

# Asymptotic Extremal Distribution for Non-Stationary, Strongly-Dependent Data

Carolina Crisci, Gonzalo Perera

Departamento Modelización Estadística de Datos e Inteligencia Artificial (MEDIA), CURE, Rocha, Universidad de la República, Montevideo, Uruguay

Email: [gperera@cure.edu.uy](mailto:gperera@cure.edu.uy)

**How to cite this paper:** Crisci, C. and Perera, G. (2022) Asymptotic Extremal Distribution for Non-Stationary, Strongly-Dependent Data. *Advances in Pure Mathematics*, 12, 479-489.

<https://doi.org/10.4236/apm.2022.128036>

**Received:** July 16, 2022

**Accepted:** August 12, 2022

**Published:** August 15, 2022

Copyright © 2022 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Fisher-Tippet-Gnedenko classical theory shows that the normalized maximum of  $n$  *iid* random variables with distribution  $F$  belonging to a very wide class of functions, converges in law to an extremal distribution  $H$ , that is determined by the tail of  $F$ . Extensions of this theory from the *iid* case to stationary and weak dependent sequences are well known from the work of Leadbetter, Lindgreen and Rootzén. In this paper, we present a very simple class of random processes that runs from *iid* sequences to non-stationary and strongly dependent processes, and we study the asymptotic behavior of its normalized maximum. More interesting, we show that when the process is strongly dependent, the asymptotic distribution is no longer an extremal one, but a mixture of extremal distributions. We present very simple theoretical and simulated examples of this result. This provides a simple framework to asymptotic approximations of extremes values not covered by classical extremal theory and its well-known extensions.

## Keywords

Extreme Events, Strongly Dependent Data, Fisher-Tippet-Gnedenko Theory

---

## 1. Introduction

The statistical analysis of extreme values has a wide and vast domain of applications on many disciplines. Extreme wind speeds are a key input for design in Structural Engineering. Maximum levels of traffic are crucial for design and operation in Telecommunications Networks. Maximum tides are essential for any policy concerning coast resources management. Extreme events on chemical, physical or biological conditions may affect dramatically very sensitive ecosystems, etc. [1]-[8].

The classical Fisher-Tippet-Gnedenko theory determines the asymptotic behavior of the maximum of a sample of size  $n$  of an iid sequence of random variables, showing that, for  $n$  tending to infinity, the distribution limit may be degenerated, and in any other case, it must be an extremal distribution [9] [10].

Since three decades ago, we know that classical theory applies to a stationary and weakly dependent sequence of random variables, thanks to the work of Leadbetter, Lindgreen and Rootzén [11]-[17].

However, in many real examples, a very big collection of maximal registers of a large series of measurements, does not fit to any extremal distribution. This is often associated to phenomena where the observed system may assume different states that produce drastical changes and that may introduce strong dependence on data. Think, for instance, of the classical series of data of the Nile River with its very ancient regimen of annual floods, and other large series of hydrological data, in particular those related to the impact of climate change [18] [19] [20].

At a theoretical level, we will apply in the extremal context a method that has been used for different purposes for non-stationary and strongly dependent data, where a random covariable indicates the state of the system, and the global behavior may be represented by a mixture of models. For instance, that is the case of Compound-Poisson approximation of High-Level Exceedances of time series [21], asymptotic of averages [22], and Nadaraya-Watson regression for functional data [23].

In this paper, we will consider data that depend of two independent components: on one hand, a categorical covariable process that describes the state of the system, that may be neither stationary nor weakly dependent, and that only satisfies that the mean frequency of each state has a (possibly random) limit, and, on the other hand, an *iid* noise. We will assume that for a given state  $j$  on the covariable, the maximal asymptotic distribution is non-degenerated and depends on  $j$ . In the main result of the paper, we will show that if the covariable process is stationary and weak-dependent, the maximal asymptotic distribution of our data is still an extremal distribution, consistently with Leadbetter, Lindgreen and Rootzén results. But we will also show that if this covariable process has a strong dependence structure, then this maximal asymptotic distribution is no longer an extremal one, but a mixture of extremal distributions.

The paper is organized as follows: we begin in Section 2 with a very brief summary of classical Fisher-Tippet-Gnedenko theory, in particular the characterization of the Maximal Domains of Attraction of extremal distributions and an elementary Lemma that we will use later. We then present in Section 3 our model, its hypotheses, some examples and the main result. In Section 4, we will fit some simulated strongly dependent data to a mixture of extremal distributions, showing that they do not fit extremal distributions. In particular, we will show the impact on return times of misfitting data to a single extremal distribution. Finally in Section 5, we present the conclusions and some further work in progress.

Therefore, what we provide here is a statistical method that is a step towards a

proper extremal values analysis for: 1) data that are suspected to be strongly dependent; 2) data of unknown dependence structure whose extremes do not fit to classical extremal distributions.

## 2. Brief Review of Classical Extreme Values Theory

The classical Fisher-Tippet-Gnedenko theory states that a distribution  $F$  belongs to the Maximal Domain of Attraction of an extremal distribution (Weibull, Gumbel, Fréchet)  $H$  if there exists an *iid* sequence  $X_1, \dots, X_n, \dots$  of random variables with distribution  $F$  such that for some deterministic sequences  $d_n$  and  $c_n > 0$ , we have that

$$\frac{\max(X_1, \dots, X_n) - d_n}{c_n} \xrightarrow[n]{w} H \quad (1)$$

Recall that

$$M_F = \sup\{t \in R : F(t) < 1\}, F^{-1}(p) = \inf\{x \in R : F(x) \geq p\} \quad \forall p \in (0, 1).$$

Then the three Maximal Domains of Attraction (MDA) are fully described as follows.

1) Fréchet of order  $\alpha$ :  $F$  belongs to the  $MDA(\Phi_\alpha)$ , where  $\Phi_\alpha(x) = \exp\{-x^{-\alpha}\}$   $\forall x > 0$  and  $\alpha > 0$  is called the order parameter, if and only if,  $M_F = +\infty$  and for  $x$  tending to infinity,  $1 - F(x) = \frac{L(x)}{x^\alpha}$  for some slowly varying function  $L$ .

In that case, in (1), the deterministic sequences are  $d_n = 0$ ,  $c_n = n^{1/\alpha}$ .

2) Weibull of order  $\alpha$ :  $F$  belongs to the  $MDA(\Psi_\alpha)$ , where  $\Psi_\alpha(x) = \exp\{-(-x)^\alpha\}$   $\forall x < 0$ , if and only if  $M_F < \infty$  and for  $x$  tending to  $M_F^-$ ,  $1 - F(x) = (M_F - x)^\alpha L\left(\frac{1}{M_F - x}\right)$  with  $L$  a slowly varying function. In this case  $d_n = M_F$  and  $c_n = n^{\frac{-1}{\alpha}}$ .

3) Gumbel:  $F$  belongs to the  $MDA(\Lambda)$ , where  $\Lambda(x) = \exp\{-e^{-x}\}$   $\forall x \in R$ , if and only if there exist an  $a < M_F$  (which may be finite or infinite), some  $c > 0$  and a positive function  $h$ , with density  $h'$ , such that  $\lim_{x \rightarrow M_F^-} h'(x) = 0$

and, for  $x$  tending to  $M_F^-$ ,  $1 - F(x)$  is equivalent to  $c \exp\left\{-\int_a^x \frac{1}{h(t)} dt\right\}$ , and

$$d_n = F^{-1}(1 - 1/n), \quad c_n = h(d_n).$$

## 3. Main Result

We will use the following elementary result, whose proof follows from simple analytical computations based on the characterizations of the involved sequences given in the previous section [24].

**Lemma 1:** If we denote by  $c_n(\Phi_\alpha)$ ,  $c_n(\Psi_\alpha)$ ,  $d_n(\Psi_\alpha)$ ,  $c_n(\Lambda)$ ,  $d_n(\Lambda)$ , the deterministic sequences corresponding to each MDA, we then have:

- i) If  $\alpha_1 < \alpha_2$ ,  $\frac{c_n(\Phi_{\alpha_2})}{c_n(\Phi_{\alpha_1})} \bar{n} > 0$ , and  $\forall \alpha > 0$ ,  $\frac{c_n(\Psi_\alpha)}{c_n(\Phi_{\alpha_1})} \bar{n} > 0$ ,  $\frac{c_n(\Lambda)}{c_n(\Phi_\alpha)} \bar{n} > 0$ .
- ii)  $\forall \alpha > 0$ ,  $\frac{c_n(\Psi_\alpha)}{c_n(\Lambda)} \bar{n} > 0$ .
- iii) If  $\alpha_1 < \alpha_2$ ,  $\frac{c_n(\Phi_{\alpha_1})}{c_n(\Phi_{\alpha_2})} \bar{n} > 0$ .

We will assume that the process  $Y$  satisfies:

**(H1)** For any state  $j = 1, \dots, k$ , there exists a (possibly random)  $b_j > 0$  such that

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n 1\{Y_{i=j}\} = b_j \text{ a.s.}$$

If  $I(t) = \sigma\{Y_i : i \geq t\}$ , and  $I(\infty) = \bigcap_{t>1} I(t)$ , then, since for any  $j$ ,  $b_j$  is  $I(\infty)$ -measurable, if  $I(\infty)$  is trivial,  $b_1, \dots, b_k$  are deterministic, but, if  $I(\infty)$  is not trivial, for some  $j$   $b_j$  may be non-deterministic, corresponding to strong dependence on the process  $Y$ . Let us show this in a very simple case.

**Example 1:**

Let  $U$  be a random variable such that  $P(U = 1) = p$ ,  $P(U = 2) = 1 - p$ . Let  $\sigma_1, \dots, \sigma_n, \dots$  an *iid* sequence of random variables on  $\{1, 2\}$  independent of  $U$  such that  $P(\sigma_i(1) = 1) = \delta$ ,  $P(\sigma_i(1) = 2) = 1 - \delta$ ,  $P(\sigma_i(2) = 1) = \eta$ ,  $P(\sigma_i(2) = 2) = 1 - \eta$ , with  $\sigma_i(1)$ ,  $\sigma_i(2)$  independent for any  $i$ ,  $0 < \delta < 1$ ,  $0 < \eta < 1$ .

Set  $Y_i = \sigma_i(U)$ .

Thus,  $\frac{1}{n} \sum_{i=1}^n 1\{Y_i = 1\} / U = 1$  has the same distribution as

$$\frac{1}{n} \sum_{i=1}^n 1\{\sigma_i(1) = 1\} \xrightarrow{\text{a.s.}} P(\sigma_i(1) = 1) = \delta \text{ (by the Strong Law of Large Numbers).}$$

On the other hand  $\frac{1}{n} \sum_{i=1}^n 1\{Y_i(1) = 1\} / U = 2$  has the same distribution as

$$\frac{1}{n} \sum_{i=1}^n 1\{\sigma_i(2) = 1\} \xrightarrow{\text{a.s.}} P(\sigma_i(2) = 1) = \eta.$$

Therefore if we assume that  $\delta \neq \eta$ , we have that  $b_1 = \begin{cases} \delta & \text{if } U = 1 \\ \eta & \text{if } U = 2 \end{cases}$

Hence,  $b_1$  is not-deterministic and  $I(\infty)$  is not trivial. Similar treatment applies to  $b_2$ .

Indeed, for the sake of simplicity, we will assume:

**(H2)** for any  $j$ ,  $b_j$  may only assume a finite number of values.

Our data will be of the form:

$$X_i = f(\varepsilon_i, Y_i) \tag{2}$$

where  $f$  is unknown, and we will assume:

**(H3)** The three following conditions are fulfilled.

- (i)  $\varepsilon_1, \dots, \varepsilon_n, \dots$  *iid*

(ii)  $Y_1, \dots, Y_n, \dots$  satisfies (H1) and (H2)

(iii) The processes  $\varepsilon_1, \dots, \varepsilon_n, \dots$  and  $Y_1, \dots, Y_n, \dots$  are independent.

Finally, we will assume that the state of the system is observable, and it affects the extremal behavior of our data. For instance,  $X$  may be a measure of tide exceedence of a baseline on a given coast, which depends on the type of wind affecting the area, that may be classified into a finite number of categories. Given the category of wind, the extreme tides behave in a significant different manner. Then,  $Y$  determines the type of wind on the coast, that may be observed from aerodynamical registers. In addition,  $X$  also depends on a series of random effects that may be considered as a white noise.

More precisely, we shall assume:

**(H4)** There exist an integer  $f, 1 < f < k$  and an integer  $g, 1 < g, f + g < k$  such that

a) For  $j = 1, \dots, f$  the *iid* process  $f(\varepsilon_1, j), \dots, f(\varepsilon_n, j), \dots$  belongs to  $\text{MDA}(\Phi_{\alpha_j})$ , where  $0 < \alpha_1 < \dots < \alpha_f$

b) For  $j = f + 1, \dots, f + g$  the *iid* process  $f(\varepsilon_1, j), \dots, f(\varepsilon_n, j), \dots$  belongs to  $\text{MDA}(\Lambda)$

c) For  $j = f + g + 1, \dots, k$ , the *iid* process  $f(\varepsilon_1, j), \dots, f(\varepsilon_n, j), \dots$  belongs to  $\text{MDA}(\Psi_{\alpha_j})$ , with  $0 < \alpha_{f+g+1} < \dots < \alpha_k$

Then we have our main result

**Theorem 1:**

Under (H3) there exists a random variable  $Z$  such that

$$\frac{\max(X_1, \dots, X_n)}{n^{1/\alpha_1}} \xrightarrow[n]{w} Z$$

In addition

a) If  $I(\infty)$  is trivial, then the distribution of  $Z$  is  $F_z(x) = \Phi_{\alpha_1} \left( \frac{x}{b_1^{1/\alpha_1}} \right)$ .

b) If  $I(\infty)$  is not trivial and  $b_1$  assumes the values  $v_1, \dots, v_r$  with probabilities  $p_1, \dots, p_r$ , then the distribution of  $Z$  is  $F_z(x) = \sum_{i=1}^r p_i \Phi_{\alpha_1} \left( \frac{x}{v_i^{1/\alpha_1}} \right)$  (Mixture of Fréchet distributions).

**Proof:**

Let us consider a fixed real  $x$  and set  $S = \{1, \dots, k\}^\infty$  the space of sequences taking values in  $1, \dots, k$

Consider  $g_n(j_1, \dots, j_n) = P \left( \frac{\max(X_1, \dots, X_n)}{n^{1/\alpha_1}} \leq x / Y_1 = j_1, \dots, Y_n = j_n \right)$  for any  $j_1, \dots, j_n$  in  $1, \dots, k$ .

Therefore,

$$P \left( \frac{\max(X_1, \dots, X_n)}{n^{1/\alpha_1}} \leq x \right) = \int_S g_n(j_1, \dots, j_n) dPY(j_1, \dots, j_n, \dots) \quad (3)$$

Re-ordering the maximum according to the state  $j$  taken by each  $Y_i$ , we get

$$g_n(j_1, \dots, j_n) = P \left( \max \left\{ \frac{\max \{X_1 1\{Y_1 = 1\}, \dots, X_n 1\{Y_n = 1\}\}}{n^{1/\alpha_1}}, \dots, \frac{\max \{X_1 1\{Y_1 = k\}, \dots, X_n 1\{Y_n = k\}\}}{n^{1/\alpha_1}} \right\} \leq x / Y_1 = j_1, \dots, Y_n = j_n \right) \tag{4}$$

Among the blocks of maximum values taken with  $Y$  fixed on a given state, the MDA is different, and taking into account Lemma 1, and (H4), all the blocks in (4) after the first, tends in probability to zero, and therefore, the limit as  $n$  tends to infinity of (4) is the same as the limit, for  $n$  tending to infinity of

$$P \left( \frac{\max \{f(\varepsilon_1, 1), \dots, f(\varepsilon_{N_1}, 1)\}}{n^{\alpha_1}} \leq x / Y_1 = j_1, \dots, Y_n = j_n \right) \tag{5}$$

where  $N_1 = \sum_{i=1}^n 1\{Y_i = 1\}$  (a random variable).

Then, the limit as  $n$  tends to infinity, of (3) is the same as the limit, for  $n$  tending to infinity of

$$\int_S P \left( \frac{\max \{f(\varepsilon_i, 1), \dots, f(\varepsilon_{N_1}, 1)\}}{n^{\alpha_1}} \leq x / Y_1 = j_1, \dots, Y_n = j_n \right) dP^Y(j_1, \dots, j_n, \dots) \tag{6}$$

Using that  $\varepsilon_1, \dots, \varepsilon_n$  are *iid*, and taking into account (5) and (6), the limit of (3) for  $n$  tending to infinity, is the same as

$$\lim_n \int_S P \left( \frac{\max \{f(\varepsilon_1, 1), \dots, f(\varepsilon_{N_1}, 1)\}}{n^{1/\alpha_1}} \leq x / Y_1 = j_1, \dots, Y_n = j_n \right) dP^Y(j_1, \dots, j_n, \dots) \tag{7}$$

Since  $\frac{N_1}{n} = \frac{1}{n} \sum_{i=1}^n 1\{Y_i = 1\} \xrightarrow{a.s.} b_1$ , and by Dominated Convergence Theorem, the limit in (7) equals

$$\int_0^1 \lim_n P \left( \frac{\max \{f(\varepsilon_1, 1), \dots, f(un, 1)\}}{(un)^{1/\alpha_1}} u^{1/\alpha_1} \leq x \right) dP^{b_1}(u) \tag{8}$$

By Fisher-Tippet-Gnedenko theorem and (H4)

$$\frac{\max \{f(\varepsilon_1, 1), \dots, f(un, 1)\}}{(un)^{1/\alpha_1}} \xrightarrow{w} \Phi_{\alpha_1} \quad \forall u \in (0, 1) \tag{9}$$

and therefore, if  $I(\infty)$  is trivial,  $b_1$  is deterministic and (8) equals to  $P(b_1^{1/\alpha_1} \Gamma \leq X)$ , with  $\Gamma \sim \Phi_{\alpha_1}$  and part a) of Theorem 1 follows.

On the other hand, if  $b_1$  is random, using (H2), part b) of Theorem 1 follows  $\diamond$

**Remark 1:**

It is clear that (H2) may be removed, leading in (b) to an integral with respect to the distribution of  $b_1$  instead of a sum.

**Remark 2:**

It is easy to obtain a similar result with Gumbel or Gumbel mixtures, if we modify (H4) removing part a) (taking  $f = 0$ ), and with Weibull or Weibull

mixtures if we remove parts a) and b) (taking  $f = g = 0$ ).

**Example 2:**

Following the ideas of Example 1, consider  $\sigma(1), \sigma(2)$  independent, such that  $P(\sigma(1)=1) = \delta, P(\sigma(1)=2) = 1-\delta, P(\sigma(2)=1) = \eta, P(\sigma(2)=2) = 1-\eta, 0 < \delta < 1, 0 < \eta < 1$  and  $\delta \neq \eta$ .

Taking  $(\sigma_1(1), \sigma_1(2)), \dots, (\sigma_n(1), \sigma_n(2)), \dots$  a sequence of independent copies of  $(\sigma(1), \sigma(2))$  it turns out that if  $U$  is a fixed random variable such that  $P(U=1) = p, P(U=2) = 1-p, 0 < p < 1$ , then if  $Y_i = \sigma_i(U)$ , we have that  $Y_1, \dots, Y_n, \dots$  fulfills (H1), (H2) with  $b_1, b_2$  random variables such that

$$b_1 = \begin{cases} \delta & \text{if } U = 1 \\ \eta & \text{if } U = 2 \end{cases}$$

and

$$b_2 = \begin{cases} 1-\delta & \text{if } U = 1 \\ 1-\eta & \text{if } U = 2 \end{cases}$$

Thus, if we assume  $0 < \alpha_1 < \alpha_2$  and consider two independent sequences  $V_1^{(1)}, \dots, V_n^{(1)}, \dots, iid \sim F^{(1)}, V_1^{(2)}, \dots, V_n^{(2)}, \dots, iid \sim F^{(2)}, F^{(i)} \in MDA(\Phi_{\alpha_i}), i = 1, 2$  and we set:

- a) If  $\sigma_i(U) = 1, X_i = V_i^{(1)}$
- b) If  $\sigma_i(U) = 2, X_i = V_i^{(2)}$

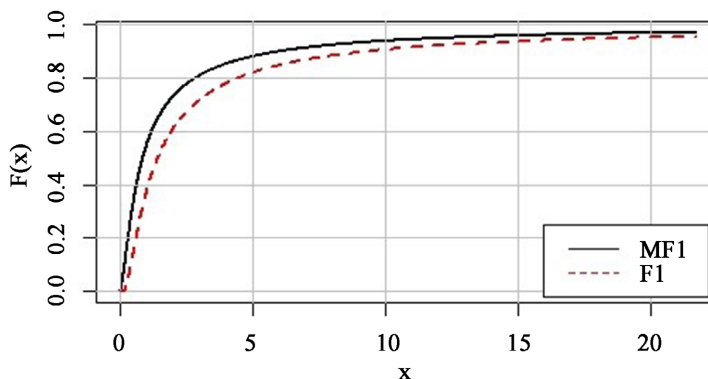
then, by Example 1 and part b) of Theorem 1,  $\frac{\max(X_1, \dots, X_n)}{n^{1/\alpha_1}} \xrightarrow{w} MF$ , with

$$MF(x) = p\Phi_{\alpha_1}\left(\frac{x}{\delta^{1/\alpha_1}}\right) + (1-p)\Phi_{\alpha_1}\left(\frac{x}{(1-\delta)^{1/\alpha_1}}\right) \quad \forall x > 0, \text{ a mixture of Fréchet}$$

distributions of order  $\alpha_1$ .

**Figure 1** shows the difference between a Fréchet model with  $\alpha_1 = 1$  (F1) and a mixture with the same  $\alpha_1$  (MF1),  $p = 0.25$  and  $\delta = 0.20$

It is clear that the tail of MF1 even for moderate values like 5, is smaller than the tail of F1, which means that the F1 leads to wrong pessimistic predictions for high levels. Indeed in **Table 1**, we present return times of MF1 and F1 for a series of levels, where the wrong pessimist results of F1 are very clear.



**Figure 1.** Fréchet model with  $\alpha_1 = 1$  (F1) and a mixture with the same  $\alpha_1$  (MF1),  $p = 0.25$  and  $\delta = 0.20$ .

**Table 1.** Return times for Fréchet and Mixture models.

Level	F1	MF1
50	50.50	77.50
100	100.50	154.43
500	500.50	769.81
1000	1000.50	1539.04
10000	10000.50	15385.20

### 4. Application to Simulated Data

We will now simulate data leading to the mixture model of the preceding example, and see if data fits only to the mixture model.

If we consider a distribution function of the form:

$$F(x)^{(\alpha)} = 1 - \frac{1}{(1+x)^\alpha} \quad \forall x > 0 \tag{10}$$

then  $1 - F(x)^{(\alpha)} = \frac{1}{(1+x)^\alpha} = \frac{1}{x^\alpha} L(x)$ , with  $L(x) = \left(\frac{x}{1+x}\right)^\alpha$  which is a slowly varying function and therefore  $F^{(\alpha)} \in \text{MDA}(\Phi_\alpha)$ .

Its inverse function is

$$F^{(\alpha)-1}(y) = \frac{1}{(1-y)^{1/\alpha}} - 1 \quad \forall y \in (0,1) \tag{11}$$

Consider  $(\sigma(1), \sigma(2))$  a random variable such that  $\sigma(1), \sigma(2)$  are independent and  $P(\sigma(1)=1) = \delta$ ,  $P(\sigma(1)=2) = 1 - \delta$ ,  $P(\sigma(2)=1) = \eta$ ,  $P(\sigma(2)=2) = 1 - \eta$ , with  $0 < \delta < 1$ ,  $0 < \eta < 1$ ,  $\delta \neq \eta$ .

We shall take  $\delta = 0.20$  and  $\eta = 0.80$ .

Consider  $n = 600$ ,  $N = 200$  and  $(\sigma_{i,j}(1), \sigma_{i,j}(2))$  (with  $1 \leq i \leq N$  and  $1 \leq j \leq n$ ) a matrix of independent copies of  $(\sigma(1), \sigma(2))$ . As seen in Example 1, if for each  $i, j$  we define  $Y_{i,j} = \sigma_{i,j}(U_j)$ , where  $U_1, \dots, U_N$  iid such that  $P(U_j = 1) = p$ ,  $P(U_j = 2) = 1 - p$  (we will take  $p = 0.25$ ), then, for each fixed  $i$ ,  $Y_{i,1}, \dots, Y_{i,n}$  is a sample of a process that fulfills (H1) and (H2).

We consider now two independent iid matrices  $\varepsilon_{i,j}^{(1)}$  (with  $1 \leq i \leq N$  and  $1 \leq j \leq n$ ) and  $\varepsilon_{i,j}^{(2)}$  (with  $1 \leq i \leq N$  and  $1 \leq j \leq n$ ) such that  $\varepsilon_{i,j}^{(1)} \sim F^{(1)}$  (the distribution of (10) for  $\alpha = 1$ ),  $\varepsilon_{i,j}^{(2)} \sim F^{(2)}$  (the distribution of (10) for  $\alpha = 2$ ), and we finally set for each  $i = 1, \dots, N$ ,

$$X_{i,j} = \begin{cases} \varepsilon_{i,j}^{(1)} & \text{if } Y_{i,j} = 1 \\ \varepsilon_{i,j}^{(2)} & \text{if } Y_{i,j} = 2 \end{cases}$$

Hence, for each  $i$ ,  $X_{i,1}, \dots, X_{i,n}$  is a sample where Theorem 1 applies, and

therefore,  $M_{n,i} = \frac{\max\{X_{i,1}, \dots, X_{i,n}\}}{n}$  must be close to the distribution

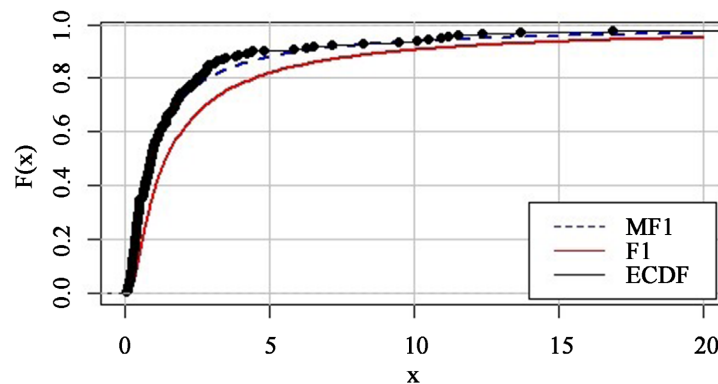


$$MF1(x) = 0.25\Phi_1\left(\frac{x}{0.2}\right) + 0.75\Phi_1\left(\frac{x}{0.8}\right).$$

We will now compute the empirical cumulative distribution function (*ecdf*) of  $M_1, \dots, M_N$  and see if it fits better to MF1 or to a simple Fréchet model (F1).

First, in **Figure 2**, we will display both models (MF1 and F1) and *ecdf* graphically, and afterwards, we will perform a Kolmogorov-Smirnov goodness of fit test (K-S test) to  $M_1, \dots, M_N$  with respect to MF1 and F1.

It can be seen that the *ecdf* is closer to the theoretical MF1 function. The K-S test support this result since  $H_0$  is retained with respect to MF1 (K-S test statistic = 0.088, critical value = 0.096 for a significance level of 0.05 and  $n = 200$ ), while it is rejected with respect to F1 (K-S test statistic = 0.177).



**Figure 2.** Theoretical cumulative distribution function of MF1 and F1, and *ecdf* of  $M_1, \dots, M_N$ .

## 5. Discussion & Conclusions

Our main interest is to provide theoretical results that guide practitioners trying to perform extremal analysis for data with complex dependence structure.

This paper shows that when the maximum of a large sample does not fit to an extremal distribution, it may be due to a strong-dependent structure that may be solved by fitting data to a mixture of extremal distributions.

Indeed, here we deal with the case when data depends on a covariable  $Y$  that may be strongly dependent, but which is observable. This is a reasonable assumption in many cases.

For future research, we also have to consider situations where  $Y$  may be hidden, and in a work in progress we are improving methods for estimation and fitting of a mixture of extremal distributions in such cases.

On the other hand, the very simple model used here to represent strongly-dependent samples, may also be applied to other techniques used in statistical analysis of extreme values. In a work in progress, strongly dependent structures are considered for Peaks Over a Manifold (POM), an approach that includes classical Peaks Over Threshold (POT) technique as a particular case [25].

## Acknowledgements

Sincere thanks to the participants of the “Jornadas de Estadística Aplicada 2019”

in La Paloma-Uruguay, for suggestions on a previous version of these results. We are also grateful to anonymous reviewers for their valuable comments on this paper.

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Batt, R.D., Carpenter, S.R. and Ives, A.R. (2017) Extreme Events in Lake Ecosystem Time Series. *Limnology and Oceanography Letters*, **2**, 63-69. <https://doi.org/10.1002/lol2.10037>
- [2] Bousquet, N. and Bernardara, P. (2017) Extreme Value Theory with Applications to Natural Hazards: From Statistical Theory to Industrial Practice. Springer, Cham.
- [3] Embrechts, P., Klüppelberg, C. and Mikosch, T. (2008) Modelling Extremal Events: For Insurance and Finance. Stochastic Modelling and Probability. 2nd Edition, Springer, Berlin.
- [4] Hannesdóttir, Á., Kelly, M.C. and Dimitrov, N. (2019) Extreme Wind Fluctuations: Joint Statistics, Extreme Turbulence, and Impact on Wind Turbine Loads. *Wind Energy Science*, **4**, 325-342. <https://doi.org/10.5194/wes-4-325-2019>
- [5] Jiménez, E., Cabanas, B. and Lefebvre, G. (2015) Environment, Energy and Climate Change I: Environmental Chemistry of Pollutants and Wastes. Springer, Berlin. <https://doi.org/10.1007/978-3-319-12907-5>
- [6] Katz, R.W., Brush, G.S. and Parlange, M.B. (2005) Statistics of Extremes: Modeling Ecological Disturbances. *Ecology*, **86**, 1124-1134. <https://doi.org/10.1890/04-0606>
- [7] Reiss, R.-D. and Thomas, M. (2007) Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields. Birkhäuser, Basel.
- [8] Rootzén, H. and Tajvidi, N. (1997) Extreme Value Statistics and Wind Storm Losses: A Case Study. *Scandinavian Actuarial Journal*, **1997**, 70-94. <https://doi.org/10.1080/03461238.1997.10413979>
- [9] Fisher, R.A. and Tippett, L.H.C. (1928) Limiting Forms of the Frequency Distribution of the Largest or Smallest Member of a Sample. *Mathematical Proceedings of the Cambridge Philosophical Society*, **24**, 180-190. <https://doi.org/10.1017/S0305004100015681>
- [10] Gnedenko, B. (1943) Sur la distribution limite du terme maximum d'une serie aleatoire. *Annals of Mathematics*, **44**, 423-453. <https://doi.org/10.2307/1968974>
- [11] Leadbetter, M.R. (1983) Extremes and Local Dependence in Stationary Sequences. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, **65**, 291-306. <https://doi.org/10.1007/BF00532484>
- [12] Leadbetter, M.R., Lindgren, G. and Rootzén, H. (1983) Extremes and Related Properties of Random Sequences and Processes. Springer, Berlin. <https://doi.org/10.1007/978-1-4612-5449-2>
- [13] Leadbetter, M.R. and Rootzén, H. (1988) Extremal Theory for Stochastic Processes. *The Annals of Probability*, **16**, 431-478. <https://doi.org/10.1214/aop/1176991767>
- [14] Rootzén, H. (1978) Extremes of Moving Averages of Stable Processes. *The Annals of Probability*, **6**, 847-869. <https://doi.org/10.1214/aop/1176995432>

- 
- [15] Rootzén, H. (1986) Extreme Value Theory for Moving Average Processes. *The Annals of Probability*, **14**, 612-652. <https://doi.org/10.1214/aop/1176992534>
- [16] Rootzén, H. (1988) Maxima and Exceedances of Stationary Markov Chains. *Advances in Applied Probability*, **20**, 371-390. <https://doi.org/10.2307/1427395>
- [17] Leadbetter, M.R., Rootzén, H. and de Haan, L. (1998) On the Distribution of Tail Array Sums for Strongly Mixing Stationary Sequences. *The Annals of Applied Probability*, **8**, 868-885. <https://doi.org/10.1214/aoap/1028903454>
- [18] Knox, J.C. (1993) Large Increases in Flood Magnitude in Response to Modest Changes in Climate. *Nature*, **361**, 430-432. <https://doi.org/10.1038/361430a0>
- [19] Mostafa, H., Roushdi, M. and Kheireldin, K. (2016) Statistical Analysis of Rainfall Change over the Blue Nile Basin. *18th International Conference on Environment and Climate Change (ICECC 2016)*, Zurich, January 2016.
- [20] Willems, P. (1998) Hydrological Applications of Extreme Value Analysis. In: Wheater, H. and Kirby, C., Eds., *Hydrology in a Changing Environment*, John Wiley & Sons, Chichester, 15-25.
- [21] Bellanger, L. and Perera, G. (2003) Compound Poisson Limit Theorems for High-Level Exceedances of Some Non-Stationary Processes. *Bernoulli*, **9**, 497-515. <https://doi.org/10.3150/bj/1065444815>
- [22] Perera, G. (2002) Irregular Sets and Central Limit Theorems. *Bernoulli*, **8**, 627-642.
- [23] Aspirot, L., Bertin, K. and Perera, G. (2009) Asymptotic Normality of the Nadaraya-Watson Estimator for Nonstationary Functional Data and Applications to Telecommunications. *Journal of Nonparametric Statistics*, **21**, 535-551. <https://doi.org/10.1080/10485250902878655>
- [24] Perera, G., Segura, A. and Crisci, C. (2021) Estadística de datos extremos: Teoría y ejemplos en R. <https://www.maren.cure.edu.uy/>
- [25] Perera, G. and Segura, A.M. (2022) Peaks over Manifold (POM): A Novel Technique to Analyze Extreme Events over Surfaces. *Advances in Pure Mathematics*, **12**, 48-62. <https://doi.org/10.4236/apm.2022.121004>