Extremes 1:1, 7–45 (1998)

Sea and Wind: Multivariate Extremes at Work

LAURENS DE HAAN ldehaan@few.eur.nl
Erasmus University Rotterdam, P.O. Box 1738, 3000 DR Rotterdam

JOHN DE RONDE j.g.dronde@rikz.rws.minvenw.nl
National Institute for Coastal and Marine management, P.O. Box 20907, 2500 EX The Hague

Abstract. Multivariate extreme value theory is used to estimate the probability of failure of a sea-wall near the town of Petten in Noord Holland, The Netherlands. The sample consists of 828 observations of still water levels and wave heights collected during storm events over a 13-year period. The paper sketches the probabilistic and statistical theory behind the estimation procedures used.

Key words. failure probability, multidimensional extremes in hydrology

Introduction

This is an account of theoretical and applied work done by the Dutch participants of the European Union project ''Neptune'' 1995–1997, grant MAS (marine science) 2-CT94-0081. The project as a whole aimed at creating a model for transferring extreme conditions in weather patterns onto sea state conditions offshore and then onto sea state conditions near shore which could threaten coastal areas. Partners were: British Maritime Technology, Lancaster University and University of East Anglia (all in the United Kingdom), National Institute for Coastal and Marine Management—RIKZ, Delft Hydraulics—WL, and Erasmus University Rotterdam—EUR (all in The Netherlands), and GKSS Forschungszentrum Geesthacht in Germany.

The project has been carried out jointly by all partners. The Dutch partners in particular have been working in close cooperation. This note reports about the work done by these people. The group consisted mainly of: J.G. de Ronde and J. van Marle (RIKZ), C.F. de Valk (WL/Argoss), D. Hurdle (WL/Alkyon), G. Draisma, L. de Haan, L. Peng and A.K. Sinha (EUR).

1. The problem

Figure 1 shows two-dimensional observations, wave heights (Hm0) and still water level (SWL) during 828 storm events spread over 13 years in front of the Dutch coast near the town of Petten. They can be considered independent and all following the same probability

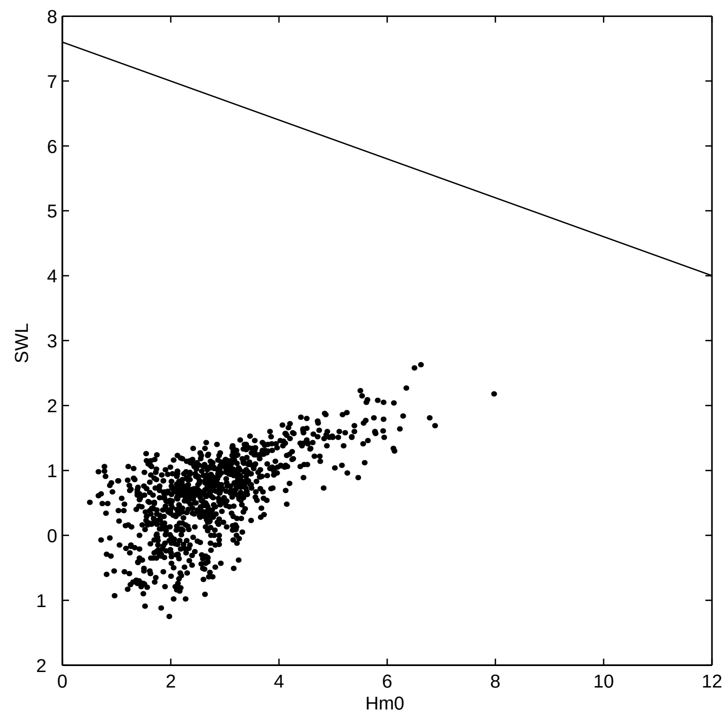

Figure 1. Wave height HmO and sea level SWL recorded during 828 storm events for the Dutch Coast. The area above the solid line represents a possible failure area.

distribution. These observations are relevant for a small stretch of sea dike that protects a gap in the natural coast protection formed by sand dunes near Petten. The dike is called ‘‘Pettemer zeedijk’’.

The area above the solid line in Figure 1 represents the failure region. Any event in that region could lead to the collapse of the dike. The failure region has been obtained by coastal engineers and depends on the structure of the dike and the adjoining part of the sea bottom.

The questions are:

- What is the probability of getting a future event in the failure region?
- Can we give a margin of reliability for our estimate of the failure probability?

The present paper will offer an asymptotic method to determine this probability. In doing so we shall also provide some insight into the theory of multidimensional extremes (probabilistic and statistical).

Some remarks about the problem can be made at this stage.

None of the observations falls in the failure area: there has not been a dangerous situation at Petten during the observation period. The design criterium being 10^{-4} per year and the observation period 13 years this is no big surprise. It means that the empirical distribution function is of no immediate use in this problem; one has to extrapolate outside the range of the observations and for this one needs smoothness conditions in the tail of the distribution.

These conditions could be parametric (i.e. one fits a completely specified parametric family of probability distributions). But since we are going to estimate the tail of the distribution, the central observations may not be of much use. So we should restrict ourselves to modeling the tail. A parametric approach to the tail is possible as we shall see, but we shall concentrate on a semi-parametric approach.

In fact it is one of the essential features of the problem as shown in Figure 1, that there is no observation near the failure region. We are going to develop an estimation procedure which is ''asymptotically correct''. In the asymptotic analysis we have to retain the feature that none of the observations is near the failure region. So, if p is the unknown failure probability, we must have $p \ll 1/n$ with n the number of observations. This means that, when n tends to infinity, p must go to zero, in particular p depends on $n : p = p_n$.

The probability of failure that we want to estimate is of the order of 10^{-4} or even less. It seems an impossible task to estimate such low probability from a sample of just 13 years. However existing theories of meteorology, hydrology and wave mechanics are insufficient to assess risk of dike failure or to model still water level, waves and wave periods over time. Since nevertheless a decision has to be reached, the only thing left is a statistical approach. Of course we are not interested in what is going to happen 10,000 years from now, but in the probability of failure ''next year''. This is how one should interpret the term ''return period''. It should also be stressed that any conclusions from the statistical analysis should be confronted with the physical conditions and modified if necessary.

2. Data and failure region definition

2.1. Data

The data set for this project was supplied by RIKZ. It contains date, time and wind and sea characteristics recorded from 1979 till 1991 at 3-hourly intervals at the Eierland station, 20 km off the Dutch coast (*offshore*). In the dataset 1090 storm events were identified (de Valk, 1994). In our contribution to the Neptune project we considered three variables: wave height, HmO, measured in meters, wave period, Tpb, measured in seconds, and still water level, SWL, measured in meters. In order to construct a set of independent observations we selected the maximum of each of the variables, recorded during each storm (note that these values were not necessarily observed at the same time). Due to missing values this resulted in a set of 828 independent observations (the cause of missing was not related to weather or sea conditions).

2.2. Failure conditions

In the Neptune project we wanted to estimate the probability of failure of the ''Pettemer zeedijk''. Failure is defined as a load (caused by onshore sea conditions) exceeding a design load (determined by the design of the dike). We considered only one failure mechanism: failure by overtopping. For this failure type a simple criterium is that more than 2% of the waves reaches the crest of the dike, in other words that the 2%-runup level exceeds the crest height (12.75 m) of the dike. This 2%-runup level is a function of the onshore sea state (wave height, period and direction and still water level). The runup ( ) function was supplied by WL (de Valk, 1994, 1996).

The onshore sea state itself is a function of the *offshore* sea state and the local topography of the sea bottom. The onshore ( ) function was also supplied by WL (Hurdle, 1995, 1996). These runup and onshore functions together determine a load function in terms of the *offshore* variables given in the dataset:

$$L(\mathrm{HmO}, \mathrm{Tpb}, \theta, \mathrm{SWL}) := \mathrm{runup}(\mathrm{onshore}(\mathrm{HmO}, \mathrm{Tpb}, \theta, \mathrm{SWL})).$$

To simplify the analysis wave direction θ was taken fixed: perpendicular to the dike (due to refraction the load is not very sensitive to offshore wave direction).

To sum up, load in terms of offshore wave height, wave period and sea water level is defined as

$$L(\mathrm{HmO}, \mathrm{Tpb}, \mathrm{SWL}) := \mathrm{runup}(\mathrm{onshore}(\mathrm{HmO}, \mathrm{Tpb}, \theta_0, \mathrm{SWL})) \tag{2.1}$$

with $\theta_0 = (2 - 5/12)\pi$, perpendicular to the dike. The corresponding failure region is defined by

$$\{(\mathrm{HmO}, \mathrm{Tpb}, \mathrm{SWL}) \mid L(\mathrm{HmO}, \mathrm{Tpb}, \mathrm{SWL}) > 12.75\}. \tag{2.2}$$

The failure area shown in Figure 1 is a further simplification. During extreme storms wave period exceeds 15 seconds and runup is almost independent of the wave period. The resulting two-dimensional failure region can be approximated by

$$\{(\mathrm{HmO}, \mathrm{SWL}) \mid 0.3\,\mathrm{HmO} + \mathrm{SWL} > 7.6\}. \tag{2.3}$$

(van Marle, 1994, 1996).

3. Outline

The problem of estimating the failure probability can only be solved within the framework of extreme value theory. If the extreme value conditions apply, then the tail of the distribution away from the data can be modeled by using the observed data. This is explained in Section 4.

In the same section it will become clear that the modeling of the two-dimensional distribution tail can be separated into modeling the tails of the two marginal distributions and modeling the dependence structure: the easiest way to study the dependence structure is to standardize the marginal tails first. The standardized tails of the marginal distributions should follow a one-dimensional extreme value distribution. Various choices are possible. We shall use the uniform distribution when estimating the dependence structure (see Section 4.1; in that case one can just use the ranks for standardization) and the distribution function $\exp -1/x$ when estimating the failure probability (see Section 4.2; in that case the tail extends to infinity which is a bit more in line with what the data suggest).

After the transformation of the marginals the tail of the distribution has approximately a very useful homogeneity property that allows us to estimate probabilities in an area where there are no observations (see Sections 4 and 5.2).

Let us say for simplicity that we want to estimate

$$P\{X > x \text{ or } Y > y\}$$

with (x, y) away from the sample in the tail (cf. Figure 1, which however depicts a different failure region, see (2.3)). First we standardize the marginal distributions. The mentioned probability then becomes

$$P\left\{\frac{1}{1-F_1(X)} > \tilde{x} \text{ or } \frac{1}{1-F_2(Y)} > \tilde{y}\right\}$$

with F_1, F_2 the marginal distribution functions and $\tilde{x} = 1/(1-F_1(x))$ and $\tilde{y} = 1/(1-F_2(y))$. Still the vector $(\tilde{x}, \tilde{y})$ is outside the range of the transformed sample (cf. figure 6, dotted line, again with the failure region (2.3)). But extreme value theory tells us that (thanks to the initial transformation)

$$P\left\{\frac{1}{1-F_1(X)} > \tilde{x} \text{ or } \frac{1}{1-F_2(Y)} > \tilde{y}\right\} \approx tP\left\{\frac{1}{1-F_1(X)} > t\tilde{x} \text{ or } \frac{1}{1-F_2(Y)} > t\tilde{y}\right\} \tag{3.1}$$

for $t > 0$ and we can choose t so small that $(t\tilde{x}, t\tilde{y})$ is indeed within the range of the sample (cf. figure 6 solid line). As a consequence the latter probability can be estimated via the empirical measure (counting the number of sample points in the set). Then also the required probability can be estimated via equation (3.1). This is the estimation method for the failure probability in a nutshell.

The dependence structure of multivariate extremes is of interest in its own right. For example we want to know if a catastrophic situation is likely to be caused by high still water level or by waves or by their joint impact. The dependence structure is estimated in two ways (see Section 4, cf. also Section 5.1) and one of the methods also serves as a check on the extreme value model.

For the first method one draws lines of equal probability for extreme quadrants (quantile curves) using the transformed data. These level curves should be convex and should have the same shape for various levels (check on the model). See for example Figure 4. Also the curves reveal the amount of asymptotic dependence which however can not be presented in one number: there is no analogue of the correlation coefficient in extreme value theory.

In the second method the so-called spectral measure is estimated. This measure serves as an (infinite-dimensional) parameter governing the dependence. Once again the amount of asymptotic dependence can be read off from the estimate of the spectral measure (Figure 5).

In Section 6 concrete estimation procedures and the resulting estimated failure probabilities and dependence structures are discussed. One alternative way of estimation is by assuming a parametric model for the spectral measure so that the entire model (including marginal distributions) can be described by just 5 parameters which can be estimated by maximum likelihood methods (cf. Section 6.2).

Special issues are addressed in Section 6.3 (three variables approach), Section 6.4 (the load variable—or how to turn a higher dimensional problem into a one-dimensional one) and Section 6.5 (the asymptotic independence issue).

4. Probability theory

In order to be able to explain the basis of the statistical procedures, we first have to look at the theory of multivariate extremes. For simplicity we shall give an exposition of the theory in $\mathbb{R}^2$. Generalization to higher dimensions is immediate.

Let $(X_1, Y_1), (X_2, Y_2), \ldots$ be a sequence of i.i.d. random vectors with probability distribution function F. Suppose that there exist sequences of constants $a_n, c_n > 0,\ b_n$ and d_n and a distribution function G with non-degenerate marginals such that

$$\begin{aligned}&\lim_{n\to\infty} \Pr\left(\frac{\max(X_1, X_2, \ldots, X_n) - b_n}{a_n} \le x, \frac{\max(Y_1, Y_2, \ldots, Y_n) - d_n}{c_n} \le y\right)\\ &= \lim_{n\to\infty} F^n(a_n x + b_n, c_n y + d_n) = G(x, y) \quad \text{weakly.}\end{aligned} \tag{4.1}$$

Of course the limit distribution is defined up to scale and shift constants, but we can choose these in a particular way as follows.

Relation (4.1) implies marginal convergence, i.e. for example

$$\lim_{n\to\infty} \Pr\left\{\frac{\max(X_1, X_2, \ldots, X_n) - b_n}{a_n} \le x\right\} = G(x, \infty)$$

and hence $G(x, \infty)$ and $G(\infty, y)$ are necessarily one-dimensional extreme-value distributions. Now we choose a_n, b_n, c_n and d_n such that

$$G(x, \infty) = \exp\{-(1 + \gamma_1 x)^{-1/\gamma_1}\} \tag{4.2}$$

$$G(\infty, y) = \exp\{-(1 + \gamma_2 y)^{-1/\gamma_2}\} \tag{4.3}$$

with γ_1 and γ_2 real parameters, the *extreme value indices*.

Clearly (4.1) implies

$$\lim_{n\to\infty} n\{1 - F(a_n x + b_n, c_n y + d_n)\} = -\log G(x, y) \tag{4.4}$$

weakly. Define a sequence of measures $v_1, v_2, \ldots$ by

$$v_n\{(s, t) \mid s > x \text{ or } t > y\} := n\{1 - F(a_n x + b_n, c_n y + d_n)\}$$

for x, y real. This relation, valid for all x and y, determines the measures $v_1, v_2, \ldots$. Relation (4.4) says that the sequence of measures v_n converges vaguely (as $n \to \infty$). This means that there must be a measure v, the limit measure, such that

$$v\{(s, t) \mid s > x \text{ or } t > y\} = -\log G(x, y) \tag{4.5}$$

for all x, y. It also follows that

$$\lim_{n\to\infty} v_n(B) = v(B)$$

for any Borel set B for which $v(\partial B) = 0$. Note that $v(B)$ is finite provided B^c contains an interval of the type $(-\infty, x) \times (-\infty, y)$ with x and y sufficiently large. But $v(\mathbb{R}^2) = \infty$. The question is: what class of measures v can occur in (4.5)?

4.1. Transformation to Fréchet marginals

In order to answer this question we apply a transformation to the marginal distributions, related to the limiting extreme-value distributions of the marginal distributions. Note that the convergence of marginals implies that for $s > 0$

$$\frac{a_{[ns]}}{a_n} \to s^{\gamma_1}, \frac{b_{[ns]} - b_n}{a_n} \to \frac{s^{\gamma_1} - 1}{\gamma_1}, \frac{c_{[ns]}}{c_n} \to s^{\gamma_2}, \frac{d_{[ns]} - d_n}{c_n} \to \frac{s^{\gamma_2} - 1}{\gamma_2}, \tag{4.6}$$

where $[x]$ is the integer part of x (cf. de Haan, 1984). Note that for any $s > 0$ by (4.4)

$$\lim_{n\to\infty} [ns]\{1 - F(a_{[ns]} x + b_{[ns]}, c_{[ns]} y + d_{[ns]})\} = -\log G(x, y). \tag{4.7}$$

Combining (4.6) and (4.7) we get by local uniformity in (4.4)

$$\lim_{n\to\infty} n\left\{1 - F\left(a_n\left[s^{\gamma_1}x + \frac{s^{\gamma_1}-1}{\gamma_1}\right] + b_n, c_n\left[s^{\gamma_2}y + \frac{s^{\gamma_2}-1}{\gamma_2}\right] + d_n\right)\right\}$$
$$= -s^{-1}\log G(x,y) \tag{4.8}$$

Or, more neatly, by substituting $(x^{\gamma 1}-1)/\gamma_1$ for x and $(y^{\gamma 2}-1)/\gamma_2$ for y in (4.8)

$$\lim_{n\to\infty} n\left\{1 - F\left(a_n\frac{(sx)^{\gamma_1}-1}{\gamma_1} + b_n, c_n\frac{(sy)^{\gamma_2}-1}{\gamma_2} + d_n\right)\right\}$$
$$= -s^{-1}\log G\left(\frac{x^{\gamma_1}-1}{\gamma_1}, \frac{y^{\gamma_2}-1}{\gamma_2}\right) \tag{4.9}$$

for $x, y, s > 0$. But a direct application of (4.4) gives

$$\lim_{n\to\infty} n\left\{1 - F\left(a_n\frac{(sx)^{\gamma_1}-1}{\gamma_1} + b_n, c_n\frac{(sy)^{\gamma_2}-1}{\gamma_2} + d_n\right)\right\}$$
$$= -\log G\left(\frac{(sx)^{\gamma_1}-1}{\gamma_1}, \frac{(sy)^{\gamma_2}-1}{\gamma_2}\right). \tag{4.10}$$

Comparing (4.9) and (4.10) we get for $x, y, s > 0$

$$-s^{-1}\log G\left(\frac{x^{\gamma_1}-1}{\gamma_1}, \frac{y^{\gamma_2}-1}{\gamma_2}\right) = -\log G\left(\frac{(sx)^{\gamma_1}-1}{\gamma_1}, \frac{(sy)^{\gamma_2}-1}{\gamma_2}\right).$$

This can be expressed more conveniently if we define

$$G_0(x,y) := G\left(\frac{x^{\gamma_1}-1}{\gamma_1}, \frac{y^{\gamma_2}-1}{\gamma_2}\right). \tag{4.11}$$

Then $G_0(x,\infty) = G_0(\infty,x) = \exp(-1/x)$ and for all $x, y, s > 0$,

$$G_0^s(sx, sy) = G_0(x,y). \tag{4.12}$$

This relation is the basic homogeneity relation of multidimensional extreme-value theory which will allow us to use semi-parametric methods in estimating the tail: write

$$A_{x,y} := \{(s,t) \mid s > x \text{ or } t > y\}. \tag{4.13}$$

If we want to estimate the probability of a set $A_{x,y}$ with (x,y) beyond the sample in the tail, we can use (4.12) and estimate instead $A_{sx,sy}$ with s so small that there are some sample points in $A_{sx,sy}$so that the empirical distribution can be used for estimating the probability of $A_{sx,sy}$. A relation like (4.5) holds for G_0 instead of G: there exists a measure v_0 such that

$$v_0(A_{x,y}) = -\log G_0(x,y). \tag{4.14}$$

The measure v_0 is called the *exponent measure* of G. From (4.12) it follows that

$$v_0(A_{sx,sy}) = v_0(sA_{x,y}) = s^{-1}v_0(A_{x,y})$$

(recall that by definition $w \in sA_{x,y}$ if and only if $w/s \in A_{x,y}$) for all $s>0$. This equality obviously extends to any Borel set: for any Borel set $B \in \mathbb{R}_+ \times \mathbb{R}_+$ and any $s>0$

$$v_0(s \cdot B) = s^{-1} \cdot v_0(B). \tag{4.15}$$

Relation (4.15) gives the exact class of measures v_0 that can occur in (4.14). In order to investigate v_0 a little more we apply (4.15) to the set

$$B_{\theta,r} = \left\{ (u,v) \,\middle|\, u^2 + v^2 > r^2, \arctan\left(\frac{u}{v}\right) \leq \theta \right\}$$

for some $\theta \in [0, \pi/2]$ and $r>0$. Since $B_{\theta,r} = rB_{\theta,1}$ relation (4.15) says that

$$v_0(B_{\theta,r}) = r^{-1} \cdot v_0(B_{\theta,1}).$$

In words: after transformation to polar coordinates the measure v_0 is a product measure. The function $v_0(B_{\theta,1})$ is in fact the distribution function of a finite measure on $[0, \pi/2]$. This measure is called the *spectral measure* of the extreme value distribution. Clearly the measure v_0 is determined by its value on all sets of the form $B_{\theta,r}$. This means that the distribution function

$$\Phi(\theta) := v_0(B_{\theta,1}) \tag{4.16}$$

of the spectral measure determines v_0, hence G_0, and hence G apart from the marginal parameters γ_1 and γ_2. So G is characterized by three quantities: γ_1, γ_2 and the spectral measure Φ. The spectral measure Φ is the exact analogue of (and in fact related to) the spectral measure of a stable distribution (see Feller, 1966, Section XVII.12). The distribution function G can be expressed in terms of γ_1, γ_2 and Φ in the following way (cf. e.g. Resnick, 1987)

$$G\left(\frac{x^{\gamma_1}-1}{\gamma_1}, \frac{y^{\gamma_2}-1}{\gamma_2}\right) = \exp\left\{-\int_0^{\pi/2}\left(\frac{\cos\theta}{x} \vee \frac{\sin\theta}{y}\right)\Phi(d\theta)\right\} \tag{4.17}$$

for $x, y>0$. This representation is unique in the sense that there is a one-to-one correspondence between G_0 and Φ. The representation is not unique in the sense that instead of the transformation to polar coordinates

$$r^2 := x^2 + y^2, \theta := \arctan\left(\frac{y}{x}\right)$$

that we used, one can choose a transformation based on any other norm, for instance

$$\begin{aligned} &r := x + y \\ &r := \max(x, y), \end{aligned}$$

(and we shall use the latter one later on). Also in the representation given, the marginal distributions of G_0 are of the form $\exp(-1/x)$ (i.e. $\gamma = 1$), but any other extreme-value distribution could be chosen as basic distribution, e.g. the ones with $\gamma = 0$ or $\gamma = -1$ (exponential, resp. uniform). Modulo these two changes the representations of de Haan and Resnick (1977), Deheuvels (1979) and Pickands (1981) (the ‘‘Pickands representation’’) are all the same. We can summarize the line of thinking by saying that after a preliminary transformation of the marginal distributions, the distribution function has an asymptotic homogeneity property. A slightly different way of obtaining the asymptotic homogeneity property is the following.

4.2. *Transformation to uniform marginals*

Assume once again that (4.1) holds. The sequences b_n and d_n can be taken as

$$\begin{aligned} b_n &:= (1 - F_1)^{\leftarrow}\left(\frac{1}{n}\right) \\ d_n &:= (1 - F_2)^{\leftarrow}\left(\frac{1}{n}\right) \end{aligned}$$

with F_1 and F_2 the two marginal distribution functions of F and the arrow indicating the left continuous inverse of a function. Then

$$\begin{aligned} &n\{1 - F(b_{[ns]}, d_{[nt]})\} \\ &\quad = n\left\{1 - F\left(b_n + a_n \frac{b_{[ns]} - b_n}{a_n}, d_n + c_n \frac{d_{[nt]} - d_n}{c_n}\right)\right\} \\ &\quad \to -\log G\left(\frac{s^{\gamma_1} - 1}{\gamma_1}, \frac{t^{\gamma_2} - 1}{\gamma_2}\right) = -\log G_0(s, t) \end{aligned}$$

as $n \to \infty$ (use (4.6)). So (by the definition of b_n and d_n)

$$\lim_{n\to\infty} n\left\{1 - F\left((1-F_1)^{\leftarrow}\left(\frac{x}{n}\right), (1-F_2)^{\leftarrow}\left(\frac{y}{n}\right)\right)\right\}$$
$$= -\log G_0\left(\frac{1}{x}, \frac{1}{y}\right) =: l(x,y), \tag{4.18}$$

or

$$\lim_{n\to\infty} n \Pr\left\{1 - F_1(X) < \frac{x}{n} \text{ or } 1 - F_2(Y) < \frac{y}{n}\right\} = l(x,y). \tag{4.19}$$

where the random vector (X, Y) has distribution function F.
The function l will play a role later on. It is a convex function. Note that the basic marginal distribution in this set-up is the extreme-value distribution with $\gamma = -1$ (in contrast to $\gamma = 1$ chosen before). This is obvious since $1 - F_1(X)$ and $1 - F_2(Y)$ have a *uniform* $(0, 1)$ distribution. More generally

$$\lim_{n\to\infty} n \Pr\left\{(1 - F_1(X), 1 - F_2(Y)) \in n^{-1}B\right\} = \tilde{\nu}_0(B) \tag{4.20}$$

where $\tilde{\nu}_0$ is the unique measure with $\tilde{\nu}_0(\{(s,t) \mid s<x \text{ or } t<y\}) = l(x,y)$ for $0<x,y<1$. Note that $\tilde{\nu}_0$ is not the same as ν_0. Again we get a homogeneity property for the measure $\tilde{\nu}_0$: for $s>0$ and any Borel set B

$$\tilde{\nu}_0(sB) = s\tilde{\nu}_0(B). \tag{4.21}$$

In particular

$$l(ax, ay) = a l(x,y) \quad \text{for } a, x, y > 0. \tag{4.22}$$

A good way to remember the role of the l-function is by combining (4.19) and (4.22): for large values of x and y we have

$$1 - F(x,y) \approx l(1 - F_1(x), 1 - F_2(y)). \tag{4.23}$$

As in the case of the measure ν_0 relation (4.21) leads to a spectral measure:

$$\tilde{\nu}_0\left\{(s,t) \middle| s \wedge t < r, \frac{s}{t} \le \tan\theta\right\} = r\tilde{\Phi}(\theta)$$

for some monotone $\tilde{\Phi}$, with $\tilde{\Phi}$ the distribution function of a finite measure (the spectral measure) on $[0, \pi/2]$.Once again G is characterized by γ_1, γ_2 and the distribution function $\tilde{\Phi}$. Note that in the development leading to $\tilde{\Phi}$, the preliminary transformation of the marginals is non-parametric whereas in the case of Φ it was parametric (using

$\gamma_1, \gamma_2, a_n, c_n, b_n$ and d_n). Two remarks can be made. Firstly there are two extreme cases for Φ (or $\tilde{\Phi}$). If Φ (or $\tilde{\Phi}$) is concentrated on the two end points $\{0\}$ and $\{\pi/2\}$, then $G(x,y) = G(x,\infty)G(\infty,y)$, i.e. the components are independent. If Φ (or $\tilde{\Phi}$) is degenerate and is concentrated on some point in $(0,\pi/2)$, then there is full dependence, i.e. one component determines the other:

$$G(x,y) = G(x,\infty) \wedge G(\infty,y).$$

Secondly a way to visualize the dependence structure is as follows. Define the level set

$$Q_c := \{(x,y) \mid l(x,y) = c\}$$

with l the limit function in (4.18). The curve Q_c has the following properties:

1. It connects the points $(c,0),(0,c)$.
2. It is the graph of a non-increasing function.
3. It is a concave function.
4. The shape of Q_{c_1} is the same as that of Q_{c_2} for $c_1 \neq c_2$.

Conversely any set of curves with the mentioned properties serves as the level set of some l-function (Huang, 1992, Chapter 3). The special cases of independence and full dependence of the marginals are reflected in the Q-curve as follows. In case of independence the Q_c-curve is given by $\{(x,y) \mid x+y=c\}$ and in case of full dependence by $\{(x,y) \mid x \vee y = c\}$ Since in a practical situation it is important to know if there is asymptotic independence, a graphical test like that of the Q-curve is helpful. We finish this section by remarking that in the two basic relations

$$\lim_{n\to\infty} n \Pr\left\{X > a_n \frac{x^{\gamma_1}-1}{\gamma_1} + b_n \text{ or } Y > c_n \frac{y^{\gamma_2}-1}{\gamma_2} + d_n\right\} = v_0(A_{x,y}) \tag{4.24}$$

and (4.20) the running integer variable n can be replaced by a continuous variable t: for some functions $a, c > 0$ and b and d real

$$\lim_{t\to\infty} t \Pr\left\{\left(\left(1+\gamma_1 \frac{X-b(t)}{a(t)}\right)^{1/\gamma_1}, \left(1+\gamma_2 \frac{Y-d(t)}{c(t)}\right)^{1/\gamma_2}\right) \in B\right\} = v_0(B) \tag{4.25}$$

and

$$\lim_{t\downarrow 0} t^{-1} \Pr\{(1-F_1(X), 1-F_2(Y)) \in tB\} = \tilde{v}_0(B) \tag{4.26}$$

where γ_1 and γ_2 are the parameters of the two marginal extreme-value distributions and B

any Borel set with $v_0(\partial B) = 0$ or $\tilde{v}_0(\partial B) = 0$ respectively; (X, Y) is a random vector with distribution function F. We shall also need

$$\lim_{t \downarrow 0} t^{-1}\{1 - F((1-F_1)^{\leftarrow}(tx), (1-F_2)^{\leftarrow}(ty))\} = l(x, y). \tag{4.27}$$

5. Statistical theory

5.1. Estimation of the extreme value distribution

The aim is to develop estimators for the function l (cf. (4.18)) or equivalently for the distribution function G_0 as well as for the distribution function Φ of the spectral measure, based on observations $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ from the initial distribution F. The estimators will be consistent and asymptotically normal. First of all remark that $-\log G_0$ is an approximation only to the far tail $(1 - F(a_n x + b_n, c_n y + d_n))$ of the original distribution. Hence only ''larger'' observations from F tell us anything about G_0 or l. So only a small number, say k, of the original n observations will be used for estimation, namely those observations for which at least one of the components exceeds a certain threshold. The number k (depending on $n : k = k(n)$) should go to infinity with n in order to enable us to use the law of large numbers for proving the consistency but should be small in relation to n, since only the tail is relevant. Hence we require

$$k = k(n) \to \infty, \quad \frac{k(n)}{n} \to 0 \quad (\text{as } n \to \infty).$$

Let us now start from relation (4.27) and replace t by k/n

$$\lim_{n\to\infty} \frac{n}{k} \Pr\left\{1 - F_1(X) < \frac{k}{n}x \text{ or } 1 - F_2(Y) < \frac{k}{n}y\right\} = l(x, y).$$

We want to turn the left hand side into an estimator for l. In order to do so we replace Pr by the empirical measure P_n and F_1 and F_2 by their empirical counterparts $F_1^{(n)}$ and $F_2^{(n)}$:

$$\begin{aligned}
&\hat{l}(x, y)\\
&\quad = \frac{n}{k}\frac{1}{n}\sum_{i=1}^{n} \mathbf{I}\{n(1 - F_1^{(n)}(X_i)) < kx \text{ or } n(1 - F_2^{(n)}(Y_i)) < ky\}\\
&\quad = \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\{X_i > X_{n-[kx],n} \text{ or } Y_i > Y_{n-[ky],n}\}\\
&\quad = \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\{R_i^X > n - kx \text{ or } R_i^Y > n - ky\}
\end{aligned} \tag{5.1}$$

where $X_{1,n} \leq X_{2,n} \leq \cdots \leq X_{n,n}$ are the order statistics from $(X_1, X_2, \ldots, X_n)$ and $Y_{1,n} \leq Y_{2,n} \leq \cdots \leq Y_{n,n}$ the order statistics from $(Y_1, Y_2, \ldots, Y_n)$; $R_i^X = rank(X_i)$ among $(X_1, X_i, \ldots, X_n)$ and $R_i^Y = rank(Y_i)$ among $(Y_1, Y_2, \ldots, Y_n)$. The estimator $\hat{l}$ is consistent for l. Under additional assumptions the stochastic process

$$\sqrt{k}\{\hat{l}(x,y) - l(x,y)\}$$

converges in $D([0,1] \times [0,1])$ to a zero mean Gaussian process with known covariance structure (Huang, 1992). The convergence rate $1/\sqrt{k}$ will be relevant to the rest of our story, but not the limit process. In a similar way we can construct an estimator for the distribution function $\tilde{\Phi}$ of the spectral measure. Recall that for $0 < x, y < 1$

$$l(x,y) = \tilde{v}_0(([x,1] \times [y,1])^c) \quad (\text{in}[0,1]^2)$$

and that

$$\tilde{\Phi}(\theta) = r^{-1}\tilde{v}_0\left\{(s,t) \mid s \wedge t < r, \frac{s}{t} \leq \tan\theta\right\}.$$

Proceeding in a way analogous to what we did for the l-estimator, we get (take $r = k/n$).

$$\hat{\tilde{\Phi}}(\theta) := \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\left\{n - R_i^X \vee R_i^Y < k, \frac{n - R_i^X}{n - R_i^Y} \leq \tan\theta\right\}. \tag{5.2}$$

The estimator $\hat{\tilde{\Phi}}$ is consistent for $\tilde{\Phi}$. Under additional assumptions the stochastic process

$$\sqrt{k}(\hat{\tilde{\Phi}}(\theta) - \tilde{\Phi}(\theta))$$

converges to a zero mean Gaussian process with known covariance structure (see Einmahl, de Haan, Piterbarg, 1997). An alternative way of estimating l and $\tilde{\Phi}$ is via relation (4.25)

$$\hat{l}'(x,y) := \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\left\{\hat{X}_i\left(\frac{n}{k}\right) > \frac{1}{x} \text{ or } \hat{Y}_i\left(\frac{n}{k}\right) > \frac{1}{y}\right\} \tag{5.3}$$

and

$$\hat{\tilde{\Phi}}'(\theta) := \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\left\{\hat{X}_i\left(\frac{n}{k}\right) \vee \hat{Y}_i\left(\frac{n}{k}\right) > 1, \arctan\left(\hat{Y}_i\left(\frac{n}{k}\right)/\hat{X}_i\left(\frac{n}{k}\right)\right) \leq \theta\right\} \tag{5.4}$$

where

$$\hat{X}_i\left(\frac{n}{k}\right) := \left(1+\hat{\gamma}_1\frac{X_i-\hat{b}(n/k)}{\hat{a}(n/k)}\right)^{1/\hat{\gamma}_1}, \hat{Y}_i\left(\frac{n}{k}\right) := \left(1+\hat{\gamma}_2\frac{Y_i-\hat{d}(n/k)}{\hat{c}(n/k)}\right)^{1/\hat{\gamma}_2}. \tag{5.5}$$

The estimators $\hat{\gamma}_1, \hat{\gamma}_2$, $\hat{a}(n/k)$, $\hat{b}(n/k)$, $\hat{c}(n/k)$ and $\hat{d}(n/k)$ are well-known from the solution of the corresponding one-dimensional problem. As estimators for $\gamma_1, \gamma_2, a(n/k), b(n/k), c(n/k)$ and $d(n/k)$ we have used the ones suggested in Dekkers, Einmahl and de Haan, 1989: define the functions

$$\bar{t} := t \wedge 0, \quad \rho_1(t) := \frac{1}{1-\bar{t}}, \quad \rho_2(t) := \frac{2}{(1-\bar{t})(1-2\bar{t})},$$

and for $r = 1, 2$

$$H_{1,r} := k^{-1}\sum_{i=0}^{k-1}\{\log X_{n-i,n} - \log X_{n-k,n}\}^r,$$

$$H_{2,r} := k^{-1}\sum_{i=0}^{k-1}\{\log Y_{n-i,n} - \log Y_{n-k,n}\}^r.$$

Now we define estimators as

$$\begin{aligned}
\hat{\gamma}_j :=& H_{j,1} + 1 - \frac{1}{2}\left(1 - \frac{H_{j,1}^2}{H_{j,2}}\right)^{-1}\\
\hat{b}\left(\frac{n}{k}\right) :=& X_{n-k,n}\\
\hat{a}\left(\frac{n}{k}\right) :=& \frac{X_{n-k,n}\sqrt{3H_{1,1}^2 - H_{1,2}}}{\sqrt{(e_1(\hat{\gamma}_1))^2 - \rho_2(\hat{\gamma}_1)}}\\
\hat{d}\left(\frac{n}{k}\right) :=& Y_{n-k,n}\\
\hat{c}\left(\frac{n}{k}\right) :=& \frac{Y_{n-k,n}\sqrt{3H_{2,1}^2 - H_{2,2}}}{\sqrt{(e_1(\hat{\gamma}_2))^2 - \rho_2(\hat{\gamma}_2)}}.
\end{aligned}$$

The estimators $\hat{l}\,'$ and $\hat{\tilde{\Phi}}'$ are consistent.

Under additional assumptions

$$\sqrt{k}\{\hat{l}\,'(x,y) - l(x,y\} \text{ and } \sqrt{k}\{\hat{\tilde{\Phi}}'(\theta) - \Phi(\theta)\}$$

converge to a zero mean Gaussian stochastic processes with known covariance structure (see de Haan and Resnick, 1993, Einmahl, de Haan and Sinha, 1996, Sinha, 1997, Chapter 4).

5.2. *Estimating an exceptional set*

We have now finished the necessary preparation for tackling the problem set out in Section 1: estimating the probability of an exceptional set.

We are given n i.i.d. two-dimensional observations $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ from an unknown distribution function F which we assume to be in the domain of attraction of an extreme-value distribution i.e. for which (4.1) holds. Also we are given some set C, the exceptional or *failure set*. None of our observations fall in the set C and we are asked to estimate the probability of C.

Since none of the observations fall in the set C, a first approximation for its probability is $\Pr(C) < 1/(n+1)$. Since we are going to apply asymptotic theory (pretending that the number of observations n grows indefinitely) and since the fact that none of the observations fall into C is an essential feature of the problem, we have to assume that $\lim_{n\to\infty} \Pr(C) \le \lim_{n\to\infty} 1/(n+1) = 0$ so we have to assume that the set C in fact depends on n. So we now write $C = C_n$ and assume $\Pr(C_n) \to 0$ (as $n \to \infty$).

In order to make the method of estimation more transparent, we shall first look at the special case

$$C_n = \{(s,t) \mid s > x_n \text{ or } t > y_n\}, \tag{5.6}$$

i.e. C_n is the complement of a rectangle. Recall that the tail of F has a certain useful approximate homogeneity property after a transformation of the marginal distributions. So we are going to express everything including C_n in the transformed coordinates. Hence define sequences $q_n, r_n > 0$ via the equalities

$$\begin{cases} x_n = b\left(\dfrac{n}{k}\right) + a\left(\dfrac{n}{k}\right) \cdot \dfrac{q_n^{\gamma_1} - 1}{\gamma_1} \\ y_n = d\left(\dfrac{n}{k}\right) + c\left(\dfrac{n}{k}\right) \cdot \dfrac{r_n^{\gamma_2} - 1}{\gamma_2}. \end{cases} \tag{5.7}$$

In order to derive an estimator for $\Pr(C_n)$ we consider the following (approximate) equalities (use (4.25) with $B = (0, q_n] \times (0, r_n])$

$$\begin{aligned}\Pr(C_n) &= 1 - F(x_n, y_n)\\ &= 1 - F\left(b\left(\frac{n}{k}\right) + a\left(\frac{n}{k}\right)\frac{q_n^{\gamma_1} - 1}{\gamma_1}, d\left(\frac{n}{k}\right) + c\left(\frac{n}{k}\right)\frac{r_n^{\gamma_2} - 1}{\gamma_2}\right)\\ &\approx \frac{k}{n}\{-\log G_0(q_n, r_n)\}.\end{aligned}$$

We are tempted to use this as a basis for an estimator for $\Pr(C_n)$. However q_n and r_n tend to zero, so it is of no use. We have to use the homogeneity property of the function $-\log G_0$ again in order to get the arguments of $-\log G_0$ in the right range. This is done as follows. According to one-dimensional extreme value theory

$$1 - F_1(x_n) \approx \frac{k}{n}\left(1 + \gamma_1 \frac{x_n - b(n/k)}{a(n/k)}\right)^{-1/\gamma_1} = \frac{k}{n}\frac{1}{q_n}$$

and

$$1 - F_1(y_n) \approx \frac{k}{n}\left(1 + \gamma_2 \frac{y_n - d(n/k)}{c(n/k)}\right)^{-1/\gamma_2} = \frac{k}{n}\frac{1}{r_n}.$$

Now from the above

$$\begin{aligned}P_r(C_n) &\approx \frac{k}{n}\{-\log G_0(q_n, r_n)\}\\ &= \frac{k\sqrt{q_n^{-2} + r_n^{-2}}}{n}\{-\log G_0(q_n\sqrt{q_n^{-2} + r_n^{-2}}, r_n\sqrt{q_n^{-1} + r_n^{-2}})\}.\end{aligned} \tag{5.8}$$

We have to suppose now that $\lim_{n\to\infty} q_n\sqrt{1/q_n^2 + 1/r_n^2}$ exists $\in (0, 1)$, i.e. that

$$\lim_{n\to\infty} \frac{1 - F_1(x_n)}{1 - F_2(y_n)}$$

exists in $(0, \infty)$. Then the arguments of $-\log G_0$ in (5.8) stay away from zero and infinity as they should.

The estimator for $\Pr(C_n)$ then becomes (via (5.6) and(5.8))

$$\frac{k\sqrt{\hat{q}_n^{-2} + \hat{r}_n^{-2}}}{n}\{-\log \hat{G}(\hat{q}_n\sqrt{\hat{q}_n^{-2} + \hat{r}_n^{-2}}, \hat{r}_n\sqrt{\hat{q}_n^{-2} + \hat{r}_n^{-2}})\}$$

where $-\log \hat{G}_0(x,y) = \hat{l}'(1/x, 1/y)$ (see (3.18) and (5.3)) and

$$\hat{q}_n := \frac{k}{n}\left(1+\hat{\gamma}_1 \frac{x_n - \hat{b}(n/k)}{\hat{a}(n/k)}\right)^{1/\hat{\gamma}_1}, \hat{r}_n := \frac{k}{n}\left(1+\hat{\gamma}_2 \frac{y_n - d(n/k)}{c(n/k)}\right)^{1/\hat{\gamma}_2}.$$

The asymptotic behavior of $1-\tilde{F}$ is as follows:

$$\tau(\hat{q}_n, \hat{r}_n, \hat{\gamma}_1, \hat{\gamma}_2, n)\left(\frac{1-\tilde{F}(x_n, y_n)}{1-F(x_n, y_n)} - 1\right)$$

is asymptotically standard normal where τ is a known completely specified function tending to infinity (see Sinha, 1996, Chapter 2 or de Haan, 1994; for the one-dimensional analog see Dijk and de Haan, 1992). This leads to an asymptotic confidence interval for $\Pr(C_n)$.

Now after finishing this simple case, we are ready to consider the problem of estimating $\Pr(C_n)$ for a set C_n of quite arbitrary form. In complete analogy to the previous case we suppose that

$$C_n = \mathbf{a}\left(\frac{n}{k}\right)\frac{(c_n A)^{\gamma} - 1}{\gamma} + \mathbf{b}\left(\frac{n}{k}\right)$$

where $\mathbf{a}(n/k) = (a(n/k), c(n/k)), \mathbf{b}(n/k) = (b(n/k), d(n/k)), \gamma = (\gamma_1, \gamma_2)$ and the sums, products and powers are defined component wise. Recall that by $c_n A$ we mean all elements of A, each multiplied by c_n.

Here A is a fixed set and c_n a sequence of positive numbers with $c_n \to \infty,\ n \to \infty$ (the sequence c_n plays the role of the sequence $(q_n^{-2} + r_n^{-2})^{-1/2}$ and the set A that of the set

$$\left\{(s,t) \mid s < \lim_{n\to\infty} q_n \sqrt{\frac{1}{q_n^2} + \frac{1}{r_n^2}} \text{ or } t < \lim_{n\to\infty} r_n \sqrt{\frac{1}{q_n^2} + \frac{1}{r_n^2}}\right\}$$

in the example above).

Now note that by (4.15) and (4.25)

$$\begin{aligned}
p_n : &= \Pr(C_n) = \Pr\left\{(x,y) \in \mathbf{a}\left(\frac{n}{k}\right)\frac{(c_n A)^{\gamma} - 1}{\gamma} + \mathbf{b}\left(\frac{n}{k}\right)\right\} \\
&= \Pr\left\{\left(\left(1+\gamma_1 \frac{X - b(n/k)}{a(n/k)}\right)^{1/\gamma_1}, \left(1+\gamma_2 \frac{Y - d(n/k)}{c(n/k)}\right)^{1/\gamma_2}\right) \in c_n A\right\} \\
&\approx \frac{k}{n} v_0(c_n A) \\
&= \frac{k}{nc_n} v_0(A).
\end{aligned}$$

So the proposed estimator of p_n is

$$\hat{p}_n := \frac{k}{n\hat{c}_n}\hat{v}_0(\hat{A})$$

where (extending (5.3) for each Borel set B)

$$\hat{v}_0(B) := \frac{1}{k}\sum_{i=1}^{n}\mathbf{I}\left\{\left(\left(1+\hat{\gamma}_1\frac{X_i-\hat{b}(n/k)}{\hat{a}(n/k)}\right)^{1/\hat{\gamma}_1},\left(1+\hat{\gamma}_2\frac{Y_i-d(n/k)}{c(n/k)}\right)^{1/\hat{\gamma}_2}\right)\in B\right\}.$$

Next we are going to introduce the estimators $\hat{c}_n$ for c_n and $\hat{A}$ for A. First of all note that c_n and A are not uniquely determined. In order to determine them exactly we require that A satisfies $\min\{s \mid (s,s)\in A\}=1$, i.e. the point (1.1) is on the boundary of A.

At this point, in order to proceed we have to specify our problem a little further. The failure region in Figure 1 is in fact determined by a function: its functional form is

$$\{(s,t)\mid (0.3)s+t\geq 7.6\}.$$

So we shall assume that

$$C_n=\{(s,t)\mid f_n(s,t)\geq 1\}$$

and more specifically, that

$$f_n(s,t)=f\left(\frac{s}{x_n},\frac{t}{y_n}\right)$$

where x_n and y_n are positive numbers and f some fixed known function.

Now

$$\begin{aligned}A&=\frac{1}{c_n}\left(1+\gamma\frac{C_n-\mathbf{b}(n/k)}{\mathbf{a}(n/k)}\right)^{1/\gamma}\\&=\left\{(s,t)\middle| f\left(\frac{1}{x_n}\left\{a\left(\frac{n}{k}\right)\frac{(c_ns)^{\gamma_1}-1}{\gamma_1}+b\left(\frac{n}{k}\right)\right\},\right.\right.\\&\qquad\left.\frac{1}{y_n}\left\{c\left(\frac{n}{k}\right)\frac{(c_nt)^{\gamma_2}-1}{\gamma_2}+d\left(\frac{n}{k}\right)\right\}\right)\geq 1\right\}.\end{aligned}$$

But according to our choice of A, the point (1.1) must be on the boundary of A. So c_n must be the solution of

$$f\left(\frac{1}{x_n}\left\{a\left(\frac{n}{k}\right)\frac{u^{\gamma_1}-1}{\gamma_1}+b\left(\frac{n}{k}\right)\right\},\frac{1}{y_n}\left\{c\left(\frac{n}{k}\right)\frac{u^{\gamma_2}-1}{\gamma_2}+d\left(\frac{n}{k}\right)\right\}\right)=1.$$

This determines $c_n = u(\mathbf{a}(n/k), \mathbf{b}(n/k), \gamma)$. But we do not know $(\mathbf{a}(n/k), \mathbf{b}(n/k), \gamma)$. So we can only find $\hat{c}_n = u(\hat{\mathbf{a}}(n/k), \hat{\mathbf{b}}(n/k), \hat{\gamma})$ which is the solution of

$$f\left(\frac{1}{x_n}\left\{\hat{a}\left(\frac{n}{k}\right)\frac{u^{\hat{\gamma}_1}-1}{\hat{\gamma}_1}+\hat{b}\left(\frac{n}{k}\right)\right\},\frac{1}{y_n}\left\{\hat{c}\left(\frac{n}{k}\right)\frac{u^{\hat{\gamma}_2}-1}{\hat{\gamma}_2}+\hat{d}\left(\frac{n}{k}\right)\right\}\right)=1.$$

We have now defined $\hat{c}_n$. Finally we define $\hat{A}$ as follows:

$$\hat{A} := \frac{1}{\hat{c}_n}\left(1+\hat{\gamma}\frac{C_n-\hat{\mathbf{b}}(n/k)}{\hat{\mathbf{a}}(n/k)}\right)^{1/\hat{\gamma}}.$$

Now the estimator $\hat{p}_n$ is fully defined. Under various conditions concerning the underlying distribution function, the sequence $k(n)$, the function f and the sequences x_n and y_n, we have the following asymptotic result for $\hat{p}_n$:

$$\frac{\sqrt{k}}{q_{\hat{\gamma}_1\wedge\hat{\gamma}_2}(\hat{c}_n)}\left(\frac{\hat{p}_n}{p_n}-1\right)$$

is asymptotically normal with mean zero and known variance, where q is a known completely specified function such that $\sqrt{k}/q_{\hat{\gamma}_1\wedge\hat{\gamma}_2}(\hat{c}_n)\to\infty, n\to\infty$ (see de Haan and Sinha, 1997).

6. Estimation procedures and results

We are now going to apply the described semi-parametric estimation procedures and also maximum likelihood procedures to the two-dimensional data set (still water level and wave height) and later to the three-dimensional data (still water level, wave height and wave period).

6.1. Semi-parametric estimation procedures

The steps to be considered are:

- Estimation of the one-dimensional shape parameters γ_1 and γ_2.

- Estimation of the dependence structure via the spectral measure $\Phi(\cdot)$ or the set $Q(\cdot)$. This step is not essential for the estimation of the failure probability, but it provides insight in the asymptotic dependence between the two variables. The main purpose is to check that there is indeed asymptotic dependence: in case of asymptotic independence the estimation of p_n has to be done in a different way. See Subsection 6.5.
- Estimation of failure probability and confidence interval.

6.1.1. Univariate analysis. Analysis of the marginals starts with estimation of the parameters of the extreme value distribution, $a(n/k), b(n/k), c(n/k), d(n/k), \gamma_1, \gamma_2$. Figure 2 shows the estimate of the extreme value indices $\hat{\gamma}$, for various values of k. Unfortunately, the plot of HmO suggests that only a limited number of order statistics may be used. We decided to use $k = 27$ upper order statistics for our analysis. Table 1 presents the resulting estimates.

Using these estimates we can fit an extreme-value distribution to the tail of the empirical distribution function. As Figure 3 shows, this fitted curve allows us to estimate probabilities outside the range of observations. Note that for the wave height the correspondence between empirical and parametric exceedence probabilities is only fair in the tail. Note also that the sea level seems to have a finite upper limit. This is confirmed by the analysis, based on a much larger dataset, of Dillingh et al. (1993).

6.1.2. The dependence structure of bivariate extremes. We are going to visualize the asymptotic dependence structure via the Q-curve and the function $\tilde{\Phi}$. Figure 4 shows the estimated Q-curve: in the upper figure the set $\{(x, y) \mid \hat{l}(x, y) = c\}$ is shown (based on ranks) for $c = 0.042, 0.034, 0.025, 0.017$ and 0.0085 and in the lower figure the set $\{(x, y) \mid \hat{l}'(x, y) = c\}$ is shown (based on $(\hat{X}_i(n/k), \hat{Y}_i(n/k))$ for the same values of c (cf. Subsection 5.1). The two pictures are not much different although the first one is neater. The characterizing properties of the Q-curve from Section 4, concavity and equality of shape of Q_c for different values of c, seem to be approximately true for the curves of Figure 4. This gives some confidence in the extreme value model. Also even for small values of c the curve seems to differ significantly from a straight line so that indeed there is asymptotic dependence (cf. Section 4).

Figure 5 displays the estimate of the spectral measure $\tilde{\Phi}$ using ranks ($\hat{\tilde{\Phi}}$, top) and using $(\hat{X}_i(n/k), \hat{Y}_i(n/k))$ ($\hat{\tilde{\Phi}}'$, bottom). In both pictures $\hat{\tilde{\Phi}}(\pi/2)$ is scaled down from $\frac{39}{27}$ to 1. On the horizontal axis the angles of the individual transformed tail observations are displayed. The estimate for $\tilde{\Phi}$ is just the empirical distribution function based on those observations.

Table 1. Extreme value distribution parameters.

Parameter	HmO	SWL
γ_1 γ_2	−0.0074	−0.12
b, d	5.53	1.69
a, c	0.53	0.2915

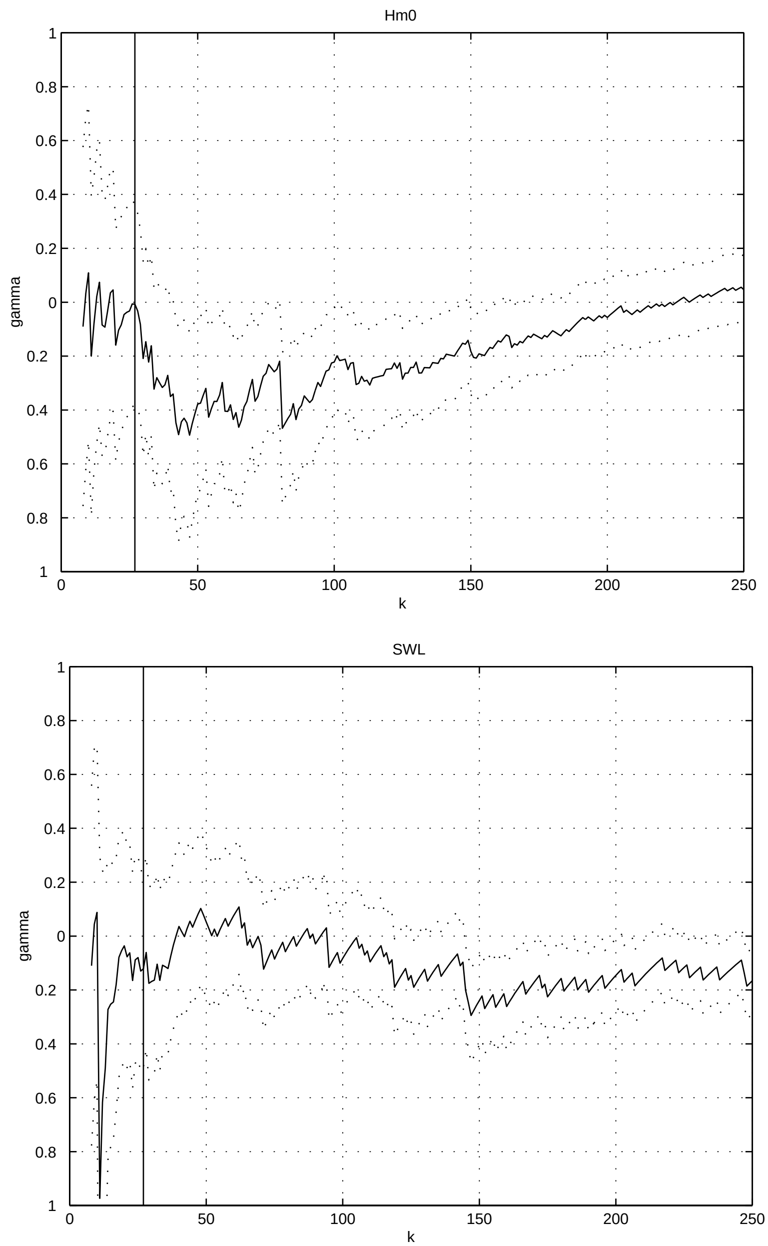

Figure 2. Moment estimates $\hat{\gamma}_i$ using various numbers of order statistics k. The vertical lines indicate the actual number on which the estimates are based ($k = 27$). (top: wave height; bottom: sea level)

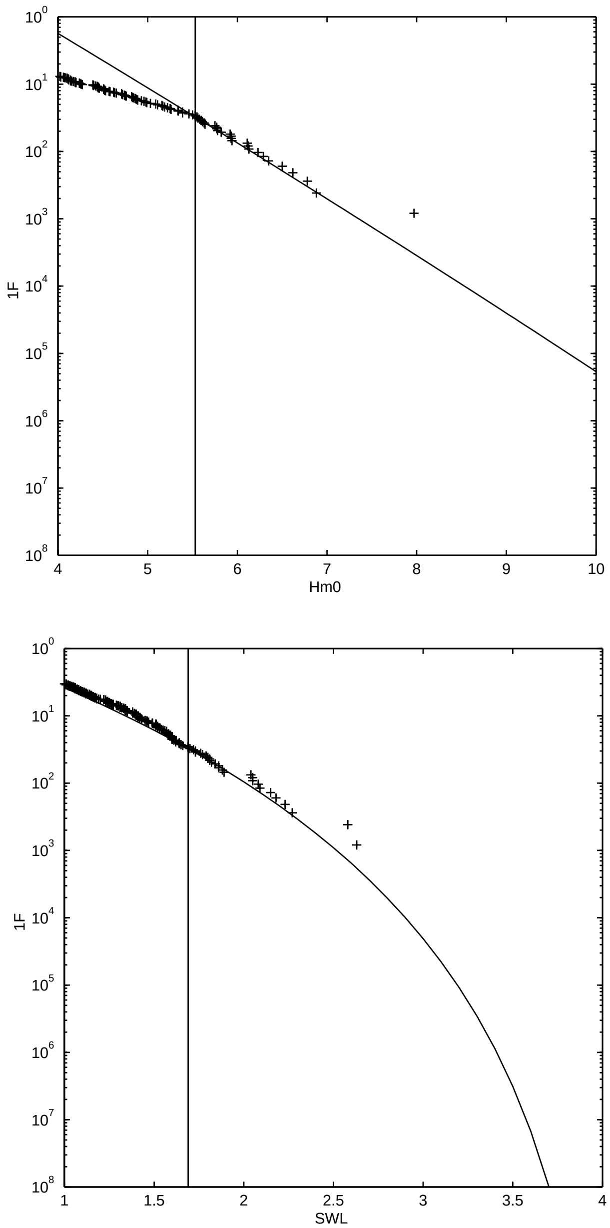

Figure 3. Estimated marginal exceedence probabilities. The solid line represents the parametric estimate. The '+' symbols represent the empirical estimate. The vertical line represents the $(n-27)$th order statistic. (top wave height HmO; bottom sea level SWL)

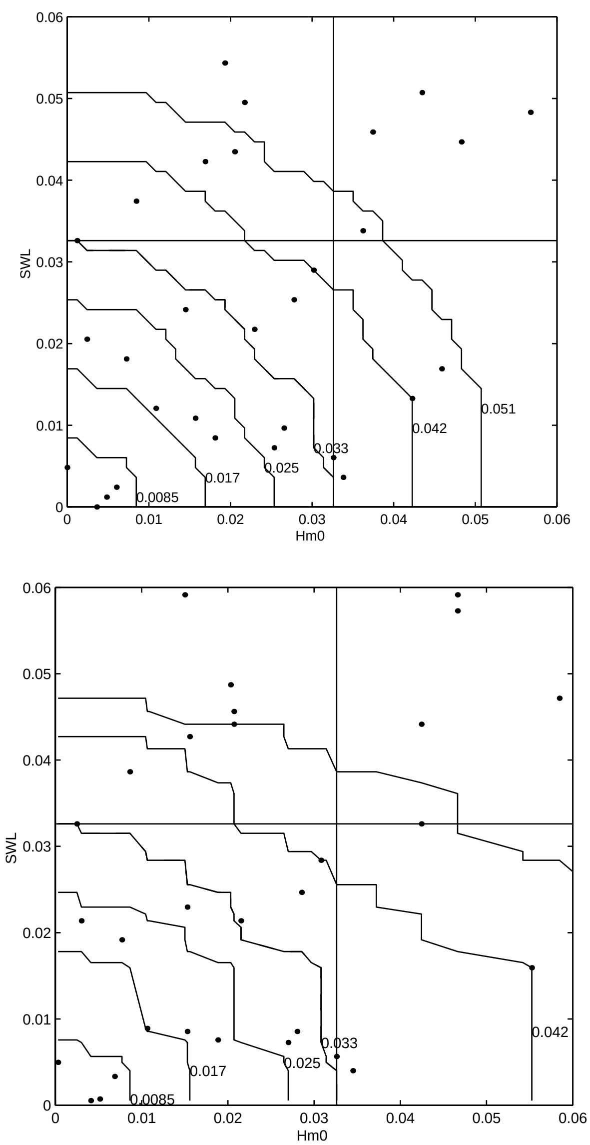

Figure 4. Q-curves. Level sets of the functions $\hat{l}$ (top) and $\hat{l}'$ (bottom). The quantile curves correspond to $k = 7$, 14, 21, 28, 35 and 42.

Since these observations are not clustered in the neighborhood of the points 0 and $\pi/2$, we once again conclude that there is indeed asymptotic dependence.

6.1.3. Estimating the probability of a flood. For p_n, the probability of having an observation in the failure region, we have the following (cf. Section 5).

Define for $i = 1, 2, \ldots, n$

$$\widetilde{X}_i\left(\frac{n}{k}\right) := \left[1 + \gamma_1 \frac{X_i - b(n/k)}{a(n/k)}\right]^{1/\gamma_1}, \widetilde{Y}_i\left(\frac{n}{k}\right) := \left[1 + \gamma_2 \frac{Y_i - d(n/k)}{c(n/k)}\right]^{1/\gamma_2}$$

and

$$\hat{X}_i\left(\frac{n}{k}\right) := \left[1 + \hat{\gamma}_1 \frac{X_i - \hat{b}(n/k)}{\hat{a}(n/k)}\right]^{1/\hat{\gamma}_1}, \hat{Y}_i\left(\frac{n}{k}\right) := \left[1 + \hat{\gamma}_2 \frac{Y_i - \hat{d}(n/k)}{\hat{c}(n/k)}\right]^{1/\hat{\gamma}_2}.$$

Then

$$\begin{aligned} p_n &= \Pr((X, Y) \in C_n) \\ &= \Pr\left(\left(\tilde{X}\left(\frac{n}{k}\right), \tilde{Y}\left(\frac{n}{k}\right)\right) \in \left(\mathbf{1} + \gamma \frac{C_n - \mathbf{b}}{\mathbf{a}}\right)^{1/\gamma}\right) \\ &= \Pr\left(\left(\tilde{X}\left(\frac{n}{k}\right), \tilde{Y}\left(\frac{n}{k}\right)\right) \in c_n A\right) \\ &\approx \frac{k}{n} v_0(c_n A) \\ &= \frac{k}{n c_n} v_0(A) \end{aligned}$$

and the latter can be estimated by

$$\hat{p}_n := \frac{k}{n\hat{c}_n} \hat{v}_0(\hat{A}).$$

The various steps are illustrated in Figures 1 and 6. Figure 1 shows the observations (X_i, Y_i) and the failure set C_n. Figure 6 shows the transformed data $(\hat{X}_i(n/k), \hat{Y}_i(n/k))$ [resembling $(\tilde{X}_i(n/k), \tilde{Y}_i(n/k))$] and the transformed failure set (indicated by the dotted line)

$$\left(\mathbf{1} + \gamma \frac{C_n - \mathbf{b}}{\mathbf{a}}\right)^{1/\gamma} = c_n A.$$

For convenience both variables are displayed on a logarithmic scale.

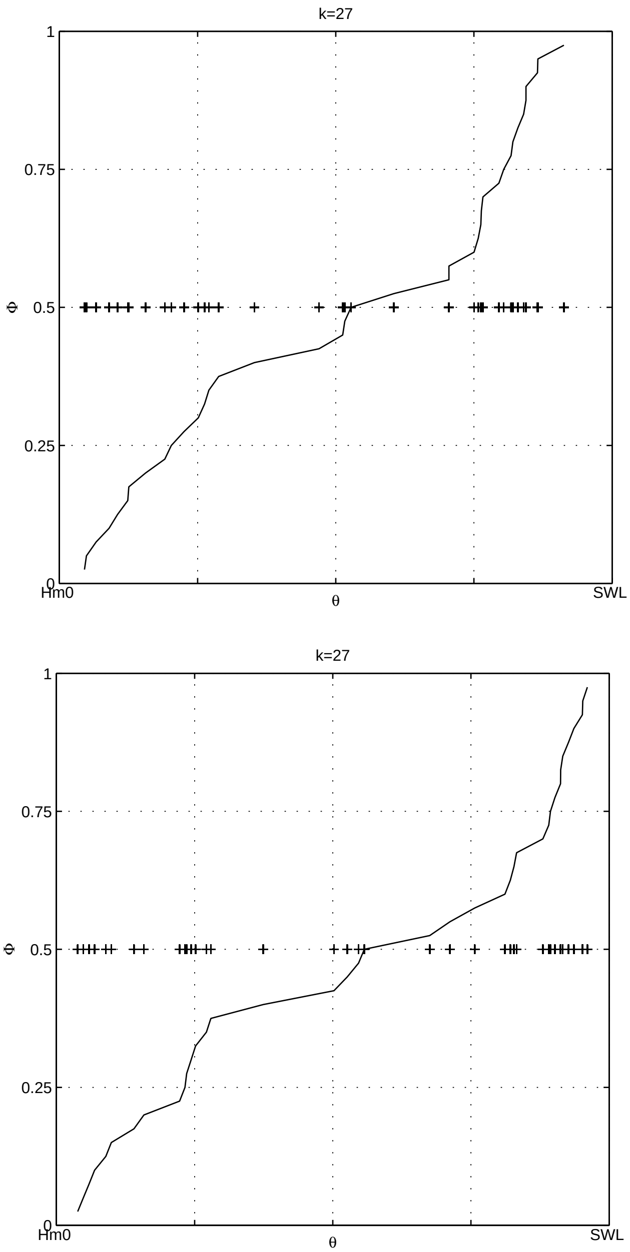

Figure 5. The spectral measure. Individual angular coordinates θ_i are shown as '+' signs. The solid line represents the corresponding distribution function Φ scaled down from $39/27$ to 1. The top figure shows estimates using ranks; the bottom figure shows estimates using $(\hat{X}_i(n/k), \hat{Y}_i(n/k))$. The horizontal axis runs from 0 (HmO) to $\pi/2$ (SWL).

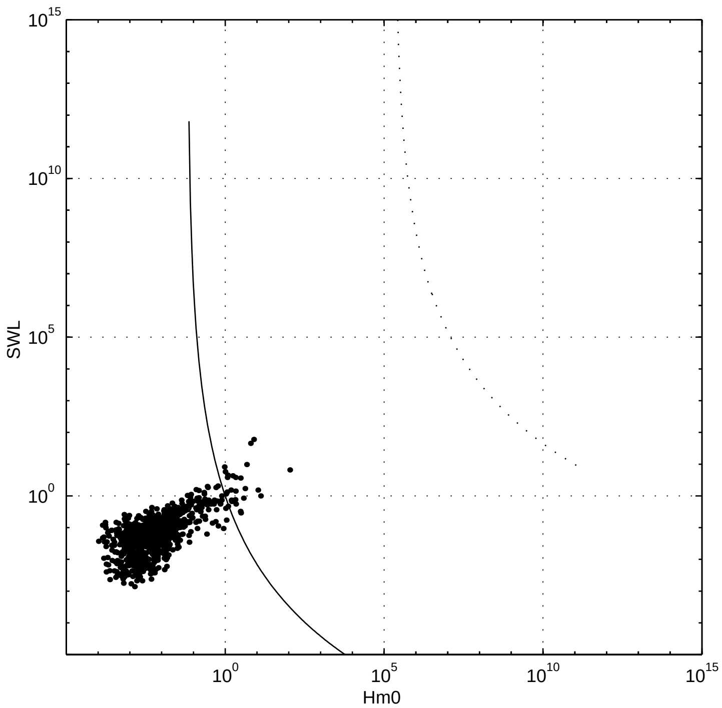

Figure 6. Transformed data and transformed failure region. The latter is represented by the area under the dotted line. The area under the solid line respresents the transformed failure area multiplied by $c_n = 2.9772 \cdot 10^6$.

Figure 6 also shows the transformed data $(\hat{X}_i(n/k), \hat{Y}_i(n/k))$ and the transformed and shifted failure set (indicated by the solid line)

$$A = \frac{1}{c_n}\left(\mathbf{1} + \gamma \frac{C_n - \mathbf{b}}{\mathbf{a}}\right)^{1/\gamma}$$

on logarithmic scales.

Note that the point (1.1) is indeed on the boundary of A. The number of points $(\hat{X}_i(n/k), \hat{Y}_i(n/k))$ in A is 26 and the estimated c_n is 2.9772×10^6. This results in an estimated failure probability

$$\hat{p} := \frac{k}{n\hat{c}_n} \hat{\nu}_0(\hat{A}) = \frac{27}{(828)(2.9772 \times 10^6)} \frac{26}{27} = 1.0547 \times 10^{-8}.$$

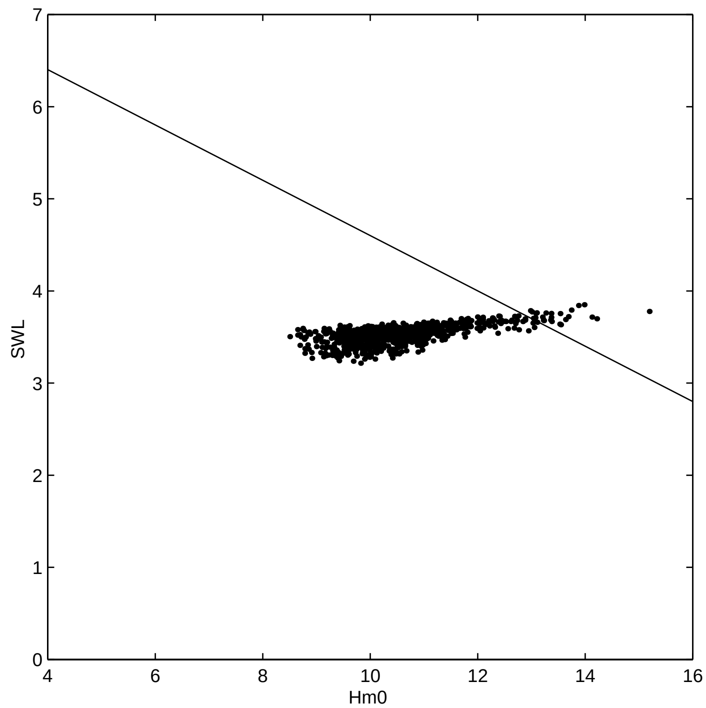

Figure 7. Analog of figure 6 in the original scale.

This is the failure probability per storm. The failure probability per year is approximately 0.88×10^{-6}.

We can also give an analog of the last picture, Figure 6, in the original scale, i.e. just the original failure set but variables transformed, shifted and then transformed back. This is done in Figure 7.

6.1.4. Confidence interval. In order to get a confidence interval for p_n we reformulate the result of Subsection 5.2 as follows:

$$\frac{\sqrt{k}}{q_{\hat{\gamma}_1 \wedge \hat{\gamma}_2}(\hat{c}_n)}(\log \hat{p}_n - \log p_n)$$

is asymptotically normal with mean zero and variance $\sigma^2(\gamma_1, \gamma_2, v_0(\cdot), \Phi(\cdot), f, x_n, y_n)$. For the application of this result we still have to estimate the spectral measure Φ. This can be done in a non-parametric or in a semi-parametric way as for $\tilde{\Phi}$ (cf. Subsection 5.1).

Preliminary calculations with the non-parametric estimation procedure lead for the probability of failure per year to the 95% confidence interval

$$(0.286 \times 10^{-6}, 2.73 \times 10^{-6})$$

and with the semi-parametric estimation procedure to the 95% confidence interval

$$(0.48 \times 10^{-6}, 1.63 \times 10^{-6})$$

However these seem unrealistically narrow and need to be checked.

Finally we investigate the influence of the choice of c_n on the result. Remember we choose c_n such that the point (1,1) is on the boundary of A. Changing the c_n does not affect the estimated failure probability much.

6.2. Maximum likelihood method

6.2.1. Full extreme value model. Maximum likelihood methods in two-dimensional extreme value theory have been introduced by Smith (see e.g. Smith et al., 1990) and expanded by Tawn and collaborators (see e.g. Coles and Tawn, 1991 and 1994). The general assumption is that the underlying distribution function F actually coincides in the tail with a multivariate extreme value distribution. Moreover it is often assumed that the l-function (Section 4) is a known function up to one parameter:

$$l(p, q) = (p^{1/\beta} + q^{1/\beta})^{\beta}$$

for some $\beta \in [0, 1]$ (logistic dependence structure: Gumbel, 1960, Smith, 1990). So for $x > X_{n-k,n}$

$$1 - F(x, \infty) = \frac{k}{n}\left(1 + \gamma_1 \frac{x - X_{n-k,n}}{a_1(n/k)}\right)^{-1/\gamma_1}$$

for $y > Y_{n-k,n}$

$$1 - F(\infty, y) = \frac{k}{n}\left(1 + \gamma_2 \frac{y - Y_{n-k,n}}{a_2(n/k)}\right)^{-1/\gamma_2}$$

and for $x > X_{n-k,n}$ and $y > Y_{n-k,n}$

$$1 - F(x, y) = \frac{k}{n}\left\{\left(1 + \gamma_1 \frac{x - X_{n-k,n}}{a_1(n/k)}\right)^{-1/(\beta\gamma_1)} + \left(1 + \gamma_2 \frac{y - Y_{n-k,n}}{a_2(n/k)}\right)^{-1/(\beta\gamma_2)}\right\}^{\beta}.$$

Table 2. Separately and jointly estimated distribution parameters; standard deviations in brackets.

	Estimated seperately	Estimated jointly	
β		0.56	(0.07)
x_0	5.53		
a_1	0.54	0.49	(0.14)
γ_1	−0.016	0.16	(0.23)
y_0	1.69		
a_2	0.33	0.33	(0.09)
γ_2	−0.19	−0.043	(0.24)

The contribution of each observation to the likelihood is determined according to the threshold censored likelihood approach of Smith (1994).

$$L(X_i, Y_i) = \begin{cases} (\partial^2 F/\partial x \partial y)(X_i, Y_i) & \text{for } x > X_{n-k,n}, y > Y_{n-k,n} \\ (\partial F/\partial x)(X_i, Y_{n-k,n}) & \text{for } x > X_{n-k,n}, y \le Y_{n-k,n} \\ (\partial F/\partial y)(X_{n-k,n}, Y_i) & \text{for } x \le X_{n-k,n}, y > Y_{n-k,n} \\ F(X_{n-k,n}, Y_{n-k,n}) & \text{for } x \le X_{n-k,n}, y \le Y_{n-k,n}. \end{cases}$$

The estimators of the model parameter vector $(\beta, a_1, \gamma_1, a_2, \gamma_2)'$ are obtained by maximizing the resulting log-likelihood function. Results are shown in Table 2. Note that the estimated parameters γ_1 and γ_2, when estimated jointly with β, are somewhat different from when they are estimated in a purely one-dimensional context. We shall come back to this point later.

The failure probability can be estimated by integrating the resulting parametric density over the failure region, or rather the conditional density of X, given that (X, Y) is in the failure region. For details see Bruun and Tawn (1995, 1996) or Draisma e.a. (1996, EUR-10). This leads to a point estimate plus confidence interval:

$$\hat{p} = 1.14 \times 10^{-3}$$

for the failure probability per year; 95% confidence interval:

$$(0, 8.75 \times 10^{-3}).$$

This means that zero probability is a real possibility.

Alternatively, once the model parameters $(\beta, a_1, \gamma_1, a_2, \gamma_2)'$ are estimated, one can find the failure probability by using the shift method from Section 5. This leads to a point estimate for p but no confidence interval:

$$\hat{p} = 1.3 \times 10^{-3}.$$

The procedure can be modified in yet another way. Since, as we have remarked in Section 5, in the semi-parametric set-up the uncertainty in the estimation of l does not enter in the final result, we can use a non-parametric estimator of l instead of the logistic one mentioned before. Since Huang's (1992) non-parametric estimator is not a homogeneous function we have to do this via the spectral measure Φ. But since a discrete measure does not fit into the maximum likelihood procedure, we smooth the estimator $\hat{\Phi}$ or $\hat{\tilde{\Phi}}$ of Section 5 by convolution with a uniform distribution over a short interval. For details see Draisma e.a. (1996, EUR-11).

6.2.2. Mixed model. A time series of 13 years is very short if one wishes to design a dike. Return periods in The Netherlands for coastal defences vary between 2000 and 10,000 years (central part of The Netherlands). The return period of 10,000 year implies a failure frequency of 10^{-4} per year. Because especially for water levels the periods of observations are much longer than 13 years a method has been developed in order to make use of this. As a first step the marginal statistics of water level (Dillingh et al., 1993), wave height (de Ronde et al., 1995) and wave period (Roskam et al.,) have been derived taking the longest period of observation and the geographical distribution of the extremes into account. Also mathematical models of storm surges and of extreme wave heights have been used to check the geographical distribution and the physical limits. For still water levels the usual extreme value modeling has been used (Dillingh et al., 1993) but not so for the wave heights. In that case a conditional Weibull distribution has been fitted to the larger observations and used for extrapolation (de Ronde et al., 1995). In order to stay in line with this previous analysis a two-dimensional analysis has been undertaken using extreme value (for still water) and conditional Weibull (for wave height) marginals.

This is possible only in the framework of maximum likelihood methods. The estimated failure probability per year is 2.0×10^{-4} and the 95% confidence interval $(0, 2.0 \times 10^{-3})$. However the desired connection with previous one-dimensional results has not been achieved: once again the multidimensional maximum likelihood context changes the marginal parameters. For details see Draisma et al. (1997, EUR/RIKZ-96.2+3).

6.3. Three variables approach

The three variables involved are still water level SWL, wave height HmO (both measured in meters) and wave period Tpb (in seconds). The offshore failure region is given in (2.2). The probabilistic and mathematical-statistical theory is completely analogous to the two-dimensional case.

One-dimensional marginal parameters and two-dimensional dependence structures are estimated as before. The results are shown in Table 3, and Figures 8 and 9. The trivariate dependence structure is shown in two ways.

The empirical l-function is calculated:

$$\begin{aligned}&\hat{l}(x, y, z)\\ &\quad = \frac{1}{k}\sum_{i=1}^{n} \mathbf{I}\{X_i \geq X_{n-[kx]+1,n} \text{ or } Y_i \geq Y_{n-[ky]+1,n} \text{ or } Z_i \geq Z_{n-[kz]+1,n}\}\end{aligned}$$

Table 3. Extreme value distribution parameters.

Parameter	HmO	Tpb	SWL
k	27	27	27
γ	− 0.0074	0.1183	− 0.1215
b	5.53	11.57	1.69
a	0.53	1.03615	0.2915

and represented graphically in the following way: for each direction (x, y, z) the multiple $(\lambda_x, \lambda_y, \lambda_z)$ with the property $\hat{l}(\lambda_x, \lambda_y, \lambda_z) = 1$ is presented. The result is Figure 8; in the three pictures various numbers of upper order statistics (chosen such that at least one of the three components exceeds a certain bound) have been used $(k = 54, 27, 14)$. Once again the two properties of the l-function discussed in Section 3: convexity and similarity of shape and correct boundary conditions, seem to hold approximately for $\hat{l}$.

Figure 9 represents the dependence structure in yet another way, namely via the scatter plot of the spectral measure obtained, that is, of the directions of the larger transformed observations. The directions are indicated by their projection on the plane $x + y + z = 1$.

From both representations it is quite obvious that there is no asymptotic independence (this would mean a flat convex function in Figure 8 or concentration of points near the three vertices in Figure 9).

The estimation of the failure probability is completely analogous to the two-dimensional situation. However we have not yet completed the theoretical calculations necessary for the construction of a confidence interval in three-dimensional space. The estimated failure probability in this model is effectively zero: an offshore sea-level of 4 m is necessary for the sea wall to fail—a height that is about the maximum of the estimated marginal distribution of SWL (see Figure 3). Only if we adopt the γ-value for SWL that results from a much longer series $(\hat{\gamma} = -0.067)$, the estimated failure probability becomes positive: 8.4×10^{-6} per year. In this case only still water level is significant. For details see Draisma et al. 1996, EUR-7.

6.4. *The load variable—reduction to the one-dimensional case*

A direct approach that however is less flexible and does not provide insight in the asymptotic dependence structure, is the following: for each combination (X_i, Y_i, Z_i) of offshore variables we calculate the corresponding load $L(X_i, Y_i, Z_i)$ on the sea wall (for the function L see Subsection 2.2).

Since the failure region is defined as

$$\{(x, y, z) \mid L(x, y, z) > 12.75\},$$

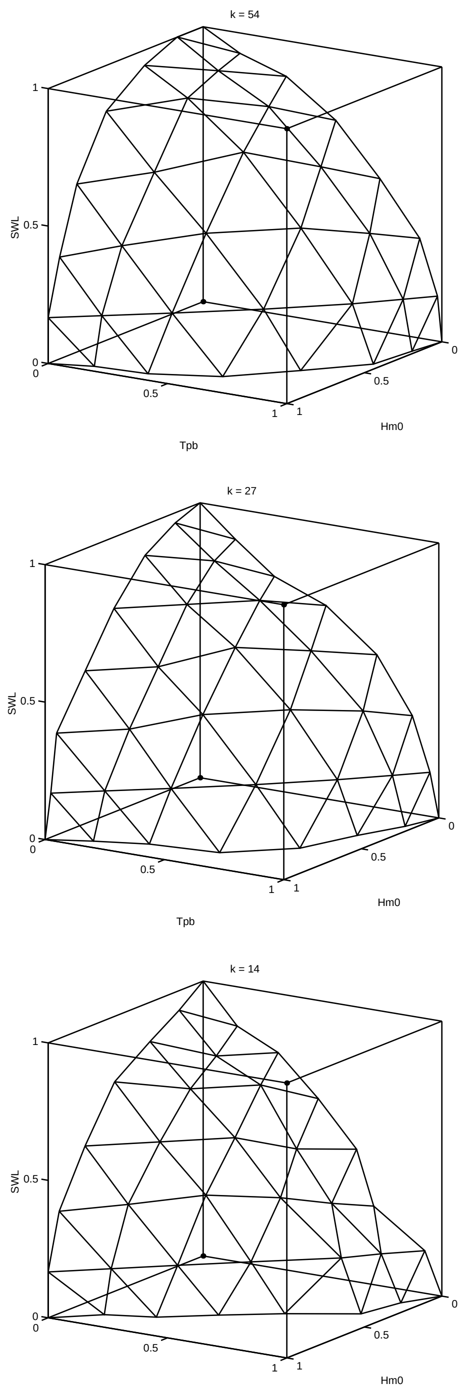

Figure 8. Trivariate Q-surfaces: level sets of $\hat{l}$. Contours shown correspond to $k = 54$, 27 or 14 observations respectively.

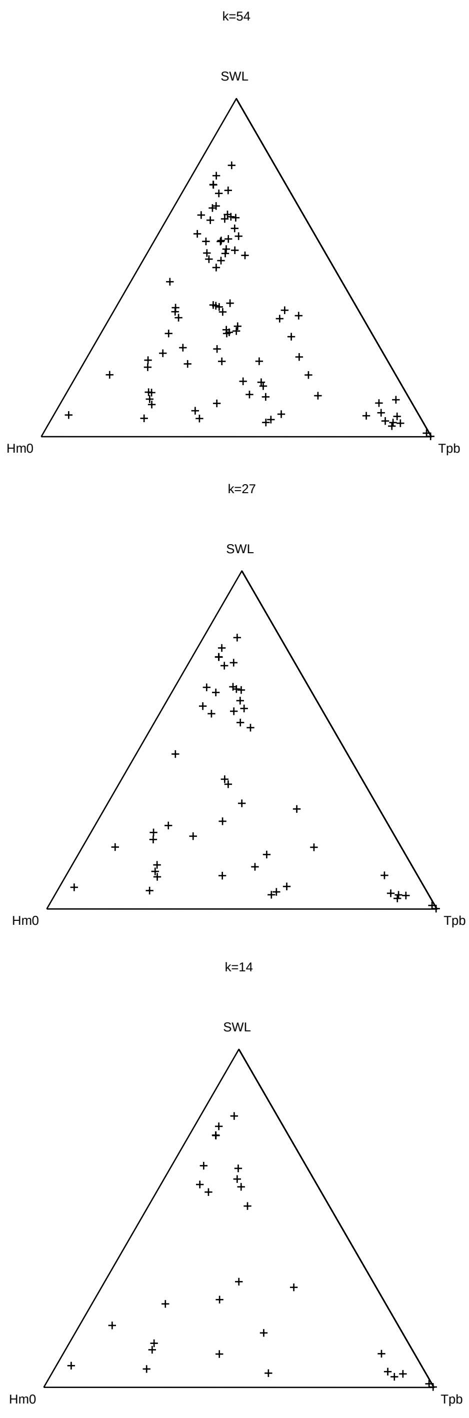

Figure 9. Trivariate angular coordinates representing the spectral measure. Scatter plots shown correspond to $k = 54$, 27 or 14 observations respectively.

we have now reduced the three-dimensional problem to an exceedence probability problem in one-dimensional space, which is easy to handle (Dijk and de Haan, 1992):

$$\Pr\{L(X,Y,Z) > 12.75\} \approx \frac{k}{n}\left(\max\left\{0, 1+\hat{\gamma}_n \frac{12.3-\hat{b}(n/k)}{\hat{a}(n/k)}\right\}\right)^{-1/\hat{\gamma}_n}$$

where the estimators $\hat{\gamma}_n, \hat{a}(n/k)$ and $\hat{b}(n/k)$ are obtained from the L-data set. Figure 10 shows the γ-estimator and the corresponding distribution tail. Figure 10 shows that the failure probability is effectively zero. This approach is also called the ''structure variable approach''.

6.5. *The asymptotic independence issue*

A special position in multidimensional extreme value theory is occupied by the situation when the two components of the vector of componentwise maxima are independent in the limit. This is a situation that occurs often in practice. Criteria for this to happen have been given by Geffroy (1958, 1959) and Sibuya (1960). As remarked earlier this phenomenon is reflected in the l-function (the graph of the function and hence also the Q-graph are completely flat) and the spectral measure Φ (which is concentrated in the extreme points).

Asymptotic dependence has little to do with the amount of independence of the component of the original random vector: consider i.i.d. random vectors $(U_i, 1-U_i)$ where U_i's are i.i.d. *uniform*(0,1) random variables $(i = 1, 2, \ldots)$. Then the limit distribution of the sequence $(n(\max_{i\le n} U_i - 1), n\max_{i\le n}(1-U_i))$ has independent components, although the original components are completely determined by each other. But of course, if the original components are independent, this is also true for the limiting distributions.

Perhaps the best way to formulate Sibuya's criterion in our context is to start from a version of the domain of attraction criterion, namely (4.4):

$$\lim_{n\to\infty} n\{1-F(a_n x + b_n, c_n y + d_n)\} = -\log G(x,y)$$

or equivalently,

$$\lim_{n\to\infty} n\Pr\left(\frac{X-b_n}{a_n} > x \text{ or } \frac{Y-d_n}{c_n} > y\right) = -\log G(x,y) \tag{6.1}$$

where (X, Y) has distribution function F. Now there are two possibilities: either

$$\lim_{n\to\infty} n\Pr\left(\frac{X-b_n}{a_n} > x \text{ and } \frac{Y-d_n}{c_n} > y\right) \tag{6.2}$$

is positive for all x and y in the range of G or it is zero for all those x and y. The latter case

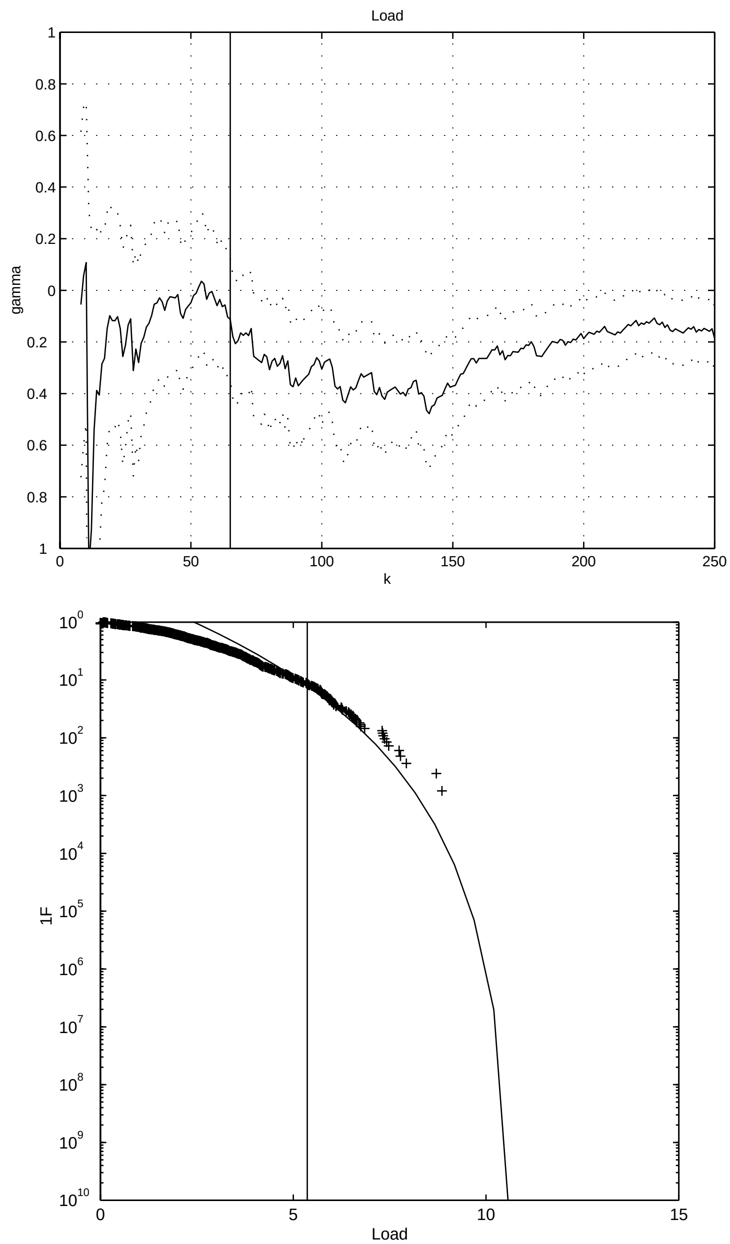

Figure 10. The univariate approach. The top figure illustrates the estimation of the extreme value index γ for the load variable. The bottom figure shows the exceedence probabilities for the load variable. The solid line shows the adapted tail distribution function. The vertical lines indicate the order statistics used to calculate the estimate ($k = 65$).

represents asymptotic independence. This means that in this case the probability mass of the set in (6.2) is of lower order than that of the set in (6.1) as $n \to \infty$. The probability of the set (6.1) (or more generally sets of similar form, i.e. that do not extend to $x = -\infty$ or $y = -\infty$) is completely unpredictable (i.e. impossible to estimate) in our model. So if the failure area is of that particular form, we are in trouble. That is the reason why a preliminary check on asymptotic independence is of primary importance.

Fortunately our data set does not exhibit asymptotic independence, but corresponding data sets for the British coast do. This is why the British statistical group in the Neptune project (J. Tawn and J. Bruun) has paid much attention to this problem and has in fact come up with a more refined model that accommodates for this complication (see e.g. Ledford and Tawn, 1997 and Bruun and Tawn, 1998).

Of course another and easy way out is to use the load variable approach of Subsection 6.4. But the model by Ledford and Tawn is much more interesting and it can be introduced in the following way.

Let us start from a special case (for simplicity) of the domain of attraction condition (cf. (4.25)):

$$\lim_{t\to\infty} t\{1 - F(tx, ty)\} = -\log G(x, y)$$

or if (X, Y) is a random vector with distribution function F,

$$\lim_{t\to\infty} t \Pr\{X > tx \text{ or } Y > ty\} = -\log G(x, y). \tag{6.3}$$

For many of the results mentioned in e.g. Section 5 we need a second order strengthening of this condition. Suppose there exists a positive function A such that for all $x, y > 0$

$$\lim_{t\to\infty} \frac{t \Pr\{X > tx \text{ or } Y > ty\} + \log G(x, y)}{A(t)} =: H(x, y) \tag{6.4}$$

exists, is finite and not identically zero. It is easy to see that the function A must be a regularly varying function with exponent less or equal to zero (cf. de Haan and Resnick, 1993). Suppose that in fact this exponent, ρ, is negative. Specialization of (6.4) to the two components X and Y gives

$$\lim_{t\to\infty} \frac{t \Pr\{X > tx\} - x^{-1}}{A(t)} = H(x, \infty) \tag{6.5}$$

$$\lim_{t\to\infty} \frac{t \Pr\{Y > ty\} - y^{-1}}{A(t)} = H(\infty, y). \tag{6.6}$$

Combination of (6.4), (6.5) and (6.6) gives

$$\lim_{t\to\infty} \frac{t\,\Pr\{X > tx, Y > ty\} - \log G(x,y) - x^{-1} - y^{-1}}{A(t)}$$
$$= H(x,\infty) + H(\infty,y) - H(x,y). \tag{6.7}$$

Now in the case of asymptotic independence

$$G(x,y) = e^{-1/x}e^{-1/y},$$

so that (6.7) reads as

$$\lim_{t\to\infty} \frac{t\,\Pr\{X > tx, Y > ty\}}{A(t)} = H(x,\infty) + H(\infty,y) - H(x,y) \neq 0. \tag{6.8}$$

Supposing that the limit is non-zero and comparing this relation with (6.3), i.e.

$$\lim_{t\to\infty} t\,\Pr\{X > tx \text{ or } Y > ty\} = -\log G(x,y),$$

we see that relation (6.8) indeed allows us to evaluate asymptotically the probability of the previously completely unpredictable tail events. The extra parameter $\eta := 1 - \rho$ is added to the model. See for further details on the model and its applications Ledford and Tawn (1997) and Bruun and Tawn (1998).

Acknowledgment

Comments by G. Draisma, H. Rootzén, J. Tawn, L. Peng and two anonymous referees have significantly improved the structure of the paper. Support for the first named author by the European Union (grant ERBFMBICT961465) and the the hospitality of the Departamento de Estatistística, University of Lisbon during a sabbatical period are gratefully acknowledged. Most pictures were produced by G. Draisma.

References

1. Bruun, J.T. and Tawn, J.A., A comparison of joint probability models and structure variable approaches for the estimation of coastal design levels, Technical Report. Neptune project T400: A Bivariate Simulation Study, (1995).
2. Bruun, J.T. and Tawn, J.A., ''Comparison of approaches for estimating the probability of coastal flooding,'' *Applied Statistics*, to appear (1998).
3. Coles, S.G. and Tawn, J.A., ''Modelling extreme multivariate events,'' *J. Royal Statist. Soc*. series B, (1991).
4. Coles, S.G. and Tawn, J.A., ''Statistical methods for multivariate extremes: an application to structural design,'' *Appl. Statist*. 43, 1–48, (1994).

5. Dekkers, A., Einmahl, J., and Haan, L. de, ''A moment estimator for the index of an extreme-value distribution,'' *Ann. Statist.* 17, 1833–1855, (1989).
6. Dijk, V. and Haan, L. de, ''On the estimation of the exceedence probability of a high level.'' Order Statistics and Nonparametrics: Theory and Applications. (P.K. Sen and I.A. Salama, ed), Elsevier, Amsterdam, 1992.
7. Dillingh, D. et al., De basispeilen langs de Nederlandse kust, Technical Report DGW-93.023, Rijks Instituut voor Kust en Zee, Rijkswaterstaat, 1993.
8. Draisma, G., Haan, L. de, Peng, L., and Sinha, A.K., Reports Neptune T400 EUR 4, 5, 6, 7, 8, 9, 10, 12; Reports EUR/RIKZ 96.2 and 96.3, 1996, 1997. Available at http://www.cs.few.eur.nl/few/people/ldehaan.
9. Einmahl, J., Haan, L. de, and Piterbarg, V., ''A nonparametric estimator for the spectral measure for extremes,'' Technical Report, 1997.
10. Einmahl, J., Haan, L. de, and Sinha, A.K., Estimating the spectral measure of an extreme value distribution to appear in Stoch Proc. Appl. (1997).
11. Feller, W., *An Introduction to Probability Theory and Its Applications*, Vol. II. Wiley, New York, 1966.
12. Geffroy, J., Contributions à la théorie des valeurs extrêmes. *Publ. Inst. Statist. Univ. Paris* 7/8, 37–185. 1958, (1959).
13. Gumbel, E.J., ''Distribution des valeurs extrêmes en plusieurs dimensions.'' *Publ. Inst. Statist. Univ. Paris* 9, 171–173, (1960).
14. Haan, L. de, ''Slow variation and characterization of domains of attraction,'' in Statistical Extremes and Applications (J. Tiago de Oliveira, ed), Reidel, Dordrecht, 1984.
15. Haan, L. de, Estimating Exceedance Probabilities in Higher-Dimensional Space. *Commun. Statist Stoch Models* 10, 765–780, (1994).
16. Haan, L. de and Resnick, S.I., Estimating the limit distribution of multivariate extremes. *Commun. Statist. Stoch Models* 9, 275–309. A corrected version (1995) is available from the authors, 1993.
17. Haan, L. de and Sinha, A.K., Estimation of the Failure Probability. Technical Report, 1997.
18. Huang, X., ''Statistics of bivariate extremes.'', Thesis, Erasmus University Rotterdam. Tinbergen Institute Research'' series no. 22, 1992.
19. Hurdle, D., Matlab routines for offshore to onshore transformation of wave parameters at the ''Pettemer zeewering'', Distributed through the Delft Hydraulics ftp site, 1995.
20. Hurdle, D., ''Prediction of the wave conditions at Petten.'' Alkyon: file DAVID.WOR, November 10, 1996.
21. Ledford, A.W. and Tawn, J.A., ''Statistics for near independence in multivariate extreme values,'' *Biometrika*, 83, 169–187, (1997).
22. Marle, J.G.A. van, Personal communication. December 30, 1994.
23. Marle, J.G.A. van, Hydraulic design conditions. Memo RIKZ /Rijkswaterstaat, June 17, 1996.
24. Meer, J.W. van der, Golfoploop en golfverslag bij dijken. Technical Report H638, WL Delft, in opdracht van Rijkswaterstaat, 1993.
25. Resnick, S.I., *Extreme Values, regular variation and point processes*, Springer, New York, 1987.
26. Roskam, B. and Hoekema, D., ''Readmemx.dos: Description of recorded data on waves, wind and water level at five locations in the North sea.'' Readme file accompanying RIKZ data files, May 1996.
27. Ronde, J.G. de et al., ''Wave conditions along the Dutch coast at considerable water depth.'' Ministerie van Verkeer en Waterstaat, rapport RIKZ-95.024, 1995.
28. Sibuya, M., ''Bivariate extreme statistics,'' *Ann. Inst. Stat. Math.* 11, 195–210, (1960).
29. Sinha, A.K., ''Estimating Failure Probability when failure is rare: multidimensional case.'' Thesis, Erasmus University Rotterdam. Tinbergen Institute Research series no. 165, 1997.
30. Smith, R.L., ''Extreme Value theory,'' in Handbook of Applicable Mathematics (W. Ledermann, ed), vol. 7. Wiley, New York, 1990.
31. Smith, R.L., ''Multivariate threshold methods''. in Galambos et al: Extreme value theory and applications, Kluwer, Dordrecht, 225–248, 1994.
32. Smith, R.L., Tawn, J.A., and Yuen, H.K., ''Statistics of multivariate extremes,'' *Int. Statist. Rev.* 58, 47–58, (1990).
33. Valk, C.F. de, ''Selection of storm events and estimation of exceedence frequencies of significant wave height for five North sea locations,'' Technical Report H1931, Delft Hydraulics, 1994.
34. Valk, C.F. de, ''Estimation of frequencies of failure of coastal structures from offshore data of environmental loads,'' Technical Report, Delft Hydraulics, February 1996.