Jordan Journal of Mathematics and Statistics.
Yarmouk University

DOI:https://doi.org/10.47013/19.2.13

A New Regression Model for Multivariate Extremes Using the Peaks-Over-Threshold Method

Amina EL BERNOUSSI[1,*], *Mohamed EL ARROUCHI*[1]

Ibn Tofail University, Faculty of Sciences, Laboratory Analysis, Geometry and Applications, Kenitra, Morocco

Received: Dec. 29, 2024 Accepted:Oct. 13, 2025

Abstract: This paper presents a new regression model tailored for analyzing multivariate extremes, particularly when both the response variable and the covariates exhibit extreme behavior. The methodology combines the Peaks-Over-Threshold (POT) approach with ARIMA-based preprocessing to ensure the stationarity of climatic time series data before extreme value modeling. A key innovation of our approach is the use of a Logistic-Normal prior for spectral density estimation, which provides a more flexible and expressive alternative to the classical Dirichlet priors typically used in Peaks-Over-Threshold (POT) modeling of multivariate extremes. This framework enables the construction of smooth regression manifolds that describe conditional quantiles of extreme responses given extreme covariates, offering improved interpretability and adaptability to complex dependence structures. The model is applied to temperature and precipitation data from Meknes, Morocco (1973–2024), revealing meaningful extremal dependence patterns. Validation is conducted through quantile regression manifolds and bootstrap-based uncertainty estimates, demonstrating the practical relevance and robustness of the proposed method in environmental applications.

Keywords: Spectral measure; Bivariate extreme value distribution; Joint distribution; Quantile regression; Peaks-Over-Threshold; ARIMA method.

2010 Mathematics Subject Classification. 26A25; 26A35.

1 Introduction

In recent years, there has been significant interest in statistical methods for simulating extremes using time series data. This involves analyzing a sequence of data points collected at uniform intervals over time to examine trends, patterns, and other temporal dynamics. It is more beneficial in various applied disciplines, particularly in methodology related to earth science, as seen in studies by Cheng et al. [9], Katz et al. [23], and Mentaschi et al. [31]. It makes sense for us to be cautious of the harm and disasters that can result from climate extreme occurrences that have an impact on society and the economy. Since both are dynamic, time series can be utilized to capture the fluctuations in climate.

One of the primary challenges in modeling extreme risks is the limited number of extreme observations, making it difficult to accurately estimate the probabilities of rare events. Extreme Value Theory (EVT) addresses this issue by providing asymptotic results that extend beyond observed values. EVT has a substantial body of literature and is increasingly applied in fields such as insurance and finance. The two main EVT approaches are the Block Maxima and Peaks-Over-Threshold (POT) methods. The Block Maxima method focuses on the distribution of the maximum values within blocks of data, while the POT method examines the distribution of values exceeding a specified threshold. A major limitation of using the component-wise maximum for modeling multivariate extremes is that it typically leads to a mismatch in the timing of the component-wise maximum vectors within blocks. More specifically, likely, the annual maximum temperature and annual maximum precipitation in the data application (and, in fact, more generally in most places on Earth) occur at different times. To this end, we focus on the bivariate POT method to analyze the joint distribution of temperature and precipitation extremes, which requires the threshold to be determined.

* Corresponding author e-mail: amina.elbernoussi@uit.ac.ma

To determine the appropriate threshold, we analyze the function of empirical means and create a Mean Excess plot. Typically, thresholds identified from this plot correspond to about 10% extreme values exceeding these limits. This approach aligns with recommendations from Gavin [18] and Neftci [32], who suggest using a maximum of 10% of the distribution tails for defining extreme values. It's important to note, as emphasized by Longin [29] and Gavin [18], that choosing a threshold can be somewhat subjective, but it must maintain a balance between variance and bias to ensure the generalized Pareto distribution (GPD) estimator converges effectively. To verify the fit of the empirical distribution of extreme observations with the theoretical distribution, we employed several methods, including QQ-Plot, the Kolmogorov-Smirnov test, and the GPD test. In our analysis, the Kolmogorov-Smirnov test yielded favorable results, leading us to retain 10% of the tails of the precipitation and temperature distributions to define the extreme returns.

The main result in our paper is that we have fitted a new model in this research that explicitly develops a regression procedure for the case where both the response and the variables are severe. It has to do with quantile regression, which was initially created by Koenker and Bassett[26]. The original goal of the technique was to model the conditional quantile of a response T given a covariate $S = (S_1, ..., S_p)^T$, that is, where

$$F^{-1}(\alpha|s) = \inf\{t : F(t|s) \geq \alpha\}, \quad 0 < \alpha < 1, \tag{1}$$

and $F(t|s)$ is the distribution function of $T|S = s$. The concept of regression manifolds involves a group of regression lines that illustrate the relationship between extreme values. It helps to understand how an extreme value covariate can impact an extreme value response, as explained by de Carvalho et al. [13].

For this useful purpose, we have proposed a prior in the regression manifold space by resorting to a Logistic-Normal distribution on the space of spectral densities, which was proposed by Aitshison et al. [3]. Hanson et al. [22] and Carvalho et al.[13] proposed a Dirichlet distribution as a prior. Moreover, the Logistic-Normal distribution is a valuable tool for analyzing temperature and precipitation data in environmental research. Its flexibility, ability to model dependencies, and support for Bayesian analysis make it suitable for understanding complex relationships in climatic data, ultimately aiding in better decision-making and predictions in environmental management and policy. Aitshison [3] proved that it offers a richer parameterization compared to the Dirichlet distribution, which is often used for compositional data. This flexibility allows researchers to capture complex relationships and correlations between temperature and precipitation data, which may not be independent. Also, it facilitates straightforward estimation and hypothesis testing for environmental data. In addition, it supports Bayesian methods, allowing for the incorporation of prior knowledge about temperature and precipitation patterns. This is particularly useful in predictive modeling and assessing future climate scenarios. While the Dirichlet distribution is commonly used for modeling proportions, it assumes independence among components, which may not be realistic for temperature and precipitation data. The Logistic-Normal distribution, with its ability to model dependencies, provides a more accurate representation of the underlying processes.

Ultimately, we implemented our novel approach using actual data of methodology from local Meknes-Morocco, which we considered the extreme precipitations as a response and the extreme temperature as covariates during the period 1973-2024. For data preprocessing, we use the AutoRegressive Integrated Moving Average (ARIMA) method, which is a widely established statistical technique used for forecasting and analyzing time series data. Developed by statisticians George E. P. Box and Gwilym M. Jenkins [7], ARIMA combines three key components: the autoregressive (AR) aspect, which models the relationship between current observations and their past values; the integrated (I) component, which involves differencing the data to achieve stationarity; and the moving average (MA) aspect, which accounts for the influence of past forecast errors on current values.

The main theoretical contribution of this work is the construction of a quantile regression framework for multivariate extremes using a Logistic-Normal prior on the spectral density, which enables flexible and smooth regression manifolds. From a methodological perspective, the model integrates the POT method to extract extremes and the ARIMA model to ensure stationarity in the time series before applying extreme value techniques. The applied contribution lies in demonstrating the model's effectiveness on real climatic data from Meknes, Morocco (1973–2024), where precipitation and temperature extremes are jointly analyzed. This framework captures interpretable dependence structures between extreme precipitation and temperature and fills a gap in classical models that typically neglect the joint extremal behavior of covariates and responses.
This work aligns closely with the scope of the Jordan Journal of Mathematics and Statistics, which emphasizes both theoretical advancements and applied methodologies in mathematical statistics. By proposing a new regression model for analyzing extremes in both response and covariate variables using the multivariate Peaks-Over-Threshold framework, our study contributes to the development of statistical tools for non-stationary and dependent data—topics that are

central to the journal's mission. The integration of extreme value theory, spectral modeling, and likelihood-based inference reflects a rigorous mathematical foundation, while the application to environmental data demonstrates the model's relevance to real-world phenomena, consistent with the journal's interest in bridging theory and application.
The rest of the paper is structured as follows. Section 2 reviews some fundamental theories in EVT with a focus on the multivariate setting and the main result. Section 3 presents a simulation study and results of our study based on Meknes-Morocco meteorology data, where we consider the precipitation as a response and the temperature as a covariate, and finally, we summarize with a little discussion for our new approach.

2 Methodology

In this section, we outline the key steps of our modeling approach. We begin by collecting raw time series data on extreme precipitation and temperature, where precipitation is treated as the response variable and temperature as the covariate. The POT method is applied to extract the extreme values, using a high threshold u that will be selected based on empirical diagnostics.
To ensure stationarity and remove trends, we employ the ARIMA model, a widely used technique in time series analysis. Once the data are detrended and stabilized, they are transformed to unit Fréchet margins, enabling the use of multivariate extreme value theory.
Our main contribution lies in constructing a new quantile regression model tailored for extremes. This involves modeling the spectral measure with a Logistic-Normal prior distribution, which allows for greater flexibility in capturing dependence structures. Before delving into the technical formulation, we first review the foundational elements of the proposed framework.

2.1 The Extreme Value Theory and Peak-Over-Threshold Approach

EVT offers a robust framework for analyzing climate extremes and determining their corresponding return levels. While Coles [11] advocates a pragmatic enhancement of classical extreme value models to address the challenges posed by non-stationary environmental processes, Katz et al. [24] emphasize the integration of covariates and statistical downscaling—especially through maximum likelihood estimation—to improve the physical relevance and methodological rigor of hydrologic applications under climate variability and change. EVT posits that, under certain conditions, the distribution of extreme values—encompassing both maxima and minima—approaches one of three limiting distributions: Gumbel, Fréchet, or Weibull. Leadbetter [27] demonstrates that even without the local dependence condition—which typically controls clustering of extreme events—the limiting distribution of maxima in stationary sequences remains unchanged under broad assumptions, and that the extent of clustering can be effectively captured by a parameter known as the "extremal index". Together, these distributions comprise the Generalized Extreme Value (GEV) distribution. A variety of studies have utilized GEV to explore extreme events such as Aghakouchak et al. [2], Beniston et al. [6], El Adlouni et al. [15], Gumbel [20], Katz [23], Kharin et al. [25], Smith [35], Towler et al. [36], Villarini et al. [37] and Zhang et al. [39]. This methodology is often called the block maxima approach. An alternative method within EVT, called the POT approach, centers on assessing extreme values that surpass a defined high threshold, employing a generalized Pareto distribution. Smith [34] develops asymptotic methods for estimating the tail of a distribution using threshold exceedances, introducing a new estimator for the index of regular variation that often outperforms Hill's estimator, and demonstrating improved efficiency—especially within the domain of attraction of the Gumbel distribution—through maximum likelihood fitting of the generalized Pareto distribution. Both the annual maxima and POT techniques are extensively applied in the study of extreme climate phenomena, such as Davison et al.[12], Li et al. [28], and Villarini et al. [38] studies.
Consider a sequence $(T_i)_{i\geq 1}$ of independent and identically distributed random variables with a common distribution F. Let $\mathbf{M}_n$ represent the maximum of this sequence of n random variables, i.e., $\mathbf{M}_n = \max(T_1,\dots,T_n)$. According to the Fisher–Tippett theorem (Fisher and Tippett [17]; Gnedenko [19]), there exist sequences of constants $c_n > 0$ and d_n such that the normalized random variable $\mathbf{M}_n$ converges in distribution to a random variable with a non-degenerate distribution function G, i.e.

$$Pr\{(\mathbf{M}_n - d_n)/c_n \leq t\} \to G(t), \qquad as \quad n \to \infty. \tag{2}$$

We represent the cumulative distribution function of the GEV as follows Coles [11]

$$F_{\mu,\sigma,\xi}(s) = \exp\left\{-\left(1+\xi\left(\frac{s-\mu}{\sigma}\right)\right)^{-1/\xi}\right\}, \quad \left(1+\xi\left(\frac{s-\mu}{\sigma}\right)\right) > 0. \tag{3}$$

The GEV distribution is a versatile model for capturing extreme behaviors and is defined by three parameters: μ, σ, and ξ. The location parameter μ indicates the center of the distribution; the scale parameter σ measures the extent of deviations from μ while the shape parameter ξ affects the distribution's tail characteristics. Specifically, as ξ approaches 0, the distribution converges to the Gumbel distribution; when $\xi < 0$, it corresponds to the Weibull distribution, and for $\xi > 0$, it represents the Fréchet distribution.

Since extreme events are by nature rare, relying solely on the block maximum method can be inefficient, as it overlooks less extreme events that can still provide valuable insights into tail behavior. The peaks over threshold method is an alternative to the block maximum approach, which focuses on the asymptotic distribution of exceedances beyond a high fixed threshold. Balkema and de Haan's [4] result enables us to approximate the conditional distribution of exceedances above a specified high threshold. If there exist normalizing sequences $c_n > 0$ and d_n such that the expression (2) holds, meaning F is in the max-domain of attraction of an $F_{\mu,\sigma,\xi}$ distribution, then for a sufficiently high threshold v, we can model the limiting distribution of the exceedances $T - v|T > v$ using a GPD $G_{\hat{\sigma}_v,\xi}$,

$$Pr\{(T_n - d_n)/c_n > v + t|(T_n - d_n)/c_n > v\} \overrightarrow{n\to\infty} \begin{cases} (1+\xi t/\hat{\sigma}_v)_+^{-1/\xi}, & \text{if } \xi \neq 0 \\ \exp(-t/\hat{\sigma}_v), & \text{if} \xi = 0. \end{cases} \tag{4}$$

where $\hat{\sigma}_v = \sigma + \xi(v - \mu)$ and the shape parameter ξ corresponds to the relevant GEV distribution. The limiting distribution function $G_{\hat{\sigma}_v,\xi}$, in model (4) is defined on $\{t : t > 0\}$ and $(1+\xi t)/\hat{\sigma}_v$, and the case of $\xi = 0$ is interpreted as the limit. By a slight abuse of notation, we say that $T|T > v \sim GPD(v, \hat{\sigma}_v, \xi)$ whenever $T - v|T > v \sim GPD(\hat{\sigma}_v, \xi)$ i.e. $GPD(v, \hat{\sigma}_v, \xi)$ is a $GPD(\hat{\sigma}_v, \xi)$ shifted by the threshold v. We apply the asymptotic result (4) to support the use of the GPD as the model for exceedances $T_i - v$ where $T_i > v$ for a chosen high threshold v. The threshold is selected based on a bias–variance trade-off: a lower threshold results in a greater number of exceedances and reduces variance, but this increases bias since the asymptotic approximation for the tail of the distribution may be inaccurate. The GPD can be fitted using maximum likelihood estimation (Coles [11]) or other estimation techniques, as discussed in Beirlant et al. [5], section 5.3.

In general, our interest extends beyond merely identifying the maxima of observations; we are also concerned with the behavior of observations that significantly surpass a high threshold. One approach to extracting extremes from a sample of observations $(S_i)_{1\leq i\leq n}$ (where the distribution function is $F(s) = Pr(S_i \leq s)$) involves focusing on exceedances above a specified high threshold v. An exceedance occurs when $S_i > v$. Therefore, the excess over v is defined as $T_i = S_i - v$. This methodology is referred to as the POT approach.

In this famous approach, selecting the appropriate threshold is crucial as it involves a trade-off between variance and bias. The mean residual life plot and the parameter stability plot can help identify a suitable threshold. Specifically, these plots allow us to determine a range of thresholds where the empirical mean excess function is likely to exhibit a linear relationship with the threshold v. Therefore, the tail of the potential distribution F for $s > v$ is approximated as follows (recall that $\bar{F} = 1 - F$)

$$\bar{F}(s) = \varepsilon Pr\{S > s|S > v\} = \varepsilon\left(1 + \xi\frac{s-v}{\sigma}\right)^{-1/\xi}, \quad s > v, \tag{5}$$

where $\varepsilon = Pr\{S > v\}$. In practice, the exceedance probability $\bar{F}(s)$ provides valuable insight into potential risk. Its estimate can be derived using the extrapolation approach outlined in Equation (5). This involves approximating the tail probability of the Generalized Pareto (GP) model through maximum likelihood estimation of σ and ξ, based on the exceedances $s_i > v$ with $s_1 \geq \ldots \geq s_{n_v}$ that exceed the threshold v. The estimate of $\varepsilon = \bar{F}(v)$ can be calculated as n_v/n. When employing the POT method, selecting an appropriate threshold is crucial, as it involves a trade-off between variance and bias. The mean residual life plot and parameter stability plot can aid in identifying a suitable threshold by establishing a range where the empirical mean excess function is expected to vary linearly for the threshold v

$$e_n(v) = \frac{1}{n_v}\sum_{i=1}^{n_v}(s_i - v). \tag{6}$$

As previously noted, we will concentrate on the dependence structure of a bivariate vector $(S,T) = F(s,t)$ with marginal distributions F_S and F_T. We transform these marginal distributions into a common unit Fréchet distribution using the following approach. For appropriate thresholds v_s and v_t, the margins can be approximated as described in Equation (5), using the parameter sets $(\varepsilon_s, \sigma_s, \xi_s)$ and $(\varepsilon_t, \sigma_t, \xi_t)$.

Deanship of Research and Graduate Studies, Yarmouk University, Irbid, Jordan.

Let's define by

$$\hat{S} = -\left(\log\left\{1-\varepsilon_s\left[1+\frac{\xi_s(S-v_s)}{\sigma_s}\right]^{-1/\xi_s}\right\}\right)^{-1}, \tag{7}$$

and

$$\hat{T} = -\left(\log\left\{1-\varepsilon_t\left[1+\frac{\xi_t(T-v_t)}{\sigma_t}\right]^{-1/\xi_t}\right\}\right)^{-1}. \tag{8}$$

As a result, the transformed vector $(\hat{S},\hat{T}) = \hat{F}(\hat{s},\hat{t})$ has a marginal distribution function that is approximately a standard Fréchet distribution for $s > v_s$ and $t > v_t$, for mor details see Cai et al. [8].

2.2 The Multivariate Extreme Value Theory and Regression Manifolds

Now, we need to lay the groundwork for bivariate non-stationary extremes. In the statistical modeling of extreme value dependence, the spectral measure of a bivariate extreme value distribution is crucial. This paper introduces a model for the spectral measure that applies to any number of dimensions and allows for a generalization that includes mass on the simplex boundaries. To establish the foundation, let $\{(S_i,T_i)\}_{i=1}^n$ with $t \in [i-\frac{n}{2}, i+\frac{n}{2}]$ be a sequence of independent identically distributed random vectors with unit Fréchet marginal distributions, i.e. $\exp(-1/x)$, for $x > 0$. In our case, T_i is considered as a response, but S_i is considered as a covariate. Let the componentwise block maxima be $\mathbf{M}_n = (\mathbf{M}_{s,n},\mathbf{M}_{t,n})$ with $\mathbf{M}_{t,n} = \max\limits_{i=1,\ldots,n}\{T_i\}$ and $\mathbf{M}_{s,n} = \max\limits_{i=1,\ldots,n}\{S_i\}$. According to the componentwise maxima method proposed by Coles [11], the representation of the bivariate model becomes simple by assuming both margins have the standard Fréchet distribution. The vector of normalized componentwise maxima $\mathbf{M}_n/n$ converges in distribution to a random vector (S,T) which adheres to a bivariate extreme value distribution characterized by a joint distribution function G. This distribution is non-degenerate and takes the following form

$$G(s,t) = \exp\{-L(s,t)\}, \qquad (s,t) \in (0,\infty)^2,$$

with

$$L(s,t) = 2\int_0^1 \max\left(\frac{w}{s}, \frac{1-w}{t}\right) H(dw). \tag{9}$$

L is the exponent measure; for more information, it is preferable to see Coles [11], de Haan et al. [14] and Pickands [33]. Additionally, H is known as a spectral measure on $[0,1]$ which has the best key to control the dependence between the extreme values, and it's a probability measure and obeys the mean constraint

$$\int_0^1 wH(dw) = \frac{1}{2}. \tag{10}$$

If H is absolutely continuous with respect to the Lebesgue measure, its density can be represented as the Radon-Nikodym derivative $h = \frac{dH}{dw}$, where w is defined over the interval $[0,1]$.

Now, we characterize the regression manifold as the collection of regression lines,

$$\mathbf{R} = \{\mathbf{R}_\alpha : 0 < \alpha < 1\}, \quad \text{with} \quad \mathbf{R}_\alpha = \{t_{\alpha|s} : s \in (0,\infty)\}, \tag{11}$$

where

$$t_{\alpha|s} = \inf\{t > 0, \quad G_{T|S}(t|s) \geq \alpha\}, \tag{12}$$

is a conditional quantile of a bivariate extreme value distribution, with $\alpha \in (0,1)$ and $s \in (0,\infty)$. The concept of regression manifolds refers to the collection of conditional extreme quantiles $t_{\alpha|s}$, which describe how a covariate extreme $S = s$ influences the response variable T at a given extremal level $\alpha \in (0,1)$. More precisely, for each fixed value of s, the quantity $t_{\alpha|s}$ represents the α-quantile of the conditional distribution of $T|S = s$, derived under the multivariate extreme value framework. In general, $t_{\alpha|s}$ varies with α, capturing different levels of extremity. However, in special cases such as the logistic extreme value model—which assumes perfect dependence—the regression lines become independent of α. This behavior reflects the fact that, under perfect dependence, the conditional distribution becomes degenerate along a fixed line, resulting in quantile curves that collapse into a single straight line. Therefore, any claim of "no

variation with α" should be interpreted carefully and only in light of the underlying dependence structure and model assumptions.
Although the conditional quantile function $t_{\alpha|s}$ is defined via integration over the estimated spectral density $h(w)$, this formulation inherently involves uncertainty due to estimation variability in h, and the threshold selection process. In the current study, we do not provide a full sensitivity analysis or variance propagation framework for $t_{\alpha|s}$, which would require a Bayesian posterior sampling or a frequentist delta-method approach. However, to partially address this limitation, we performed a bootstrap-based quantile band estimation, which provides empirical uncertainty bounds around the estimated regression manifolds. A formal sensitivity analysis (e.g., via perturbations in the prior distribution of h or via Sobol indices) remains a direction for future research to improve interpretability and robustness.

We denote $G_{T|S}(t|s) = \mathbb{P}(T \leq t|S = s)$ the conditional distribution function of the response T given covariate $S = s$, derived from the joint bivariate extreme value distribution as defined by

$$G_{T|S}(t|s) = 2\exp\left\{-2\int_0^1 \max\left(\frac{w}{s}, \frac{1-w}{t}\right) h(w)dw + s^{-1}\right\} \int_{w(s,t)}^1 wh(w)dw. \tag{13}$$

We will now examine certain parametric instances of regression manifolds as recently defined.

Example 1(Logistic Model). An instance of the Logistic regression manifold is shown in Fig. 1. This follows from the Logistic bivariate extreme value distribution function given by

$$G(s,t) = \exp\{-(s^{-1/d} + t^{-1/d})^d\}, \qquad s,t > 0, \tag{14}$$

where $d \in (0,1]$ represents the strength of dependence between extremes. As d approaches 0, the dependence becomes stronger, with $d \to 0$ indicating perfect dependence. The conditional distribution of T given S is

$$G_{T|S}(t|s) = G(s,t)(s^{-1/d} + t^{-1/d})^{d-1} s^{1-1/d} \exp(1/s), \qquad s,t > 0, \tag{15}$$

For regression manifolds representation, we use the approximation of the exact Logistic regression manifolds defined by Carvalho et al. [13] as follows for $\alpha \in (0,1)$ and $s >> 1$

$$\tilde{t}_{\alpha|s} = \frac{d}{1-d}\{\alpha^{1/(d-1)} - 1\}^{-d-1}\{\alpha^{d/(1-d)} - 1\}\alpha^{1/(d-1)} + \{\alpha^{-1/(1-d)} - 1\}^{-d} s. \tag{16}$$

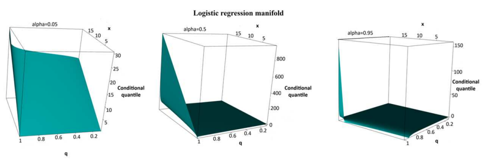

Fig. 1: Regression manifold for bivariate Logistic

Example 2(Husler-Reiss Model). An instance of Husler-Reiss regression manifold is shown in Fig. 2. It arises from the Husler-Reiss bivariate extreme value distribution function, which has the following form

$$G(s,t) = \exp\left\{-s^{-1}\Phi\left(a + \frac{1}{2a}\log\frac{t}{s}\right) - t^{-1}\Phi\left(a + \frac{1}{2a}\log\frac{s}{t}\right)\right\}, \qquad s,t > 0, \tag{17}$$

where Φ denotes the standard normal distribution function and $a \in (0,\infty]$ is a parameter that modulates the dependency between extremes: as $a \to 0$, the dependence becomes perfect, while in the limit as $a \to \infty$, the dependency reaches complete independence. The collection of regression lines $\mathbf{R}_\alpha$ for this model lacks explicit forms and is derived using (12) with

$$G_{T|S}(t|s) = \left[\Phi\left(a + \frac{1}{2a}\log\frac{t}{s}\right) + \frac{1}{2a}\phi\left(a + \frac{1}{2a}\log\frac{t}{s}\right) - \frac{st^{-1}}{2a}\phi\left(a + \frac{1}{2a}\log\frac{s}{t}\right)\right] \times G(s,t)\exp(1/s), \tag{18}$$

for $s,t > 0$ where ϕ is the standard Normal density function.

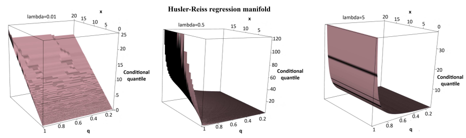

Fig. 2: Regression manifold for bivariate Husler-Reiss

Example 3(Coles-Tawn Model). An instance of Coles-Tawn regression manifold is shown in Fig. 3 (bottom). This idea is derived from the Coles-Tawn bivariate extreme value distribution function, which is defined as follows

$$G(s,t) = \exp[-s^{-1}\{1 - Be(q;a+1,b)\} - t^{-1}Be(q;a,b+1)], \quad s,t > 0, \tag{19}$$

where $Be(q;\gamma,\beta)$ represents the cumulative distribution function of a *Beta* distribution with parameters $\gamma,\beta > 0$. In this context, q is defined as $q = at^{-1}/(at^{-1}+bs^{-1})$ with $a,b > 0$ serving as parameters that control the dependence between extremes. The scenario where $a = b = 0$ indicates complete independence, while the case of $a = b \to \infty$ signifies perfect dependence. For fixed values of $a(b)$, an increase in $b(a)$ leads to a stronger dependence. The family of regression lines $\mathbf{R}_\alpha$ for this model does not have a straightforward representation and is determined through calculations using (12) for $s,t > 0$, with

$$G_{T|S}(t|s) = \left[1 - Be(q;a+1,b) + \frac{(a+1)b}{\zeta} be(q;a+2,b+1) - \frac{s}{t}\frac{a(b+1)}{\zeta} be(q;a+1,b+2)\right] G(s,t)\exp(1/s), \tag{20}$$

where $be(q;\gamma,\beta)$ is the density function of the Beta distribution with parameters $\gamma,\beta > 0$ and $\zeta = (a+b+2)(a+b+1)$.

It is preferable to see de Carvalho et al. [13].

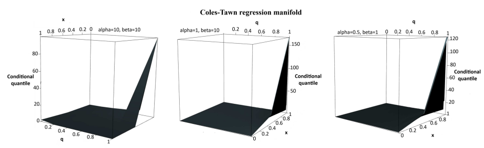

Fig. 3: Regression manifold for bivariate Coles-Tawn.

To illustrate the construction and reliability of regression manifolds, we provide a comparative analysis of three dependence models—Coles–Tawn, Logistic, and Husler–Reiss—based on bootstrap simulations. The regression manifold is defined as the set of conditional extreme quantiles $t_{\alpha|s}$, which describe the evolution of the response variable under varying levels of extremal dependence. Figures 1, 2, and 3 display the smooth variation of $t_{\alpha|s}$ across quantile levels, while Table 1 summarizes the bootstrap estimates, including mean values, standard deviations, and 95% confidence intervals.

Model	Mean	SD	95% Confidence Interval
Coles–Tawn	15.09	0.54	[14.04 ; 16.10]
Logistic	14.97	0.53	[13.96 ; 16.02]
Hüsler–Reiss	15.05	0.51	[14.03 ; 16.00]

Table 1: Bootstrap-based estimates of conditional quantiles obtained from 1,000 boots and 7,000 samples per model.

These results demonstrate the stability and consistency of the regression manifold across models. Notably, the Logistic model, which previously showed signs of overestimation, now aligns closely with the Coles–Tawn and Husler–Reiss models, indicating robust calibration. Confidence bands derived from bootstrap replicates are used to quantify uncertainty around the manifold, offering a transparent view of model variability. Before defining our new angular density model, we cite some properties of the Logistic-Normal distribution.

2.3 The Logistic-Normal Distribution

Aitchison [3] proposed a technique to approximate a Dirichlet distribution with a Logistic-Normal distribution to minimize their Kullback-Leibler divergence (KL), which is minimized

$$K(dir, h_{\mu,\Sigma}) = \int_{\Delta_d} dir(s|\beta) \log\left(\frac{dir(s|\beta)}{h_{\mu,\Sigma}(s)}\right) dx, \tag{21}$$

with dir is the density function of the Dirichlet distribution and $h_{\mu,\Sigma}$ is the density function of the Logistic-Normal distribution. This K is minimized by $\mu_i = \delta(\beta_i) - \delta(\beta_{d+1})$, $\sigma_{ii} = \varepsilon(\beta_i) + \varepsilon(\beta_{d+1})$ for $(i = 1,...,d)$, $\sigma_{ij} = \varepsilon(\beta_{d+1})$ for $(i \neq j)$ where $\delta(s) = \Gamma'(s)/\Gamma(s)$, $\varepsilon(s) = \delta'(s)$, and $\Delta_d = \{u \in P^d : w_1 + ... + w_d < 1\}$.

Applying the logistic transformation to a normal distribution of dimension d results in a distribution on the d-dimensional simplex, which can aptly be termed a Logistic-Normal distribution.
The Logistic-Normal distribution offers greater flexibility compared to the Dirichlet distribution, as it can account for correlations between components of probability vectors. This makes it a valuable tool for simplifying statistical analyses of compositional data, as it enables researchers to examine log ratios of data vector components, which are often of interest in comparison to absolute component values.
The probability simplex, being a bounded space, poses challenges for applying standard techniques typically used for vectors in $\mathbb{R}^n$. Aitchison highlighted the issue of spurious negative correlations when these methods are directly applied to simplicial vectors. However, transforming compositional data in the probability simplex $\mathscr{S}^D$ through the inverse of the additive logistic transformation results in real-valued data in $\mathbb{R}^{D-1}$, where standard techniques can be effectively employed. This method supports the application of the Logistic-Normal distribution, often referred to as the "Gaussian of the simplex."

The Logistic-Normal distribution extends the logit-normal distribution to D-dimensional probability vectors by applying a logistic transformation to a multivariate normal distribution.
Let $\mathbb{R}^d$ represent a d-dimensional real space, P^d denote the positive orthant of $\mathbb{R}^d$ and Δ_d indicate the d-dimensional positive simplex defined by $\Delta_d = \{u \in P^d : w_1 + ... + w_d < 1\}$.

For any d-dimensional vector u and any real-valued function g, let $g(w)$ signify the d-dimensional vector where the ith component is given by $g(w_i)(i = 1,...,d)$. Assuming v follows a multinormal distribution $\mathbf{N}_d(\mu, \Sigma)$ over $\mathbb{R}^d$. The exponential transformation from $\mathbb{R}^d$ to P^d, expressed as $w = e^v$, alongside its inverse logarithmic transformation $v = \log w$, serves as the standard method for establishing a corresponding log-normal distribution, where w distributed as $\Lambda_d(\mu, \Sigma)$. Similarly, a logistic transformation from $\mathbb{R}^d$ to Δ_d, along with its inverse log-ratio transformation

$$w = e^v / \left(1 + \sum_{j=1}^{d} e^{v_j}\right), \qquad v = \log(w/w_{d+1}), \tag{22}$$

where

$$w_{d+1} = 1 - \sum_{j=1}^{d} w_j, \tag{23}$$

can be used to define a Logistic-Normal distribution over Δ_d, and we can denote that w follows this distribution $L_d(\mu, \Sigma)$. The density function of $L_d(\mu, \Sigma)$ is then for $w \in \Delta_d$

$$h_{\mu,\Sigma}(w) = |2\pi\Sigma|^{-\frac{1}{2}} \left(\prod_{j=1}^{d+1} w_j\right)^{-1} \exp[-\frac{1}{2}\{\log(w/w_{d+1}) - \mu\}^T \Sigma^{-1} \{\log(w/w_{d+1}) - \mu\}]. \tag{24}$$

Even though all positive order moments $\mathbb{E}(w_j^a)$ $(a > 0)$ and the geometric moment $\exp\{\mathbb{E}(\log w_j)\}$ exists, but their integral representations cannot be simplified into a straightforward form. However, this is not a significant drawback since practical analysis often focuses more on the ratios w_j/w_k or their logarithms. Based on normal-logarithmic theory, let σ_{jk} represent the (j,k)th element of the covariance matrix Σ. Additionally, we adopt the convention that $\mu_{d+1} = 0$ and $\sigma_{j,d+1} = 0$ for $(j = 1,...,d+1)$. Consequently, we have that

$$\mathbb{E}\{\log(w_j/w_k)\} = \mu_j - \mu_k, \quad \mathrm{cov}\{\log(w_j/w_k), \log(w_l/w_m)\} = \sigma_{jl} + \sigma_{km} - \sigma_{jm} - \sigma_{kl}, \tag{25}$$

$$\mathbb{E}(w_j/w_k) = \exp\{\mu_j - \mu_k + \frac{1}{2}(\sigma_{jj} - 2\sigma_{jk} + \sigma_{kk})\}, \tag{26}$$

$$\mathrm{cov}(w_j/w_k, w_l/w_m) = \mathbb{E}(w_j/w_k)\mathbb{E}(w_l/w_m)\{\exp(\sigma_{jl} + \sigma_{km} - \sigma_{jm} - \sigma_{kl}) - 1\}. \tag{27}$$

2.4 Main Result

Before introducing our modeling framework, we clarify the key concepts involved. In the context of multivariate extremes, the regression manifold refers to the collection of conditional quantiles $t_{\alpha|s}$, which describe how an extreme response variable depends on an extreme covariate. This generalizes classical quantile regression to the tail of the distribution, where both variables lie in the domain of attraction of an extreme value distribution.

To construct such regression manifolds, we rely on a spectral representation of multivariate extreme value distributions. The dependence structure between extremes is captured via a spectral density $h(w)$, where $w \in (0,1)$ parametrizes the contribution of each variable to the extremal event. In classical approaches, this density is either fixed or modeled with a Dirichlet prior. In contrast, we adopt a Logistic-Normal prior, which offers greater flexibility by allowing for smooth, continuous modeling of $h(w)$ while satisfying the standard constraint $\int_0^1 wh(w)dw = \frac{1}{2}$. This modeling choice allows us to define a family of regression manifolds $t_{\alpha|s}$ that vary with the quantile level α, except in the limiting case of perfect dependence.
In this section, we explore the process of understanding regression manifolds through data analysis. In this context, de Carvalho et al. [13] introduced a prior related to the space of spectral measures put forth by Hanson et al. [22]. The model is a Bernstein polynomial, and its expression contains the Dirichlet distribution. In addition, Aitchison [3] proposed the famous relationship between the Dirichlet distribution and Logistic-Normal distribution and proved the flexibility of this last. So, we introduce our new model.

Consider a sequence $\{(S_i,T_i)\}_{i=1}^n$ of independent random vectors with unit Fréchet marginal distributions. Define $R_i = T_i + S_i$ and $W_i = (S_i,T_i)/R_i$ as the pseudo-angular decomposition of the observations. De Haan and Resnick [14] demonstrated that the convergence of normalized componentwise maxima to G is equivalent to a specific weak convergence of measures, and they generalized this result for time series

$$\mathbb{P}(W \in .|R > v) \xrightarrow{d} H(.), \qquad \text{as } v \to \infty.$$

This implies that when the radius R is large enough, the pseudo-angles W become almost independent of R and resemble a distribution linked to the spectral measure H. Therefore, to learn about $\mathbf{R}_\alpha$, we initially need to investigate H by analyzing $t = |\{W_i : R_i > v, i = 1,...,n\}|$ which counts the number of exceedances above a high threshold v, where $t = o(n)$. We will outline the methodology for this analysis next.

In this paper, we model the spectral density h via the spectral measure via the Logistic-Normal distribution for obtaining a more admirable and more important result, which can thus be considered as the ”simplex Gaussian”. Let's find the right simplex to define our fabulous spectral density h.

Firstly, we know the Logistic-Normal distribution is defined on the d-dimensional positive simplex defined by $\Delta_d = \{w \in P^d : w_1 + ... + w_d < 1\}$. We add at this simplex $w_{d+1} = 1 - \sum_{i=1}^d w_i$. We obtain $\Delta_{d+1} = \{w \in P^d : w_1 + ... + w_d + w_{d+1} = 1\}$. So, we consider the simplex $S_d = \{(w_1,...,w_d) \in [0,1]^d, \sum_{i=1}^d w_i = 1\} \subset \mathbb{R}^d$. We can work for that because the $\{w_i\}_{i=1}^d$ and w_{d+1} are linear. So, in our case, we take $d = 1$.

For modeling our spectral density h via the Logistic-Normal distribution, we must first estimate μ and σ, the Logistic-Normal parameters. Therefore, the next theorem presents a good result for estimating the parameter μ.

Theorem 1. *We consider the Logistic-Normal density $h_{\mu,\sigma}$ defined on the interval $(0,1)$ by*

$$h_{\mu,\sigma}(u) = \frac{\exp\left(-\frac{\log^2\left(\frac{u}{1-u}-\mu\right)}{2\sigma^2}\right)}{\sigma\sqrt{2\pi}u(1-u)}, \qquad for\ u \in (0,1), \quad \mu \in \mathbb{R} \quad and \quad \sigma > 0.$$

Then

$$\int_0^1 uh_{\mu,\sigma}(u)dw = \frac{1}{2} \Leftrightarrow \mu = 0.$$

Proof. We consider $\mathbb{E}(u) = \int_0^1 th_{\mu,\sigma}(t) = \frac{1}{2}$ and $\mathscr{Z} \sim \mathscr{N}(0,1)$.

So

$$\mathbb{E}\left(\frac{e^{\sigma\mathscr{Z}+\mu}}{1+e^{\sigma\mathscr{Z}+\mu}}\right) = \frac{1}{2} = \mathbb{E}\left(\frac{e^{\sigma\mathscr{Z}}}{1+e^{\sigma\mathscr{Z}}}\right).$$

We know $\mathscr{Z} \stackrel{\mathscr{L}}{=} -\mathscr{Z}$. Then

$$\mathbb{E}\left(\frac{e^{-\sigma\mathscr{Z}+\mu}}{1+e^{-\sigma\mathscr{Z}+\mu}}\right) = \mathbb{E}\left(\frac{1}{1+e^{\sigma\mathscr{Z}}}\right).$$

Thus

$$\mathbb{E}\left(\frac{e^{\sigma\mathscr{Z}+\mu}}{1+e^{\sigma\mathscr{Z}+\mu}}\right) = \mathbb{E}\left(\frac{1}{1+e^{\sigma\mathscr{Z}}}\right)$$
$$\Leftrightarrow \mathbb{E}\left(\frac{1}{1+e^{\sigma\mathscr{Z}+\mu}}\right) - \mathbb{E}\left(\frac{1}{1+e^{\sigma\mathscr{Z}}}\right) = 0$$
$$\Leftrightarrow \mathbb{E}\left(\frac{e^{\sigma\mathscr{Z}}(1-e^{\mu})}{(1+e^{\sigma\mathscr{Z}+\mu})(1+e^{\sigma\mathscr{Z}})}\right) = 0$$
$$\Leftrightarrow (1-e^{\mu})\mathbb{E}\left(\frac{e^{\sigma\mathscr{Z}}}{(1+e^{\sigma\mathscr{Z}+\mu})(1+e^{\sigma\mathscr{Z}})}\right) = 0$$
$$\Leftrightarrow \mu = 0$$

because $\frac{e^{\sigma\mathscr{Z}}}{(1+e^{\sigma\mathscr{Z}+\mu})(1+e^{\sigma\mathscr{Z}})} > 0$ almost surely.

If $\mu = 0$, we solve the integral $\int_0^1 wh_{0,\sigma}(w)dw$. After all calculation, we find that the integral is $1/2$ hence the result.

Now, we model the spectral measure density h by the expression

$$h(w) = h_{0,\sigma}(w), \qquad w \in [0,1]. \tag{28}$$

The graphical representation of $h(w)$ provides valuable insight into its behavior; hence, referring to Figure 6 is recommended for enhanced interpretability. So, we obtain

$$G(s,t) = \exp\left(\frac{-2}{s}\int_{w(s,t)}^{1} wh_{0,\sigma}dw - \frac{2}{t}\int_{w(t,s)}^{1} wh_{0,\sigma}dw\right), \qquad s,t > 0. \tag{29}$$

Note the parameter σ is estimated by obeying at the moment constraint $\int_0^1 wH(dw) = \frac{1}{2}$. We use the logistic transformation $w = e^v/(1+e^v)$ and $v = \log(w/1-w)$ with v follows a Normal distribution $\mathscr{N}(0,\sigma)$.
We computed the estimator using two approaches: the classical maximum likelihood method and the approach introduced by Stephenson and Tawn [1]. Both methods lead to the same likelihood expression, given by:

$$\log L(s,t) = \sum_{i=1}^{n}\log\left[\frac{2}{s_i^2t_i^2\sigma^2\pi}\int_{w(s,t)}^{1}\frac{\exp\left(-\frac{\log^2\left(\frac{w}{1-w}\right)}{2\sigma^2}\right)}{1-w}dw\int_{w(t,s)}^{1}\frac{\exp\left(-\frac{\log^2\left(\frac{w}{1-w}\right)}{2\sigma^2}\right)}{1-w}dw + \sqrt{\frac{2}{\pi}}\times\frac{1}{\sigma(s_i+t_i)s_it_i}\exp\left(-\frac{\log^2\left(\frac{s_i}{t_i}\right)}{2\sigma^2}\right)\right]$$
$$-\frac{2}{s_i\sigma\sqrt{2\pi}}\int_{w(s,t)}^{1}\frac{\exp\left(-\frac{\log^2\left(\frac{w}{1-w}\right)}{2\sigma^2}\right)}{1-w}dw - \frac{2}{t_i\sigma\sqrt{2\pi}}\int_{w(t,s)}^{1}\frac{\exp\left(-\frac{\log^2\left(\frac{w}{1-w}\right)}{2\sigma^2}\right)}{1-w}dw. \tag{30}$$

and we maximize the function to obtain a new estimator $\hat{\sigma}$. The estimation of the scale parameter σ is based on the maximization of a log-likelihood function that includes several terms involving integrals over the joint distribution of the observed variables. These integrals are evaluated numerically using deterministic adaptive quadrature (via R's integrate function), which ensures stable and accurate approximation of the involved expressions without relying on Monte Carlo sampling. The optimization is performed using a bounded one-dimensional search algorithm—Brent's method—implemented via R's optimize function. This method is well-suited for smooth and unimodal objective functions. The search interval for σ is chosen to ensure numerical stability and to reflect plausible values based on the data. From a computational standpoint, the main challenge lies in repeatedly evaluating the integrals at each iteration of the optimization. To address this, we rely on efficient adaptive quadrature, which ensures that the estimation of σ remains tractable and reproducible, while avoiding the need for stochastic integration methods such as Monte Carlo. The result is a robust and efficient estimation procedure that aligns with the theoretical framework of non-stationary extreme value analysis.

Now, to incorporate a prior into the space of regression manifolds, we substitute the spectral density

$$h(w) = h_{0,\hat{\sigma}}(w), \qquad w \in [0,1], \tag{31}$$

into (13) with $w(s,t) = s/(s+t)$ for $s,t > 0$.

This prior induces a prior on the space of regression lines $\mathbf{R}_\alpha = \{t_{\alpha|s} : s \in (0,\infty)\}$, where $t_{\alpha|s}$ is a solution to equation, $G_{T|S}(t|s) = \alpha$, for $\alpha \in (0,1)$, where

$$G_{T|S}(t|s) = 2\exp\left\{-2\int_0^1 \max\left(\frac{w}{s}, \frac{1-w}{t}\right) h_{0,\hat{\sigma}}(w)dw + s^{-1}\right\} \int_{w(s,t)}^1 wh_{0,\hat{\sigma}}(w)dw, \tag{32}$$

with $w(s,t) = s/(s+t)$ for $s,t > 0$.
Motivated by the limitations of existing priors, this work adopts a Logistic-Normal prior for spectral density estimation in the modeling of multivariate extremes. While prior studies—such as Hanson et al. [22] and de Carvalho et al. [13]—commonly use Dirichlet distributions, our approach builds on the Logistic-Normal formulation of Aitchison [3], which provides greater flexibility and richer parametrization. Unlike the Dirichlet, which imposes independence among components, the Logistic-Normal prior captures complex dependence structures between extremes, such as temperature and precipitation. As demonstrated in El Bernoussi and El Arrouchi [16], this choice allows for the construction of smooth regression manifolds and enhances the interpretability of conditional quantile estimates, making it particularly suited to environmental applications. Moreover, this framework is embedded within a POT-ARIMA pipeline, ensuring both stationarity and relevance in extreme value contexts.

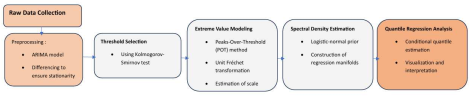

Fig. 4: Flowchart of the Complete Modeling Pipeline: From Preprocessing to Regression Manifold Estimation

. The flowchart 4 outlines a comprehensive pipeline for analyzing extreme events using advanced statistical techniques. Beginning with raw data collection, the process incorporates ARIMA-based preprocessing to ensure stationarity, followed by threshold selection via the Kolmogorov-Smirnov test. The extreme value modeling stage applies the Peaks-Over-Threshold method and transforms data using the unit Fréchet distribution to estimate scale parameters. Spectral density estimation is then performed using a Logistic-Normal prior to construct regression manifolds. Finally, conditional quantile regression enables precise estimation and visualization of tail behavior, offering valuable insights for risk assessment and predictive modeling.

3 Simulation Study and Results

In this section, we will apply our approach to the raw data of extreme precipitation as a variable response and temperature as a covariates from the local Meknes-Morocco during the period 1973-2024. The estimate of the spectral density in figure 6 is almost symmetric, suggesting that a simple parametric model would perform well here, such as a famous ARIMA method, which, after data collection, we will test the stationarity using this method and then remove it in the case of non-stationarity of the data. The algorithmic steps are detailed in Algorithm 1, which outlines the simulation procedure used for both precipitation and temperature time series.

3.1 Preprocessing

For both precipitation and temperature time series, we performed ARIMA modeling using the auto.arima function from the forecast package in R. Additionally, ARIMA models were fitted using XLSTAT software to cross-validate the results and ensure robustness across different platforms. The optimal models selected were ARIMA(1,0,0) with a non-zero mean, indicating that both series exhibit stationary behavior without significant temporal autocorrelation or trend. For temperature, the estimated mean is 41.42 with a standard error of 0.53, and residual variance 3.72. The model fit yields AIC = 52.77 and BIC = 56.20, confirming the adequacy of this simple model. Similarly, the precipitation series resulted in a mean of 40.74 (standard error 3.36) with residual variance 440, AIC = 342.13, and BIC = 345.65. These results justify the stationarity assumption and validate our preprocessing steps before the application of extreme value modeling. Figure 5 presents a comprehensive diagnostic analysis of the temperature and precipitation time series prior to the application of extreme value modeling. The top row illustrates the ARIMA model fit for both maximum temperature and precipitation, including point forecasts and 95% confidence intervals. These plots confirm the presence of temporal structure and justify the use of ARIMA to remove trends and autocorrelation, thereby ensuring stationarity—a prerequisite for subsequent extreme value analysis.

The middle row displays the residuals from the fitted ARIMA models. The absence of visible patterns or autocorrelation in the residuals suggests that the models adequately capture the underlying dynamics of the series. This supports the validity of the differencing and modeling steps applied during preprocessing.

The bottom row contains Q-Q plots comparing the residuals to a Normal distribution. The alignment of points along the diagonal indicates that the residuals are approximately normally distributed, further confirming the adequacy of the ARIMA fit and the appropriateness of the transformation for statistical inference.

Overall, this figure validates the preprocessing pipeline, demonstrating that the ARIMA modeling step successfully stabilizes the series and prepares the data for threshold selection and transformation to unit Fréchet margins in the subsequent POT framework.

3.2 Results

A major limitation of using the component-wise maximum for modeling multivariate extremes is that it typically leads to a mismatch in the timing of the component-wise maximum vectors within blocks. More specifically, likely, the annual maximum temperature and annual maximum precipitation in the data application (and, in fact, more generally in most places on Earth) occur at different times. To this end, we work with the POT method, which requires the threshold to be determined.

To do this, we plot the function of the empirical means to determine the correct threshold. However, thresholds that are easily determined on the Mean Excess plot generally have 10% extreme values above these thresholds. This choice is based on the results obtained from the Mean Excess plot and the proposals made by Gavin [18] and Neftci [32]. These authors recommend choosing a maximum of 10% of the tails of the distribution to form the extreme values. We recall that, as indicated by Longin [29] and Gavin [18], the choice of threshold may be subjective but must necessarily respect the balance between variance and bias to allow convergence of the GPD estimator. This is why we test the conformity of the empirical distribution of extreme observations and the theoretical distribution using different methods: QQ-Plot, Kolmogorov-Smirnov, and the GPD test. For our study, the Kolmogorov-Smirnov test provided us with good results: we retained 10% of the tails of the precipitation and temperature distributions to form the extreme returns.

Next, we fit the GPD model using the scale and shape estimators obtained by the maximum likelihood method (see Coles [11] and Beirlant et al. [5]). Thus, we convert our data into unit Fréchet margins via equations (7) and (8).

Series	Test	Statistic	p-value	Conclusion
Temperature	Augmented Dickey-Fuller (ADF)	-1.1632	0.9043	Non-stationary (fail to reject H_0)
	KPSS	0.49453	0.0429	Non-stationary (reject H_0 of stationarity)
	Phillips-Perron (PP)	-38.311	0.0100	Stationary (reject H_0 of unit root)
Precipitation	Augmented Dickey-Fuller (ADF)	-3.3704	0.07013	Likely non-stationary (borderline)
	KPSS	0.47791	0.04664	Non-stationary (reject H_0 of stationarity)
	Phillips-Perron (PP)	-30.262	0.0100	Stationary (reject H_0 of unit root)
POT Threshold	Kolmogorov-Smirnov (K-S)	$D = 0.20952$	0.9087	Threshold appropriate (fail to reject H_0)
Dependence	Pearson's Chi-squared (χ^2)	$\chi^2 = 16.452$ (df = 5)	0.005666	Dependent variables (reject H_0 of independence)

Table 2: Summary of Stationarity and Dependence Tests

The results of the Augmented Dickey-Fuller (ADF) test and the KPSS test for the time series of temperature indicate conflicting conclusions regarding its stationarity. The ADF test, with a statistic of -1.1632 and a p-value of 0.9043, suggests failing to reject the null hypothesis, implying that the series is likely non-stationary. Conversely, the KPSS test shows a statistic of 0.49453 with a p-value of 0.0429, leading to the rejection of its null hypothesis and indicating non-stationarity as well. This discrepancy arises from the differing null hypotheses of the two tests: ADF assumes non-stationarity while KPSS assumes stationarity. To resolve this, it is recommended to conduct a visual analysis of the time

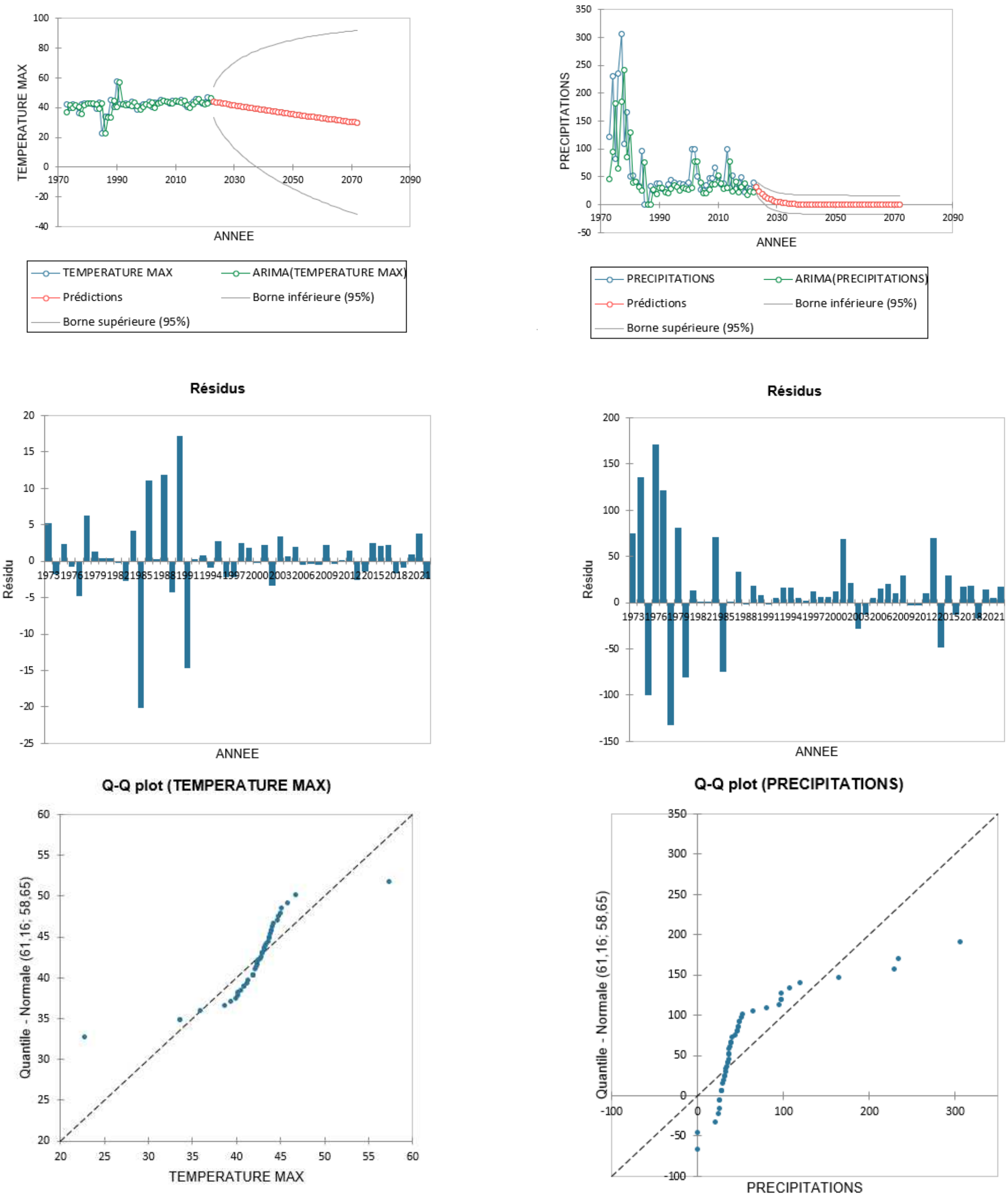

Fig. 5: ARIMA-Based Preprocessing and Diagnostic Assessment of Temperature and Precipitation Time Series

series for trends or seasonality, differencing the series, and performing additional tests like the Phillips-Perron test, as well as examining potential transformation methods to achieve stationarity for effective modeling. The results of the Phillips-Perron unit root test for the time series of temperature add further insight into its stationarity. The test yielded a Dickey-Fuller Z(alpha) statistic of -38.311 with a p-value of 0.01, which indicates strong evidence against the null hypothesis of a unit root, suggesting that the series is likely stationary. This finding contrasts with the results of the ADF test, which had a statistic of -1.1632 and a p-value of 0.9043, implying that the series is non-stationary. Similarly, the KPSS test showed a statistic of 0.49453 with a p-value of 0.0429, leading to the rejection of its null hypothesis and also indicating non-stationarity. The conflicting conclusions arise from the different null hypotheses of these tests: the ADF test assumes non-stationarity, while the KPSS test assumes stationarity. Given these mixed results, it is advisable to

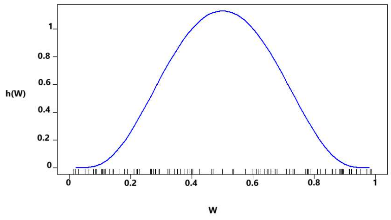

Fig. 6: Spectral density estimate with 90% credible band along with a rug of pseudo-angles

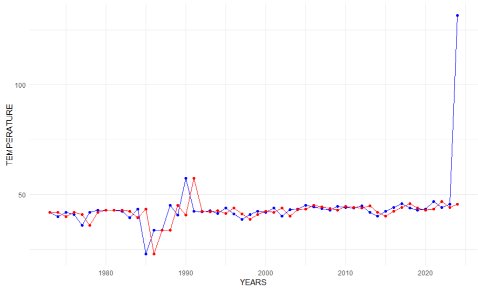

Fig. 7: True extreme raw temperature data (blue) and adjusted with the ARIMA method (red) from the local Meknes-Morocco.

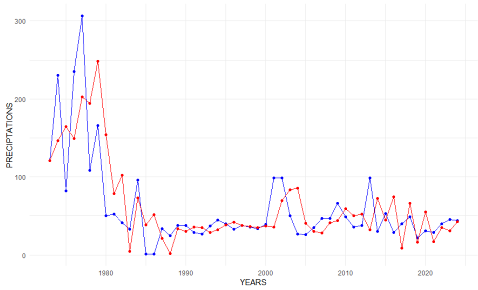

Fig. 8: True extreme raw precipitation data (blue) and adjusted with the ARIMA method (red) from the local Meknes-Morocco.

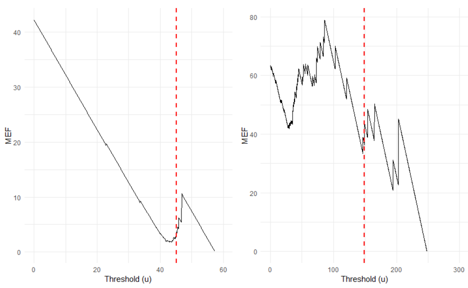

Fig. 9: Mean Excess plot for temperature (left) and precipitations(right) with the threeshold which returned 10% for the tails.

conduct a visual analysis of the time series to identify any trends or seasonal patterns, consider differencing the series,

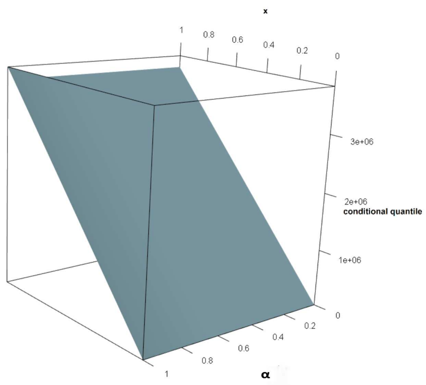

Fig. 10: Regression Manifold for Precipitation Based on Temperature on the Original Scale.

and explore transformations to achieve stationarity. Additionally, relying on the Phillips-Perron test results, which suggest stationarity, may guide further modeling efforts effectively. The figure 7 also shows the same result, with the blue curve representing the raw data and the red one representing their adjustment. The results of the ADF test and the KPSS test for the time series of precipitation suggest its likely non-stationarity. The ADF test yielded a Dickey-Fuller statistic of -3.3704 with a p-value of 0.07013, indicating that we fail to reject the null hypothesis of a unit root, suggesting the series is non-stationary, albeit close to the significance threshold. Conversely, the KPSS test revealed a statistic of 0.47791 and a p-value of 0.04664, which leads us to reject its null hypothesis of stationarity, further indicating non-stationarity. These results illustrate a slight discrepancy, with the ADF test suggesting the series is marginally non-stationary, while the KPSS test provides stronger evidence against stationarity. Visual analysis, where the blue curve represents the raw data and the red curve indicates adjustments, can help reveal underlying trends or seasonal patterns. To address the non-stationarity, it is recommended to consider differencing the series, apply potential transformations, and conduct further testing to clarify the stationarity characteristics before proceeding with modeling. The results of the Phillips-Perron unit root test for the time series of precipitation provide additional clarity regarding its stationarity. The test yielded a Dickey-Fuller Z(alpha) statistic of -30.262 with a p-value of 0.01, indicating strong evidence against the null hypothesis of a unit root, suggesting that the series is likely stationary. This finding contrasts with the earlier results from the ADF test, which showed a statistic of -3.3704 and a p-value of 0.07013, implying that the series is non-stationary, albeit close to the significance threshold. Similarly, the KPSS test indicated non-stationarity with a statistic of 0.47791 and a p-value of 0.04664, leading to the rejection of its null hypothesis of stationarity. The conflicting conclusions from the ADF and KPSS tests highlight the complexity of assessing stationarity, while the Phillips-Perron test strongly supports the idea of stationarity. Given these mixed results, it is advisable to conduct a thorough visual analysis of the series to identify trends or seasonal patterns, consider differencing or transforming the data, and perform additional tests to ensure a robust understanding of the stationarity characteristics before proceeding with any modeling efforts. To this end, the figure 8 shows the stationarity of the time series. To select a threshold for the peak threshold method, we use the Kolmogorov-Smirnov (K-S) test. In this context, the K-S test produced a statistic $D = 0.20952$ with a p-value of 0.9087. This high p-value indicates that we cannot reject the null hypothesis that the distribution of our data is similar to the specified theoretical distribution, which means that the threshold proposed by the peak threshold method is appropriate for modelling the data. These three unit root tests rely on different null hypotheses: the ADF and PP tests test the null hypothesis of a unit root (non-stationarity), and the KPSS test assumes stationarity under the null.

While ADF and KPSS sometimes yield conflicting results, the PP test, known for its robustness to heteroskedasticity and serial correlation, confirmed the absence of a unit root. Based on this, and to ensure consistency in preprocessing, we applied ARIMA-based detrending for both series. Even though the PP test indicated stationarity, we opted to apply ARIMA to both series to ensure uniform treatment, eliminate residual trends, and facilitate comparability across extreme modeling steps. Figure 9 displays the Mean Excess Function (MEF) plots for temperature (left) and precipitation (right), which are instrumental in selecting appropriate thresholds for the Peaks-Over-Threshold (POT) method. The MEF, defined as the average excess over a threshold u, is expected to exhibit a linear trend beyond a suitable threshold if the data follow a Generalized Pareto Distribution (GPD).

In the temperature plot, the MEF shows a relatively stable behavior with a slight upward trend near the selected threshold, suggesting that the tail behavior is consistent with the GPD assumption. This supports the choice of threshold corresponding to the upper 10% of the data, as indicated by the red dashed line.

Conversely, the precipitation plot exhibits more erratic fluctuations, which may reflect higher variability or the presence of multiple regimes in the tail behavior. Despite this, the selected threshold still aligns with the empirical recommendation of retaining approximately 10% of the extreme values, balancing bias and variance in the estimation process.

These plots provide visual validation for the threshold selection and confirm the suitability of the POT framework for modeling extremes in both temperature and precipitation series. The chosen thresholds will be used to extract exceedances and fit the GPD, forming the basis for subsequent scale parameter estimation and regression manifold construction. Also, the associated figure 9 illustrates the stationarity of the series, reinforcing the idea that the threshold chosen is relevant to the analysis, and allows significant variations in the behaviour of the data to be captured without introducing non-stationary bias. This approach is useful for identifying and modelling peak structures in time series, ensuring better management of extreme phenomena.

As we can see from the figure 10, regression lines do not depend on α. This shows that the extremes are perfectly dependent, so the spectral measure H assigns all its mass to the barycenter of the simplex. Based on the results of the Pearson's Chi-squared test (χ^2), we can conclude that there is a statistically significant relationship between the variables in your contingency table. The test yielded a chi-squared statistic of 16.452 with 5 degrees of freedom, and a p-value of 0.005666. Since the p-value is less than the common significance level of 0.05, we reject the null hypothesis of independence. This indicates that the variables are not independent and there is an association between them. This finding is crucial for understanding the underlying patterns in your data and can guide further analysis and decision-making. This proves the same result obtained by our study with our new quantile regression technique. All simulations and statistical analyses were conducted using the R programming language, complemented by XLSTAT software for additional data processing and visualization tasks.

4 Discussion

We develop a new regression model specifically designed for situations where both the response variable and the covariate exhibit extreme behavior. The model builds upon the Peaks-Over-Threshold (POT) method by selecting exceedances above a high threshold and transforming them into unit Fréchet margins. Under this framework, the joint distribution of exceedances is approximated by a bivariate extreme value distribution. This allows us to capture the joint extremal dependence structure, which is often overlooked in classical univariate approaches. Unlike previous studies that have primarily relied on the Generalized Extreme Value (GEV) distribution to model extremes (e.g., Carvalho et al. [13]), our contribution adopts the POT framework, which provides a more flexible and efficient use of available extreme data by focusing on exceedances rather than block maxima.

The construction of the model is inspired by quantile regression, in that it studies how the conditional behavior of an extreme response varies with an extreme covariate. However, it extends this idea by incorporating asymptotic results from extreme value theory (EVT) and defining a regression manifold — a family of regression lines that respect the constraints of bivariate EVT. This manifold provides a structured way to interpret how covariates influence the response in the tail region.

To flexibly model the spectral measure, which encodes the dependence structure between extremes, we introduce a Logistic-Normal prior on the space of spectral densities. This choice is particularly suitable for compositional data, where components are proportions summing to one, and allows for modeling log-ratio relationships and correlations. The parameter σ of the Logistic-Normal distribution is estimated via maximum likelihood. Although theoretical properties such as consistency and asymptotic normality of the estimator are not yet formally proven, they represent promising directions for future research.

Our numerical analyses demonstrate strong performance, including goodness-of-fit and independence tests such as the chi-square test. Instead, we emphasize that the model performs well within the scope of the datasets considered. Compared to standard alternatives like the multivariate Generalized Pareto Distribution (MGPD), our approach offers a more structured regression interpretation while still honoring tail dependence. Future comparisons will help quantify its advantages in terms of interpretability and predictive power.

In contrast to extremal quantile regression methods such as those in Chernozhukov [10], which estimate high quantiles of the response given general covariates, our model is tailored to scenarios where both variables are extreme. It explicitly accounts for their joint extremal structure, offering a more coherent framework for analyzing co-occurring extremes. This distinction is crucial in applications where the covariate itself is also subject to extreme behavior.

Finally, we acknowledge several limitations. The current framework is restricted to the bivariate case, and extending it to higher dimensions is non-trivial due to the complexity of high-dimensional spectral measures. Moreover, the assumption that exceedances follow the asymptotic form may not hold in small samples. Future work should address these limitations and explore applications in fields such as climate science, geology, and economics, where understanding joint extremes is essential for risk assessment and decision-making.

Algorithm 1 Extreme Quantile Regression with Logistic-Normal Prior

Input:

–Time series of temperature (covariate) and precipitation (response)
–Thresholds u_s and u_t for POT method
–Quantile level $\alpha \in (0,1)$

Steps:

1. **Preprocessing:**
   –Apply ARIMA to each series to ensure stationarity
   –Extract residuals and assess normality via Q-Q plots
2. **Extreme Value Selection:**
   –Use Mean Excess Plot to determine thresholds u_s and u_t
   –Retain top 10% values as extremes (validated by K-S test)
3. **Transformation to Unit Fréchet Margins:**
   –Estimate GPD parameters for each series
   –Apply transformations from equations 7 and 8
4. **Pseudo-Angle Construction:**
   –Compute $R_i = S_i + T_i$
   –Derive $W_i = (S_i, T_i)/R_i$ for exceedances
5. **Spectral Density Estimation:**
   –Model $h(w)$ using Logistic-Normal distribution:
   –Apply logistic transformation: $w = \frac{e^v}{1+e^v}$
   –Estimate parameter σ via MLE
   –Ensure moment constraint: $\int wh(w)dw = 1/2$
6. **Conditional Quantile Regression:**
   –Define conditional distribution $G_{T|S}(t|s)$ using equation 32
   –Solve $G_{T|S}(t|s) = \alpha$ to obtain $t_{\alpha|s}$
7. **Confidence Band Estimation:**
   –Apply bootstrap method to estimate uncertainty around $t_{\alpha|s}$

Output:

–Conditional quantile curves $t_{\alpha|s}$ for various α
–Estimated spectral density $h(w)$
–Bootstrap confidence intervals

Declarations

Competing interests: The authors declare that they have no competing interests.
Authors' contributions: A. El Bernoussi conceptualized the study, developed the methodology, performed the simulations, and wrote the manuscript. M. El Arrouchi contributed to the theoretical framework, supervised the research, and reviewed the final version of the manuscript. Both authors read and approved the final manuscript.
Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.
Availability of data and materials: The datasets analyzed in this study can be accessed from Meteociel and Infoclimat free websites': the period 1973-1979 of these data are downloaded from https://www.meteociel.fr/climatologie/obs_villes.php?code2=60150&mois=12&annee=1973 but the other period 1980-2024 is downloaded from https://www.infoclimat.fr/climatologie/annee/1980/meknes/valeurs/60150.html, which have authorized us to use them.
Acknowledgments: The authors would like to thank the editorial team and reviewers of the Jordan Journal of Mathematics and Statistics for their valuable comments and suggestions, which helped improve the quality and clarity of this manuscript. We also acknowledge the support and encouragement provided by our respective institutions throughout the development of this research.

References

[1] A.M. Stephenson, J.A. Tawn, Exploiting occurrence times in likelihood inference for componentwise maxima, Biometrika, **92** (2005), 213–227.
[2] A. AghaKouchak, N. Nasrollahi, Semi-parametric and Parametric Inference of Extreme Value Models for Rainfall Data, Water Resour. Manage., **24(6)** (2010), 1229–1249.DOI :https://doi.org/10.1007/s11269-009-9493-3
[3] J. Aitchison, S.M. Shen, Logistic-Normal Distributions: Some Properties and Uses, Biometrika, **67(2)** (1980), 261–272. DOI : https://doi.org/10.2307/2335470
[4] A.A. Balkema, L. de Haan, Residual life time at great age, Ann. Probab., **2** (1974), 792–804.
[5] J. Beirlant, Y. Goegebeur, J. Segers, J. Teugels, D. De Waal, C. Ferro, Statistics of Extremes: Theory and Applications, Wiley, New York, 2004.
[6] M. Beniston, D.B. Stephenson, O.B. Christensen, C.A.T. Ferro, C. Frei, S. Goyette, K. Halsnaes, T. Holt, K. Jylhä, B. Koffi, J. Palutikof, R. Schöll, T. Semmler, K. Woth, Future extreme events in European climate: an exploration of regional climate model projections, Climatic Change, **81(1)** (2007), 71–95. DOI : https://doi.org/10.1007/s10584-006-9226-z
[7] G.E.P. Box, G.M. Jenkins, Time Series Analysis: Forecasting and Control, Holden-Day, San Francisco, 1970.
[8] Z. Cai et al., Joint extremes in precipitation and infectious disease in the USA: A bivariate POT study, One Health, **17** (2023), 100636.
[9] L. Cheng, A. AghaKouchak, E. Gilleland, R.W. Katz, Non-stationary extreme value analysis in a changing climate, Climatic Change, **127(2)** (2014), 353–369. DOI : https://doi.org/10.1007/s10584-014-1254-5
[10] V. Chernozhukov, Extremal quantile regression, Ann. Statist., **33(2)** (2005), 806–839. DOI : https://doi.org/10.1214/009053604000001165
[11] S. Coles, Extremes of Non-stationary Sequences, in An Introduction to Statistical Modeling of Extreme Values, Springer, 2001, 105–123.
[12] A.C. Davison, R.L. Smith, Models for Exceedances Over High Thresholds, J. R. Stat. Soc. B, **52(3)** (1990), 393–425. DOI : https://doi.org/10.1111/j.2517-6161.1990.tb01796.x
[13] M. de Carvalho, A. Kumukova, G. dos Reis, Regression-type analysis for multivariate extreme values, Extremes, **25(4)** (2022), 595–622. DOI : https://doi.org/10.1007/s10687-022-00446-6
[14] L. de Haan, S.I. Resnick, Limit theory for multivariate sample extremes, Z. Wahrsch. Verw. Gebiete, **40(4)** (1977), 317–337. DOI : https://doi.org/10.1007/BF00533086
[15] S. El Adlouni et al., Generalized maximum likelihood estimators for the nonstationary generalized extreme value model, Water Resour. Res., **43(3)** (2007). DOI : https://doi.org/10.1029/2005WR004545
[16] A. El Bernoussi, M. El Arrouchi, A New Regression Model for Analyzing Non-Stationary Extremes in Response and Covariate Variables with an Application in Meteorology, Journal of Applied Probability and Statistics, **20(3)** (2025) (Accepted).
[17] R.A. Fisher, L.H.C. Tippett, Limiting forms of the frequency distribution of the largest or smallest member of a sample, Math. Proc. Camb. Phil. Soc., **24** (1928), 180–190.
[18] J. Gavin, Extreme value theory—an empirical analysis of equity risk, UBS Warburg Working Paper, 2000.
[19] B. Gnedenko, Sur la distribution limite du terme maximum d'une série aléatoire, Ann. Math., **44** (1943), 423–453.
[20] E.J. Gumbel, On the Frequency Distribution of Extreme Values in Meteorological Data, Bull. Amer. Meteorol. Soc., **23(3)** (1942), 95–105. DOI : https://doi.org/10.1175/1520-0477-23.3.95
[21] E.J. Gumbel, Statistics of Extremes, Columbia Univ. Press, New York, 1958.
[22] T.E. Hanson, M. de Carvalho, Y. Chen, Bernstein polynomial angular densities of multivariate extreme value distributions, Stat. Probab. Lett., **128** (2017), 60–66. DOI :https://doi.org/10.1016/j.spl.2017.03.030
[23] R.W. Katz, Statistical Methods for Nonstationary Extremes, in Extremes in a Changing Climate, Springer, Dordrecht, 2013, 15–37.
[24] R.W. Katz, M.B. Parlange, P. Naveau, Statistics of extremes in hydrology, Adv. Water Resour., **25(8)** (2002), 1287–1304. DOI : https://doi.org/10.1016/S0309-1708(02)00056-8
[25] V.V. Kharin, F.W. Zwiers, Estimating Extremes in Transient Climate Change Simulations, J. Climate, **18(8)** (2005), 1156–1173. DOI : https://doi.org/10.1175/JCLI3320.1
[26] R. Koenker, G. Bassett, Regression Quantiles, Econometrica, **46(1)** (1978), 33–50. DOI : https://doi.org/10.2307/1913643
[27] M.R. Leadbetter, Extremes and local dependence in stationary sequences, Z. Wahrsch. Verw. Gebiete, **65(2)** (1983), 291–306. DOI : https://doi.org/10.1007/BF00532484
[28] Y. Li, W. Cai, E.P. Campbell, Statistical Modeling of Extreme Rainfall in Southwest Western Australia, J. Climate, **18(6)** (2005), 852–863. DOI : https://doi.org/10.1175/JCLI-3296.1
[29] F. Longin, Value at Risk: Une nouvelle approche fondée sur les valeurs extrêmes, Ann. Écon. Stat., **51** (1998), 23–51.
[30] F. Longin, From value at risk to stress testing: The extreme value approach, Working Paper, 2000.
[31] L. Mentaschi et al., The transformed-stationary approach: A generic and simplified methodology for non-stationary extreme value analysis, Hydrol. Earth Syst. Sci., **20(9)** (2016), 3527–3547. DOI : https://doi.org/10.5194/hess-20-3527-2016
[32] S.N. Neftci, Value at Risk Calculations, Extreme Events, and Tail Estimation, J. Derivatives, Spring (2000), 1–15.
[33] J. Pickands, Multivariate extreme value distributions, Proc. 43rd Sess. Int. Statist. Inst., 1981, 859–878.
[34] R.L. Smith, Estimating Tails of Probability Distributions, Ann. Statist., **15(3)** (1987), 1174–1207. DOI : http://dx.doi.org/10.1214/aos/1176350499
[35] R.L. Smith, Extreme value statistics in meteorology and the environment, Environ. Stat., **8** (2001), 300–357.

[36] E. Towler et al., Modeling hydrologic and water quality extremes in a changing climate: A statistical approach based on extreme value theory, Water Resour. Res., **46(11)** (2010). DOI : https://doi.org/10.1029/2009WR008876

[37] G. Villarini et al., Flood frequency analysis for nonstationary annual peak records in an urban drainage basin, Adv. Water Resour., **32(8)** (2009), 1255–1266. DOI : http://dx.doi.org/10.1016/j.advwatres.2009.05.003

[38] G. Villarini et al., Annual maximum and peaks-over-threshold analyses of daily rainfall accumulations for Austria, J. Geophys. Res., **116(D5)** (2011). DOI : https://doi.org/10.1029/2010JD015038

[39] X. Zhang, W.D. Hogg, É. Mekis, Spatial and temporal characteristics of heavy precipitation events over Canada, J. Climate, **14(9)** (2001), 1923–1936. DOI : http://dx.doi.org/10.1175/1520-0442(2001)014%3C1923:SATCOH%3E2.0.CO;2