

## Article

# Analyzing the Gaver—Lewis Pareto Process under an Extremal Perspective

Marta Ferreira <sup>1,2,3,\*</sup> and Helena Ferreira <sup>4</sup><sup>1</sup> Centro de Matemática da Universidade do Minho, Campus de Gualtar 4710-057 Braga, Portugal<sup>2</sup> Centro de Matemática Computacional e Estocástica, Departamento de Matemática-Instituto Superior Técnico Av. Rovisco Pais 1, 1049-001 Lisboa, Portugal<sup>3</sup> Centro de Estatística e Aplicações, Faculdade de Ciências, Universidade de Lisboa, 1749-016 Lisboa, Portugal<sup>4</sup> Universidade da Beira Interior, Centro de Matemática e Aplicações (CMA-UBI), Avenida Marquês d'Ávila e Bolama, Covilhã 6200-001, Portugal; [helena.ferreira@ubi.pt](mailto:helena.ferreira@ubi.pt)\* Correspondence: [msferreira@math.uminho.pt](mailto:msferreira@math.uminho.pt)

Academic Editor: Mogens Steffensen

Received: 10 April 2017; Accepted: 16 June 2017; Published: 27 June 2017

**Abstract:** Pareto processes are suitable to model stationary heavy-tailed data. Here, we consider the auto-regressive Gaver–Lewis Pareto Process and address a study of the tail behavior. We characterize its local and long-range dependence. We will see that consecutive observations are asymptotically tail independent, a feature that is often misevaluated by the most common extremal models and with strong relevance to the tail inference. This also reveals clustering at “penultimate” levels. Linear correlation may not exist in a heavy-tailed context and an alternative diagnostic tool will be presented. The derived properties relate to the auto-regressive parameter of the process and will provide estimators. A comparison of the proposals is conducted through simulation and an application to a real dataset illustrates the procedure.

**Keywords:** extreme value theory; autoregressive processes; extremal index; asymptotic tail independence

**MSC:** 60G70; 60G10

## 1. Introduction

Increased exposure to catastrophic losses and the complexity of financial instruments require sophisticated risk assessment tools in areas such as (re) insurance, banking, finance, among others. Extreme value theory plays an important methodological role in risk management by providing appropriate instruments to deal with values as large as or even higher than those ever observed. These techniques include heavy-tailed models and measures to evaluate tail dependence, namely to infer to what extent the occurrence of a risk value in some variable influences an analogous occurrence in another variable.

Linear ARMA (autoregressive moving average) with heavy-tailed noise may be suitable to model time series presenting peaks of observations. However, in place of a summation, a maximum operator is more propitious to derive extremal properties. Max-autoregressive and moving maximum models were developed within this spirit, such as MARMA (max-autoregressive moving average) introduced in [Davis and Resnick \(1989\)](#) and M4 (multivariate maxima of moving maxima) processes presented in [Smith and Weissman \(1996\)](#). The Pareto model, which is closed under geometric multiplication and minimization, also motivated the first-order Pareto processes presented in [Arnold \(2001\)](#). Further analysis may be found in [Ferreira \(2016\)](#) and references therein.

A random variable (rv) is modeled by Pareto( $\sigma, \alpha$ ) if it has distribution function (df)

$$F_X(x) = 1 - (x/\sigma)^{-\alpha}, \quad x > \sigma, \quad \sigma > 0, \quad \alpha > 0. \quad (1)$$

This model is a particular case of Pareto-type tail models, the class of regular varying tail distributions, that is,

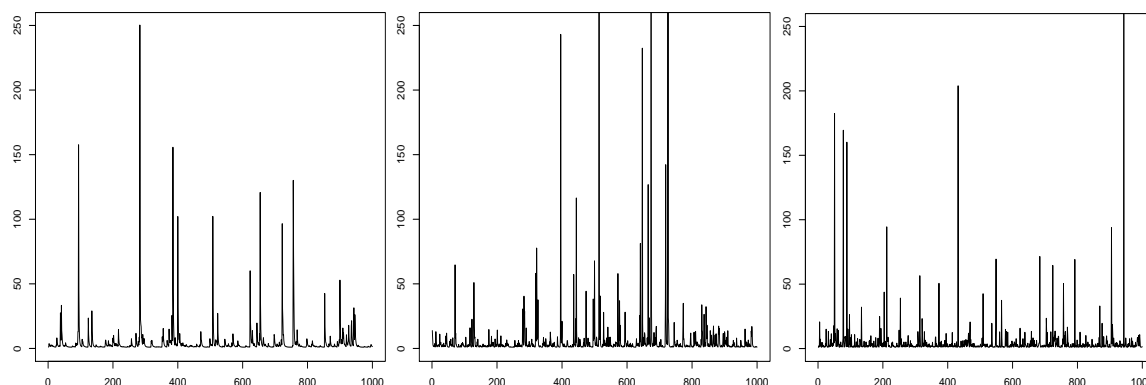
$$1 - F(x) := \bar{F}(x) = x^{-\alpha} L(x), \quad (2)$$

where  $L(x)$  is a slowly varying function (i.e.,  $L(x)$  is a real function of positive real values satisfying  $L(tx)/L(x) \rightarrow 1$ , as  $x \rightarrow \infty, \forall t > 0$ ). Parameter  $1/\alpha$  is usually denoted as the tail index of the Pareto rv.

Consider  $\{X_n\}_{n \geq 1}$  a Gaver–Lewis Pareto (GLP) process, presented in [Arnold \(2001\)](#). More precisely, for each positive integer  $n$ , we have

$$X_n = \sigma^p X_{n-1}^{1-p} \epsilon_n^{U_n}, \quad (3)$$

where  $\{\epsilon_n\}_{n \geq 1}$  is an independent and identically distributed (iid) sequence with common df Pareto( $\sigma, \alpha$ ) given in Equation (1) and independent of the Bernoulli( $p$ ) iid sequence  $\{U_n\}_{n \geq 1}$  with  $0 < p < 1$ , and  $\epsilon_i$  independent of  $X_j \forall i > j$ . If  $X_0 \sim \text{Pareto}(\sigma, \alpha)$ , then  $\{X_n\}_{n \geq 1}$  is stationary also with marginal df Pareto( $\sigma, \alpha$ ). The GLP process corresponds to the exponentiated version of the Gaver–Lewis process introduced in [Gaver and Lewis \(1980\)](#), and hence its name. Simulated samples from the GLP process with marginals Pareto(1, 1) and  $p = 0.25, 0.5, 0.75$  are plotted in Figure 1.



**Figure 1.** Simulated sample paths of the GLP process with marginals Pareto(1, 1) for  $p = 0.25$  (left),  $p = 0.5$  (middle) and  $p = 0.75$  (right).

This is a model within the heavy-tailed class where mean values (of different orders) may not exist. For instance, in this case, the mean exists only if  $\alpha > 1$  and the variance/covariance exists whenever  $\alpha > 2$ . In [Arnold \(2001\)](#), the autocorrelation was derived as

$$\rho(X_{n-1}, X_n) = \frac{(1-p)(\alpha-2)}{\alpha+p-2}, \quad (4)$$

for  $\alpha > 2$ . Moreover, in heavy-tailed models, the extremal observations are important, and a dependence analysis based on central measures like the most common autocorrelation may be misleading if the dependence in the tails presents a different structure from the remaining.

Here, we focus on the extremal behavior of the GLP process, namely the tail dependence structure. Despite being a heavy-tailed model, it is practically unknown in modeling extreme values. We shall see that it has interesting properties, such as asymptotic tail independence, i.e., the probability that one observation exceeds an increasing large value given that the previous one has already exceeded it, approaches zero. The rate of the convergence, usually denoted *coefficient of asymptotic tail independence*  $\eta$  ([Ledford and Tawn \(1996\)](#); [Wadsworth and Tawn \(2012\)](#)) captures a residual tail dependence, revealing a kind of “penultimate” clustering, i.e., an aggregation of not so high values. This is a not so fortuitous behavior in real applications and can be observed in the well-known Gaussian

processes (see [Bortot and Tawn \(1998\)](#), [Ramos and Ledford \(2013\)](#), and references therein). In practice, ignoring this phenomena will result in misleading inference (see, e.g., [Poon et al. \(2003\)](#)). However, not all series associated with environmental, social or economic phenomena are susceptible to Gaussian modeling, especially when they have heavy tails. The most common extremal models MARMA and M4, as well as, the Yeh–Arnold–Robertson Pareto (III) and the Lawrence–Lewis Pareto processes (see, respectively, [Ferreira \(2012\)](#) and [Ferreira \(2016\)](#) and references therein) are not tail independent and new processes have been appearing ([Heffernan et al. \(2007\)](#); [Ferreira and Canto e Castro \(2010\)](#); [Ferreira and Ferreira \(2014\)](#) and [Ferreira and Ferreira \(2015\)](#)). The GLP is an additional contribution within this class. Coefficient  $\eta$  will also be extended to observations that are lag- $m$  apart, providing an alternative to the autocorrelation function (acf).

This paper is organized as follows. The tail dependence measures and conditions to be analyzed are detailed in Section 2 and applied to the GLP process in Section 3. The tail characterization provides us with methods to estimate the autoregressive parameter  $p$ , which will be compared through simulation in Section 4. An illustration with a real dataset is addressed in Section 5.

## 2. Tail Dependence

A stationary sequence  $\{X_n\}_{n \geq 1}$  has extremal index  $\theta \in [0, 1]$  if, for each  $\tau > 0$ , there is a sequence of normalized levels  $\{u_n \equiv u_n^{(\tau)}\}_{n \geq 1}$ , i.e.,

$$n(1 - F(u_n)) \rightarrow \tau,$$

as  $n \rightarrow \infty$ , such that

$$P(\max(X_1, \dots, X_n) \leq u_n) \rightarrow e^{-\theta\tau} \quad (5)$$

([Leadbetter et al. \(1983\)](#)). The extremal index is a measure for the clustering propensity, being interpreted as the arithmetic inverse of the mean number of exceedances of an increasing threshold per independent cluster. The null case is often ignored, corresponding to a degenerate limiting distribution for the maximum. The value  $\theta = 1$  is associated to iid sequences but not only these. Below, we shall see that it is a form of asymptotic independence of extremes.

Some dependence conditions allow us to derive the extremal index through the joint distribution of  $s$  consecutive terms of  $\{X_n\}_{n \geq 1}$ .

The long-range condition  $D(u_n)$  of [Leadbetter \(1974\)](#) states that  $\alpha_{n,l_n} \rightarrow 0$ , as  $n \rightarrow \infty$ , for some sequence  $l_n = o(n)$ , where

$$\alpha_{n,l} = \sup \left\{ |P(M_{i_1, i_1+p} \leq u_n, M_{j_1, j_1+q} \leq u_n) - P(M_{i_1, i_1+p} \leq u_n) P(M_{j_1, j_1+q} \leq u_n)| : \right. \\ \left. 1 \leq i_1 < i_1 + p + l \leq j_1 < j_1 + q \leq n \right\}, \quad (6)$$

with  $M_{i,j} = \max_{s=i+1}^j (X_s)$ ,  $M_{0,j} = M_j$  and  $M_{i,j} = -\infty$  for  $i \geq j$ . Consider  $\{k_n\}_{n \geq 1}$  such that,

$$k_n \rightarrow \infty, k_n \alpha_{n,l_n} \rightarrow 0, k_n l_n / n \rightarrow 0, \text{ as } n \rightarrow \infty. \quad (7)$$

Observe that  $D(u_n)$  is a milder condition than the usual mixing, such as strong-mixing.

Under condition  $D(u_n)$ , we say that the local dependence condition  $D^{(s)}(u_n)$  of [Chernick et al. \(1991\)](#) holds for  $\{X_n\}_{n \geq 1}$ , if for some  $\{k_n\}_{n \geq 1}$  satisfying Equation (7), we have

$$nP(X_1 > u_n, M_{1,s} \leq u_n < M_{s,r_n}) \xrightarrow{n \rightarrow \infty} 0,$$

with  $\{r_n = \lfloor n/k_n \rfloor\}_{n \geq 1}$  ( $\lfloor x \rfloor$  denoting the integer part of  $x$ ). The validation of  $D^{(s)}(u_n)$  implies that  $D^{(s')}(u_n)$  holds for  $s' > s$ .

If  $D^{(s)}(u_n)$  holds, the extremal index exists and can be computed through (Chernick et al. (1991))

$$\theta = \lim_{n \rightarrow \infty} P(M_{1,s} \leq u_n | X_1 > u_n). \quad (8)$$

Observe that, under  $D^{(1)}(u_n)$ , we have  $\theta = 1$ . Condition  $D^{(s)}(u_n)$  is also implied by

$$n \sum_{j=s+1}^{r_n} P(X_1 > u_n, M_{1,s} \leq u_n < X_j) \xrightarrow{n \rightarrow \infty} 0.$$

This corresponds to condition  $D'(u_n)$  of Leadbetter et al. (1983) whenever  $s = 1$ , which locally restricts the occurrence of clusters of exceedances and thus leads to  $\theta = 1$ . Condition  $D''(u_n)$  of Leadbetter and Nandagopalan (1989) is obtained with  $s = 2$  and locally restricts upcrossing clustering.

Observe that, under  $D''(u_n)$ , we can write

$$\theta = \lim_{n \rightarrow \infty} P(X_2 \leq u_n | X_1 > u_n)$$

and thus state

$$\theta(u_n) \sim 1 - P(X_2 > u_n | X_1 > u_n), \quad n \rightarrow \infty,$$

where  $a(x) \sim b(x)$  means  $a(x)/b(x) \rightarrow c$ , as  $x \rightarrow \infty$ , for some constant  $c$ . Observe that, if  $\theta < 1$ , then  $P(X_2 > u_n | X_1 > u_n) > 0$ , meaning that consecutive observations are tail dependent. On the other hand, under a unit extremal index, we have  $P(X_2 > u_n | X_1 > u_n) \rightarrow 0$ , as  $n \rightarrow \infty$ , and thus consecutive observations are asymptotically tail independent. This feature has been observed in some real data and theoretical examples (Ledford and Tawn (1996); Bortot and Tawn (1998); Ramos and Ledford (2013)).

Ledford and Tawn (1996) introduced the asymptotic tail independence coefficient,  $\eta$ , in order to measure the rate of convergence of  $P(X_2 > F^{-1}(1-t) | X_1 > F^{-1}(1-t))$  towards 0, where  $F^{-1}$  is the quantile function, capturing a kind of pre-asymptotic dependence. More precisely, the asymptotic tail independence coefficient,  $\eta \in [0, 1]$ , exists whenever it holds

$$P(X_1 > F^{-1}(1-t), X_2 > F^{-1}(1-t)) \sim t^{1/\eta} L(1/t), \quad t \downarrow 0. \quad (9)$$

Thus, under Equation (9), we can state,

$$\theta(u_n) \sim 1 - P(X_2 > u_n | X_1 > u_n) \sim 1 - u_n^{1-1/\eta} L(u_n), \quad n \rightarrow \infty.$$

Observe that, if  $\eta = 1$  and  $L(u_n) \not\rightarrow 0$ , we have  $\theta < 1$  and thus an effect of clustering of high values. Under a unit extremal index, the coefficient  $\eta < 1$  measures the rate of convergence of  $\theta$  towards 1, capturing a kind of pre-asymptotic clustering, despite a resembling of the process to an iid sequence at increasingly high thresholds.

Analogous with the acf, we extend the coefficient  $\eta$  and state the tail dependence within random pairs that are lag- $m$  apart,  $(X_i, X_{i+m})$ ,  $i \geq 1$ , through the coefficient  $\eta_m$ , i.e.,

$$P(X_1 > F^{-1}(1-t), X_{1+m} > F^{-1}(1-t)) \sim t^{1/\eta_m} L(1/t), \quad t \downarrow 0,$$

where  $\eta_1 \equiv \eta$ .

The tail dependent class has been greatly enhanced within the methodology of extreme values. However, this approach results in the overestimation of extremal dependence if the series is actually

asymptotically independent. An illustration with financial data may be seen in [Poon et al. \(2003\)](#). The most recent literature has been addressing this issue, namely, with the introduction of new models comprising asymptotic tail independence ([Heffernan et al. \(2007\)](#); [Ferreira and Canto e Castro \(2010\)](#); [Ferreira and Ferreira \(2014, 2015\)](#)). In the next section, we will show that the GLP belongs to this latter class of models.

### 3. The Tail Dependence of the Gaver–Lewis Process

In the following, and without loss of generality, we will take  $\sigma = 1$ .

“Mixing” conditions roughly state that two rvs become increasingly independent as they get more apart in time. One of its forms is the  $\beta$ -mixing condition, defined by

$$\beta(l) := \sup_{p \in \mathbb{N}} E \left( \sup_{B \in \mathcal{F}(X_{p+l+1}, \dots)} |P(B|\mathcal{F}(X_1, \dots, X_p)) - P(B)| \right) \xrightarrow{l \rightarrow \infty} 0,$$

with  $\mathcal{F}(\cdot)$  denoting the  $\sigma$ -field generated by the indicated random variables ([Bradley \(2005\)](#)).

**Proposition 1.** *The GLP process is  $\beta$ -mixing.*

**Proof.** The  $\beta$ -mixing condition will be proved through the sufficient conditions of regeneration and aperiodicity (see, e.g., [Bradley \(2005\)](#); Corollary 3.6).

In the following, consider notation  $Q^m(x, ]1, y]) = P(X_{1+m} \leq y | X_1 = x)$ , with  $Q(x, ]1, y]) \equiv Q^1(x, ]1, y])$ .

Observe that

$$Q(x, ]1, y]) = P(X_2 \leq y | X_1 = x) = P\left(\epsilon_2^{U_2} \leq \frac{y}{x^{1-p}}\right) = F_\epsilon\left(\frac{y}{x^{1-p}}\right) p + 1 - p, \quad y \geq x^{1-p}.$$

First, we show that GLP is regenerative, that is, it has a regeneration set, i.e., a recurrent set  $R$  such that, for some  $m \in \mathbb{N}$ , a distribution  $\varphi$  and  $\kappa \in (0, 1)$ , we have

$$Q^m(x, B) \geq \kappa \varphi(B), \quad x \in R$$

for all Borel set  $B$  over  $\mathbb{R}$ . If, for any regeneration set  $R$  and any Borel set  $B$  over  $\mathbb{R}$ , we have

$$Q^{m+1}(x, B) \geq \kappa_1 \varphi(B) \quad \text{and} \quad Q^m(x, B) \geq \kappa_2 \varphi(B), \quad \forall x \in R, \quad (10)$$

for some  $m \in \mathbb{N}$  and  $\kappa_1, \kappa_2 \in (0, 1)$ , then the process is said to be aperiodic ([Asmussen \(1987\)](#)).

Consider  $R = ]1, r[$  (and thus recurrent since it is in the state space  $]1, \infty[$  of the process) and  $B$  a Borel set over  $\mathbb{R}$ . Let  $x \in R$ ,  $S = [r, r^{1/(1-p)})$  and  $V \sim \text{Pareto}(r^{1-p}, \alpha)$ . For all  $x \in R$ , we have

$$Q(x, B) \geq \int_{B \cap S} dQ(x, z) \geq P(V \in B \cap S)p,$$

and thus regeneration holds by considering  $m = 1$ ,  $\varphi(B) = P(V \in B | V \in S)$  and  $\kappa = P(V \in S)p$ . Observe that  $S$  is also regenerative since it is recurrent ( $S \subset ]1, \infty[$ ) and,  $\forall z \in S$ ,  $z^{1-p} \leq r < y$ , for any  $y \in S$ , and thus  $Q(z, B) \geq \kappa \varphi(B)$ , with  $\varphi(B)$  and  $\kappa$  as above. Now, we have

$$Q^2(x, B) \geq \int_S Q(z, B) dQ(x, z) \geq \kappa \varphi(B) Q(x, S) \geq \kappa \varphi(B) p P(V \in S).$$

Therefore, the aperiodicity condition in Equation (10) is satisfied if we take  $\kappa_1 = \kappa p P(V \in S)$  and  $\kappa_2 = \kappa$ .  $\square$

Note that condition  $D(u_n)$  given in Equation (6) is weaker than  $\beta$ -mixing and thus holds for GLP by the previous result.

**Proposition 2.** The GLP process satisfies condition  $D'(u_n)$  for sequences  $\{k_n\}_{n \geq 1}$  satisfying Equation (7) and such that  $(2-p)^{n/k_n}/n^p \rightarrow 0$ , as  $n \rightarrow \infty$ .

**Proof.** We have successively, for  $r_n = \lfloor n/k_n \rfloor$ ,  $n \geq 1$ ,

$$\begin{aligned}
 & n \sum_{j=2}^{r_n} P(X_1 > u_n, X_j > u_n) \\
 &= n \sum_{j=2}^{r_n} P(X_1 > u_n, X_j > u_n, U_2 = \dots = U_j = 0) + n \sum_{j=2}^{r_n} \sum_{k=1}^{j-1} P\left(X_1 > u_n, X_j > u_n, \sum_{i=2}^j \mathbb{1}_{\{U_i=1\}} = k\right) \\
 &\leq \sum_{j=2}^{r_n} \left( \frac{\tau(1-p)^{j-1}}{(n/\tau)^{1/(1-p)^{j-1}-1}} + \sum_{k=1}^{j-1} \sum_{2 \leq s_1 < \dots < s_k \leq j} \frac{\tau}{(n/\tau)^{1-(1-p)^{j-1}}} \prod_{i=0}^{k-1} \frac{p(1-p)^{j-1-k}}{1-(1-p)^{s_k-s_i}} \right) \\
 &\leq \sum_{j=2}^{r_n} \left( \frac{\tau(1-p)^{j-1}}{(n/\tau)^{1/(1-p)^{j-1}-1}} + \sum_{k=1}^{j-1} \sum_{2 \leq s_1 < \dots < s_k \leq j} \frac{\tau(1-p)^{j-1-k}}{(n/\tau)^p} \right) \\
 &\leq \frac{\tau}{(n/\tau)^p} \sum_{j=2}^{r_n} \sum_{k=0}^{j-1} \binom{j-1}{k} (1-p)^{j-1-k} \\
 &\leq \frac{\tau^{p+1}}{1-p} \frac{(2-p)^{r_n}}{n^p}.
 \end{aligned} \tag{11}$$

□

**Corollary 1.** The GLP process has  $\theta = 1$ .

This result reveals that high observations of the GLP process behave similar to an iid scenario. However, there is a weak dependence that may be evaluated through the Ledford and Tawn coefficient  $\eta$  in Equation (9). Moreover, we will see that it relates with parameter  $p$  of the process.

**Proposition 3.** The GLP process has  $\eta = 1/(1+p)$ .

**Proof.** Consider  $a_t = F_X^{-1}(1-t)$  and take  $t \downarrow 0$ . Observe that

$$\begin{aligned}
 & P(X_1 > a_t, X_2 > a_t) \\
 &= pP(X_1 > a_t, X_1^{1-p} \epsilon_2 > a_t) + (1-p)P(X_1 > a_t, X_1^{1-p} > a_t) \\
 &= p \int_{a_t}^{\infty} \bar{F}_{\epsilon_2}(a_t/x^{1-p}) dF_X(x) + (1-p)\bar{F}_X(a_t^{1/(1-p)}) \\
 &= a_t^{-\alpha(1+p)} + (1-p)a_t^{-\alpha/(1-p)} \sim t^{p+1} + t^{1/(1-p)}(1-p), \text{ as } n \rightarrow \infty.
 \end{aligned}$$

□

The fluctuation probability in the GLP process, given by

$$f := P(X_{n-1} < X_n) = pP(X_{n-1}^p < \epsilon_n) = p \int_1^{\infty} F_X(x^{1/p}) dF_{\epsilon}(x) = p/(1+p),$$

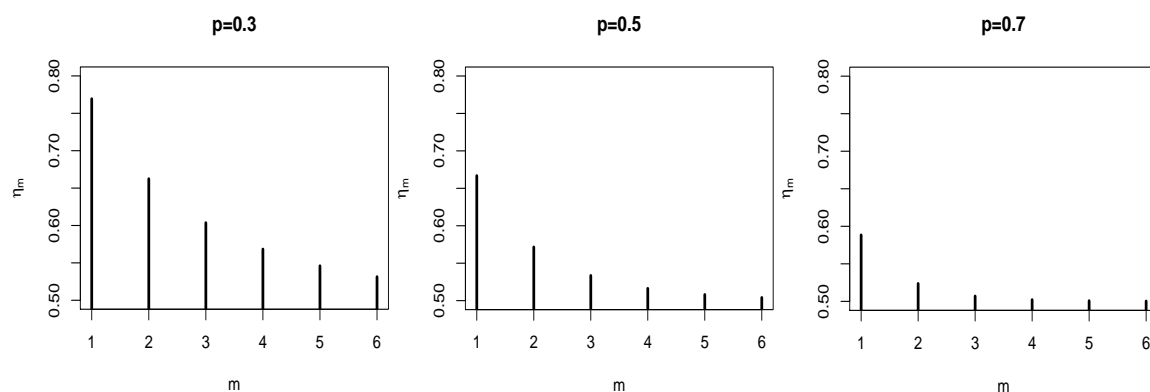
is a simple measure that will be useful in the following.

**Corollary 2.** *The GLP process verifies the following equalities:*

- (i)  $p = 1/(1 - f) - 1$ ;
- (ii)  $p = 1/\eta - 1$ ;
- (iii)  $\eta = 1 - f$ .

This result states a characterizing feature that can be helpful in model specification. Moreover, in order to satisfy  $0 < p < 1$ , we must have  $0 < f < 1/2$  and  $1/2 < \eta < 1$ .

Another interesting property for model identification is based on the lag- $m$  coefficient  $\eta_m$ , analogous with the acf for linear models. The plots in Figure 2 exhibit a power decay as the acf of AR(1) processes. Observe also that the smaller the value of  $p$ , the higher we must choose the lag- $m$  in order to have "almost" independent observations, i.e.,  $\eta_m \approx 1/2$ .



**Figure 2.** Plots of  $\eta_m$  for the GLP process with  $p = 0.3$  (left);  $p = 0.5$  (middle) and  $p = 0.7$  (right), for lags  $m = 1, \dots, 6$ .

**Proposition 4.** *The GLP process has lag- $m$  coefficient  $\eta_m$  given by*

$$\eta_m = 1 / \left( 2 - (1 - p)^m \right). \quad (12)$$

**Proof.** Consider  $a_t = F_X^{-1}(1 - t)$ . The product of powered Pareto rvs is still Pareto-type tail distributed (see, e.g., Arnold (2001)) and thus, applying Equation (2) and the theorem of the dominated convergence, we have successively, as  $t \downarrow 0$ ,

$$\begin{aligned} P(X_1 > a_t, X_{1+m} > a_t) &= \int_{a_t}^{\infty} P\left(x^{(1-p)^m} \prod_{j=0}^{m-1} \epsilon_{m+1-j}^{U_{m+1-j}(1-p)^j} > a_t\right) dF_X(x) \\ &\sim \int_{a_t}^{a_t^{(1-p)^{-(m+1)}}} \left(a_t x^{-(1-p)^{m+1}}\right)^{-\alpha} L(x/t) dF_X(x) + \int_{a_t^{(1-p)^{-(m+1)}}}^{\infty} dF_X(x) \\ &\sim a_t^{-\alpha} L(1/t) \int_{a_t}^{a_t^{(1-p)^{-(m+1)}}} \alpha x^{-\alpha(1-(1-p)^{m+1})-1} dx + a_t^{-\alpha(1-p)^{-(m+1)}} \\ &= \frac{a_t^{-\alpha} L(1/t)}{1 - (1-p)^{m+1}} \left( a_t^{-\alpha(1-(1-p)^{m+1})} - a_t^{-\alpha((1-p)^{-(m+1)}-1)} \right) + a_t^{-\alpha(1-p)^{-(m+1)}} \\ &\sim L^*(1/t) \left( t^{2-(1-p)^{m+1}} \right), \end{aligned} \quad (13)$$

where  $L(1/t)$  and  $L^*(1/t)$  are slowly varying functions.  $\square$

#### 4. Estimation

Relations (i) and (ii) stated in Corollary 2 will provide us with estimators for the autoregressive parameter  $p$ . More precisely, from (i), we have

$$\hat{p}^{(F)} = \frac{1}{1 - \hat{f}} - 1, \quad (14)$$

with  $\hat{f}$  corresponding to the empirical counterpart of  $f$ ,

$$\hat{f} = \frac{1}{n-1} \sum_{j=2}^n \mathbf{1}_{\{X_{j-1} < X_j\}},$$

provided that  $\hat{f} < 1/2$  (notation  $\mathbf{1}_{\{\cdot\}}$  means the indicator function). From the iid property of the generating process  $\varepsilon_t$  (and with  $X_0$  as specified for stationarity), we have ergodicity and thus consistence of the proposed estimators. In addition,  $\hat{f}$  corresponds to the mean of Bernoulli trials with Markov dependence. From Klotz (1973), we have that  $\sqrt{n}(\hat{f} - f)$  converges in distribution to a centered Gaussian model, and thus  $\sqrt{n}(\hat{p}^{(F)} - p)$  by the Delta Method, as  $n \rightarrow \infty$ . For more details, see Ferreira (2012).

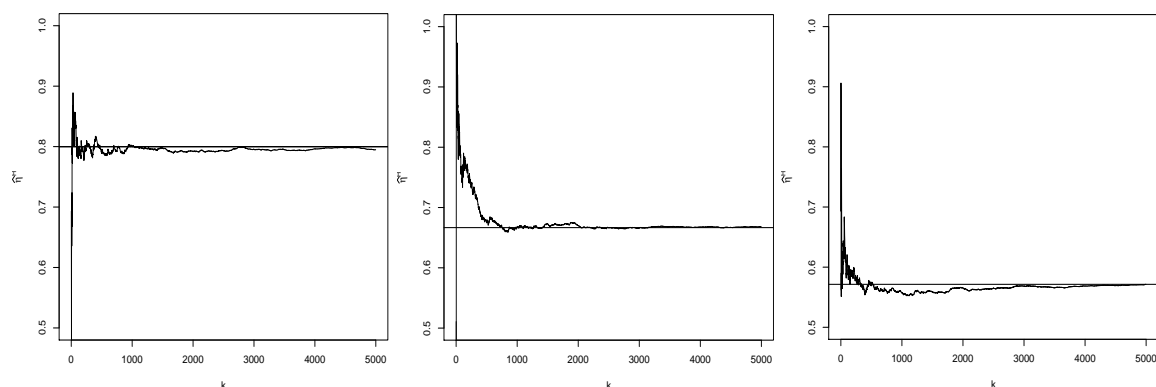
From (ii), the estimation of  $p$  is based on the estimation of  $\eta$  through

$$\hat{p}^{(H)} = \frac{1}{\hat{\eta}^{(H)}} - 1, \quad (15)$$

as long as  $\hat{\eta}^{(H)} > 1/2$ . Observe that  $\eta$  corresponds to the tail index of  $T = \min(1/(1 - F(X_1)), 1/(1 - F(X_2)))$ . The most common method developed in literature is the Hill estimator (Hill (1975))—thus the superscript “H”. More precisely, we have

$$\hat{\eta}^{(H)} = \frac{1}{k} \sum_{i=1}^k \log T_{n-i+1:n} - \log u,$$

where  $T_{n-k+1:n}, \dots, T_{n:n}$  are the  $k$  larger order statistics of  $T$  that exceed  $u$ . It is usual to consider  $u \in [T_{n-k:n}, T_{n-k+1:n}[$  and plot  $\hat{\eta}^{(H)}$  as a function of  $k$ . In Figure 3, we can see the Hill trajectories of  $\hat{\eta}$  for the respective GLP models considered in Figure 1. The paths are quite stable around the true value of  $\eta$  for a large range of values of  $k$ . Indeed, variable  $T$  corresponds to the minimum of unit Pareto rvs, where the Hill estimator behaves particularly well. Consistency and asymptotic normality of the Hill estimator  $\hat{\eta}^{(H)}$  can be seen in Draisma et al. (2004).



**Figure 3.** Hill plots of  $\hat{\eta}^{(H)}$  for the GLP process with marginals Pareto(1, 1) and  $p = 0.25$  (left);  $p = 0.5$  (middle) and  $p = 0.75$  (right).



Expecting to observe time series data behaving exactly as the GLP functional Equation (3) is not realistic. At best, we might observe perturbed versions of the GLP process, for instance “noisy” processes of the form  $X_n^{(\delta)} = X_n + \delta Z_n$ ,  $n \geq 1$ , where  $\{Z_n\}_{n \geq 1}$  is an iid sequence of standard Gaussian rvs and  $\delta > 0$ . Thus, the simulations cover the GLP and “noisy” GLP sample paths for  $\delta = 0, 0.1, 1$ . We consider 1000 replicas of sizes  $n = 100, 1000, 5000$  for  $p = 0.25, 0.5, 0.75$ , and marginals Pareto(1, 1). The computed estimates of the root mean squared error (rmse) and absolute bias (abias) are reported in Table 1, where the estimator  $\hat{p}^{(H)}$  is based on thresholds  $u$  corresponding to the sample minimum ( $q_0$ ), the median ( $q_{50}$ ) and the percentile 80 ( $q_{80}$ ). We also register the number of fails resulting, respectively, from  $\hat{f} \geq 1/2$  and  $\hat{\eta}^{(H)} \leq 1/2$ . Not surprisingly, they are more associated to small sample sizes, where the case  $p = 0.75$  seems particularly sensitive. Indeed, the results tend to be slightly worse under large  $p = 0.75$ , where the process approximates to independence. In practice, the difficulty in deciding between tail dependence ( $\eta = 1$ ) and asymptotic independence ( $\eta < 1$ ) is well known. For a survey on this topic, see, e.g., Poon et al. (2003) and Beirlant et al. (2004). The results get better as the sample sizes increase. We observe that estimator  $\hat{p}^{(F)}$  is the best for the GLP process but not so robust for “noisy” GLP. In what concerns estimator  $\hat{p}^{(H)}$ , it seems to present an overall better performance under  $u = q_0$ .

The estimation of the tail index parameter  $\alpha$  may be conducted through the Hill estimator  $\hat{\alpha}^{(H)}$  (Hill (1975)), which is consistent and asymptotically normal under strong mixing conditions (see Rootzn et al. (1990)).

**Table 1.** Simulation results of the root mean squared error (RMSE) and of the absolute bias (abias). The last three columns correspond to the number of fails (nf) in each case.

		RMSE			abias					
$n = 100$		$\delta = 0$	$\delta = 0.1$	$\delta = 1$	$\delta = 0$	$\delta = 0.1$	$\delta = 1$	$nf_{\delta=0}$	$nf_{\delta=0.1}$	$nf_{\delta=1}$
$p = 0.25$	$\hat{p}^{(F)}$	0.0548	0.1844	0.5050	0.0005	0.1693	0.4940	0	0	24
	$\hat{p}_{q_0}^{(H)}$	0.0775	0.0837	0.2302	0.0547	0.0070	0.2127	0	0	1
	$\hat{p}_{q_{50}}^{(H)}$	0.1225	0.1140	0.1095	0.0801	0.0130	0.0202	0	0	1
	$\hat{p}_{q_{80}}^{(H)}$	0.2280	0.2236	0.2098	0.1587	0.0500	0.0842	8	7	35
$p = 0.5$	$\hat{p}^{(F)}$	0.0707	0.1225	0.3317	0.0013	0.0943	0.3200	0	0	68
	$\hat{p}_{q_0}^{(H)}$	0.0894	0.0894	0.1789	0.0492	0.0557	0.1576	0	0	0
	$\hat{p}_{q_{50}}^{(H)}$	0.1342	0.1414	0.1342	0.0586	0.0633	0.0197	2	2	1
	$\hat{p}_{q_{80}}^{(H)}$	0.2191	0.2121	0.2049	0.0816	0.0783	0.0182	100	83	68
$p = 0.75$	$\hat{p}^{(F)}$	0.0894	0.0548	0.1483	0.0001	0.0265	0.1287	10	0	222
	$\hat{p}_{q_0}^{(H)}$	0.1000	0.0949	0.1265	0.0419	0.0422	0.0917	21	22	93
	$\hat{p}_{q_{50}}^{(H)}$	0.1265	0.1225	0.1342	0.0116	0.0120	0.0028	143	123	121
	$\hat{p}_{q_{80}}^{(H)}$	0.1975	0.1897	0.1975	0.0611	0.0526	0.0686	311	274	267
$n = 1000$	rmse	$\delta = 0$	$\delta = 0.1$	$\delta = 1$	$\delta = 0$	$\delta = 0.1$	$\delta = 1$	$nf_{\delta=0}$	$nf_{\delta=0.1}$	$nf_{\delta=1}$
$p = 0.25$	$\hat{p}^{(F)}$	0.0000	0.1673	0.5010	0.0003	0.1656	0.4998	0	0	0
	$\hat{p}_{q_0}^{(H)}$	0.0775	0.0000	0.1449	0.0547	0.0169	0.1423	0	0	0
	$\hat{p}_{q_{50}}^{(H)}$	0.1225	0.0316	0.0775	0.0801	0.0080	0.0742	0	0	0
	$\hat{p}_{q_{80}}^{(H)}$	0.2280	0.0548	0.0775	0.1587	0.0209	0.0494	8	0	0
$p = 0.5$	$\hat{p}^{(F)}$	0.0000	0.0949	0.3302	0.0005	0.0943	0.3285	0	0	0
	$\hat{p}_{q_0}^{(H)}$	0.0316	0.0316	0.1095	0.0058	0.0122	0.1068	0	0	0
	$\hat{p}_{q_{50}}^{(H)}$	0.0447	0.0447	0.0632	0.0087	0.0072	0.0508	0	0	0
	$\hat{p}_{q_{80}}^{(H)}$	0.0775	0.0775	0.0949	0.0158	0.0143	0.0459	0	0	0
$p = 0.75$	$\hat{p}^{(F)}$	0.0316	0.0548	0.1732	0.0006	0.0459	0.1695	0	0	10
	$\hat{p}_{q_0}^{(H)}$	0.0316	0.0316	0.0707	0.0053	0.0075	0.0636	0	0	0
	$\hat{p}_{q_{50}}^{(H)}$	0.0548	0.0548	0.0548	0.0084	0.0062	0.0186	0	0	0
	$\hat{p}_{q_{80}}^{(H)}$	0.1000	0.0949	0.1000	0.0120	0.0047	0.0287	16	16	7

Table 1. Cont..

		RMSE			abias					
$n = 5000$	rmse	$\delta = 0$	$\delta = 0.1$	$\delta = 1$	$\delta = 0$	$\delta = 0.1$	$\delta = 1$	$\text{nf}_{\delta=0}$	$\text{nf}_{\delta=0.1}$	$\text{nf}_{\delta=1}$
$p = 0.25$	$\hat{p}^{(F)}$	0.0000	0.1643	0.5000	0.0002	0.1649	0.4994	0	0	0
	$\hat{p}_{q_0}^{(H)}$	0.0000	0.0000	0.1378	0.0015	0.0107	0.1360	0	0	0
	$\hat{p}_{q_{50}}^{(H)}$	0.0000	0.0000	0.0837	0.0027	0.0004	0.0836	0	0	0
	$\hat{p}_{q_{80}}^{(H)}$	0.0316	0.0316	0.0707	0.0054	0.0044	0.0618	0	0	0
$p = 0.5$	$\hat{p}^{(F)}$	0.0000	0.0949	0.3302	0.0002	0.0939	0.3302	0	0	0
	$\hat{p}_{q_0}^{(H)}$	0.0000	0.0000	0.1049	0.0004	0.0061	0.1025	0	0	0
	$\hat{p}_{q_{50}}^{(H)}$	0.0000	0.0000	0.0632	0.0007	0.0004	0.0567	0	0	0
	$\hat{p}_{q_{80}}^{(H)}$	0.0316	0.0316	0.0632	0.0020	0.0021	0.0563	0	0	0
$p = 0.75$	$\hat{p}^{(F)}$	0.0000	0.0447	0.1703	0.0002	0.0451	0.1685	0	0	0
	$\hat{p}_{q_0}^{(H)}$	0.0000	0.0000	0.0632	0.0011	0.0034	0.0585	0	0	0
	$\hat{p}_{q_{50}}^{(H)}$	0.0316	0.0316	0.0316	0.0013	0.0000	0.0257	0	0	0
	$\hat{p}_{q_{80}}^{(H)}$	0.0447	0.0447	0.0548	0.0037	0.0026	0.0367	0	0	0

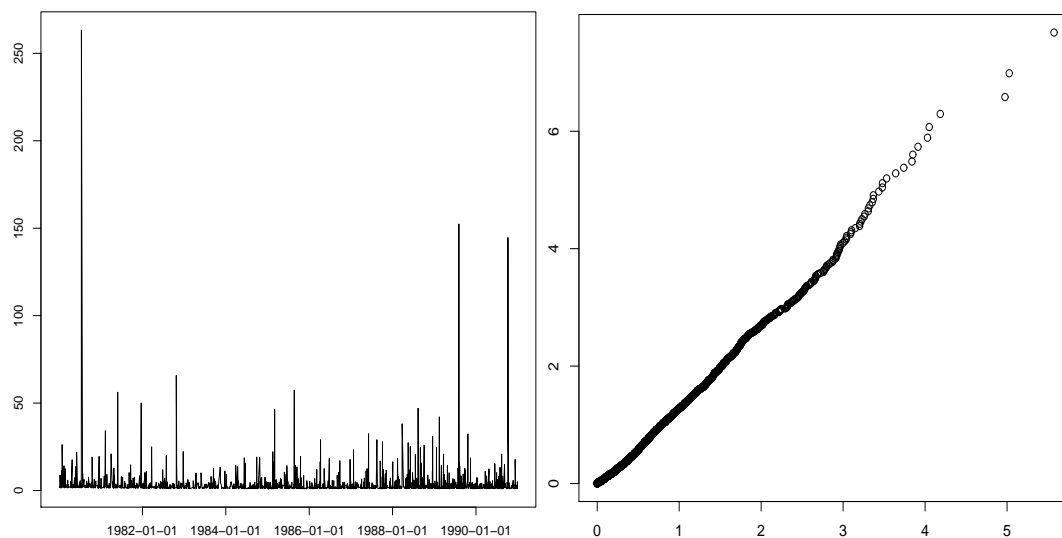
## 5. Application

Insurance loss data is typically well modeled by heavy-tailed processes. We consider the daily closing values for the Danish fire losses registered from January 1980 to December 1990, plotted in Figure 4 (left). Observe the high values that appear suddenly, similar to the GLP simulated sample paths (see Figure 1), as well as the close linearity of the Pareto quantile-quantile plot (right panel of Figure 4). In Figure 5 (left), the almost plane region of Hill's sample path led us to the estimate  $\hat{\alpha}^{(H)} \approx 1.4$ . Thus, we cannot assume the existence of the acf and should avoid an analysis based on this tool. We conduct the estimate of the GLP parameter  $p$  through  $\hat{p}^{(F)}$  and  $\hat{p}^{(H)}$ . More precisely, based on Equation (14), we obtain  $\hat{p}^{(F)} = 0.9945$  and, using Equation (15), we derive  $\hat{p}_{q_0}^{(H)} = 0.9610$ ,  $\hat{p}_{q_{50}}^{(H)} = 0.9715$  and  $\hat{p}_{q_{80}}^{(H)} = 0.9996$ , where the quantiles  $q_0$ ,  $q_{50}$  and  $q_{80}$  were considered according to the simulation study (see also the sample path estimates of Hill in the right panel of Figure 5). Formula (12) of the lag- $m$  coefficient  $\eta_m$  of Proposition 4 is a similar tool to the role of the acf in identifying linear models. Table 2 presents the estimates of  $\eta_m$ , for  $m = 1, 2, 3$ , obtained from estimator  $\hat{\eta}_m^{(H)}$ , which consists of the Hill estimator, respectively applied to lag- $m$  apart random pairs  $(X_i, X_{i+m})$ , as well as, estimates of  $\hat{\eta}_m(\hat{p}^{(F)})$  and  $\hat{\eta}_m(\hat{p}^{(H)})$  derived from Equation (12) by replacing  $p$ , respectively, by  $\hat{p}^{(F)}$  and  $\hat{p}^{(H)}$ . The closeness between the two type of estimates,  $\hat{\eta}_m^{(H)}$  and  $\hat{\eta}_m(\hat{p}^{(\cdot)})$ , shows a further contribution in favor of the model.

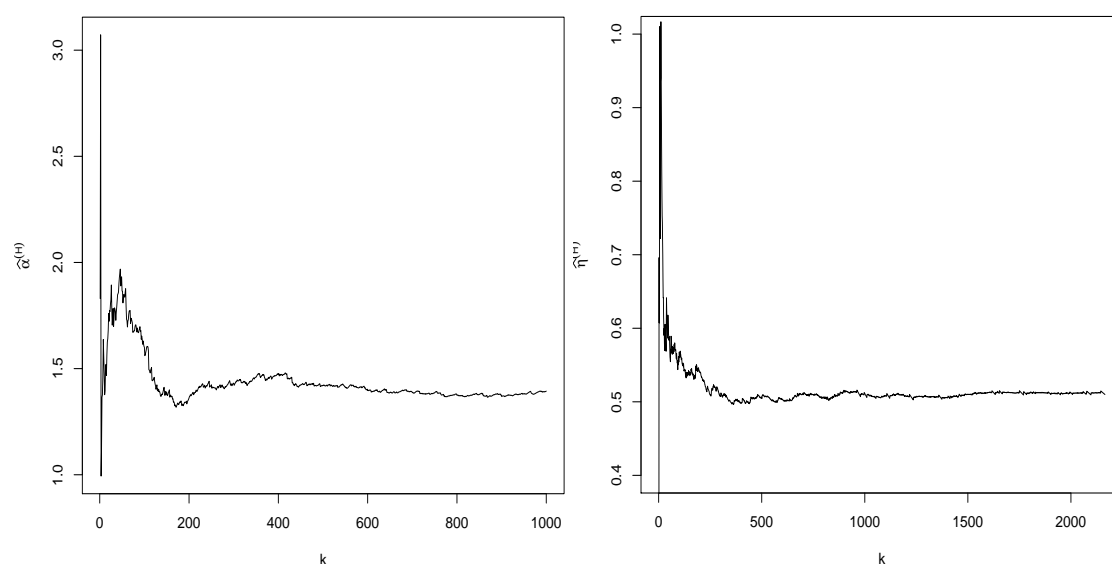
The GLP process thus seems to have potential in the modeling of this type of data. More tools regarding goodness-of-fit analysis will be addressed in a future work.

Table 2. Danish fire losses: estimates of the lag- $m$  coefficient  $\eta_m$ .

$\hat{\eta}_1(\hat{p}^{(F)})$	0.5014		
$\hat{\eta}_2(\hat{p}^{(F)})$	0.5000		
$\hat{\eta}_3(\hat{p}^{(F)})$	0.5000		
$\hat{\eta}_m$	$q_0$	$q_{50}$	$q_{80}$
$\hat{\eta}_1^{(H)} \equiv \hat{\eta}_1(\hat{p}^{(H)})$	0.5099	0.5072	0.5001
$\hat{\eta}_2^{(H)}$	0.5094	0.4995	0.5209
$\hat{\eta}_2(\hat{p}^{(H)})$	0.5004	0.5002	0.5000
$\hat{\eta}_3^{(H)}$	0.5081	0.4985	0.4820
$\hat{\eta}_3(\hat{p}^{(H)})$	0.5000	0.5000	0.5000



**Figure 4.** Danish fire losses: daily closing values from January 1980 to December 1990 (**left**); Pareto quantile-quantile plot (**right**).



**Figure 5.** Danish fire losses: trajectory of Hill estimates of  $\hat{\alpha}^{(H)}$  (**left**) and  $\hat{\eta}^{(H)}$  (**right**).

**Acknowledgments:** The authors wish to thank the reviewers for their important comments that have improved this work. The first was financed by Portuguese Funds through FCT—Fundação para a Ciência e a Tecnologia within the Project UID/MAT/00013/2013 and by the research center CEMAT (Instituto Superior Técnico, Universidade de Lisboa) through the Project UID/Multi/04621/2013. The second author’s research was partially supported by the research unit UID/MAT/00212/2013.

**Author Contributions:** The authors contributed equally to the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Arnold, Barry C. 2001. Pareto Processes. In *Handbook of Statistics*. Edited by D. N. Shanbhag and C. R. Rao. Vol. 19. Amsterdam: Elsevier Science B.V.
- Asmussen, Soren. 1987. *Applied Probability and Queues*. Hoboken: John Wiley & Sons.

- Beirlant, Jan, Yuri Goegebeur, Johan Segers, and Jozef Teugels. 2004. *Statistics of Extremes: Theory and Applications*. Hoboken: John Wiley & Sons.
- Bortot, Paola, and Jonathan A. Tawn. 1998. Models for the extremes of Markov chains. *Biometrika* 85: 851–67.
- Bradley, Richard C. 2005. Basic Properties of Strong Mixing Conditions. A Survey and Some Open Questions. *Probability Surveys* 2: 107–44.
- Chernick, Michael R., Tailen Hsing, and William P. McCormick. 1991. Calculating the extremal index for a class of stationary sequences. *Advance in Applied Probability* 23: 835–50.
- Davis, Richard A., and Sidney I. Resnick. 1989. Basic properties and prediction of max-ARMA processes. *Advance in Applied Probability* 21: 781–803.
- Draisma, Draisma, Holger Dress, Ana Ferreira, and Laurens de Haan. 2004. Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli* 10: 251–80.
- Ferreira, Marta. 2016. The Lawrence-Lewis Pareto process: an extremal approach. *Electronic Journal of Applied Statistical Analysis* 9: 68–82.
- Ferreira, Marta. 2012. On the extremal behavior of a Pareto process: An alternative for armax modeling. *Kybernetika* 48: 31–49.
- Ferreira, Marta, and Luísa Canto e Castro. 2010. Modeling rare events through a pRARMAX process. *Journal of Statistical Planning and Inference* 140: 3552–66.
- Ferreira, Helena, and Marta Ferreira. 2015. Extremes of scale mixtures of multivariate time series. *Journal of Multivariate Analysis* 137: 82–99.
- Ferreira, Helena, and Marta Ferreira. 2014. Extremal behavior of pMAX processes. *Statistics & Probability Letters* 93: 46–57.
- Gaver, Donald, and P. A. W. Lewis. 1980. First-Order Autoregressive Gamma Sequences and Point Processes. *Advances in Applied Probability* 12: 727–45.
- Heffernan, Janet E., Jonathan A. Tawn, and Zhengjun Zhang. 2007. Asymptotically (in)dependent multivariate maxima of moving maxima processes. *Extremes* 10: 57–82.
- Hill, Bruce M. 1975. A Simple General Approach to Inference About the Tail of a Distribution. *The Annals of Statistics* 3: 1163–74.
- Klotz, Jerome. 1973. Statistical inference in Bernoulli trials with dependence. *The Annals of Statistics* 1: 373–79.
- Leadbetter, M. Ross. 1974. On extreme values in stationary sequences. *Zeitschrift Für Wahrscheinlichkeitstheorie Und Verwandte Gebiete* 28: 289–303.
- Leadbetter, Malcolm R., Georg Lindgren, and Holger Rootzén. 1983. *Extremes and Related Properties of Random Sequences and Processes*. New York: Springer.
- Leadbetter, M. Ross, and S. Nandagopalan. 1989. On exceedance point processes for stationary sequences under mild oscillation restrictions. *Lecture Notes in Statistics* 51: 69–80.
- Ledford, Anthony W., and Jonathan A. Tawn. 1996. Statistics for near independence in multivariate extreme values. *Biometrika* 83: 169–87.
- Poon, Ser-Huang, Michael Rockinger, and Jonathan Tawn. 2003. Modelling Extreme-Value Dependence in International Stock Markets. *Statistica Sinica* 13: 929–53.
- Ramos, Alexandra, and Anthony Ledford. 2013. Estimation of the extremal index function in case of asymptotically independent Markov chains and its application to stock market indexes. In *Studies in Theoretical and Applied Statistics: Subseries B: Recent Developments in Modeling and Applications in Statistics (SPE2010)*. Edited by P. Oliveira, M. G. Temido, C. Henriques and M. Vichi. New York: Springer, pp. 89–96.
- Rootzen, Holger, Malcolm R. Leadbetter, and Laurens De Haan. 1990. *Tail and Quantile Estimation for Strongly Mixing Stationary Sequences*. Series: Report 9024/A; Rotterdam: Erasmus University.
- Smith, Richard L., and Ishay Weissman. 1996. *Characterization and Estimation of the Multivariate Extremal Index*. Technical Report; Chapel Hill: University of North Carolina.
- Wadsworth, Jennifer L., and Jonathan A. Tawn. 2012. Dependence modelling for spatial extremes. *Biometrika* 99: 253–72.

