# Modelling multivariate extremes

Paul Embrechts
Department of Mathematics, ETH Zürich

Laurens de Haan and Xin Huang
Econometrisch Instituut
Erasmus Universiteit Rotterdam

## 1 Introduction

Both in insurance and finance, the stochastic modelling of extremes is of importance. Think for instance of such notions as large (catastrophic) claims, value at risk, probable maximal loss or the so–called Pareto "law" which says that for portfolios with large claims typically "20% of the claims is responsible for 80% of the total claim amount". For a discussion on the latter, see for instance Aebi, Embrechts and Mikosch [1]. A review paper on extremes in insurance and financce is Embrechts and Schmidli [9]. A comprehensive textbook treatment of extremes in insurance and finance is Embrechts, Klüppelberg and Mikosch [6], where also a very extensive list of references is to be found. See also Beirlant, Teugels and Vynckier [2] and Reiss and Thomas [24]. Within the realm of extreme value theory, most of the material published concerns the one dimensional case, the main reason being that in higher dimensional Euclidean space there is no standard notion of order, and consequently, there is no standard notion of extremes. A good discussion on possible approaches is to be found in Resnick [25]. An interesting recent review is Joe, Smith and Weisman [19] in which threshold methods play a key role, see also Sinha [26] and Pickands [22]. Both [25] and [19] give references for further reading. The main aim of our paper is to present one possible approach for modelling tail events in the multivariate case. For reasons of notational simplicity, we concentrate on two dimensions. It is important to stress at this point the fact that current multivariate extreme value theory, from an applied point of view, only allows for a treatment of fairly low–dimensional problems. The latter often suffices in insurance where two- or three–line (i.e. dimension two or three) products already are fairly advanced. In finance however, a typical investment portfolio involves several hundred if not more than a thousand instruments. The truly multivariate extreme value analysis of such problems for the moment is well outside the reach of the available theory. An approach by the second author embedding such high dimensional problems in an infinite dimensional one may prove to be useful here; see de Haan [13], de Haan and Tao Lin [16].

Suppose $F$ is the distribution function (d.f.) of the random vector $(X, Y)$, we want to discuss methods for describing $1 - F(x, y)$ with $x$, $y$ large in some sense. In most practical problems in insurance and finance one is faced with extremal events based on correlated data. Think for instance of the estimation of profit–loss distributions in finance where the underlying portfolio consists of many highly correlated financial instruments. In housing insurance, large losses can occur by a combination of various meteorological conditions (rain, wind ...) resulting in an extreme event.

The notion of tail dependence function we shall introduce is by no means the only one available, by concentrating however on one particular measure for multivariate extremes, we hope to confront the applied reader (actuary, financial analyst) with some of the important issues underlying multivariate extremes in general. The references will guide the interested reader into more extensive treatments on this subject.

## 2 A multivariate tail function

Suppose $F(x, y)$ is the d.f. of the random vector $(X, Y)$ defined on some probability space $(\Omega, \mathcal{F}, P)$. We call

$$D_F(u_1, u_2) = F\left(F_1^{\leftarrow}(u_1),\, F_2^{\leftarrow}(u_2)\right), \quad 0 \le u_1, u_2 \le 1, \tag{1}$$

the *tail dependence function* of $F$, where $F_1$ and $F_2$ are the marginal distribution functions of $X$ and $Y$, and

$$F_i^{\leftarrow}(u_i) = \inf\{y \in R : F_i(y) \ge u_i\}, \quad 0 \le u_i \le 1, \quad i = 1, 2,$$

are the *generalised inverse functions* of $F_1$, $F_2$. By definition

$$F_i^{\leftarrow}(F_i(x)) = x, \quad x \in R, \quad i = 1, 2,$$

and for $x, y \in R$,

$$F(x, y) = D_F\left(F_1(x),\, F_2(y)\right).$$

In order to avoid some non–essential technicalities, we assume that $F_1$ and $F_2$ are continuous. For a discussion on generalised inverses and their use in probability, see [25]. Through the transformation (1) we have "uniformised" the marginal distributions of $F$, indeed $D_F(u_1, u_2) = Pr(U_1 \le u_1, U_2 \le u_2)$ where $U_1$, $U_2$ are (possibly dependent) uniform $(0, 1)$ variables. This transformation is also referred to as the copula–transformation. For a full discussion on its use in financial and insurance risk management, see Embrechts, McNeil and Straumann [7, 8] and the references therein. For some recent publications in the realm of insurance, see Frees and Valdez [10], and Klugman and Parsa [21]. Though one could study $D_F$ in full generality, for applications in insurance and finance it turns out that some tail condition on the underlying $F$ is useful. The latter can be motivated either by a domain of attraction condition for stable r.v.'s or for extreme value distributions; see for instance de Haan [12], Huang [17], Geluk and de Haan [11], Bingham, Goldie and Teugels [3] and Smith [27].

**Definition 2.1** *Suppose $F$ is the d.f. of $(X,Y)$ and assume that there exists a function $l_F : [0,\infty) \times [0,\infty) \to [0,\infty)$, such that for all $x, y \geq 0$,*

$$\lim_{t\to 0} t^{-1}\left\{1 - D_F(1-tx,\,1-ty)\right\} \text{ (exists) } = l_F(x,y)\,, \tag{2}$$

*then $l_F$ is called the stable tail dependence function (STDF) of $F$.*

We shall often write $D$, $l$ instead of $D_F$, $l_F$. The following properties of $l$ hold:

$$\begin{array}{lll} \text{(L1)} & l(x,y) + l\,(x_1,y_1) \leq l\,(x,y_1) + l\,(x_1,y)\,, & x \leq x_1\,, \quad y \leq y_1\,; \\ \text{(L2)} & l(x,0) = l(0,x) = x\,, & x \geq 0\,; \\ \text{(L3)} & l(sx,sy) = sl(x,y)\,, & s,x,y \geq 0\,. \end{array}$$

The motivation behind Definition 2.1 (i.e. condition (2)) is to be found in [17]: it turns out that (2) holds whenever $F$ belongs to a so–called extreme domain of attraction. The latter amounts to the condition:

$$\lim_{n\to\infty} \left\{1 - F\left(a_n x + b_n\,,\; c_n y + d_n\right)\right\} \quad \text{exists, } \notin \{0,\infty\}\,, \tag{3}$$

for $x$, $y$ real, $a_n > 0$, $c_n > 0$ and $b_n$, $d_n$ real. For the precise link between (2) and (3), see [17], where also a more detailed discussion of (3) is given. In practice, typical situations where (3) may hold are so–called heavy–tailed (or Pareto–type) distributions where

$$1 - F_i(x) \sim C_i x^{-\alpha_i}\,, \quad x \to \infty\,, \quad \alpha_i > 0\,, \quad C_i > 0\,, \quad i = 1,2\,.$$

Throughout this paper, we shall write $\sim$ to denote asymptotic equivalence, i.e. $f(x) \sim g(x)$, $x \to \infty$, if $\lim_{x\to\infty} f(x)/g(x) = 1$. The symbol $\approx$ will be used to denote 'approximately' equal. In all cases where we use the latter symbol, we can make its meaning mathematically precise.

The homogeneity property (L3) implies that the function $l$ is determined by the level curve

$$Q_F = \left\{(x,y) \in R_+^2 : l_F(x,y) = 1\right\}\,.$$

For our purpose, the knowledge of $Q_F$ suffices. We shall refer to it as the $Q$–curve of $l$. Because of (L3) we can solve the equation

$$l(\rho\cos\theta, \rho\sin\theta) = 1$$

to obtain the following polar coordinate representation of $Q$ (again we drop the $F$ from $Q_F$):

$$Q = \left\{(\rho,\theta) \in [0,\infty) \times [0,\pi/2] : \rho = q(\theta) := \left(l(\cos\theta,\sin\theta)\right)^{-1}\right\}\,.$$

Then

$$l(x,y) = \sqrt{x^2+y^2}\; l\left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right) = \frac{\sqrt{x^2+y^2}}{q(\arctan(y/x))}\,.$$

The next theorem characterises $Q$–curves, for a proof see [17].

**Theorem 2.2** *For a given d.f. $F$, the $Q$–curve is concave ending at $(1,0)$ and $(0,1)$. Conversely, any curve satisfying the previous properties is the $Q$–curve of some d.f. $F$.*

Typical examples of $Q$–curves are given in Figure 1.

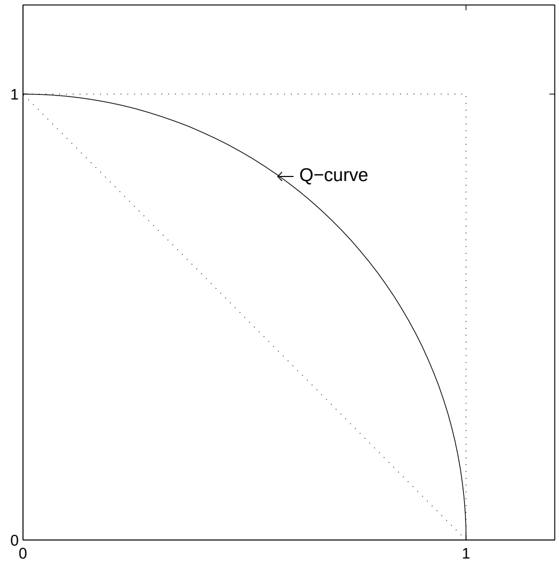


Figure 1: Typical $Q$–curve

# 3 From theory to praxis

Various concepts in extreme value theory originate from applied problems in environmental science. For an interesting series of papers on this, see for instance [20] and [23]. We will use the language of insurance and finance to indicate how the previously introduced notions of tail dependence $l$ and $Q$–curves provide useful tools in the analysis of multivariate extremes. In an insurance setting, we typically think of a two–line product $(X, Y)$ where $X$ (resp. $Y$) stands for the loss variable in the first (resp. second) line of business. Within finance, we can think of $X$ and $Y$ as representing the profit and loss variables of two financial products/portfolios/trading desks, say. For ease of discussion, we will stick to the insurance language. We therefore model the loss experience for the two lines of business by $X$, $Y$ respectively. To make the problem more concrete, we could also think of flood–damage–insurance, where $X$ and $Y$ would then correspond to the sea–level at two different locations. A catastrophic event occurs when $X$ (resp. $Y$) exceeds the dyke height $d_x$ (resp. $d_y$) at the respective locations. In the latter example, a link to the insured losses in case of a flood (at at least one of the two locations) would have to be established. Finally, one could also look for the dyke height example at the events $\{X > d_x\}$ and/or $\{Y > d_y\}$ as triggering events for a catastrophe bond, say.

We neglect time dependence in our discussion below. Suppose $F$ is the d.f. of

$(X, Y)$, then the following probabilities are of crucial interest:

$$p_1 = P(X > x) = 1 - F_1(x)$$
$$p_2 = P(Y > y) = 1 - F_2(y)$$

$$\begin{aligned} p_{12} &= P(X > x \quad \text{or} \quad Y > y) \\ &= 1 - F(x, y) \\ &= 1 - D_F\left(F_1(x), F_2(y)\right) \\ &= 1 - D_F\left(1 - p_1, 1 - p_2\right) . \end{aligned}$$

Typically, $p_1$ and $p_2$ are (should be!) very small so that we can hope to be working under the basic assumption (2), yielding the following approximation:

$$p_{12} \approx \sqrt{p_1^2 + p_2^2}\, l \left( \frac{p_1}{\sqrt{p_1^2 + p_2^2}}, \frac{p_2}{\sqrt{p_1^2 + p_2^2}} \right) = \frac{\sqrt{p_1^2 + p_2^2}}{q(\theta)}, \tag{4}$$

where $\theta = \arctan(p_2/p_1)$. Hence the calculation of the probability $p_{12}$ of the extremal event "a large (catastrophic) loss occurs in either of the two lines of business" is reduced to the calculation of the marginal probabilities $p_1$, $p_2$ (by using standard one dimensional methods) and the $Q$–curve $q(\theta)$, the latter being obtained via the tail dependence function $l$. In other words, the problem of bivariate estimation can be decomposed into two parts: one part concerns univariate tail estimation not taking dependence between the underlying variables into account, the other part makes use of the dependency. Using the function $q(\theta)$, approximations to various functionals of interest can be obtained.

Let $\tau$ for instance denote the number of lines where a catastrophic claim occurs. Using (4) we immediately get estimates for the conditional probability of such simultaneous events:

$$\begin{aligned} P(\tau = 2 \mid \tau \geq 1) &= \frac{P(X > x, Y > y)}{P(X > x \text{ or } Y > x)} \\ &= \frac{p_1 + p_2}{p_{12}} - 1 \\ &\approx (\cos\theta + \sin\theta)\, q\,(\theta) - 1 \qquad (5) \\ &= \frac{q(\theta) - q_0(\theta)}{q_0(\theta)}, \qquad (6) \end{aligned}$$

where $q_0(\theta) = (\cos\theta + \sin\theta)^{-1}$. Clearly, in (5) we have again used asymptotic estimates based on $p_1$ and $p_2$ small. Formula (6) allows for an interesting geometrical interpretation. Indeed, $q(\theta)/q_0(\theta)$ is the quotient of the distance to $(0, 0)$ from the point on the $Q$–curve corresponding to $(\theta)$ and the distance from the corresponding point on the straight line connecting $(1, 0)$ and $(0, 1)$ also to $(0, 0)$. This interpretation is made clear in Figure 2. It is precisely geometric interpretations like these which render the proposed tail dependence function $l$ and its associated $Q$–curves so elegant and useful for practical applications.

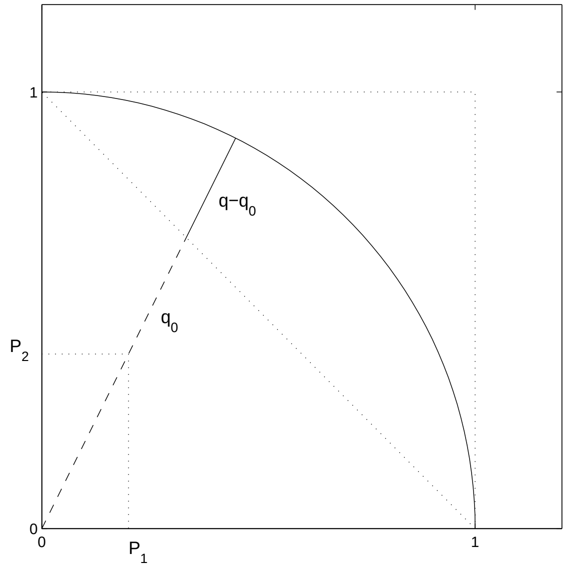


Figure 2: Conditional probability

Another quantity of interest is $E(\tau \mid \tau \geq 1)$, which is the expected number of joint catastrophic events, given that at least one such event occurs. One immediately derives that (approximately)

$$E(\tau \mid \tau \geq 1) = \frac{q(\theta)}{q_0(\theta)} \,. \tag{7}$$

Summing up (6) and (7) we can say that the conditional distribution of two simultaneous extreme events, given that at least one such event occurs, essentially depends on the ratio $p_2/p_1$ (through $\theta$) and the shape of the $Q$–curve (through $q(\theta)$). The more concave the $Q$–curve, the more likely such events will happen simultaneously. It therefore becomes interesting to use $Q$–curves as graphical measures of dangerousness, especially in comparing and contrasting various risks.

Besides the above formulae for the calculation of extreme event probabilities, one often uses these results backwards. In the dyke example, take for instance a (typically very) small threshold probability $p$; we want to minimise the combined dyke height $x + y$ (amount of sand involved for example) while keeping the flood probability $p_{12}$ equal to $p$. The crucial equation (4) becomes:

$$\sqrt{p_1^2 + p_2^2} = pq(\theta) \,.$$

We also have

$$\cos\theta \sqrt{p_1^2 + p_2^2} = p_1 \,,$$
$$\sin\theta \sqrt{p_1^2 + p_2^2} = p_2 \,.$$

Hence,

$$\begin{cases} p_1 = pq(\theta)\cos\theta \\ p_2 = pq(\theta)\sin\theta \,. \end{cases}$$

Combining this with the definition of $p_1$ and $p_2$ before relation (4) yields, for $0 \le \theta \le \pi/2$,

$$\begin{cases} x_p(\theta) = F_1^{\leftarrow}(1 - pq(\theta)\cos\theta) \\ y_p(\theta) = F_2^{\leftarrow}(1 - pq(\theta)\sin\theta) \end{cases} \tag{8}$$

therefore $\{(x_p(\theta), y_p(\theta)) : 0 \le \theta \le \pi/2\}$ is the required $1-p$ quantile curve, i.e. for all $0 \le \theta \le \pi/2$ we have

$$p_{12} = 1 - F\left(x_p(\theta), y_p(\theta)\right) = p\,.$$

The problem posed can now be solved as follows:

i) Estimate the tail of the marginal quantile functions $F_1^{\leftarrow}(1-px)$ and $F_2^{\leftarrow}(1-py)$ using univariate techniques, see for instance [4], [6] and [15].

ii) Estimate the $Q$–curve $q(\theta)$. For a discussion on this topic, see next section.

iii) Use the results from the above two steps and (8) to obtain an estimate for the quantile curve $((x_p(\theta), y_p(\theta))$.

iv) Move the line $x + y = c$, $c > 0$ by increasing $c$ till it touches the quantile line in iii) for the first time, the coordinates of the touching point being the solution to our problem. For a pictorial representation of this procedure see Figure 3 below.

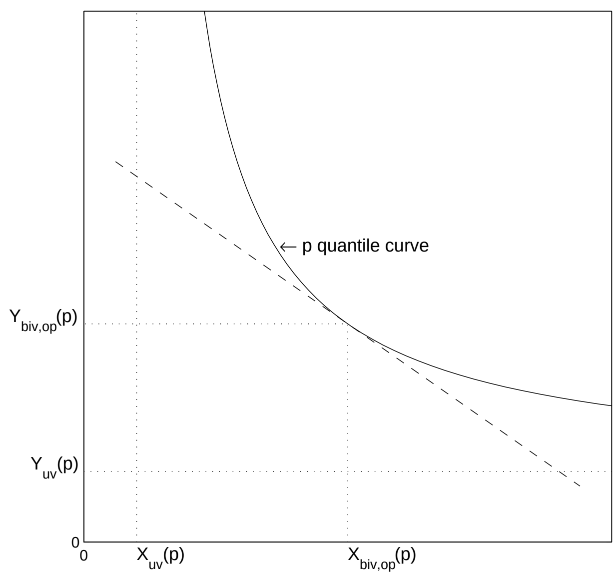


Figure 3: Optimum dyke heights at 2 places. Note that when we treat two locations together, the required dyke height for overall safety is higher than when treated separately. Something for financial people to notice as well.

## 4 Estimating the tail dependence function

Our final goal is obtaining an estimator for the $Q$–curve; since the latter is defined via the underlying d.f. $F$, a natural estimator can be based on the empirical d.f.

$$F_n(x,y) = \frac{1}{n}\sum_{i=1}^{n} I\left(X_i \le x,\ Y_i \le y\right) ,$$

where $(X_1, Y_1), \ldots, (X_n, Y_n)$ are i.i.d. r.v's with common d.f. $F$, and $I(A)$ denotes the indicator function of $A$ (taking value 1 on $A$ and 0 on $A^c$). Denote by $X_{1,n} \le \ldots \le X_{n,n}$ the order statistics of $X_1 \ldots X_n$. Let $[x]$ denote the integer part of $x$. Since

$$F_n\left(X_{[nx],n} Y_{n,n}\right) = \frac{[nx]}{n} \sim x\,, \quad 0 \le x \le 1\,, \quad n \to \infty\,,$$

we see that the empirical counter part of $F_1^{\leftarrow}(x)$ is $X_{[nx]}$ and likewise $Y_{[ny]}$ for $F_2^{\leftarrow}(y)$. This motivates the following empirical estimator for $l_F(x,y)$:

$$\begin{aligned}\hat{l}_{F,n}(x,y) &= 1 - F_n\left(X_{[n-kx],n},\ Y_{[n-ky],n}\right)\\ &= \frac{1}{k}\sum_{i=1}^{n} I\left(X_i > X_{[n-kx],n} \quad \text{or} \quad Y_i > Y_{[n-ky].n}\right)\end{aligned}$$

for $0 \le x,\ y \le 1$ and $k = k(n)$ is chosen in such a way that typically $k$ is large $(k(n) \to \infty)$, but also $k$ is small in comparison to $n$, i.e. $k(n)/n \to 0$. For a discussion of the latter condition see [17], p. 14. Using the notion of ranks, there is another way to calculate the above estimator. Let

$$R_i^x = \sum_{j=1}^{n} I\left(X_j \le X_i\right) , \qquad R_i^y = \sum_{j=1}^{n} I\left(Y_j \le Y_i\right)$$

be the ranks of $X_i$, $Y_i$, $i = 1 \ldots n$. Using this notation we immediately have that

$$X_i = X_{R_i^x,n}\,, \qquad Y_i = Y_{R_i^y,n}\,, \qquad i = 1 \ldots n\,,$$

and

$$\hat{l}_{F,n}(x,y) = \frac{1}{k}\sum_{i=1}^{n} \left(R_i^x > n - kx \quad \text{or} \quad R_i^y > n - ky\right) .$$

From this expression it immediately follows that $\hat{l}_{F,n}$ only depends on the relative order of the observation and hence is invariant under any monotone transformation of the marginals. Using the above estimator of $l_F$ we also obtain a natural estimator for the $Q$–curve:

$$\hat{q}_n(\theta) = \left(\hat{l}_{F,n}(\cos\theta, \sin\theta)\right)^{-1} , \qquad 0 \le \theta \le \pi/2\,.$$

The following is an example from hydrology, related to the flood–insurance problem briefly discussed above: see de Haan and de Ronde [14] for more details.

**Example 4.1** At a point on the northern Dutch coast, which is protected by a sea dyke called "Pettemer zeedijk", during 828 storm events spanning 13 years, two parameters were recorded, the high tide seawater levels (SWL, a kind of short term average of the water level) and wave height (HmO). Both play a role in the possible failure of the dyke. The data are represented in Figure 4. Also a failure area is indicated in Figure 4, i.e. if 0.3 HmO + SWL exceeds 7.6 meters, the dyke is in danger. The problem is to assess the probability of failure of the dyke during a future storm event.

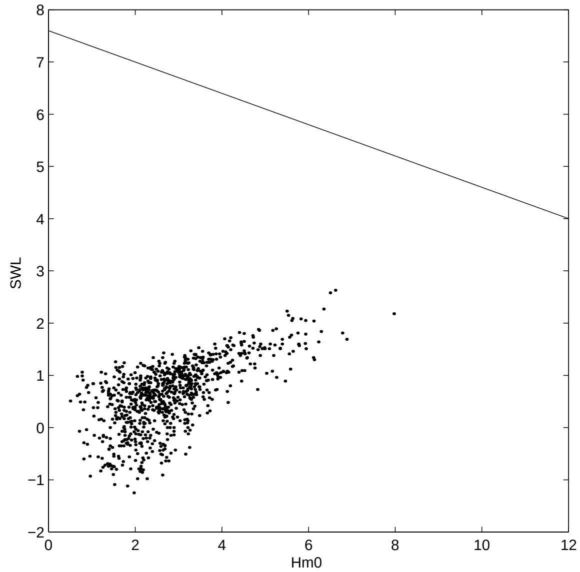


Figure 4: Wave–height (Hm0) and high–tide water level (SWL) at the Eierland station (North Sea), recorded during 828 storm events. The area above the line represents the failure area for the "Pettemer zeewering".

In order to estimate this probability, we first focus on the two marginal distributions. The tail can be modelled using an extreme value distribution. This involves estimating a shape parameter $\gamma_i$ and scale and shift constants $b_i$ and $a_i$, $i = 1, 2$ (see for example [5]). The $\widehat{\gamma}_i$'s turn out to be negative, pointing to a possible boundedness of the probability distributions. The values of $\gamma_i$, $a_i$ and $b_i$ are given in Table 1.

| parameter | HmO | SWL |
|---|---|---|
| $\gamma_1$, $\gamma_2$ | -0.0074 | -0.12 |
| $b, d$ | 553 | 169 |
| $a, c$ | 49.5 | 27.46 |

Table 1: Extreme value distribution parameter estimates

Next we turn our attention to the $Q$–curve representing the dependence in the tail. Figure 5 gives estimates of the $Q$–curve using smaller and smaller portions of the tail observations. One sees that the shapes of the various estimates are not too different (see property L3) and that the curves represent approximately convex functions (see Theorem 2.2). This inspires some confidence in our model. The curve seems to deviate somewhat from a straight line; this points to the absence of independence of the two marginals in the tail.

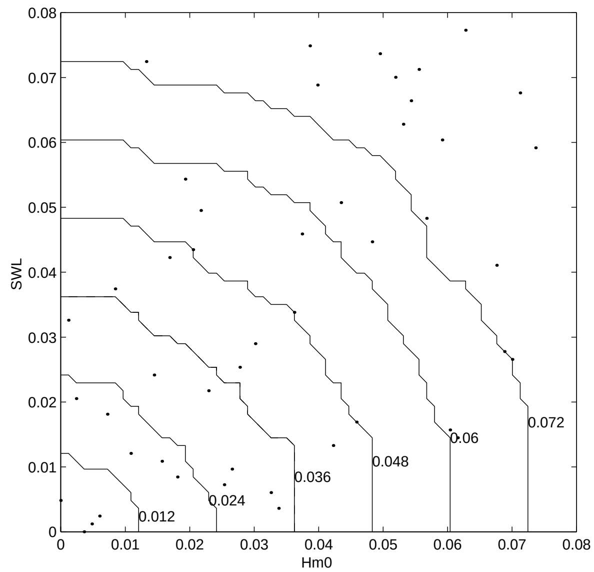


Figure 5: Q–curves for wave–height (Hm0) and high–tide water level (SWL). The axes are transformed to the empirical estimates of $1-F_i$, $i=1,2$. Curves are labeled with the empirical estimates of $1-D_f$.

In the final estimation procedure we do not use the $Q$–curve explicitly. Rather we use the homogeneity property (L3) directly: it means that if every element of a set is multiplied by $s>0$, the probability of the set is multiplied by $s$. The property is approximately true in the tail after transforming the marginal distributions to uniform ones. This is precisely what has been done in Figure 6. The observations and the failure region have been transformed to marginal uniformity. Note that the origin is now in the upper right corner.

Next the failure region is multiplied by a factor which is chosen in such a way that the boundary falls just inside the range of the observations. The result is shown in Figure 7 (note that as a result of the logarithmic scale, multiplication is now a shift). The multiplication factor used is $1.36\times 10^7$.

Finally the number of observations in the transformed failure region is counted (the number is 26).

So the estimated probability of failure is

$$\frac{1}{1.36\times 10^7}\cdot\frac{26}{828}=0.23\times 10^{-8}\,.$$

This is well below the government requirement of a failure probability of $10^{-4}$ in a year.

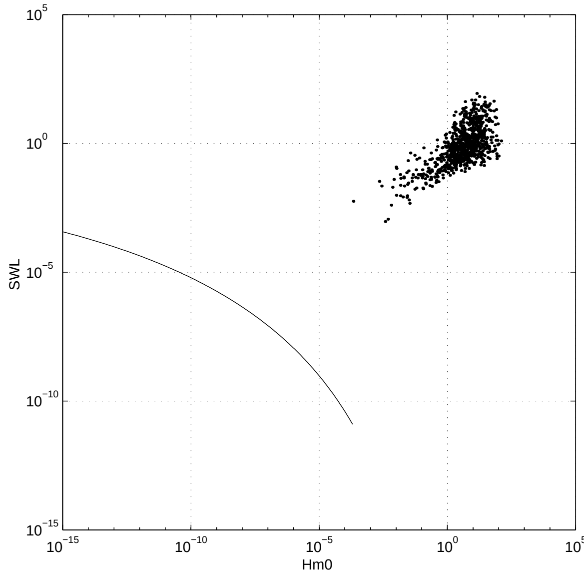


Figure 6: Data and failure region after transformation. Axes transformed are to the empirical estimates of $1 - F_i$, $i = 1, 2$.

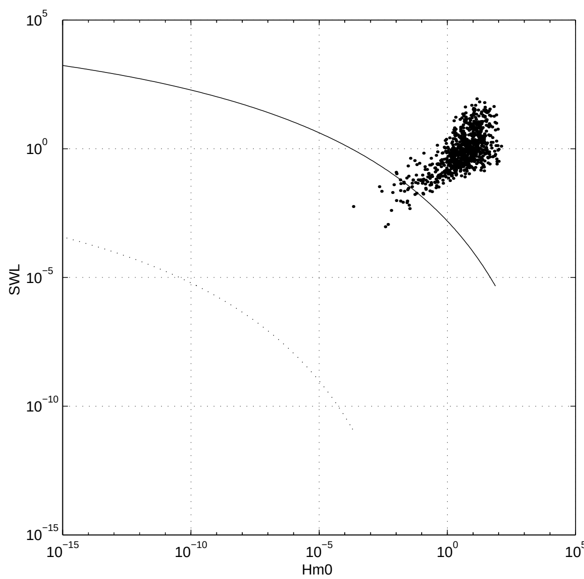


Figure 7: Data and multiplied failure region (solid line) after transformation. The original failure region is indicated by the dotted line.

# 5 Conclusion

In this paper we have discussed some of the issues underlying the theory of multivariate extremes. A brief discussion of an example from the sea dyke protection area showed how the already available theory can be used. Such analysis can be applied for instance within the construction of catastrophy bonds (alternative risk transfer). Often in these cases, the triggering event is a catastrophe expressed in terms of environmental factors. A typical example is the Tokyo Marine bond where the trigger consists of an earthquake in the Tokyo area with well defined location (place, depth) of the epicenter and a minimal strength on the Richter scale. Other applications could involve derivatives written on an extremal event of two or more underlying financial factors. A fine example of the use of these methods in finance (foreign exchange) is Starica [28]. Finally, potential applications definitely exist within the wide area of Integrated Risk Management. As already indicated in the paper, various methodological problems still have to be resolved; the core theory however is there. Future applications will guide its further development.

## Acknowledgment

The second author takes pleasure in thanking the Mathematical Research Centre (FIM) at the ETHZ for its hospitality; the third author gratefully acknowledges financial support by Swiss Re.

## References

[1] Aebi, M., Embrechts, P. and Mikosch, T. (1992) A large Claim Index. *Mitteilungen SVVM*, Heft 2.

[2] Beirlant, J., Teugels, J.L.T. and Vynckier, P. (1996) *Practical Analysis of Extreme Value.* Leuven University Press, Leuven.

[3] Bingham, N.H., Goldie, C.M. and Teugels, J.L. (1987) *Regular Variation.* Cambridge University Press, Cambridge.

[4] Dekkers, A.L.M. and Haan, L. de (1989) On the estimation of the extreme–value index and large quantile estimation. *Ann. Statist.* **17**, 1795–1832.

[5] Dekkers, A.L.M., Einmahl, J.M.J. and Haan, L. de (1989) A moment estimator for the index of an extreme–value distribution *Ann. Statist.* **17**, 1833–1855.

[6] Embrechts, P., Klüppelberg, C. and Mikosch, T. (1997) *Modelling Extremal Events for Insurance and Finance.* Springer, Berlin.

[7] Embrechts, P., McNeil, A. and Straumann, D. (1999) Correlation: Pitfalls and Alternatives. *RISK* 1999(5), 69–71.

[8] Embrechts, P., McNeil, A. and Straumann, D. (1999) Correlation and dependency in risk management: properties and pitfalls. Preprint ETH, Zürich, www.math.ethz.ch/~embrechts.

[9] Embrechts, P. and Schmidli, H. (1994) Modelling of extremal events in insurance. *ZOR–Math. Methods Oper. Res.* **39**, 1–34.

[10] Frees, W.E. and Valdez, E.A. (1998) Understanding relationships using copulas *North American Actuarial J.* **2**, 1–25.

[11] Geluk, J.L. and Haan, L. de (1987) *Regular Variation, Extensions and Tauberian Theorems.* CWI Tract **40**, Amsterdam.

[12] Haan, L. de (1970) *On Regular Variation and Its Application to Weak Convergence of Sample Extremes.* CWI Tract **32**, Amsterdam.

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

[14] Haan, L. de and Ronde, J. de (1998) Sea and Wind: Multivariate Extremes at work. *Extremes*, **1**, 7–45.

[15] Haan, L. de and Rootzén, H. (1993) On the estimation of high quantiles *J. Statist. Plann. Inference* **35**, 1–13.

[16] Haan, L. de and Tao Lin (1999) On convergence towards an extreme value distribution in $C[0,1]$. Preprint.

[17] Huang, X. (1992) *Statistics of Bivariate Extreme Values.* Thesis, Erasmus University Rotterdam, The Netherlands.

[18] Hogg, R.V. and Klugman, S.A. (1984) *Loss Distributions.* Wiley, New York.

[19] Joe, H., Smith, R.L. and Weissman, I. (1992) Bivariate threshold methods for extremes *J. Roy. Statist. Soc. Ser. B* **54**, 171–183.

[20] *Journal of Research of the National Institute of Standards and Technology* (1994). Special Issue: Extreme Value Theory, **99** (4).

[21] Klugman, S.A. and Parsa, R. (1999) Fitting bivariate loss distributions with copulas. *Insurance Math. Econom.* **24**, 139–148.

[22] Pickands, J. (1981) Multivariate extreme value distributions. *Proceedings, 43rd Session Internat. Statist. Inst. Buenos Aires*, Argentina. Book 2, 859–878.

[23] NIST special publication 866: Extreme Value Theory and Applications. U.S. Department of Commerce, 1994.

[24] Reiss, R.D. and Thomas, M. (1997) *Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and other Fields.* Birkhäuser, Basel.

[25] Resnick, S.L. (1987) *Extreme Values, Regular Variation, and Point Processes.* Springer, New York.

[26] Sinha, A.K. (1997) *Estimating Failure Probability when Failure is Rare: Multi-dimensional Case.* Tinbergen Institute Research Series 165, Erasmus University Rotterdam.

[27] Smith, R.L. (1994) Multivariate threshold methods, in J. Galambos, ed., *Extreme Value Theory and Applications*, Kluwer Academic Publishers, pp. 225–248.

[28] Starica, C. (1999) Estimation of the extreme value distribution for Constant Conditional Correlation models. *J. Empirical Finance*, to appear.