are shown first, and then I show some simulation studies approving the plausibility of the proposed models and methods. Finally, numerous 2 dimensional (2D) and 3 dimensional (3D) applications for wind data can be found in section 5, and further promising ideas in section 6.

2 Extreme Value Theory

After recalling the necessary notations and definitions I prove an interesting invariance property of the multivariate generalized Pareto distribution (MGPD). Then I show that in the proposed MGPD framework the class of eligible dependence models is rather limited, and characterize the circumstances by an easily verifiable property. Beyond the finite dimensional models, the infinite

[1] The research visit was granted by Lund University, Sweden, see website http://www.maths.lth.se/seamocs for further details.

[2] My research leading to this thesis was partially supported by the TÁMOP project 4.2.1./B – 09/KMR – 2010 – 0003.

dimensional generalization is also discussed. I show which subclasses of the max-stable processes satisfy the applicability condition and which do not.

2.1 Preliminaries

Let us denote by $\mathbf{M}_n$ the componentwise maxima of d-dimensional observations. If there exist $\mathbf{a}_n$ and $\mathbf{b}_n > \mathbf{0}$ sequences of normalizing vectors, such that

$$P\Big(\frac{\mathbf{M}_n - \mathbf{a}_n}{\mathbf{b}_n} \le \mathbf{z}\Big) \to \mathbf{G}(\mathbf{z}), \tag{1}$$

where the G_i margins of the limit distribution are non-degenerate, then $\mathbf{G}$ is called multivariate extreme value distribution (MEVD). For the characterization of MEVD it is very common to use unit Fréchet margins, let it be denoted by $\mathbf{G}_\star$. The representation in the bivariate case (BEVD) using exponent measure $\mu_\star$, exponent measure function $V_\star$, spectral measure W or dependence function A - we call all of these later as the dependence structure - is the following (see de Haan and Resnick, 1987 and Pickands, 1981)

$$-\log G_\star(y_1, y_2) = \mu_\star([\mathbf{0}, \infty) \setminus [\mathbf{0}, \mathbf{y}]) = V_\star(y_1, y_2) \tag{2}$$

$$= \int_{S_2} \bigwedge_{j=1}^{2} \{\omega_j \log G_j(y_j)\} W(d\omega) = \left(\frac{1}{y_1} + \frac{1}{y_2}\right) A\left(\frac{y_1}{y_1 + y_2}\right). \tag{3}$$

The dependence function A is necessarily convex and satisfies $(1-t) \vee t \le A(t) \le 1$ for $t \in [0,1]$, or equivalently the spectral measure W of $V_\star$ satisfies $\int_{S_2} \omega_j W(d\omega) = 1$ for $j = 1, 2$, where S_2 is the 2 dimensional unit simplex.

2.2 MGPD models and conditions for density

Definition 1 (Rootzén and Tajvidi, 2006) *A distribution function H is a multivariate generalized Pareto distribution (MGPD) if*

$$H(\mathbf{x}) = \frac{-1}{\log G(\mathbf{0})} \log \frac{G(\mathbf{x})}{G(\mathbf{x} \wedge \mathbf{0})} \tag{4}$$

for some MEVD G with non-degenerate margins and with $0 < G(\mathbf{0}) < 1$.

We call it BGPD in the bivariate and TGPD in the trivariate cases later. Let $\mathbf{X}$ be a random vector with distribution function F, $\{\mathbf{u}(t) : t \in [1, \infty)\}$ a d-dimensional curve starting at $\mathbf{0}$ and $\sigma(\mathbf{u}) = \sigma(\mathbf{u}(t)) > 0$ be a function. Then the normalized exceedances at level $\mathbf{u}$ can be defined as

$$\mathbf{X}_\mathbf{u} = \frac{\mathbf{X} - \mathbf{u}}{\sigma(\mathbf{u})}.$$

The link between MEVD and MGPD is the following.

Theorem 1 (Rootzén and Tajvidi, 2006) *Suppose, that G is a d dimensional MEVD with $0 < G(\mathbf{0}) < 1$. If F is in the domain of attraction of G then there exist an increasing continuous curve $\mathbf{u}$ with $F(\mathbf{u}(t)) \to 1$ as $t \to \infty$, and a function $\sigma(\mathbf{u}) > \mathbf{0}$ such that*

$$P(\mathbf{X}_\mathbf{u} \le \mathbf{x} | \mathbf{X}_\mathbf{u} \not\le \mathbf{0}) \to \frac{-1}{\log G(\mathbf{0})} \log \frac{G(\mathbf{x})}{G(\mathbf{x} \wedge \mathbf{0})} \tag{5}$$

as $t \to \infty$, for all $\mathbf{x}$.

An interesting property of the MGPD is presented below.

Theorem 2 (Rakonczai and Turkman, 2012) *For any (X_1, X_2) random vector of BGPD there is a continuous, increasing curve ϱ starting from point $(0, 0)$ for which the value of distribution function $H(x_1, x_2)$ at $(x_1, x_2) \in \varrho$ is invariant under changing the underlying dependence structure (see Equation 2).*

Moreover, I show that analogous considerations are valid for arbitrary dimension as well. The following statement is useful when constructing absolutely continuous MGPD models from a known absolutely continuous MEVD:

Theorem 3 (Rakonczai and Zempléni, 2011) *Let H be a MGPD represented by an absolutely continuous MEVD G having spectral measure W. H is absolutely continuous $\Leftrightarrow W(int(\mathcal{S}_d)) = d$ holds, i.e. all mass is put on the interior of the simplex. In the bivariate case, this is also equivalent with $W(\{0\}) = W(\{1\}) = 0$ and alternatively with*

$$-A'(0) = A'(1) = 1, \tag{6}$$

where A' is the derivative of the Pickands dependence function.

2.3 Infinite dimensional EVT

Max-stable processes (de Haan, 1984 and Vatan, 1985) are an infinite-dimensional generalization of extreme value distributions.

Definition 2 (de Haan, 1984) *Let T be an index set and $\{Y_i(t)\}_{t\in T}$, $i = 1, \ldots, n$ be n independent replications of a continuous stochastic process. Assume that there are sequences of continuous functions $a_n(t) \in \mathbb{R}$ and $b_n(t) > 0$ such that*

$$Z(t) = \lim_{n\to\infty} \frac{\bigvee_{i=1}^{n} Y_i(t) - a_n(t)}{b_n(t)} \text{ for any } t \in T.$$

If this limit exists, the limit process $Z(t)$ is a max-stable process.

The following theorems are useful when modeling threshold exceedances of processes.

Theorem 4 (Rakonczai and Turkman, 2012) *The BGPD model defined by the Smith max-stable process, i.e. having exponent measure as*

$$\begin{aligned} V_\star(z_1, z_2) &= \frac{1}{z_1}\Phi\left(\frac{m(h)}{2} + \frac{1}{m(h)}\log\left(\frac{z_2}{z_1}\right)\right) \\ &+ \frac{1}{z_2}\Phi\left(\frac{m(h)}{2} + \frac{1}{m(h)}\log\left(\frac{z_1}{z_2}\right)\right), \end{aligned} \tag{7}$$

is absolutely continuous.

Theorem 5 (Rakonczai and Turkman, 2012) *The BGPD model defined by the Schlather max-stable process, i.e. having exponent measure as*

$$V_\star(z_1, z_2) = \frac{1}{2}\left(\frac{1}{z_1} + \frac{1}{z_2}\right)\left(1 + \left[1 - 2(\varrho(h) + 1)\frac{z_1 z_2}{(z_1 + z_2)^2}\right]^{1/2}\right) \tag{8}$$

is not absolutely continuous.

3 Parametric families and extensions

In order to resolve the problem of the limitation of absolutely continuous MGPD models I propose a new construction method (see Rakonczai and Zempléni, 2011) which makes it possible to build further dependence models based on the existing ones. I introduce some new parametric families using the above procedure for the bivariate case (see Rakonczai, 2012) and for higher dimensions, then investigate their properties. Finally, I extend the use of copulas to the interdependence structure of time series.

3.1 Construction of new asymmetric models in 2D

Recalling that if $-A'(0) = A'(1) = 1$ is fulfilled, then the correspondent BGPD is absolutely continuous, we are looking for a suitable transformation $\Upsilon(x)$ following the next algorithm:

1. Take a parametric **baseline dependence model** satisfying the extra condition in Equation 6 as well (e.g. symmetric logistic/negative logistic/mixed, etc.);
2. Take a strictly monotonic **transformation** $\Upsilon(x) : [0,1] \to [0,1]$, such that $\Upsilon(0) = 0$, $\Upsilon(1) = 1$;
3. Construct a **new dependence model** from the baseline model $A_\Upsilon(t) = A(\Upsilon(t))$;
4. **Check** whether A_Υ is still a valid dependence function satisfying even the extra condition of Equation 6.

For the construction it is natural to assume that it has a form of $\Upsilon(t) = t + f(t)$. By choosing f such that $f'(0) = f'(1) = 0$, Equation 6 remains automatically fulfilled. although the boundary condition and convexity need to be checked. There are two examples shown. We can consider f as

$$f_{\psi_1,\psi_2}(t) = \psi_1(t(1-t))^{\psi_2}, \text{ for } t \in [0,1], \tag{9}$$

where $\psi_1 \in \mathbb{R}$ and $\psi_2 \geq 1$ are asymmetry parameters. Or we can look for f function having two parameters ϕ_1, ϕ_2 which satisfies the following constraints (Rakonczai, 2012):

- $f_{\phi_1,\phi_2}(0) = f_{\phi_1,\phi_2}(1) = 0$ and $f'_{\phi_1,\phi_2}(0) = f'_{\phi_1,\phi_2}(1) = 0$,
- $f_{\phi_1,\phi_2}\left(\frac{1}{\phi_2}\right) = 0$, $f_{\phi_1,\phi_2}\left(\frac{1}{2\phi_2}\right) = \phi_1$ and $f_{\phi_1,\phi_2}\left(\frac{\phi_2+1}{2\phi_2}\right) = -\phi_1$,

e.g. by using Hermite interpolation.

3.2 New asymmetric models in higher dimensions

The same idea of subsection 3.1 can be applied successfully in higher than 2 dimensional cases as well. Finding a feasible baseline model and a transformation preserving the necessary constraints of a (multivariate) dependence function may lead to a more flexible class of dependence structures. We may check convexity by the usual second-order condition. Expanding the idea of the $\Psi-$transformation in Equation 9 we may consider $\Psi(t_1, \ldots, t_{d-1})$ as

$$\left(t_1 + \psi_{1,1}\left(t_1\left[1 - \sum_{i=1}^{d-1} t_i\right]\right)^{\psi_{1,2}}, \ldots, t_{d-1} + \psi_{d-1,1}\left(t_{d-1}\left[1 - \sum_{i=1}^{d-1} t_i\right]\right)^{\psi_{d-1,2}}\right).$$

This approach is mentioned as subject of future work in Rakonczai and Zempléni (2011), in the thesis I applied it for 3D wind speed data. Functions for fitting these new models are published in the latest release of `mgpd`.

3.3 From Copulas to Autocopulas

To the analogy of the autocorrelation function I introduce the autocopulas for the lagged series (see Rakonczai et al., 2011), which can reveal the specifics of the dependence structure by using copula goodness-of-fit tests (see Rakonczai and Zempléni, 2007) in a much finer way. The main theorem about copulas is proven by Sklar (1959) who has shown that for any continuous d-variate distribution function H, with univariate margins F_i there exits a unique copula C, a distribution over the d-dimensional unit cube, with uniform margins, such that $H(x_1, ..., x_d) = C(F_1(x_1), ..., F_d(x_d))$. Applications of the following definition can be found at the end of the next section.

Definition 3 (Rakonczai et al., 2011) *Given a strictly stationary time series Y_t and $l \in \mathbb{Z}^+$ the l-lag autocopula $C_{Y,l}$ is the copula of the bivariate random vector (Y_t, Y_{t-l}). The l-lag autocopulas as the function of the lag l give the autocopula function.*

4 Goodness-of-Fit and Simulation

Here I propose to use the basic concept of χ^2-test for testing bivariate models, with a slight modification of allocating cells for the statistics by certain density curves instead of quantiles. I use this method intensively in next sections, where applications are shown. Different goodness-of-fit tests for copula models are presented as well, which are very useful for testing autocopulas defined in subsection 3.3. According to my best knowledge no methods are known for simulation from BGPD (apart from logistic) hence in the thesis I propose the use of an approximation which is easy to apply in practice. Finally, I investigate the efficiency of some models by simulation study.

4.1 Goodness-of-Fit

χ^2-Test for Prediction Regions. The construction of a compact prediction region based on bivariate density function is described in Hall and Tajvidi (2004). We can base the investigation of the models on χ^2-type test statistics on this kind of regions. These regions can be viewed as a partition of the plane. As we can calculate the expected number of observations in each of these partitions if the parameters of the model are known, we might compare the theoretical and the observed frequencies by using e.g. χ^2 statistic as a measure of distance between the chosen model and the observations.

Copula GoF Test. A specific test algorithm of Rakonczai and Zempléni (2007) which is a modified version of a basic approach presented by Genest et al. (2006) is based on univariate tests, after performing an appropriate dimension reduction procedure on the copula distribution. For this purpose we suggest the use of the so-called Kendall's transform as follows $\mathcal{K}(t) = P(H(X,Y) \le t) = P(C(F(X), G(Y))) \le t) = P(C(U,V)) \le t)$. The empirical version of $\mathcal{K}$ is denoted by $\mathcal{K}_n$. Although a closed formula for $\mathcal{K}(t)$ is only available for some specific copula families, it can be approximated by simulation from any given model The proposed test statistics are summarized in Table 1, where $(t_i)_{i=1}^m$ is an appropriately fine division of the interval $(0,1)$.

Table 1: Numerically approximated Cramér-von Mises type test statistics on the Kendall's process

Focused Regions	Test Statistics
Global	$S_1 = \frac{1}{m}\sum_{t_i\in[0+\varepsilon,1-\varepsilon]}(\mathcal{K}(\theta_n,t_i)-\mathcal{K}_n(t_i))^2$
Upper Tail	$S_2 = \frac{1}{m}\sum_{t_i\in[0+\varepsilon,1-\varepsilon]}\frac{(\mathcal{K}(\theta_n,t_i)-\mathcal{K}_n(t_i))^2}{1-\mathcal{K}(\theta_n,t_i)}$
Lower Tail	$S_3 = \frac{1}{m}\sum_{t_i\in[0+\varepsilon,1-\varepsilon]}\frac{(\mathcal{K}(\theta_n,t_i)-\mathcal{K}_n(t_i))^2}{\mathcal{K}(\theta_n,t_i)}$
Lower and Upper Tail	$S_4 = \frac{1}{m}\sum_{t_i\in[0+\varepsilon,1-\varepsilon]}\frac{(\mathcal{K}(\theta_n,t_i)-\mathcal{K}_n(t_i))^2}{\mathcal{K}(\theta_n,t_i)(1-\mathcal{K}(\theta_n,t_i))}$

4.2 Simulation studies

Comparison between logistic BEVD and BGPD. Simulations from the domain of attraction of BEVD and BGPD distribution shows that in the case of logistic dependence for medium and high level of association the BGPD estimates performs well (similar to BEVD) and there is a strong bias (both in estimates and regions) for low level of association and given sample size. Although for increased sample sizes and higher threshold levels the bias disappears, BEVD is more efficient for low correlation. Similar investigation assuming another dependence structures show that actually the asymptotic tail dependence has a crucial role in efficiency of BGPD. Low tail dependence implies slower convergence in thresholds and so requires more observations for the satisfactory estimation whereas from medium level of tail dependence the BGPD is similar to BEVD.

Standard Error of Asymmetric BGPD Estimates. Focusing on model construction presented in subsection 3.1, I have generated BGPD samples of the new families by using the approximative simulation and computed the standard error of their estimates for different baseline models by bootstrap methods. Small deviations from the known parameter values and the relatively small standard errors verified the capability of the model construction. (This also means that the simulator procedure we used had been accurate enough.)

Testing for Heteroscedasticity in AR Models. Time series having the same weak AR representation, i.e. complying the same AR model driven by uncorrelated innovations do have identical autocovariance structure. As a consequence, no test based on autocovariances can really make a distinction between an ARCH- (or GARCH) innovation driven AR (hereinafter AR-ARCH) and an i.i.d. one driven AR series. The identification of the autocopula may serve this end. I have made simulations (sample size = 500) and GoF tests checking the match of the autocopulas for $l = 1, ..., 7$ lags by the test statistics from Table 1. The null-hypothesis was that the autocopula of the sample arises from the AR model. Basically in the 70% of the cases the 1-lag autocopulas were detected as non-AR. This rate goes above 90% for $n = 1000$ and above 99% for $n = 1500$, and as expected, the efficiency of the test can be increased by the appropriate weights within the test statistics. Specially using test statistics focusing on the tails of the distribution performs better in detecting non-linearity.

5 Applications to Wind Speed Data

As an illustration of the presented methods I have investigated a wind speed dataset from 5 sites of North-Germany, namely Hamburg, Hanover, Bremerhaven, Fehmarn and Schleswig. Beyond fitting and comparing standard and novel extreme value models, I have also calculated prediction regions in order to interpret the model estimates. After a preliminary univariate investigation of the data I have fitted standard BEVD and BGPD models, and found that the fit of the models is good in general, although in many cases it looks as more tuned model constructions - including time dependence and/or asymmetry - could improve the fit substantially. Another novelty is that 3D applications have also been discussed.

Non-stationary BGPD Models. Here I investigate a non-stationary extension of the BGPD model which allows for the possibility that the characteristics of extreme events are changing over time, or depend upon the value of some other covariate (see Rakonczai et al., 2010). We focus on modeling threshold exceedances occurring for four pairs of stations. Several non-stationary BGPD models have been fitted assuming different number of time dependent parameters. Based on likelihood ratios I found that assuming temporal trends leads to significant improvement in the fit. Cross-validation results show relevant improvement in predicting quantiles.

Asymmetric BGPD and TGPD models. For three pairs of the sites there have been 6 different standard dependence model assumed in the BGPD model, namely logistic, negative logistic, Coles-Tawn (Dirichlet), bilogistic, negative bilogistic and Tajvidi (generalized symmetric logistic). Then the $\Psi-$ or $\Phi-$transformations have been considered for the logistic and negative logistic models leading to 4 new families. I found the new parameters significant in most of the cases and for all three pairs the new models have happened to be among the very bests. Similarly in the trivariate case, I have applied the Ψ-transformation for the 3D logistic and negative logistic dependence structure within the TGPD models. The computations have been made by the new updated version of the `mgpd` (1.99) `R`-package. The most important code segments are also included in the thesis.

6 Future Objectives

Finally I sketched some ideas for further research objectives, which is basically solving the same problems as above for several sites simultaneously. I show an attempt for the 5D case and an idea for spatial application for large grids.

Related publications with my contribution are the followings:

- Rakonczai, P. and Turkman, F. (2012) Applications of generalized Pareto processes. (Technical report under progress, OTKA outgoing mobility grant, Lisbon, Portugal)
- Rakonczai, P. (2012) Asymmetric dependence models for bivariate threshold exceedance models. *Forum Statisticum Slovacum, ISSN 1336-7420* **1**, p.25-32.
- Rakonczai, P. and Zempléni, A. (2012) Bivariate generalized Pareto distribution in practice: models and estimation. *Environmetrics*, John Wiley & Sons, **23**, p.219-227.

- Zempléni, A. and Rakonczai, P. (2011) New bivariate threshold models with hydrological applications. *Conference on Environmental Risk and Extreme Events*, Ascona, July 10-15
- Rakonczai, P., Márkus, L. and Zempléni, A. (2011) Autocopulas: investigating the interdependence structure of stationary time series. *Methodology and Computing in Applied Probability*, **14**, p.149-167.
- Rakonczai, P. (2011) `Package 'mgpd' manual`.
  see http://cran.r-project.org/web/packages/mgpd/mgpd.pdf
- Rakonczai, P. and Tajvidi, N. (2010) On prediction of bivariate extremes. *International Journal of Intelligent Technologies and Applied Statistics*, **3**(2), p.115-139.
- Rakonczai, P., Butler, A. and Zempléni, A. (2010) Modeling temporal trend within bivariate generalized Pareto models of logistic type. (Technical report, HPC-Europa2 Project, Edinburgh, UK, available at
  http://www.math.elte.hu/∼paulo/pdf/)
- Rakonczai, P. (2009) On Modelling and Prediction of Multivariate Extremes, with applications to environmental data. Centrum Scientiarum Mathematicarum, Licentiate Theses in Mathematical Sciences 2009:05
- Rakonczai, P., Márkus L. and Zempléni, A. (2008a) Goodness of Fit for Auto-Copulas: Testing the Adequacy of Time Series Models, *Proceedings of the 4th International Workshop in Applied Probability* CD-ROM, paper No.73., 6 pages, Compiegne, France
- Rakonczai, P., Márkus L. and Zempléni, A. (2008b) Adequacy of Time Series Models, Tested by Goodness of Fit for Auto-Copulas, *Proceedings of the COMPSTAT2008 conference*, Porto, Portugal
- Rakonczai, P. and Zempléni, A. (2007) Copulas and goodness of fit tests. *Recent Advances in Stochastic Modeling and Data Analysis*, World Scientific, Hackensack, NJ, p.198-206.
- Bozsó, D., Rakonczai, P. and Zempléni, A. (2005) Árvizek a Tiszán és néhány mellékfolyóján. *Statisztikai Szemle*, **83**(10-11), p.919-936.

References

[1] Genest, C. Quessy, J.-F. and Rémillard, B. (2006) Goodnes-of-fit Procedures for Copula Models Based on the Integral Probability Transformation. *Scandinavian J. of Statistics*, **33**, p.337-366.

[2] Hall, P. and Tajvidi, N. (2004) Prediction regions for bivariate extreme events. *Aust. N. Z. J. Stat.* **46**(1), p.99-112.

[3] de Haan, L. (1984) A spectral representation for max-stable processes. *Ann. Probab.*, **12**, p.1194-1204

[4] de Haan, L. and Pickands, J. (1986) Stationary min-stable stochastic processes. *Probability Theory and Related Fields*, **72**, p.477492.

[5] de Haan, L. and Resnick, S. I. (1987) On regular variation of probability densities. *Stochastic Processes and Their Applications*, **25**, p.83-93

[6] Krusper (2011) Többdimenziós extrém érték eloszlások alkalmazása biztosítási és pénzügyi adatokra. *MSc thesis*, Eötvös Loránd University, Budapest (in Hungarian) (http://www.cs.elte.hu/blobs/diplomamunkak/mat/2011/krusper_marta.pdf)

[7] Pickands, J. (1981) Multivariate extreme value distributions. *Bulletin of the International Statistical Institute, Proceedings of the 43rd Session*: p.859-878.

[8] Rootzén, H. and Tajvidi, N. (2006) The multivariate generalized Pareto distribution. *Bernoulli* **12**, p.917-930.

[9] Sklar, A. (1959) Fonctions de répartition á n dimensions et leurs marges. *Publications de lInstitut de Statistique de lUniversite de Paris'*, **8**, p.229231.

[10] Vatan, P. (1985) Max-infinite divisibility and max-stability in infinite dimensions. *Probability in Banach Spaces V.*, ed. A. Beck et al. Lecture Notes in Mathematics 1153, Springer, Berlin, p.400-425.