Discussiones Mathematicae
Probability and Statistics 33 (2013) 65–77
doi:10.7151/dmps.1149

ON THE TAIL INDEX ESTIMATION OF AN AUTOREGRESSIVE PARETO PROCESS

Marta Ferreira

Center of Mathematics of Minho University
Campus de Gualtar
4710 – 057 Braga, Portugal

e-mail: msferreira@math.uminho.pt

Abstract

In this paper we consider an autoregressive Pareto process which can be used as an alternative to heavy tailed MARMA. We focus on the tail behavior and prove that the tail empirical quantile function can be approximated by a Gaussian process. This result allows to derive a class of consistent and asymptotically normal estimators for the shape parameter. We will see through simulation that the usual estimation procedure based on an i.i.d. setting may fall short of the desired precision.

Keywords: extreme value theory, autoregressive processes, tail index estimation.

2010 Mathematics Subject Classification: Primary: 60G70; Secondary: 62G32.

1. Introduction

An extreme value analysis mainly aims the estimation of the probability of rarely observed events, specially more extreme than any previously observed. The central result in extreme value theory (EVT) states the possible probability laws that might occur in the limit of the linearly normalized maxima (minima). More precisely, for a sequence $\{X_n\}_{n\geq 1}$ of independent and identically distributed (i.i.d.) random variables (r.v.'s), having marginal cumulative distribution function (cdf) F, if there are real constants $a_n > 0$ and b_n such that,

$$P(\max(X_1, \ldots, X_n) \leq a_n x + b_n) \underset{n\to\infty}{\longrightarrow} G_\gamma(x)\,,$$

for some non-degenerate function G_γ, then this must be a generalized extreme value function,

$$G_\gamma(x) = \exp(-(1+\gamma x)^{-1/\gamma}), \quad 1+\gamma x > 0,\, \gamma \in \mathbb{R},$$

($G_0(x) = \exp(-e^{-x})$) and we say that F belongs to the max-domain of attraction of G_γ, in short, $F \in \mathcal{D}(G_\gamma)$. An analogous result is easily stated for minima by relation $\min(X_1,\dots,X_n) = -\max(-X_1,\dots,-X_n)$. The shape parameter γ also named *tail index* is of major importance as it determines the tail behavior of F. If $\gamma > 0$ then F has a heavy tail (Fréchet max-domain of attraction), $\gamma = 0$ means that F has an exponential tail (Gumbel max-domain of attraction) and $\gamma < 0$ indicates a short tail (Weibull max-domain of attraction). Here we will be interested in heavy tails. Considering the quantile function (qf), $F^{-1}(t) = \inf\{x : F(x) \geq t\}$, we have, $F \in \mathcal{D}(G_\gamma)$ for $\gamma > 0$, if and only if the tail qf satisfies

$$F^{-1}(1-tx) \sim x^{-\gamma} F^{-1}(1-t), \text{ as } t \to \infty,$$

which is also equivalent to a $-1/\gamma$-regularly varying tail at ∞, expressed as

$$1 - F(x) = x^{-1/\gamma} L_F(x), \tag{1}$$

where L_F is a slowly varying function at ∞ (i.e., $L(tx)/L(t) \sim 1$, as $t \to \infty$). The form (1) is also called a Pareto-type tail.

More recently, it has become evident that a number of phenomena requires the "heavy tail" concept and, rather frequently, extremal events (excesses over a threshold value) occur not one at time, but may cluster. Both situations are atypical for the Gaussian linear ARMA process. As an alternative, it has been considered the max-autoregressive moving average (MARMA) by replacing the summation operator by the maximum one, with marginals Fréchet and thus heavy-tailed (Davis and Resnick [3]). MARMA modeling have been widely applied to real data in several areas from environment, engineering and financial (see Lebedev [18] 2008 and references therein).

An alternative for MARMA are the not so well-known autoregressive Pareto processes, at least within an extreme values modeling context (Ferreira, [8] 2012). A Pareto process corresponds to a stochastic process whose marginal distributions are Pareto or generalized Pareto. These are heavy-tailed and satisfy (1). Generalizations of the Pareto's distribution, e.g. the Pareto(III)(μ,σ,γ) given by

$$1 - F(x) = \left[1 + \left(\frac{x-\mu}{\sigma}\right)^{1/\gamma}\right]^{-1},\ x > \mu, \tag{2}$$

have been proposed as models for economic variables (Arnold [1]). Here we consider the first order autoregressive Pareto(III) denoted Yeh-Arnold-Robertson Pareto(III) (Yeh *et al.* [25] 1988), which is defined as follows.

Considering an innovations sequence $\{\varepsilon_n\}_{n\geq 1}$ of i.i.d. r.v.'s Pareto(III)(0,σ,γ) in (2), with $\sigma,\gamma > 0$, and sequence $\{U_n\}_{n\geq 1}$ of i.i.d. r.v.'s Bernoulli(p) (independent of the ε's), we say that the process $\{X_n\}_{n\geq 1}$ is a first order Yeh-Arnold-Robertson Pareto(III), in short YARP(III)(1), if

$$X_n = \min\Big(p^{-\gamma}X_{n-1}, \frac{1}{1-U_n}\varepsilon_n\Big), \tag{3}$$

where 1/0 is interpreted as $+\infty$. By conditioning on U_n, it can be verified that the YARP(III)(1) process has a Pareto(III)(0,σ,γ) stationary distribution (X_0 has also distribution Pareto(III)(0,σ,γ)). The YARP(III)(1) have interesting extreme value properties. For instance, considering

$$T_n = \min_{0\leq i\leq n} X_i$$

and

$$M_n = \max_{0\leq i\leq n} X_i,$$

we have that $T_n \stackrel{d}{=} \min_{i\leq N}\varepsilon_i$, where ε_i, $i\geq 1$, are i.i.d. Pareto(III)$(0,\sigma,\gamma)$ and N, independent of ε_i is such that $N-1 \frown Binomial(n, 1-p)$. Moreover, we have also $n(1-p)^\gamma T_n/\sigma \stackrel{d}{\to} Weibull(1/\gamma)$ and $\frac{n^{-\gamma}}{\sigma}M_n \stackrel{d}{\to} Fréchet(0,(1-p)^{-1},1/\gamma)$. We also point out that these processes are closed for geometric minima and maxima, i.e., $T_N = \min_{1\leq i\leq N} X_i$ and $M_N = \max_{1\leq i\leq N} X_i$ where $N \frown Geometric(p)$, have also distributions, respectively, Pareto(III)(0,$\sigma p^\gamma,\gamma$) and Pareto(III)$(0,\sigma p^{-\gamma},\gamma)$. For more details, see Arnold, [2] and Krishnarani and Jayakumari, [17]. A characterization of the YARP(III)(1) process's tail behavior can be seen in Ferreira ([8] 2012) and Ferreira ([9] 2013). In particular, it was shown that it presents a β-mixing dependence structure (to be defined below). Some tail coefficients interesting for applications were also derived, e.g. the tail dependence coefficient of Sibuya ([23] 1960) and Joe ([16] 1993), tail dependence coefficients for order statistics of Ferreira and Ferreira ([10] 2012) and fragility indexes of a system introduced in Geluk *et al.*, ([12] 2007) and Ferreira and Ferreira ([11] 2012b).

Ferreira ([8] 2012a) presented a simple consistent and asymptotically normal estimator for parameter p, given by

$$\widehat{p}_n = \frac{2}{n-1}\sum_{j=2}^{n} \mathbb{1}_{\{X_{j-1}<X_j\}} - 1, \tag{4}$$

another attractive characteristic for modeling purposes.

In this work we will be concerned about the estimation of the other parameter γ. Our approach is within the EVT where this parameter is denoted tail index. There already exists estimators for this latter, e.g. Hill ([13] 1975), Pickands' ([19]

1975), maximum likelihood (Smith [24] 1987), moments (Dekkers *et al.* [4] 1987), generalized weighted moments (Hosking e Wallis [14]), among others, whose performances and properties have been analyzed in an i.i.d. context. The first studies for stationary sequences date back the nineties and concern the Hill estimator (Rootzén *et al.* [22] 1990, Hsing [15] 1991, Resnick e Stărică [20, 21] 1995, 1998). Based on results for the convergence of the tail of empirical processes, Drees ([5, 6], 1998a,b) developed a class of estimators for the tail index that includes the ones mentioned above, and proved consistency and asymptotic normality within an i.i.d. context. This approach was latter extended to a wide range of stationary sequences (Drees, [7], 2003). Examples of time series models satisfying Drees' conditions include the well-known linear ARMA and the nonlinear ARCH and MARMA. Here we shall prove that the YARP(III)(1) model satisfies the Drees' conditions too. Therefore, we will be able to state a weighted approximation for the YARP(III)(1) tail empirical process and derive consistency and asymptotic normality of several estimators of the tail parameter γ. A simulation study will illustrate that the estimates greatly improve whenever YARP(III)(1) serial dependence is taken into account.

2. YARP(III)(1): the tail index and the Drees class of estimators

In this section we address the estimation of the shape parameter γ of a YARP(III)(1) process within an EVT context, since it corresponds to the so-called tail index of the marginal distribution. This index is a very important issue in EVT and its estimation gets the attention of many researchers. Usually, the tail index estimators depend only on the k_n+1, $n \in \mathbb{N}$, largest order statistics, $X_{n:n} \geq X_{n-1:n} \geq \cdots \geq X_{n-k_n:n}$, where $\{k_n\}_{n\geq 1}$ is an intermediate sequence, i.e., a sequence of positive integers satisfying $k_n \to \infty$ and $k_n = o(n)$, as $n \to \infty$. The most used estimator within a heavy tail context is the Hill estimator

$$\hat{\gamma}_n^H = \frac{1}{k_n} \sum_{i=1}^{k_n} \log \frac{X_{n-i+1:n}}{X_{n-k_n:n}}. \tag{5}$$

In Figure 1, the solid line corresponds to a sample path of Hill estimates plotted against $k \equiv k_n$ largest order statistics. The horizontal line represents the true value. Observe the effect of an increasing bias as k increases too, whilst small values of k lead to larger variance. This is a main drawback of this type of estimators and the choose of a number k that allows the best compromise between bias and variance is of major importance.

Estimators in literature have been firstly analyzed in i.i.d. settings. For a survey, see, for instance, Drees ([5, 6], 1998a,b). However, since for applications this is often an unrealistic assumption, efforts have been made to extend the theory to stationary sequences (Rootzén *et al.* [22] 1990, Hsing [15] 1991, Resnick e Stărică [20, 21] 1995, 1998, Drees [7] 2003).

We focus on the results in Drees ([7] 2003) which are more general and can be applied under not too restricted serial dependence conditions.

In a very brief, by considering F_n the empirical cdf, $[x]$ the largest integer not exceeding x and $\{k_n\}_{n\geq 1}$ an intermediate sequence, Drees ([7] 2003) established an approximation of the weighted tail empirical qf,

$$Q_n(t) := F_n^{-1}\Big(1 - k_n t\Big) = X_{n-[k_n t]:n},\ t \in [0,1], \tag{6}$$

to a Gaussian process, for any stationary sequence $\{X_n\}_{n\geq 1}$ satisfying the conditions listed below:

- a β-*mixing* dependence structure:

$$\beta(l) := \sup_{p\in\mathbb{N}} E\big(\sup_{B\in\mathcal{F}(X_{p+l+1},\dots)} |P(B|\mathcal{F}(X_1,\dots,X_p)) - P(B)|\big) \underset{l\to\infty}{\longrightarrow} 0,$$

with $\mathcal{F}(.)$ denoting the $\sigma-$field generated by the indicated random variables. More precisely, it is assumed that there exists a sequence l_n, $n \in \mathbb{N}$, such that

$$\lim_{n\to\infty} \frac{\beta(l_n)}{l_n} n + l_n k_n^{-1/2} \log^2 k_n = 0. \tag{7}$$

- a regularity condition for the joint tail of (X_1, X_{1+m}):

$$\lim_{n\to\infty} \frac{n}{k_n} P\Big(X_1 > F^{-1}\big(1 - \frac{k_n}{n}x\big), X_{1+m} > F^{-1}\big(1 - \frac{k_n}{n}y\big)\Big) = c_m(x,y), \tag{8}$$

for all $m \in \mathbb{N}$ and $0 < x, y \leq 1 + \epsilon$.

- a uniform bound on the probability that both X_1 and X_{1+m} belong to an extremal interval:

$$\frac{n}{k_n} P\big(X_1 \in I_n(x,y), X_{1+m} \in I_n(x,y)\big) \leq (y-x)\Big(\tilde{\rho}(m) + D_1 \frac{k_n}{n}\Big), \tag{9}$$

for some constant $D_1 \geq 0$, a sequence $\tilde{\rho}(m)$, $m \in \mathbb{N}$, satisfying $\sum_{m=1}^{\infty} \tilde{\rho}(m) < \infty$ for all $m \in \mathbb{N}$, $0 < x, y \leq 1 + \epsilon$ and the extremal interval, $I_n(x,y) =]F^{-1}(1 - yk_n/n\big), F^{-1}(1 - xk_n/n)]$.

- for the sake of simplicity,

$$F^{-1}(1-t) = dt^{-\gamma}(1+r(t)) \text{ , with } |r(t)| < \Phi(t)\,. \tag{10}$$

for some constant $d > 0$ and a τ-varying function Φ at 0 for some $\tau > 0$, or $\tau = 0$ and Φ nondecreasing with $\lim_{t\downarrow 0} \Phi(t) = 0$.

- the intermediate sequence $\{k_n\}_{n\geq 1}$ is such that

$$\lim_{n\to\infty} k_n^{1/2}\Phi(k_n/n) = 0. \tag{11}$$

Theorem 1 Drees (2003)**.** *Under the conditions* (7)–(11) *with* $l_n = o(n/k_n)$, *there exist versions of the tail empirical quantile function,* $Q_n(t)$, *defined in* (6), *and a centered Gaussian process* e *with covariance function* c *given by*

$$c(x,y) := x \wedge y + \sum_{m=1}^{\infty}(c_m(x,y) + c_m(y,x)) \in \mathbb{R}\,, \tag{12}$$

such that

$$\sup_{t\in(0,1]} t^{\gamma+1/2}(1+|\log t|)^{-1/2}\left|k_n^{1/2}\Big(\frac{Q_n(t)}{F^{-1}(1-k_n/n)} - t^{-\gamma}\Big) - \gamma t^{-(\gamma+1)}e(t)\right| \to 0 \tag{13}$$

in probability.

In Drees ([5, 6], 1998a,b) it is observed that almost every estimator $\hat{\gamma}_n$ of the tail index parameter γ that are based only on the $k_n + 1$ largest order statistics can be represented as a smooth functional T (verifying some regularity conditions) applied to the tail empirical qf, i.e., $\widehat{\gamma}_n = T(Q_n)$. Some examples within this class include the above mentioned Hill, maximum likelihood, moments, Pickands' and probability weighted moments estimators. We shall denote estimators of the form $\widehat{\gamma}_n = T(Q_n)$ as the *Drees' class of tail index estimators.* Theorem 2.2 in Drees ([7], 2003) establishes the asymptotic normality of this class. More precisely,

$$k_n^{1/2}(\hat{\gamma}_n - \gamma) \underset{n\to\infty}{\longrightarrow} \mathcal{N}(0, \sigma^2_{T,\gamma})$$

weakly with

$$\sigma^2_{T,\gamma} = \gamma^2 \int_{(0,1]}\int_{(0,1]} (st)^{-(\gamma+1)} c(s,t)\nu_{T,\gamma}(ds)\nu_{T,\gamma}(dt), \tag{14}$$

where c is the function defined in (12) and $\nu_{T,\gamma}$ is a suitable signed measure. It can be proved that, in a generalized Pareto model, the Hill estimator in (5) has signed measure given by

$$\nu_{H,\gamma}(dt) = t^{\gamma} dt - \delta_1(dt),$$

where δ_1 is the Dirac measure with mass 1 at 1 (Drees [7], 2003).

Here we will prove that the YARP(III)(1) model satisfies the Drees' conditions and thus the application of the mentioned results is straightforward.

Proposition 2. *Let* $\{X_n\}_{n\geq 1}$ *be a* YARP(III)(1) *process. Then conditions* (8) *and* (9) *both hold.*

Proof. Let a_s be the qf of the YARP(III)(1) process at $1-s$, i.e., $a_s = F^{-1}(1-s)$ with

$$F^{-1}(t) = \sigma((1-t)^{-1} - 1)^{\gamma}. \tag{15}$$

We start by stating the m-step transition probability function (tpf) of the YARP(III)(1) process, given by (Ferreira, [8] 2012)

$$\begin{aligned} Q^m(x,]0, y]) &= P\big(X_{n+m} \leq y | X_n = x\big) \\ &= \begin{cases} 1 - \prod_{j=0}^{m-1} [\overline{F}_{\varepsilon}(p^{j\gamma} y)(1-p) + p], & x > y p^{m\gamma} \\ 1, & x \leq y p^{m\gamma}, \end{cases} \end{aligned}$$

which will be a fundamental tool in the following. Observe that

$$\begin{aligned} P(X_1 > a_{tx}, X_{1+m} > a_{ty}) &= \int_{a_{tx}}^{\infty} (1 - Q^m(]0, a_{ty}]))F(du) \\ &= tx - \int_{a_{tx}}^{\infty} Q^m(]0, a_{ty}])F(du). \end{aligned} \tag{16}$$

If $a_{tx} > a_{ty} p^m$, we have

$$\begin{aligned} \int_{a_{tx}}^{\infty} Q^m(]0, a_{ty}])F(du) &= \left(1 - \prod_{j=0}^{m-1} \left[\overline{F}_{\varepsilon}(p^{j\gamma} y)(1-p) + p\right]\right) \int_{a_{tx}}^{\infty} F(du) \\ &= \left(1 - \prod_{j=0}^{m-1} [\overline{F}_{\varepsilon}(p^{j\gamma} y)(1-p) + p]\right)(1 - F(a_{tx}) \end{aligned}$$

and, whenever $a_{tx} \leq a_{ty} p^m$, we have

$$\int_{a_{tx}}^{\infty} Q^m(]0,a_{ty}])F(du) = \int_{a_{tx}}^{a_{ty}p^{m\gamma}} Q^m(]0,a_{ty}])F(du) + \int_{a_{ty}p^{m\gamma}}^{\infty} Q^m(]0,a_{ty}])F(du)$$
$$= F(a_{ty}p^{m\gamma}) - F(a_{tx}) - \left(1 - \prod_{j=0}^{m-1} \left[\overline{F}_\varepsilon(p^{j\gamma}y)(1-p)+p\right]\right)(1 - F(a_{ty}p^{m\gamma})).$$

By (15), we obtain

$$F(a_{ty}p^{m\gamma}) = \frac{((ty)^{-1}-1)p^m}{1+((ty)^{-1}-1)p^m} = \frac{(1-ty)p^m}{ty+(1-ty)p^m}, \tag{17}$$

and since the innovations $\{\varepsilon_n\}_{n\geq 1}$ are i.i.d. also with df Pareto(III) $(0,\sigma,\gamma)$, then $\prod_{j=0}^{m-1}[\overline{F}_\varepsilon(p^{j\gamma}y)(1-p)+p] = ty + p^m(1-ty)$. Hence,

$$\int_{a_{tx}}^{\infty} Q^m(]0,a_{ty}])F(du) = \begin{cases} (1-ty-p^m(1-ty))tx & , \ a_{tx} > a_{ty}p^m \\ F(a_{ty}p^{m\gamma}) - F(a_{tx}) & \\ +(1-(ty-p^m(1-ty))(1-F(a_{ty}p^{m\gamma})) & , \ a_{tx} \leq a_{ty}p^m. \end{cases} \tag{18}$$

Replacing (18) in (16) and applying (17), after some simple calculations, we derive

$$P(X_1 > a_{tx}, X_{1+m} > a_{ty}) = \begin{cases} (ty+p^m(1-ty))tx & , \ a_{tx} > a_{ty}p^m \\ ty & , \ a_{tx} \leq a_{ty}p^m. \end{cases} \tag{19}$$

The limiting functions $c_m(x,y)$ in (8) are obtained by replacing t by $t_n \equiv k_n/n$ in (19):

$$c_m(x,y) = \lim_{n\to\infty} \frac{n}{k_n} P(X_1 > a_{t_nx}, X_{1+m} > a_{t_ny}) = \begin{cases} p^m x & , \ x < yp^{-m} \\ y & , \ x \geq yp^{-m}. \end{cases}$$

In what concerns condition (9), considering $I_n(x,y) = \left]a_{t_ny}, a_{t_nx}\right]$, observe that

$$\tfrac{n}{k_n} P\Big(X_1 \in I_n(x,y), X_{1+m} \in I_n(x,y)\Big) \leq \tfrac{n}{k_n} P\Big(X_1 \in I_n(x,y), X_{1+m} > a_{t_ny}\Big)$$
$$\leq \ \tfrac{n}{k_n} P\Big(X_1 \in I_n(x,y), \varepsilon_{1+m} > a_{t_ny}\Big) \leq (y-x)\tfrac{k_n}{n} y,$$

and now take $D_1 = 1+\epsilon$ (we have $y \in (0, 1+\epsilon]$) and $\tilde{\rho}(m) = 0$. ∎

In the following it is conventioned that, for $k < 1$, $\sum_{i=1}^{k} = 0$.

Proposition 3. *Let $\{X_n\}_{n\geq 1}$ be a* YARP(III)(1) *process and $\{k_n\}$ an intermediate sequence. Then the limit in* (13) *holds for the tail empirical quantile function, $Q_n(t)$, defined in* (6), *with covariance function c given by*

$$c(x,y) = \min(x,y) + \sum_{m=1}^{r}(c_m(x,y) + c_m(y,x)) + (x+y)\frac{p^{r+1}}{1-p},$$

where $r \equiv r_{x,y} = [\max\{\ln(x/y)/\ln p, \ln(y/x)/\ln p\}]$.

In the YARP(III)(1) *process, for any tail index estimator within Drees' class, $\hat{\gamma}_n = T(Q_n)$, we have that, $k_n^{1/2}(\widehat{\gamma}_n - \gamma) \to \mathcal{N}(0, \sigma^2_{T,\gamma})$ weakly with variance given in* (14).

Proof. The result in (13) and the asymptotic normality are straightforward from Proposition 2 and the β-mixing dependence structure of the YARP(III)(1) process (Ferreira [8] 2012, Proposition 2.1). Now just observe that

$$c_m(x,y) + c_m(y,x) = \begin{cases} x(1+p^m) & , \ 0 < x \leq yp^m \\ (y+x)p^m & , \ yp^m < x \leq yp^{-m} \\ y(1+c^m) & , \ yc^{-m} < x \leq 1+\epsilon, \end{cases}$$

and since $p^m \to 0$ and $p^{-m} \to \infty$, as $m \to \infty$, for fixed x and y, there exists some order $r \in \mathbb{N}$, such that, for all $m \geq r$, we have $yp^m < x \leq yp^{-m}$. Therefore,

$$\sum_{m=1}^{\infty}(c_m(x,y) + c_m(y,x)) = \sum_{m=1}^{r}(c_m(x,y) + c_m(y,x)) + \sum_{m=r+1}^{\infty}(x+y)p^m.$$

■

Observe that for the YARP(III)(1) model the asymptotic variance of the Hill estimator is given by $\gamma^2(1 + 2p(1-p)^{-1})/k_n$. We remark that the asymptotic variance within an i.i.d. context is γ^2/k_n. Thus the confidence intervals based on the i.i.d. assumption, i.e.,

$$\widehat{\gamma}_n \pm z_{1-\alpha/2}\frac{\widehat{\gamma}_n}{\sqrt{k_n}},$$

with $z_{1-\alpha/2}$ denoting the $(1-\alpha/2)$-quantile of the standard normal distribution, are shorter than the confidence intervals based on the YARP(III)(1) serial dependence, i.e.,

$$\widehat{\gamma}_n \pm z_{1-\alpha/2}\frac{\widehat{\gamma}_n}{\sqrt{k_n}}\sqrt{1 + \frac{2\widehat{p}_n}{1-\widehat{p}_n}}, \tag{20}$$

with $\widehat{p}_n$ given in (4) (see Figure 1), which may be an indication that perhaps the pretended estimation accuracy of the former is not achieved. In what follows, we analyze this issue through simulation.

We generate 1000 independent random samples of size $n = 2000$ from an unit scale YARP(III)(1) process by considering the cases $\gamma = 0.5, 1, 1.5$ and $p = 0.25, 0.5, 0.75$. The estimated empirical non-coverage probabilities of independent and serial dependent 95% confidence intervals are plotted in Figure 2 against k upper order statistics (solid line and dashed line, respectively). The horizontal line represents the nominal probability 5%. The columns correspond to $\gamma = 0.5, 1, 1.5$, respectively, and the lines correspond to $p = 0.25, 0.5, 0.75$, respectively. Observe that, in an i.i.d. setting, the non-coverage probabilities are always above the nominal level. Note also that the larger the p the stronger the serial dependence which explains the very high non-coverage probabilities obtained with $p = 0.75$. Only in the case of small values of p we may find more acceptable levels of non-coverage probabilities (see the first line plots of Figure 2 corresponding to $p = 0.25$, where the minimum non-coverage probability is about 9%). A sample fraction of $100 \lesssim k \lesssim 200$ leads to an overall good performance of confidence intervals (20) with coverage probabilities of at least 95%.

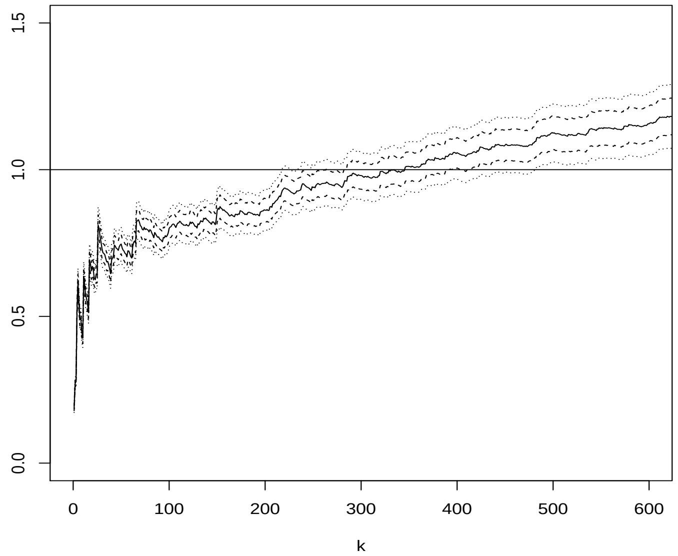

Figure 1. Sample path of the Hill estimator (solid line) applied to a YARP(III)(1) process, with $\gamma = 1$ and $p = 0.5$, and 95% confidence intervals based on the i.i.d. assumption (dashed line) and taking into account the serial dependence (dotted line).

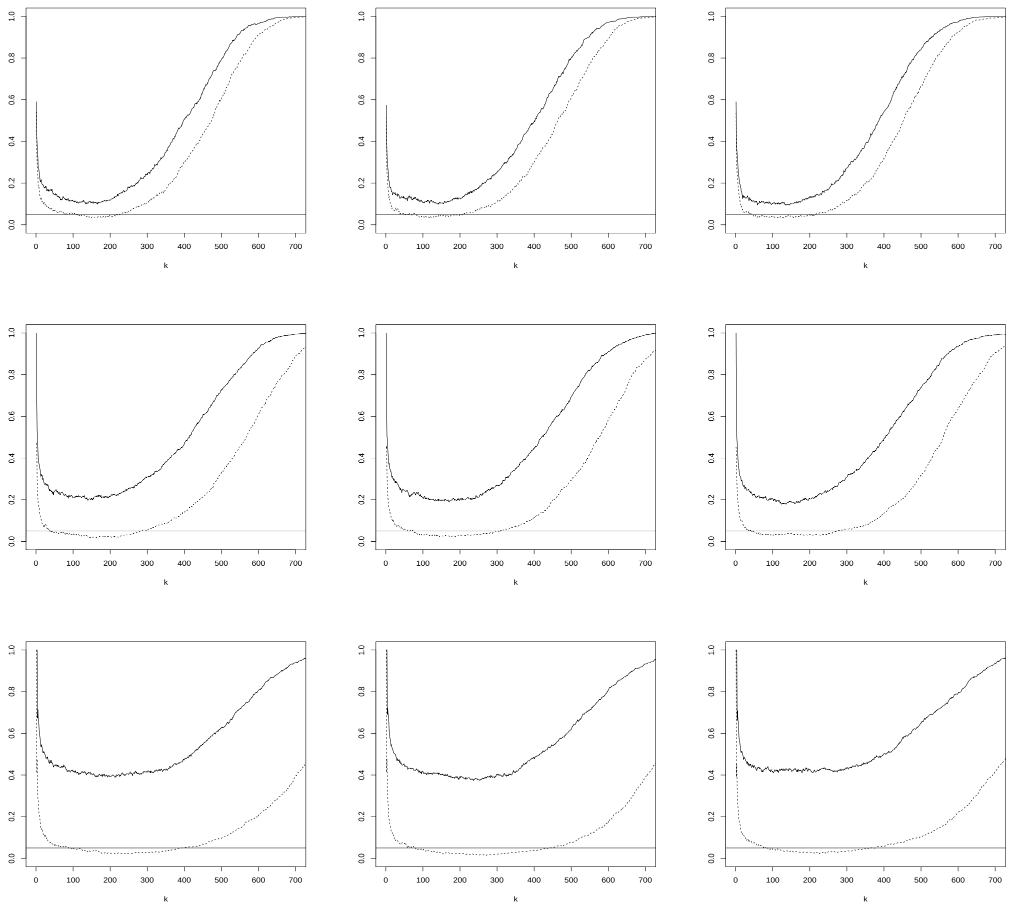

Figure 2. Empirical non-coverage probabilities of 95% confidence intervals based on the i.i.d. assumption (solid line) and taking into account the serial dependence (dashed line). The horizontal line corresponds to the probability 5%; left-to-right corresponds to $\gamma = 0.5, 1, 1.5$, respectively; top-to-bottom corresponds to $p = 0.25, 0.5, 0.75$, respectively.

Acknowledgements

The author was partially supported by the Research Centre of Mathematics of the University of Minho through the FCT Pluriannual Funding Program and by PTDC/FEDER grants from Portugal.

REFERENCES

[1] B.C. Arnold, Pareto Distributions (International Cooperative Publishing House, Fairland, 1983).

[2] B.C. Arnold, Pareto processes, in: Handbook of Statistics, D.N. Shanbhag and C.R. Rao, eds., Vol. 19, (Elsevier Science B.V., 2001).

[3] R. Davis and S. Resnick, *Basic properties and prediction of max-ARMA processes*, Adv. Appl. Probab. **21** (1989) 781–803.

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

[5] H. Drees, *On smooth statistical tail functionals*, Scand. J. Statist. **25** (1998a) 187–210.

[6] H. Drees, *A general class of estimators of the extreme value index*, J. Statist. Plann. Inference **66** (1998b) 95–112.

[7] H. Drees, *Extreme quantile estimation for dependent data with applications to finance*, Bernoulli **9** (2003) 617–657.

[8] M. Ferreira, *On the extremal behavior of a pareto process: an alternative for armax modeling*, Kybernetika **48** (2012) 31–49.

[9] M. Ferreira, Tail dependence of a Pareto process, accepted for publication in Studies in Theoretical and Applied Statistics – Selected Papers of the Statistical Societies, Springer.

[10] H. Ferreira and M. Ferreira, *Tail dependence between order statistics*, J. Multivariate Anal. **105** (2012) 176–192.

[11] H. Ferreira and M. Ferreira, *Fragility Index of block tailed vectors*, J. Statist. Plann. Inference **142** (2012) 1837–1848.

[12] J.L. Geluk, L. De Haan and C.G. De Vries, Weak and strong financial fragility. Tinbergen Institute Discussion Paper, TI 2007-023/2, 2007.

[13] B.M. Hill, *A simple general approach to inference about the tail of a distribution*, Ann. Statist. **3** (1975) 1163–1174.

[14] J.R.M. Hosking and J.R. Wallis, *Parameter and quantile estimation for the generalized Pareto distribution*, Technometrics **29** (1987) 339–349.

[15] T. Hsing, *On tail estimation using dependent data*, Ann. Statist. **19** (1991) 1574–1569.

[16] H. Joe, Multivariate Models and Dependence Concepts (Chapman & Hall, London, 1997).

[17] S.D. Krishnarani and K. Jayakumar, *A class of autoregressive processes*, Statist. Probab. Lett. **78** (2008) 1355–1361.

[18] A.V. Lebedev, *Statistical analysis of first-order MARMA processes*, Mathematical Notes **83** (2008) 506–511.

[19] J. Pickands III, *Statistical inference using extreme order statistics*, Ann. Statist. **3** (1975) 119–131.

[20] S. Resnick and C. Stărică, *Consistency of Hill's estimator for dependent data*, J. Appl. Probab. **32** (1995) 139–167.

[21] S. Resnick and C. Stărică, *Tail index estimation for dependent data*, Ann. Appl. Probab. **8** (1998) 1156–1183.

[22] H. Rootzén, M.R. Leadbetter and L. de Haan, Tail and Quantile Estimation for Strongly Mixing Stationary Sequences. Technical Report, UNC Center for Stochastic Processes, 1990.

[23] M. Sibuya, *Bivariate extreme statistics*, Ann. Inst. Statist. Math. **11** (1960) 195–210.

[24] R.L. Smith, *Estimating tails of probability distributions*, Ann. Statist. **15** (1987) 1174–1207.

[25] H.C. Yeh, B.C. Arnold and C.A. Robertson, *Pareto processes*, J. Appl. Probab. **25** (1988) 291–301.

Received 9 March 2013
Revised 16 September 2013