Journal de la Société Française de Statistique
Vol. 154 No. 2 (2013)

Environmental data: multivariate Extreme Value Theory in practice.

Titre: Données environnementales : la théorie multivariée des valeurs extrêmes en pratique.

Cai Juan Juan[1], Fougères Anne-Laure[2] and Mercadier Cécile[2]

Abstract: Let (X_t, Y_t) be a bivariate stationary time series in some environmental study. We are interested to estimate the failure probability defined as $P(X_t > x, Y_t > y)$, where x and y are high return levels. For the estimation of high return levels, we consider three methods from univariate extreme value theory, two of which deal with the extreme clusters. We further derive estimators for the bivariate failure probability, based on Draisma et al. (2004)'s approach and on Heffernan and Tawn (2004)'s approach. The comparison on different estimators is demonstrated via a simulation study. To the best of our knowledge, this is the first time that such a comparative study is performed. Finally, we apply the procedures to the real data set and the results are discussed.

Résumé : Nous nous intéressons à deux variables quantitatives mesurées au cours du temps, formant ainsi une série temporelle bivariée (X_t, Y_t) supposée stationnaire. Nous souhaitons estimer une probabilité de défaillance, définie comme la probabilité $\mathbb{P}(X_t > x, Y_t > y)$, où x et y sont deux valeurs extrêmes (trop grandes pour être observées, ou presque). Plus précisément, x et y représentent des niveaux de retour dont l'estimation sera effectuée par trois méthodes concurrentes issues de la théorie univariée des valeurs extrêmes. La théorie multivariée des valeurs extrêmes fournira des estimateurs de la probabilité de défaillance prenant en compte la dépendance, pouvant subsister ou au contraire s'effacer lorsque l'on se focalise sur les valeurs extrêmes. Nous présenterons plusieurs méthodes d'estimation, fondées sur des approches introduites par Draisma et al. (2004) d'une part, et par Heffernan and Tawn (2004) d'autre part. Nous mettrons ensuite en concurrence les estimateurs déduits sur des simulations dans un premier temps, puis sur des données réelles environnementales. Les résultats obtenus seront finalement discutés.

Keywords: Return level estimation, Extremal index, Cluster, Estimation of failure probability, Multivariate extreme values, Extremal dependence, Environnemental data
Mots-clés : Estimation du niveau de retour, Indice extrêmal, Cluster, Estimation de probabilité de défaillance, Valeurs extrêmes multivariées, Dépendance extrêmale, Données environnementales
AMS 2000 subject classifications: 62G32, 62H99, 62P12

[1] Delft Institute of Applied Mathematics (DIAM), Faculty of Electrical Engineering, Mathematics and Computer Science, Delft University of Technology, Mekelweg 4, 2628 CD Delft, The Netherlands.
E-mail: caicaij@gmail.com

[2] Université de Lyon, CNRS UMR 5208, Université Lyon 1, Institut Camille Jordan, 43 blvd. du 11 novembre 1918, F-69622 Villeurbanne cedex, France
E-mail: fougeres@math.univ-lyon1.fr
E-mail: mercadier@math.univ-lyon1.fr

1. Introduction

Extreme value theory (EVT) is now considered as a "classical" mathematical framework to evaluate some exceptional risks. See for example the books of Leadbetter et al. (1983), Resnick (1987), Embrechts et al. (1997), Coles (2001), Beirlant et al. (2004), de Haan and Ferreira (2006), or Resnick (2007). EVT models make it possible to extrapolate outside the range of the observations using the highest observations. A typical issue is to estimate a "failure probability", that is to say the probability that a given random vector $(X_1,\dots,X_d)$ belongs to a "rare set" A. Here, the failure region A corresponds to a zone where damages should be observed, so that hopefully only very few (or no) observations have been observed in this set. In the multivariate framework discussed here, a capital element is to properly take into account the dependence between the different variables $X_1,\dots,X_d$. This has been for example noticed by Patrick Galois (Météo France), who explained in the newspaper *Le Monde* of March 1st, 2010 : "Si la tempête Xynthia présente un caractère remarquable, elle n'est pas pour autant le phénomène du siècle. Elle est ainsi moins exceptionnelle que celles de 1999 et ses vents sont moins intenses qu'en 2009. Mais son issue dramatique réside dans sa conjonction à un fort coefficient de marée sur la côte atlantique, au moment même de la marée haute. Ce sont ces trois phénomènes naturels réunis qui ont provoqué les inondations des côtes et des dégâts humains et matériels." A misspecification of the dependence between the variables can induce a substancial under-estimation of the risk. See for example the contributions of Bruun and Tawn (1998), de Haan and de Ronde (1998), and Chapter 8 of de Haan and Ferreira (2006) for more on this subject.

This paper applies EVT to answer a concrete question raised in the environmental context. The work presented here is the result of a collaboration with Électricité de France. Due to confidentiality constraints, we will not describe in detail the environmental variables that have been considered, but we will start instead from a preprocessed data set, that can be considered as stationary. The data of interest have two univariate components that can be considered as stationary time series (X_t) and (Y_t). The aim of this paper is to estimate via different methods based on EVT a pre-defined failure probability of the form

$$\mathbb{P}\left(X_t > x_{t_X}, Y_t > y_{t_Y}\right),$$

where x_{t_X} and y_{t_Y} are respectively the t_X*-year return level* of X and the t_Y*-year return level* of Y, for given t_X and t_Y, as defined in Section 2.1. Due to confidentiality constraints again, we will not give the values of t_X and t_Y that have been considered.

The study has been driven in several stages, that will give the structure of the paper. Section 2 concerns the analysis of the extremal behaviors of both univariate components X and Y. In this section, the t_X-year return level of X and the t_Y-year return level of Y will be estimated by three methods, see Sections 2.2 to 2.4. A conclusion of this part is given in Section 2.5. Section 3 first introduces different estimators of the probability of a bivariate failure set, based on Draisma et al. (2004)'s approach (see Section 3.1) and on Heffernan and Tawn (2004)'s approach (see Section 3.2). A simulation study is then presented in Section 3.3 that explores the performances of the different competitors. To the best of our knowledge, this is the first time that such a comparative study is done. Finally, the analysis of the real data set is provided in Section 3.4, and some conclusions are drawn in Section 4.

2. Univariate analysis

A first step in the analysis is to consider both variables separately. So we are here concerned by the analysis of an extremal behavior in a univariate setting.

2.1. Definition of the Return Level

Let X be the quantity of interest. The t_X*-year return level*, denoted by x_{t_X}, is defined to be the $\frac{1}{t_X}$-tail quantile of the annual maximum of X ; see for instance (Coles, 2001, page 49). Namely, the level x_{t_X} satisfies

$$\mathbb{P}\left(\max_{s\in\text{year}} X_s > x_{t_X}\right) = \frac{1}{t_X} \, ,$$

where $\max_{s\in\text{year}} X_s$ denotes the annual maximum of X. In our setting, the stationary phenomenon is observed each year on a grid of size n_0. Then one can approximate x_{t_X} with the level defined by

$$\mathbb{P}\left(\max_{i=1,\ldots,n_0} X_i > x_{t_X}\right) = \frac{1}{t_X} \, .$$

Note that in the displayed formula above, the X_i's are *not necessarily independent*.

Since we are interested in the t_X-year return level of X and t_Y-year return level of Y, with both t_X and t_Y large, the empirical quantile on annual maxima are not directly tractable. Three extrapolation methods based on extreme value theory will be used instead, as presented below.

2.2. Estimation of the return level: GEV method

It is possible to fit a generalized extreme value (GEV) distribution on the sample of annual maxima. Such distribution depends on three parameters: μ denotes the location parameter, σ the scale parameter and γ the shape parameter. The cumulative distribution function is given by

$$G_{\mu,\sigma,\gamma}(x) := \exp\left(-(1+\gamma\frac{x-\mu}{\sigma})^{-1/\gamma}\right) \, ,$$

for $1+\gamma\frac{x-\mu}{\sigma} \geq 0$. The case $\gamma = 0$ is obtained as the limit of the formula. Once the model is fitted, the return levels are obtain as the quantile of the generalized extreme value distribution. Even if it is not the best choice since the number of observations is drastically reduced, this method is very well known. Figure 1 and 2 give goodness-of-fit diagnosis of the annual maxima with the GEV distribution, and show that the GEV model used fits quite well. The independence assumption of the sample is debatable, as can be seen for small values of lag on Figure 3 and 4 via Ljung-Box test. Recall that it compares the set of the first autocorrelations with the set of autocorrelations that are all zero, see Ljung and Box (1978). For a reasonable value of lag (here 10, chosen by taking into account our sample size), this test gives a p-value equal to 0.2035 for Variable X and to 0.2366 for Variable Y.

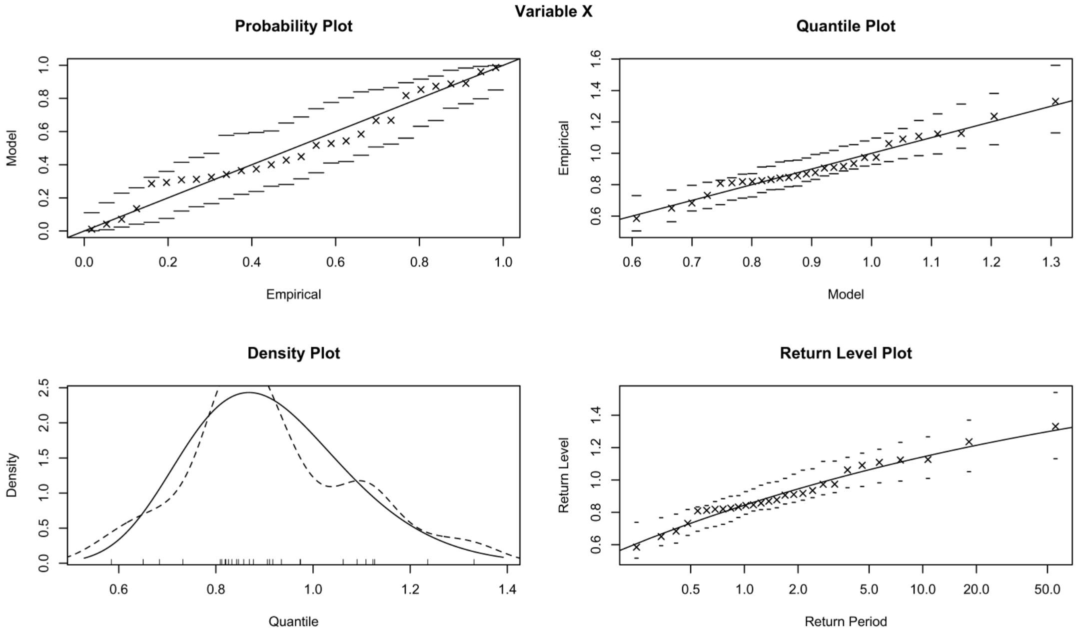


Figure 1: GEV plots for annual maxima of Variable X.

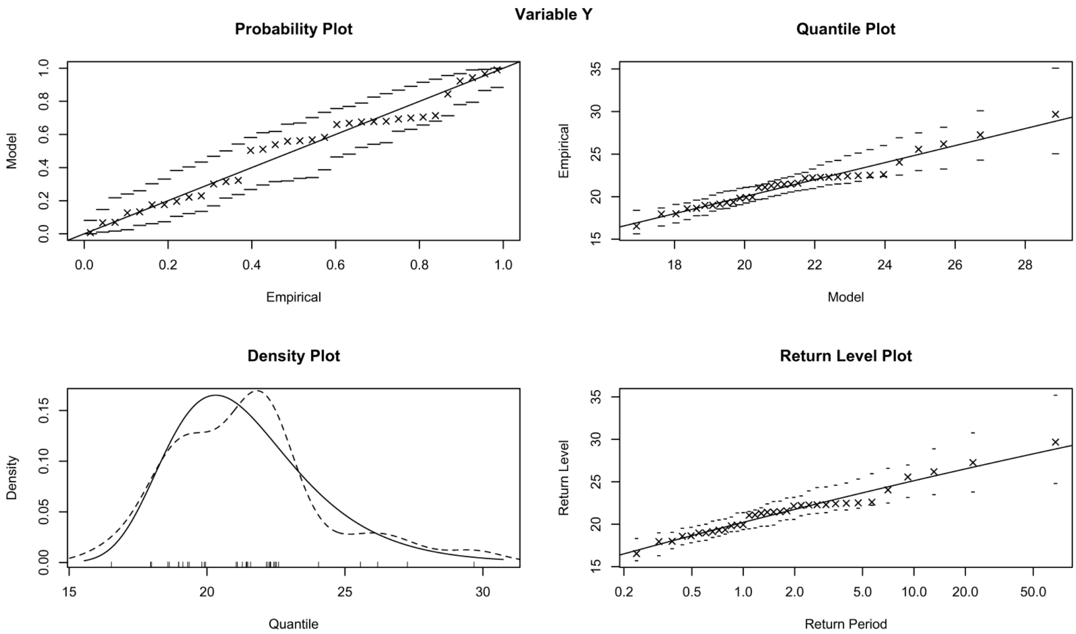


Figure 2: GEV plots for annual maxima of Variable Y.

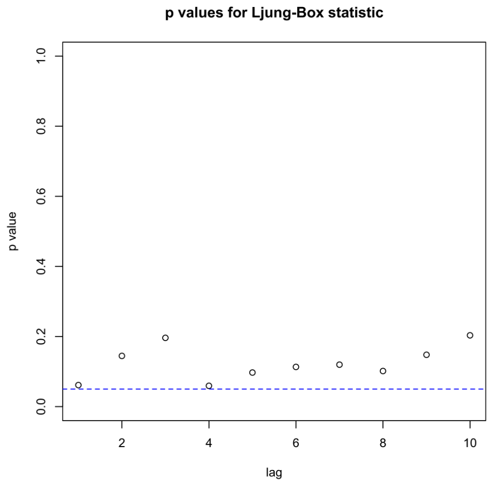


Figure 3: Ljung-Box tests for annual maxima of Variable X.

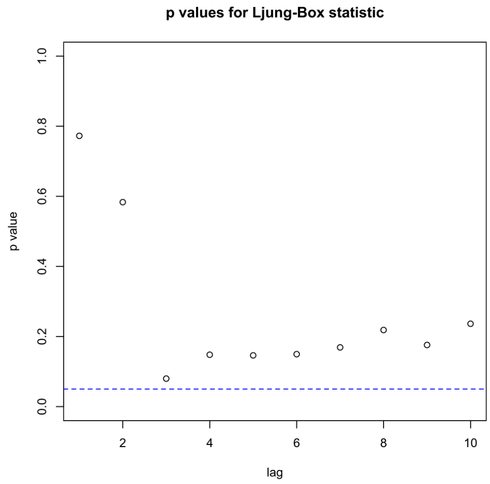


Figure 4: Ljung-Box tests for annual maxima of Variable Y.

Table 1 summarizes our estimations for t_X and t_Y-year return levels. Due to the dramatically reduced sample size and distant extrapolating, the confidence intervals (see Table 1) are too wide to be really informative. This leads to consider more sophisticated methods.

TABLE 1. *Estimation of return levels for X and Y with GEV model. The* 95% *confidence interval (denoted by* 95% *CI) is based on asymptotic normality of the estimator.*

	x_{t_X}	y_{t_Y}
Point estimate	1.506	29.585
95% CI	$[1.308, 2.681]$	$[26.933, 39.601]$

2.3. *Estimation of the return level: Extremal index method*

The serial dependence of time series affects the behavior of the extremes and requires different and more sophisticated statistical tools than those used for independent and identically distributed (i.i.d.) data. Let G be the distribution function of the maximum of the stationary collection $X_1, \dots, X_n$, and $\tilde{G}$ denote the distribution function of the maximum of an i.i.d. sequence $\tilde{X}_1, \dots, \tilde{X}_n$ with same distribution as X_1.

Thanks to Leadbetter (1983)'s theorem, the two distribution functions $\tilde{G}$ and G are linked in a simple way via $G := \tilde{G}^\theta$, where $\theta \in [0,1]$ is called the *extremal index*. Thus, the t_X-return level is the value x_{t_X} that will be calculated through the following approximation

$$\mathbb{P}\left(\max_{i=1,\dots,n_0} X_i \leq x_{t_X}\right) \simeq \mathbb{P}\left(\tilde{X}_i \leq x_{t_X}\right)^{n_0 \times \theta} .$$

In this approach, x_{t_X} is the quantile of order $(1 - 1/t_X)^{1/(n_0 \times \theta)}$ of X_1. In real applications, the order remains very high so that an extrapolation based on the extreme value theory is needed in order to estimate x_{t_X}. For a fixed threshold u, one fits the exceedances by the generalized Pareto distribution, denoted by GPD(γ,σ), which cumulative distribution function is

$$H_{\sigma,\gamma}(x) := \begin{cases} 1-(1+\gamma\frac{x-u}{\sigma})^{-1/\gamma} & if\,\gamma \neq 0 \\ 1-\exp\left(-\frac{x-u}{\sigma}\right) & if\,\gamma = 0\,, \end{cases}$$

for $x \geq u$ when $\gamma \geq 0$ and $u \leq x \leq u-\sigma/\gamma$ otherwise. As for the GEV, the parameter σ is a scale parameter whereas γ is a shape parameter. For u chosen as $X_{n-k+1,n}$ (the k-th greatest value of the observed series), the $(1-1/t_X)^{1/(n_0\times\theta)}$ quantile is estimated by:

$$\hat{x}_{t_X} = X_{n-k+1,n} + \hat{\sigma}\frac{\left(\frac{k}{\hat{\alpha}n}\right)^{\hat{\gamma}}-1}{\hat{\gamma}}\,,$$

where $\hat{\alpha} = 1-(1-1/t_X)^{1/(n_0\times\theta)} \simeq 1/(t_X\times n_0\times\theta)$.

Let summarize the different steps of the estimation procedure.

– Estimate θ of the stationary series with the method proposed in Ferro and Segers (2003). Denote the estimator with $\hat{\theta}$.
– Fit the tail of X_1 with GPD. Specifically, we apply the classical maximum likelihood method (see e.g. Beirlant et al. (2004)) to the data as if the observations were independent.
– Keep the $1-(1-1/t_X)^{1/(n_0\times\hat{\theta})}$ quantile of X as the estimate of the t_X-year return level.

Note that the second step mentioned above makes use of a method the asymptotic normality of which has been typically studied for independent and identically distributed data. Under short-term dependence, the estimators remain consistent however present different asymptotic variance (see e.g. (Coles, 2001, chapter 5)).

Figures 5 and 6 show the return level estimations for Variables X and Y respectively. In these plots, the x-axis is the *thresholding probability*, which is the empirical cumulative probability of the threshold chosen for the estimation of θ and the GPD fit. For instance, in the estimation of a large quantile by the GPD above, the thresholding probability is the empirical probability to be lower than the threshold $X_{n-k+1,n}$ and thus corresponds to $(n-k)/n$. A simple aid in the selection of a suitable threshold is to look for the threshold stability of the estimator, i.e. to choose the lowest threshold above which the estimates are approximately constant. In our case, the probability .9 seems to be an acceptable choice for the thresholding probability.

To construct confidence intervals, we resort to a bootstrap method for stationary sequences (Politis and Romano (1994)). One can find a different bootstrap method in Ferro and Segers (2003). Let λ be the parameter of interest, and $\hat{\Lambda}_m$ be its m-bootstrap estimates.

– **Bootstrap normal method (BNM)** The $(1-2\alpha)$.100% CI of λ is given by

$$\left[\hat{\lambda}-z_{1-\alpha}\hat{\sigma}_{\hat{\lambda}^*}, \hat{\lambda}+z_{1-\alpha}\hat{\sigma}_{\hat{\lambda}^*}\right]\,,$$

where $\hat{\sigma}_{\hat{\lambda}^*}$ is the sample standard deviations of $\hat{\Lambda}_m$ and where $z_{1-\alpha}$ represents the quantile of order $1-\alpha$ for the standard normal distribution.

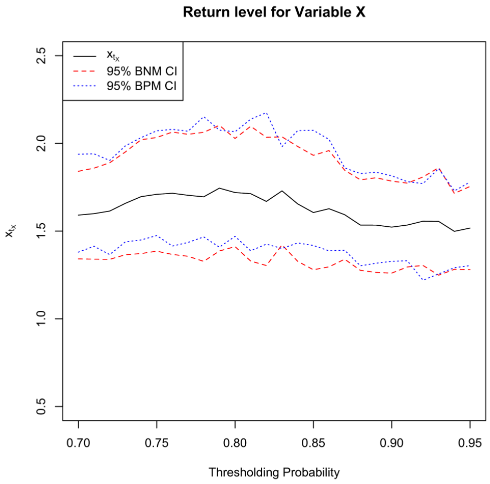


Figure 5: Return level estimates for Variable X based on the Extremal index method.

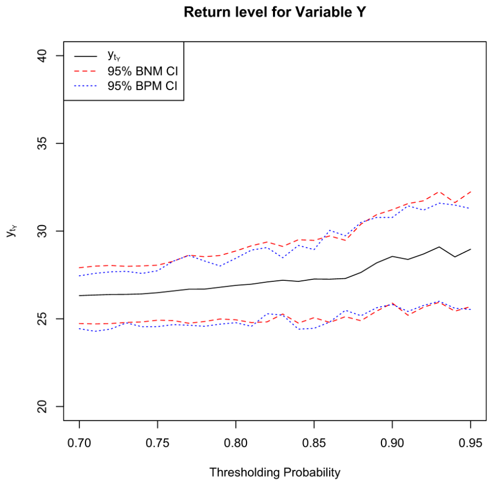


Figure 6: Return level estimates for Variable Y based on the Extremal index method.

– **Bootstrap percentile method (BPM)** The α and $1-\alpha$ sample percentiles of $\hat{\Lambda}_m$ define the $(1-2\alpha).100\%$ CI of λ by

$$[\hat{\lambda}^*_\alpha, \hat{\lambda}^*_{1-\alpha}] \, .$$

We also report the point estimates and the CI of x_{t_X} and y_{t_Y} in Table 2. The values correspond to the thresholding probability .9.

TABLE 2. *Estimation of the return level for Variable X and Variable Y with the Extremal index method.*

	x_{t_X}	y_{t_Y}
Point estimate	1.523	28.56
95% BNM CI	$[1.288, 1.757]$	$[25.569, 31.551]$
95% BPM CI	$[1.311, 1.778]$	$[25.659, 31.64]$

2.4. *Estimation of the return level: Declustering method*

In terms of extremes, the short-term dependence induces that large values will have a tendency to occur in clusters. A classical way to take this dependence into account is thus to focus on maximum per cluster. This approach is based on the following decomposition

$$\mathbb{P}\left(\max_{i=1,\dots,n_0} X_i \le x_{t_X}\right) \simeq \mathbb{P}\left(\max_{j=1,\dots,c}\max_{\ell\in\mathscr{C}_j} X_\ell \le x_{t_X}\right) ,$$

where c denotes the mean number of clusters per period, and $\mathscr{C}_j$ denotes the j-th cluster. One needs to estimate the mean number of clusters per year c, and for this, we will simply consider $\hat{c} = N_c/N$, where N_c is the total number of clusters, and N is the number of years considered in the study. Rewrite $Z_j := \max_{\ell \in \mathscr{C}_j} X_\ell$, and assume that the clusters $\mathscr{C}_j$ are defined in such a way that the Z_j's form an independent and identically distributed collection of size c. Then

$$\mathbb{P}\left(\max_{i=1,\dots,n_0} X_i \le x_{t_X}\right) \simeq \left[\mathbb{P}\left(Z_j \le x_{t_X}\right)\right]^c .$$

In this approach, x_{t_X} is therefore seen as the quantile of order $(1-1/t_X)^{1/c}$ of the cluster maxima Z_j. Again the estimation of the high order quantile x_{t_X} requires an extrapolation based on Extreme Value Theory. In the same way as in the Extremal index method described above, exceedances over a fixed threshold $Z_{n-k+1,n}$ are fitted by a GPD(γ,σ) based on the sample $\{Z_j, j = 1,\dots,c\}$, and x_{t_X} is obtained via the following formula:

$$\hat{x}_{t_X} = Z_{n-k+1,n} + \hat{\sigma}\frac{\left(\frac{k}{\hat{\alpha}n}\right)^{\hat{\gamma}} - 1}{\hat{\gamma}} ,$$

where $\hat{\alpha} = 1-(1-1/t_X)^{1/\hat{c}} \simeq 1/(t_X\hat{c})$. Note that this last method can actually be seen as a particular case of the second one, where the extremal index is estimated as the inverse of the mean size of a cluster.

Let summarize the different steps of the estimation procedure.

- Apply the automatic declustering method proposed by Ferro and Segers (2003), and exhibit a subset of cluster maxima denoted by $Z_1,\dots,Z_{N_c}$. See Beirlant et al. (2004, page 393) for a sketched presentation.
- Check that the $Z_1,\dots,Z_{N_c}$ can be considered as independent.
- Fit the tail of the Z_i's with GPD. Again, the classical maximum likelihood method (see e.g. Beirlant et al. (2004)) is used.
- Keep the $1-(1-1/t_X)^{N/N_c}$ tail quantile of Z_1 as the estimate of t_X-year return level.

Ferro and Segers (2003) automatic declustering method relies on the estimation of the extremal index of a sequence of excesses above a threshold which allows the identification of independent clusters of excesses above that threshold. See their paper for more details. This has been recently implemented by Southworth and Heffernan (2012).

The point estimators of the return level are provided on Figure 7 for Variable X and Figure 8 for Variable Y, as functions of the thresholding probability. The two confidence intervals provided are those described in Section 2.3.

According to Figures 7 and 8, empirical quantile with probability 0.9 might be considered as a suitable threshold. For this threshold, apply the first step of the algorithm. In order to check that these selected data $Z_1,\dots,Z_{N_c}$ can be considered as independent (second step of the algorithm), we conduct three independence tests; the first one (denoted by IT1) is the graphical Ljung-Box test (see e.g. Brockwell and Davis (2002)), the second one (denoted by IT2) is the serial independence test proposed by Genest et al. (2007), and the third one (denoted by IT3) is a graphical permutation test, shortly explained here: 99 pseudo-samples are built from random permutations of the $Z_1,\dots,Z_{N_c}$, for each of which the auto-correlation function (acf) is drawn. Under assumption of independence, the acf of the original sample $Z_1,\dots,Z_{N_c}$ should be within the set built by the resulting acfs.

Figures 9 and 11 provide the Ljung-Box test (IT1) for the cluster maxima built from the first step of the algorithm, for the Variables X and Y respectively. Making use of the global Cramer-von

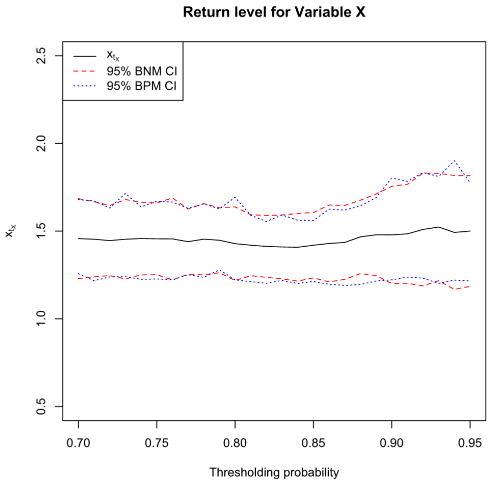


Figure 7: Return level estimator for Variable X based on the Declustering method.

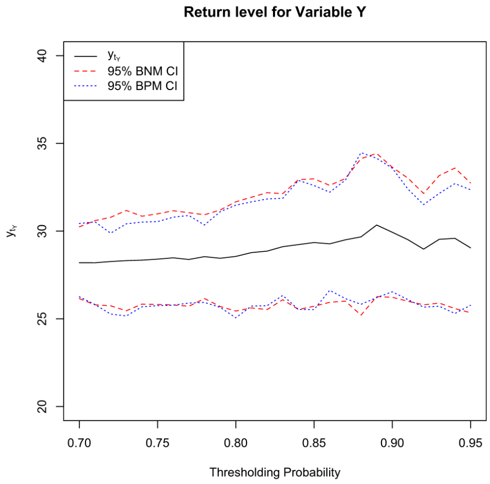


Figure 8: Return level estimator for Variable Y based on the Declustering method.

Mises statistic for IT2 gives respectively a p-value of 0.1233766 for Variable X, and of 0.453047 for Variable Y. Finally, Figures 10 and 12 provide the permutation test IT3 for Variables X and Y respectively. It consists in a visual appreciation of the independence through the acf (from lag 1) of 100 permutations of the original sample (gray lines) and the acf of the original series (black line). These tests all strengthen the hypothesis that the cluster maxima obtained from Ferro and Segers (2003) automatic declustering method (used with thresholding probability 0.9) can be considered as independent.

We report the point estimates and the CI of x_{t_X} and y_{t_Y} in Table 3. The values are obtained at the thresholding probability .9.

TABLE 3. *Estimation of the return level for Variables X and Y with the declustering method.*

	x_{t_X}	y_{t_Y}
Point estimate	1.479	29.936
95% BNM CI	$[1.231, 1.726]$	$[26.105, 33.768]$
95% BPM CI	$[1.219, 1.740]$	$[25.993, 33.543]$

2.5. Summary of the univariate studies

The return levels estimated thanks to the three previous methods (GEV method of Section 2.2, Extremal index method of Section 2.3 and Declustering method of Section 2.4) are plotted on the same graphics, see Figure 13 for Variable X and Figure 14 for Variable Y. In these plots, the x-axis is the return period. The six circles are the empirical estimations (based on annual maxima) of the T-return levels calculated for the particular values $T = 4, 9, 14, 19, 24$ and 29. This method performs well for big but observable return period T. One can check from Figure 13 that the three methods behave quite similarly on Variable X. It is less satisfying for Variable Y (see Figure 14),

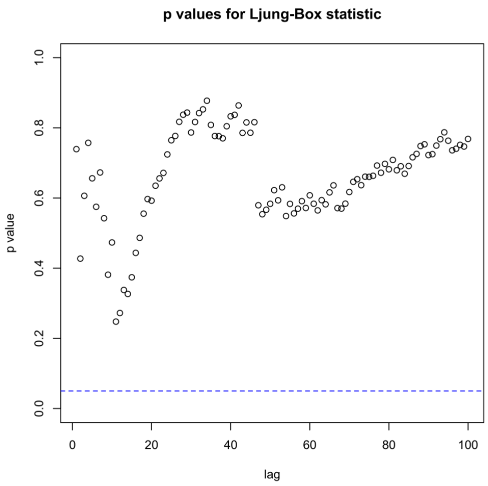


Figure 9: Ljung-Box tests (IT1) for cluster maxima of Variable X.

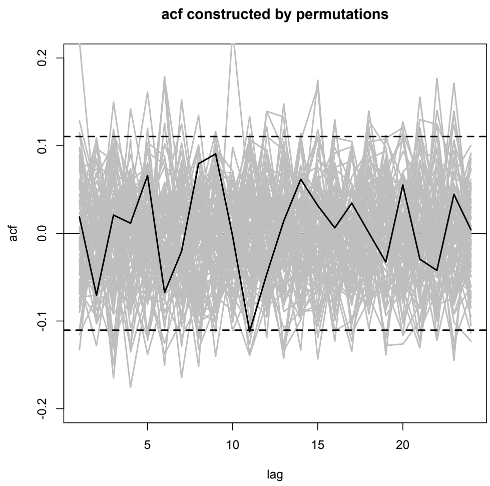


Figure 10: Permutation test (IT3) for cluster maxima of Variable X.

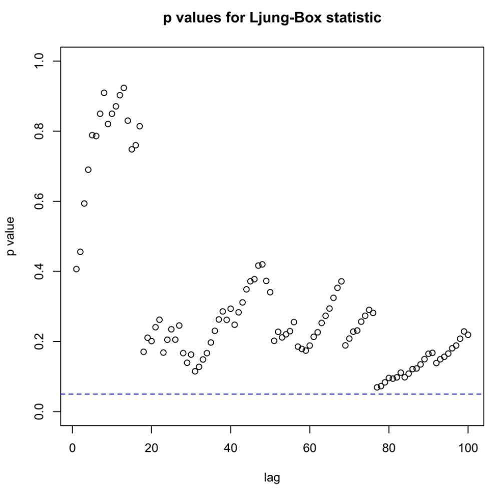


Figure 11: Ljung-Box tests (IT1) for cluster maxima of Variable Y.

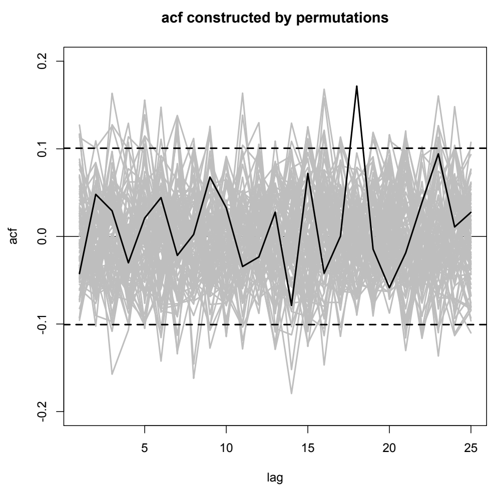


Figure 12: Permutation test (IT3) for cluster maxima of Variable Y.

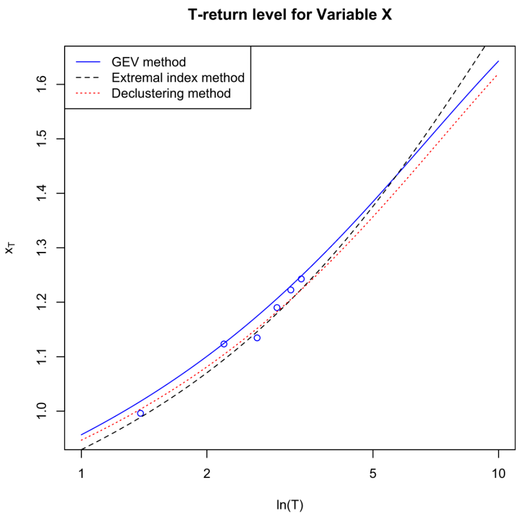


Figure 13: Return levels for Variable X: via GEV method, Extremal index method and Declustering method.

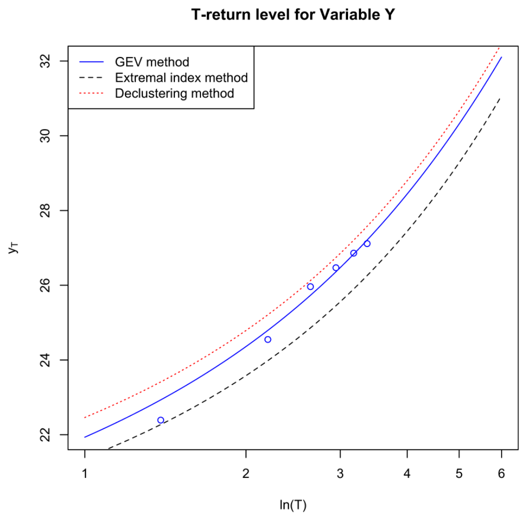


Figure 14: Return levels for Variable Y: via GEV method, Extremal index method and Declustering method.

where one observes three parallel curves. The reasons could be a choice of threshold that has to be improved, or a stationarity assumption that has to be carefully revisited. However, both confidence tubes (that are omitted here) contain the curves.

3. Bivariate analysis

In this section, we study several methods for estimating bivariate failure probability. First we describe in Sections 3.1 and 3.2 the approaches in details and then implement a simulation study to justify the performance of the methods (see Section 3.3). Afterwards we consider the real data to obtain knowledge on the extremal dependence structure of Variables X and Y, and statistical inference on the targeted failure probability is made (see Section 3.4).

Let $(X_i, Y_i), i = 1, \ldots, n$, be a stationary sequence from the distribution of (X, Y). Recall that we are interested in estimating the probability of the following set

$$\{X_t > x_{t_X}, Y_t > y_{t_Y}\}, \tag{1}$$

where x_{t_X} and y_{t_Y} are the return levels of X and Y estimated in Section 2. To this end, we need statistical techniques from bivariate extreme value theory, which involves both the study of marginal distributions and that of the tail dependence. The marginal distributions have been analyzed carefully in Section 2. Consequently, our remaining task is to model the bivariate tail dependence. For comprehensive study on multivariate extreme statistics, see for instance Chapters 6 – 8 in de Haan and Ferreira (2006) and Chapters 8 and 9 in Beirlant et al. (2004).

For our specified goal of estimating the failure probability, we will use two different approaches to model the problem. Five estimators of the failure probability will be deduced that we explain in details in the following two subsections.

3.1. Bivariate regular variation approach

This method is based on Draisma et al. (2004) and further tailored for our purpose. Let F be the distribution function of (X,Y) with marginal distributions F_X and F_Y. Suppose that F is in the domain of attraction of a bivariate extreme value distribution. This implies that firstly, the marginal distributions F_X and F_Y are in the domain of attraction of univariate extreme value distributions and secondly that

$$\lim_{t\to\infty} \frac{P(1-F_X(X) < x/t, 1-F_Y(Y) < y/t)}{P(1-F_X(X) < 1/t, 1-F_Y(Y) < 1/t)} =: c(x,y) \tag{2}$$

exists, for $x,y > 0$. The function c is thus homogeneous of order $1/\eta$ for some $\eta \in (0,1]$. The index η is called the *coefficient of tail dependence* of (X,Y). It was first studied in Ledford and Tawn (1996).

Remark: η characterises the pattern of the extremal dependence of (X,Y).

- Asymptotic dependence ($\eta = 1$): Under this situation, $\lim_{t\to\infty} P(X > x_t | Y > y_t) > 0$; in words, the extremes of X and Y tend to occur simutaneously. The tail feature of (X,Y) is often studied using the *stable tail dependence funtion* $l(x,y) := \lim_{t\to\infty} tP(1-F_X(X) < x/t \text{ or } 1-F_Y(Y) < y/t)$.
- For $\eta < 1$, the variables are asymptotically independent as $\lim_{t\to\infty} P(X > x_t | Y > y_t) = 0$. We can categorize them into three classes.
  - Positive association ($\eta \in (1/2,1)$): For this type of distributions, one has $P(X > x_t, Y > y_t) \gg P(X > x_t)P(Y > y_t)$. The joint extremes of (X,Y) happen more often than those of the distribution wih independent components X and Y.
  - Near extremal independence ($\eta = 1/2$): The joint tail of (X,Y) behaves as if X and Y were independent.
  - Negative association ($\eta \in (0,1/2)$): The joint extremes of (X,Y) happen less often than those when X and Y are independent as $P(X > x_t, Y > y_t) \ll P(X > x_t)P(Y > y_t)$.

3.1.1. Estimating the coefficient of tail dependence η

We follow the procedure suggested in Draisma et al. (2004). Write $T = \min\left(\frac{1}{1-F_X(X)}, \frac{1}{1-F_Y(Y)}\right)$. It comes from (2) that for $x > 0$,

$$\lim_{t\to\infty} \frac{P(T > tx)}{P(T > t)} = x^{-1/\eta}. \tag{3}$$

This means that the distribution of T has a heavy right tail with extreme value index η. It is natural to consider the estimator of η based on a sample of T. Write for $1 \le i \le n$,

$$T_i = \min\left(1/P_i^X, 1/P_i^Y\right),$$

where $P_i^X = 1 - F_X(X_i)$ and $P_i^Y = 1 - F_Y(Y_i)$. Note that we do not observe the sample $\{T_i, i = 1\dots,n\}$ as the marginal distributions F_X and F_Y are unknown. We obtain pseudo observations of T by estimating (P_i^X, P_i^Y), for $i = 1, \ldots, n$. The empirical counterparts are the straightforward estimators. Put

$$\widetilde{T}_i := \min\left(\frac{1}{1-\frac{R_i^X}{n+1}}, \frac{1}{1-\frac{R_i^Y}{n+1}}\right), i = 1\ldots,n, \tag{4}$$

where R_i^X is the rank of X_i among $(X_1, \ldots, X_n)$ and R_i^Y that of Y_i among $(Y_1, \ldots, Y_n)$. The collection $\{\widetilde{T}_i, i = 1, \ldots, n\}$ forms the empirical pseudo sample of T, denoted by T_{emp}.

An alternative way is to make use of the fact that the marginal distributions are in some max-domain of attraction which enables us to estimate (P_i^X, P_i^Y) with empirical or GPD method. Precisely, for $1 \le i \le n$

$$\hat{P}_i^X = \begin{cases} 1 - \frac{R_i^X}{n+1} & \text{if } X_i \le u_1; \\ p_{u_1}\left(1+\hat{\gamma}_1\frac{X_i-u_1}{\hat{\sigma}_1}\right)^{-1/\hat{\gamma}_1} & \text{otherwise;} \end{cases} \quad \text{and} \quad \hat{P}_i^Y = \begin{cases} 1 - \frac{R_i^Y}{n+1} & \text{if } Y_i \le u_2; \\ p_{u_2}\left(1+\hat{\gamma}_2\frac{Y_i-u_2}{\hat{\sigma}_2}\right)^{-1/\hat{\gamma}_2} & \text{otherwise,} \end{cases}$$

where the knowledge on $(\hat{\gamma}_i, \hat{\sigma}_i, u_i, p_{u_i})$, $i = 1,2$ comes from Section 2. Now define

$$\overline{T}_i := \min\left(\frac{1}{\hat{P}_i^X}, \frac{1}{\hat{P}_i^Y}\right), i = 1\ldots,n. \tag{5}$$

We call $T_{gpd} = \{\overline{T}_i, i = 1, \ldots, n\}$ the GPD pseudo sample of T. Then we apply Hill (1975) and ML methods to the pseudo samples of T to obtain the estimators of η, and use the following notation: $\hat{\eta}_{he}$ is the Hill estimator on T_{emp}; $\hat{\eta}_{me}$ is the ML estimator on T_{emp}; $\hat{\eta}_{hp}$ is the Hill estimator on T_{gpd}; $\hat{\eta}_{mp}$ is the ML estimator on T_{gpd}.

3.1.2. Estimating the failure probability

The probability of (1) can be written as

$$P(1 - F_X(X) < p_1, 1 - F_Y(Y) < p_2) =: p, \tag{6}$$

where $p_1 = 1 - (1 - 1/t_X)^{1/(n_0\theta_1)}$ and $p_2 = 1 - (1 - 1/t_Y)^{1/(n_0\theta_2)}$. See Section 2 for the value of θ_i, $i = 1,2$. Considering the bivariate regular variation of $(1 - F_X(X), 1 - F_Y(Y))$ as written in (2), we employ the usual extrapolation technique in extreme value theory to approximate p. Let $p_0 = p_0(n)$ be a small probability such that $p_0 \to 0$, $np_0 \to \infty$ as $n \to \infty$. Write $c_0 = \sqrt{p_1^2 + p_2^2}$. Replacing t in (2) with $\frac{1}{p_0}$ and $\frac{1}{c_0}$ respectively, we have

$$\frac{P\left(1 - F_X(X) < c_0\frac{p_1}{c_0}, 1 - F_Y(Y) < c_0\frac{p_2}{c_0}\right)}{P\left(1 - F_X(X) < c_0, 1 - F_Y(Y) < c_0\right)} \approx \frac{P\left(1 - F_X(X) < p_0\frac{p_1}{c_0}, 1 - F_Y(Y) < p_0\frac{p_2}{c_0}\right)}{P\left(1 - F_X(X) < p_0, 1 - F_Y(Y) < p_0\right)},$$

and thanks to (3), one gets

$$\frac{P\left(1 - F_X(X) < c_0, 1 - F_Y(Y) < c_0\right)}{P\left(1 - F_X(X) < p_0, 1 - F_Y(Y) < p_0\right)} \approx \left(\frac{c_0}{p_0}\right)^{1/\eta}.$$

Combining these two relations, p can be approximated as following

$$p \approx \left(\frac{c_0}{p_0}\right)^{1/\eta} P\left(1 - F_X(X) < p_0 \frac{p_1}{c_0}, 1 - F_Y(Y) < p_0 \frac{p_2}{c_0}\right).$$

In order to estimate p, it is sufficient to estimate $P\left(1 - F_X(X) < p_0 \frac{p_1}{c_0}, 1 - F_Y(Y) < p_0 \frac{p_2}{c_0}\right)$. We derive two estimators from this approximation:

- Observe that $0 < \frac{p_1}{c_0}, \frac{p_2}{c_0} < 1$ and $np_0 \to \infty$. Roughly speaking, there is a sufficient amount of observations falling in this range, which makes the empirical estimation feasible. Our first estimator of p is defined as

$$\hat{p}_e = \left(\frac{c_0}{p_0}\right)^{1/\hat{\eta}} \frac{1}{n} \sum_{i=1}^{n} I\left(1 - \frac{R_i^X}{n+1} < p_0 \frac{p_1}{c_0}, 1 - \frac{R_i^Y}{n+1} < p_0 \frac{p_2}{c_0}\right), \tag{7}$$

  where I denotes the indicator function and where $\hat{\eta}$ is one of the estimators discussed in Section 3.1.1.
- Utilizing the GPD estimators for the tail probabilities as in (5), we obtain the second estimator

$$\hat{p}_p = \left(\frac{c_0}{p_0}\right)^{1/\hat{\eta}} \frac{1}{n} \sum_{i=1}^{n} I\left(\hat{P}_i^X < p_0 \frac{p_1}{c_0}, \hat{P}_i^Y < p_0 \frac{p_2}{c_0}\right). \tag{8}$$

The third estimator of p is a *structure variable method*. Motivated by the nice property of T, we construct a structure variable T_r as following. Let $r = p_2/p_1$ and $T_r = \min\left(\frac{1}{1-F_X(X)}, \frac{r}{1-F_Y(Y)}\right)$. Note that p is linked to T_r by

$$P(1 - F_X(X) < p_1, 1 - F_Y(Y) < p_2) = P(T_r > 1/p_1).$$

In other words, p is the tail probability of T_r being above $\frac{1}{p_1}$. Similarly to (3), it follows that the distribution of T_r has a heavy right tail with extreme value index η.

- Based on the empirical pseudo sample of T_r defined as

$$\widetilde{T}_{ri} := \min\left(\frac{1}{1 - \frac{R_i^X}{n+1}}, \frac{r}{1 - \frac{R_i^Y}{n+1}}\right), i = 1, \ldots, n, \tag{9}$$

  the third estimator of p is given by

$$\hat{p}_T = \frac{k}{n}\left(p_1 \widetilde{T}_{r(n-k,n)}\right)^{1/\hat{\eta}}, \tag{10}$$

  where $k = k(n)$ is well chosen sequence of integers and $\widetilde{T}_{r(n-k,n)}$ is the $(n-k)$-th order statistics of the sample $\widetilde{T}_{ri}$.

Note that the estimators $\hat{p}_e$, $\hat{p}_p$ and $\hat{p}_T$ are all based on the assumption that the observations are independent. In the following method, we study the short term dependence of the pseudo sample of $\widetilde{T}_r$ to take into account the serial dependence of $(X_i, Y_i), i = 1, \ldots, n$.

– Based on $\widetilde{T}_{ri}, i = 1, \ldots, n$, we compute the estimators of the GPD parameters, $(\hat{\gamma}, \hat{\sigma}, u)$ and the estimator of the extremal index $\hat{\theta}$ by methods described in Section 2.3. From equation (10.7) in Beirlant et al. (2004), the tail of T_r can be fitted with a GPD, the parameters of which can be estimated by

$$\left(\hat{\gamma}, \frac{\hat{\sigma}}{\hat{\theta}^{\hat{\gamma}}}, u + \frac{\hat{\sigma}(1-\hat{\theta}^{\hat{\gamma}})}{\hat{\gamma}\hat{\theta}^{\hat{\gamma}}}\right) =: (\hat{\gamma}, \tilde{\sigma}, \tilde{u}).$$

We define the fourth estimator of the failure probability, $p = P(T_r > 1/p_1)$ as below:

$$\hat{p}_r = \frac{1}{n}\sum_{i=1}^{n} I(\widetilde{T}_{ri} > \tilde{u}) \left(1 + \hat{\gamma}\frac{1/p_1 - \tilde{u}}{\tilde{\sigma}}\right)^{-1/\hat{\gamma}}. \tag{11}$$

To conclude this section, remark that the first two estimators $\hat{p}_e$ and $\hat{p}_p$ have been studied in de Haan and de Ronde (1998) and Draisma et al. (2004), whereas the two others $\hat{p}_T$ and $\hat{p}_r$ are new estimators (also based on the bivariate regular variation).

3.2. *Conditional Probability Approach*

Heffernan and Tawn (2004) proposes a conditional approach for multivariate extreme modeling. We briefly explain how to apply this approach to estimate bivariate failure probabilities. First we have

$$P(X > x_{t_X}, Y > y_{t_Y}) = P(X > x_{t_X})P(Y > y_{t_Y}|X > x_{t_X}) = p_1 P(Y > y_{t_Y}|X > x_{t_X}).$$

As aforementioned, p_1 is considered known in this part. In order to handle the conditional probability of $P(Y > y_{t_Y}|X > x_{t_Y})$, it is assumed that there are normalizing functions $a_{|X}$ and $b_{|X}$ such that

$$\lim_{x\to\infty} P\left(\frac{Y - a_{|X}(X)}{b_{|X}(X)} \le z \mid X = x\right) = G_{|X}(z),$$

where the limit distribution $G_{|X}$ is non-degenerate. Heffernan and Tawn (2004) gives detailed procedure on estimating the normalizing functions and the limit distribution. Now,

$$\begin{aligned} P(Y > y_{t_Y}|X > x_{t_X}) &= \int_{t_X}^{\infty} P\left(\frac{Y - a_{|X}(X)}{b_{|X}(X)} > \frac{y_{t_Y} - a_{|X}(X)}{b_{|X}(X)} \mid X = x\right) \frac{dF_X(X)}{1 - F_X(x_{t_X})} \\ &\approx \int_{t_X}^{\infty} G_{|X}\left(\frac{y_{t_Y} - a_{|X}(X)}{b_{|X}(X)}\right) \frac{dF_X(X)}{1 - F_X(x_{t_X})}. \end{aligned}$$

By sampling from the simulated distributions $\hat{G}_{|X}$ and $\hat{F}_X$, a Monte Carlo procedure is suggested to estimate the targeted probability by Heffernan and Tawn (2004). Denote $\hat{p}_c$ the estimator of p by this approach:

$$\hat{p}_c = p_1 \int_{t_X}^{\infty} \hat{G}_{|X}\left(\frac{y_{t_Y} - \hat{a}_{|X}(X)}{\hat{b}_{|X}(X)}\right) \frac{d\hat{F}_X(X)}{1 - \hat{F}_X(x_{t_X})}. \tag{12}$$

The inference procedure used in Heffernan and Tawn (2004) has been recently published in Southworth and Heffernan (2012).

3.3. Simulation Study

Throughout this section we use simulated data to illustrate the performance of the estimators proposed in the previous sections.

3.3.1. The models

We consider two bivariate time series models, both constructed from the following recursive formula

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \mathbf{A} \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_t^X \\ \varepsilon_t^Y \end{pmatrix},$$

where $\mathbf{A}$ is a 2x2 diagonal matrix with entries 0.4 and 0.3, and $(\varepsilon_t^X, \varepsilon_t^Y)$, $t = 1, \ldots, n$ are independent and identically distributed positive noises. This construction leads to a stationary process with short term dependence. Then the two cases considered here differ in terms of tail behavior.

(a) Student model: the noise $(\varepsilon^X, \varepsilon^Y)$ follows a bivariate positive Student distribution with 2 degrees of freedom and correlation matrix $\begin{pmatrix} 1 & .5 \\ .5 & 1 \end{pmatrix}$; this yields asymptotically dependent data, with $\eta = 1$.

(b) Normal model: the noise $(\varepsilon^X, \varepsilon^Y)$ follows a bivariate positive normal distribution with correlation coefficient 0.5; this yields asymptotically independent data, with $\eta = 0.75$.

From each model, we simulate 300 replicate data sets with sample size $n = 6000$. From each data set, we estimate the failure probability defined as

$$p = P(X_t > x_{t_X}, Y_t > y_{t_Y}),$$

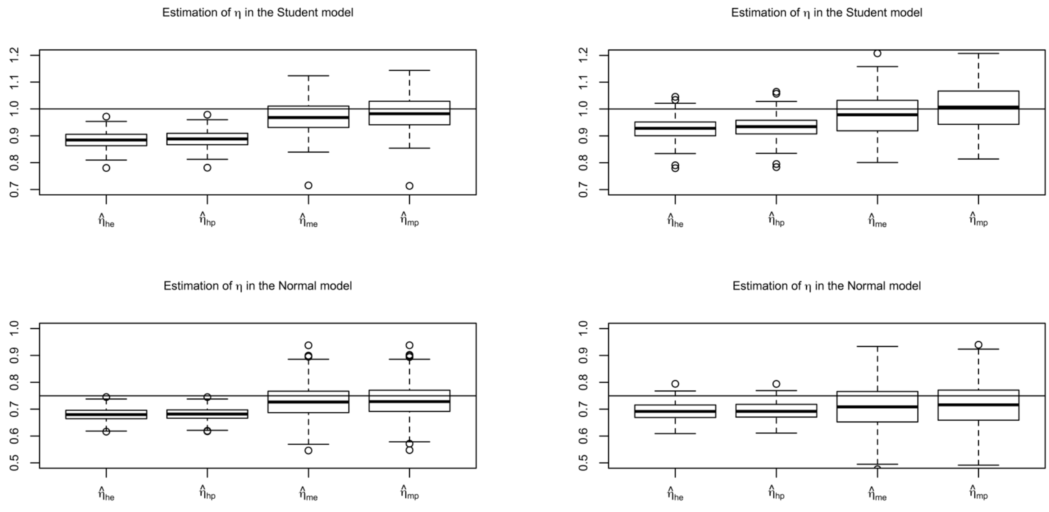
where x_{t_X} is the t_X-year return level of X_t and y_{t_Y} is the t_Y-year return level of Y_t. Intensive simulations lead to a good approximation of the probability p. In any of the simulated samples, the targeted event is extremely distant from the observations.

3.3.2. Comparison of η estimators

The estimation of η is part of the construction of the first three estimators of the failure probability of Section 3.1.2. In this subsection, we look at the performance of the four estimators of η introduced in Section 3.1.1. For this, we will consider two different thresholding probabilities .9 and .95, see Figure 15.

Two types of behaviors among the four estimators appear in the boxplots on Figure 15. The last two ($\hat{\eta}_{me}$ and $\hat{\eta}_{mp}$) behave similarly, showing a small bias, whereas the first two ($\hat{\eta}_{he}$ and $\hat{\eta}_{hp}$) show a comparable behavior with a bigger bias and a smaller variance. As a consequence, we keep $\hat{\eta}_{he}$ and $\hat{\eta}_{mp}$ as the representatives of the two groups. From this simulation study, one can conclude that $\hat{\eta}_{mp}$ outperforms its competitors in terms of bias. We will however keep both for further study, as η is only an input parameter used in the estimation of the failure probability.

Regarding the choice of thresholding probability, the results are not significantly different, so that we will keep the value .9 in the sequel, leading to a bigger sample size.



(a) η-thresholding probability .9 (b) η-thresholding probability .95

Figure 15: Estimates of η from the four methods defined in Section 3.1.1. First row is the Student model, and second row is the normal model. Each boxplot is based on 300 simulated data sets of size 6000. The horizontal lines indicate the true value of η.

3.3.3. Comparison of estimators of the failure probability

We implement the five estimators defined in Section 3.1.2: $\hat{p}_e$, $\hat{p}_p$, $\hat{p}_T$, and $\hat{p}_r$ from bivariate regular variation method and $\hat{p}_c$ from the conditional probability approach. The first three estimators will be used with two values of the input parameters η, respectively given by $\hat{\eta}_{mp}$ and $\hat{\eta}_{he}$. Notation $\hat{p}_e\_\hat{\eta}_{mp}$ and $\hat{p}_e\_\hat{\eta}_{he}$ will be used to differenciate the two $\hat{\eta}$ inputs, and analogously for $\hat{p}_p$ and $\hat{p}_T$. To assess the performance of the eight estimators, we calculate the relative error $1-\hat{p}/p$, where $\hat{p}$ denotes each estimator of the failure probability, see Figure 16. All the thresholds used in the estimation of η and the failure probability within the different methods are systematically fixed: η- and p-thresholding probabilities are .9.

Overall, the eight estimators perform decently given the challenging goal. From the first six boxplots in both Student and Normal cases, one can conclude that it is sufficient to consider the two estimators $\hat{p}_e\_\hat{\eta}_{mp}$ and $\hat{p}_e\_\hat{\eta}_{he}$, which are the best representatives among the first six estimators. The last two estimators also give satisfying results. According to this simulation study, we will keep the four estimators $\hat{p}_e\_\hat{\eta}_{mp}$, $\hat{p}_e\_\hat{\eta}_{he}$, $\hat{p}_r$ and $\hat{p}_c$ in the application on bivariate (X,Y) data.

3.4. Bivariate extreme analysis on real data

The final goal of this paper is to estimate the failure probability on the real data set of size approximately 6000, using the different methods proposed previously.

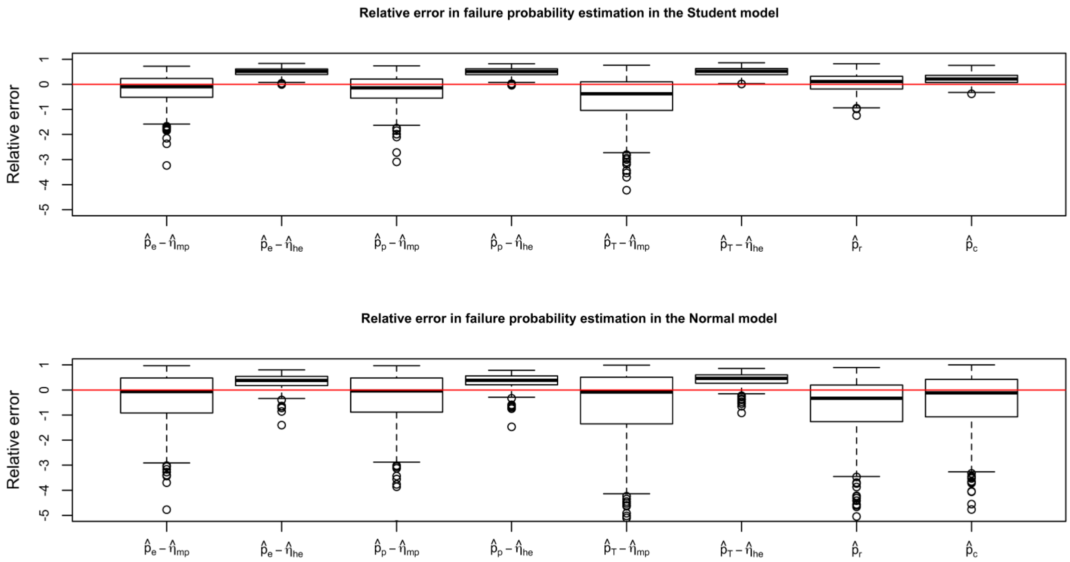


Figure 16: Relative error of the failure probability estimators. Each boxplot is based on 300 simulated data sets of size 6000.

3.4.1. Point Estimates

As a first step, we consider the estimation of η to study the extremal dependence between Variable X and Variable Y. From the upper plot in Figure 17, we can conclude that the pair (X,Y) has positive association, as both estimates strictly lie in the interval $(0.5,1)$. The lower plot of Figure 17 gives the estimates of the extremal index of the structure variable defined in (9). From the conclusion of Section 3.3.2, one would choose .9 as thresholding probability. Besides, standard visual criteria (stability zone) would lead to the choice of .95 from Figure 17. Thus, we will keep both choices to proceed.

As discussed in Section 3.1.2, the failure probability can be written as

$$P(1-F_X(X_t)<p_1, 1-F_Y(Y_t)<p_2)\ ,$$

where $p_1 = 1-(1-1/1000)^{1/(n_0\theta_1)}$ and $p_2 = 1-(1-1/100)^{1/(n_0\theta_2)}$. From Section 2.3, we have $\hat{p}_1 \approx 9.46\times 10^{-6}$ and $\hat{p}_2 \approx 1.06\times 10^{-4}$.

Now we calculate the following list of failure probability estimators:

$$\hat{p}_e\_\hat{\eta}_{he},\ \hat{p}_e\_\hat{\eta}_{mp},\ \hat{p}_r\_\hat{\theta} \text{ and } \hat{p}_c\ .$$

As motivated above, the input parameters η (in $\hat{p}_e\_\hat{\eta}_{he}$ and $\hat{p}_e\_\hat{\eta}_{mp}$) and θ (in $\hat{p}_r\_\hat{\theta}$) are estimated at two thresholding probabilities, .9 and .95. The final point estimates of the failure probability are shown in Figures 18 and 19 (on logarithmic scales). As a comparison, we keep the curve of $\hat{p}_c$ commonly in both figures.

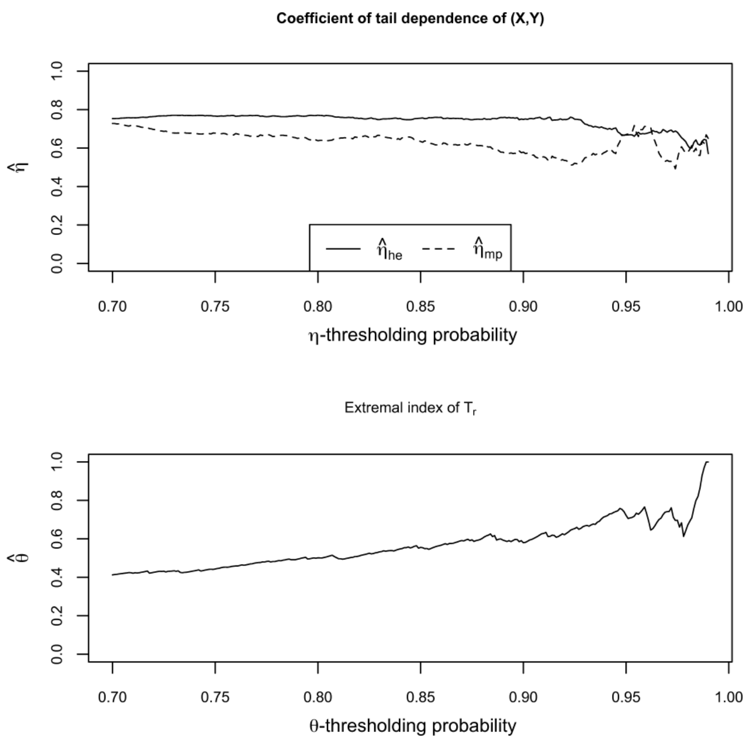


Figure 17: The upper plot gives the estimates of η by the two methods selected in Section 3.3.2. The lower plot gives the estimates of θ defined in Section 3.1.2.

Let us call *concordance zone* the thresholding probability area where most of the estimators give consistent answers. From Figure 18, the concordance zone of three estimators $\hat{p}_e\_\hat{\eta}_{he}$, $\hat{p}_r\_\hat{\theta}$ and $\hat{p}_c$ is around thresholding probability .9, which can also be identified as the stability zone of these three estimators. This yields 6.32×10^{-7} as the point estimate of the failure probability. The concordance zone for the four estimators in Figure 19 is around .95, which concludes 2.46×10^{-7} as the point estimate. There is no convincing criterium to make a choice between these two values. Indeed, both values are obtained with estimators based on completely different approaches (bivariate regular variation approach, see Section 3.1 and conditional approach, see Section 3.2). Note that $\hat{p}_e\_\hat{\eta}_{he}$ is the most stable among the four estimators and can be considered as the representative of concordance zones in both Figures 18 and 19. Some confidence intervals will be obtained for $\hat{p}_e\_\hat{\eta}_{he}$ in Section 3.4.2.

Unlike the great similarity of $\hat{p}_e\_\hat{\eta}_{he}$ and $\hat{p}_e\_\hat{\eta}_{mp}$ presented in Figure 19, there is a big gap between these two estimators in Figure 18. Recall that they only differ via η input, so that one can measure the sensitivity of the failure probability estimation in terms of η estimation from the upper plot of Figure 17. This points out the importance of getting stable and trustable estimators for the coefficient of tail dependence η.

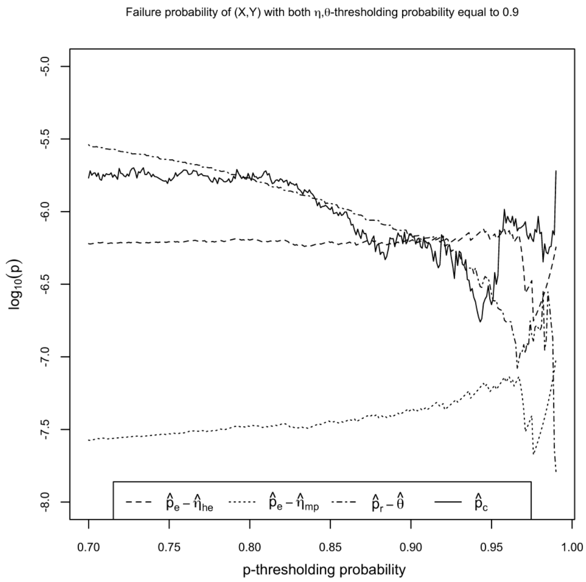


Figure 18: Estimation of the failure probability using the four estimators with η-thresholding probability .9 (in $\hat{p}_e\_\hat{\eta}_{he}$ and $\hat{p}_e\_\hat{\eta}_{mp}$) and with θ-thresholding probability .9 (in $\hat{p}_r\_\hat{\theta}$).

3.4.2. Confidence Intervals

Next we propose bootstrap procedures to construct confidence intervals of the failure probability using $\hat{p}_e\_\hat{\eta}_{he}$ at thresholding probabilities .9 and .95. Let $\hat{p}_e\_\hat{\eta}_{he}(0.9)$ denote the estimator for which thresholding probability .9 is used in both estimation of η and the failure probability. Define $\hat{p}_e\_\hat{\eta}_{he}(0.95)$ in the same way. The procedure is briefly introduced as following and the results are reported in Table 4.

– Extract the estimates of p_1, p_2 and the GPD fits of Variables X and Y from the univariate study.
– Apply the method in Politis and Romano (1994) to obtain the bootstrap samples. Based on each sample, compute the estimates of the failure probability.
– Construct confidence intervals with the two methods introduced in Section 2.3, refered as the Bootstrap normal method (BNM) and the Bootstrap percentile method (BPM).

TABLE 4. *Estimation of the failure probability using $\hat{p}_e\_\hat{\eta}_{he}$ at thresholding probabilities .9 and .95., with confidence intervals.*

	Point estimate	one-sided 95% BNM CI	95% BPM CI
$\hat{p}_e\_\hat{\eta}_{he}(0.9)$	6.32×10^{-7}	$[0, 1.21\times10^{-6}]$	$[2.45\times10^{-7}, 1.51\times10^{-6}]$
$\hat{p}_e\_\hat{\eta}_{he}(0.95)$	2.46×10^{-7}	$[0, 7.60\times10^{-7}]$	$[6.12\times10^{-8}, 1.13\times10^{-6}]$

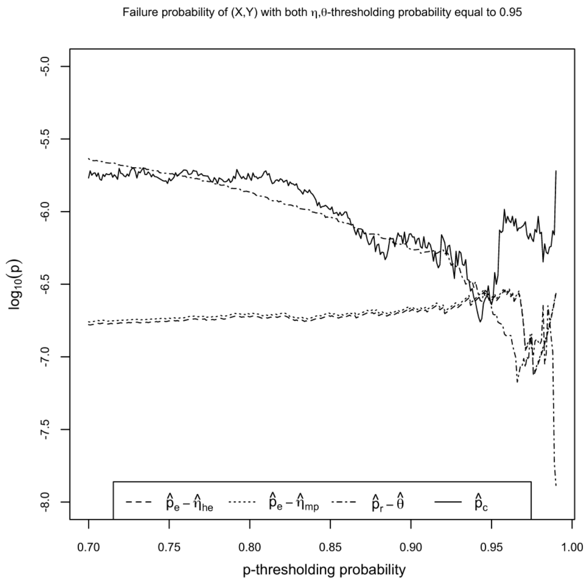


Figure 19: Estimation of the failure probability using the four estimators with η-thresholding probability .95 (in $\hat{p}_e\_\hat{\eta}_{he}$ and $\hat{p}_e\_\hat{\eta}_{mp}$) and with θ-thresholding probability .95 (in $\hat{p}_r\_\hat{\theta}$).

4. Conclusion

In the presence of short term dependence in the data, there is a tendency for the extremes to occur into clusters. Extreme Value Theory allows to take into account this short term dependence, via the estimation of the extremal index θ or via the identification of independent clusters. Adding these two methods to the classical GEV method, we have implemented and compared three estimation procedures in the univariate framework.

From Tables 1, 2 and 3, we can conclude that the extremal index method leads to the narrowest confidence intervals for the estimation of the return levels of Variables X and Y. These intervals are still rather wide, which inherits from the fact that the extrapolation asked (large t_X and t_Y values) is rather ambitious.

The results presented in both univariate and bivariate sections have been obtained under the hypothesis that the processes are stationary, and belong to a domain of attraction for the maximum. These hypotheses could be more deeply explored, and this could be an interesting issue for future research.

The main objective of the present study was to estimate the failure probability defined through return levels for Variables X and Y. We have implemented several methods that give rather coherent answers. Variables X and Y present an asymptotic independence with positive association.

One issue is the choice of the thresholding probability that produces fluctuations of the failure probability estimation. The selection of the different estimators and thresholds have been done with the help of an intensive simulation study. But this remains a delicate point.

Since the problem is difficult, it is crucial to associate confidence intervals to point estimates.

Table 4 summarizes the answer to the question designed in a specific context.

In terms of interpretation, one could say that the failure event would occur on average once among 8 700 years (based on $\hat{p}_e\_\hat{\eta}_{he}(0.9)$) or once among 22 300 years (based on $\hat{p}_e\_\hat{\eta}_{he}(0.95)$). These numbers could be compared to 5 500 000, which corresponds to the number of years obtained under assumption of total independence.

Acknowledgements

This research has been carried out under a contract partially funded by EdF R&D (MADONE). The authors wish to thank the EdF partners namely Pietro Bernadara (LNHE), Anne Dutfoy (MRI), Marie Gallois (MRI), Nicolas Malleron (LNHE), Sylvie Parey (MFEE) and Nicolas Roche (LNHE).

We would like to express our thanks to the Editor-in-Chief and the referees for their excellent suggestions for improving the paper.

References

Beirlant, J., Goegebeur, Y., Segers, J., and Teugels, J. (2004). *Statistics of Extremes: Theory and Applications*. Wiley, Chichester.

Brockwell, P. J. and Davis, R. A. (2002). *Introduction to Time Series and Forecasting*. Springer.

Bruun, J. T. and Tawn, J. A. (1998). Comparison of approaches for estimating the probability of coastal flooding. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 47:405 – 423.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London.

de Haan, L. and de Ronde, J. (1998). Sea and wind: Multivariate extremes at work. *Extremes*, 1:7–45.

de Haan, L. and Ferreira, A. (2006). *Extreme Value Theory: An Introduction*. Springer Verlag.

Draisma, G., Drees, H., Ferreira, A., and de Haan, L. (2004). Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli*, 10:251–280.

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

Ferro, C. A. T. and Segers, J. (2003). Inference for clusters of extreme values. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 65:545–556.

Genest, C., Quessy, J.-F., and Rémillard, B. (2007). Asymptotic local efficiency of Cramér-von Mises tests for multivariate independence. *Ann. Statist.*, 35(1):166–191.

Heffernan, J. E. and Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, 66:497–546.

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *Annals of Statistics*, 3:1163–1174.

Leadbetter, M. R. (1983). Extremes and local dependence in stationary sequences. *Zeitschrift fÃ¼r Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65:291–306.

Leadbetter, M. R., Lindgren, G., and Rootzén, H. (1983). *Extremes and related properties of random sequences and processes*. Springer Series in Statistics. Springer-Verlag, New York.

Ledford, A. and Tawn, J. (1996). Statistics for near independence in multivariate extreme values. *Biometrika*, 83:169–187.

Ljung, G. M. and Box, G. E. P. (1978). On a measure of a lack of fit in time series models. *Biometrika*, 65(2):297–303.

Politis, D. N. and Romano, J. P. (1994). The stationary bootstrap. *Journal of the American Statistical Association*, 89:1303–1313.

Resnick, S. I. (1987). *Extreme Values, Regular Variation, and Point Processes*. Springer, New York.

Resnick, S. I. (2007). *Heavy-Tail Phenomena: Probabilistic and Statistical Modeling*. Springer Series in Operations Research and Financial Engineering. Springer, New York.

Southworth, H. and Heffernan, J. E. (2012). *Texmex: Threshold exceedences and multivariate extremes*. R package version 1.3.